



Best Practices for Configuring DCB with VMware ESXi 5.1 and Dell EqualLogic Storage

A Dell Reference Architecture

Dell Storage Engineering
October 2013

This document has been archived and will no longer be maintained or updated. For more information go to the [Storage Solutions Technical Documents page on Dell TechCenter](#) or contact support.

Revisions

Date	Description
September 2013	Initial release
October 2013	Added information around NIC VLAN configuration

© 2013 Dell Inc.

Trademarks used in this text:

Dell™, the Dell logo, PowerEdge™, EqualLogic™, and Force10™ are trademarks of Dell Inc. Other Dell trademarks may be used in this document. Microsoft® and Windows® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. VMware®, vMotion®, vCenter® and vSphere® are registered trademarks or trademarks of VMware, Inc. in the United States or other countries. Broadcom® and NetXtreme® are registered trademarks of Broadcom Corporation. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and/or names or their products and are the property of their respective owners. Dell disclaims proprietary interest in the marks and names of others.



Table of contents

Revisions.....	2
Acknowledgements.....	5
Feedback	5
Executive summary	6
1 Introduction.....	7
1.1 Objectives.....	7
1.2 Audience.....	7
1.3 Terminology.....	7
2 Solution components overview	9
2.1 Data Center Bridging overview.....	9
2.1.1 Data Center Bridging standards	9
2.1.2 Data Center Bridging benefits	10
2.2 Server virtualization overview	10
2.2.1 vSphere vCenter.....	10
2.2.2 vSphere Client.....	10
2.2.3 vSphere vMA / vCLI.....	11
2.2.4 ESXi Hypervisor 5.1.....	11
2.3 NIC Partitioning (NPAR) overview	11
2.3.1 iSCSI initiator types	12
2.3.2 EqualLogic MEM.....	13
3 Test configuration	14
3.1 Network configuration.....	14
3.1.1 Switch DCB configuration.....	15
3.1.2 Switch interconnect and VLAN configuration	17
3.1.3 Verifying switch DCB configuration	18
3.2 Storage configuration	21
3.2.1 Data Center Bridging configuration	21
3.2.2 Volume access configuration	22
3.3 ESXi Host configuration	23
3.3.1 Configure NPAR in Dell System Setup	23
3.3.2 Server virtualization configuration.....	27



3.3.3 EqualLogic MEM installation and configuration	29
4 Test results and analysis	32
4.1 Ethernet retransmissions	32
4.2 Response time	32
5 Best practice recommendations.....	34
6 Conclusion.....	35
A Solution configuration details	36
B Workload characterizations	37
C Additional resources.....	38



Acknowledgements

This best practice white paper was produced by the following members of the Dell Storage team:

Engineering: Gregory Brent Seaton

Technical Marketing: Guy Westbrook

Editing: Margaret Boeneke

Additional contributors:

Michael Kosacek

Feedback

We encourage readers of this publication to provide feedback on the quality and usefulness of this information by sending an email to SIFeedback@Dell.com.



SIFeedback@Dell.com



Executive summary

Server virtualization is an important strategy for organizations seeking to reduce the cost and overhead of managing ever expanding data centers. VMware® vSphere® virtualization combined with the flexible and scalable iSCSI storage provided by Dell™ EqualLogic™ PS Series arrays create an environment for a dynamic and efficient data center.

This paper provides a reference architecture outlining the best practices for implementing end-to-end datacenter bridging (DCB) using VMware ESXi 5.1 hypervisor on Dell PowerEdge™ R620 servers, Dell Networking S4810 switches, and Dell EqualLogic iSCSI storage. This reference architecture showcases three PS6110 arrays servicing six virtual machines with four iSCSI volumes each, for a total of 24 iSCSI volumes under continuous load.



1 Introduction

Organizations can find significant cost savings, improved operations flexibility, and ease of implementation with server virtualization systems such as VMware vSphere. These benefits, while great, are still limited by the requirement to deploy and support separate and dedicated LAN and SAN networks; however with the introduction of Data Center Bridging (DCB), these benefits can be more fully realized when implemented on an infrastructure just as capable and flexible. An infrastructure supporting DCB provides the ability to converge LAN and SAN traffic into a single infrastructure in a controlled manner. The reference architecture laid out in this paper provides an end-to-end solution with DCB capable Compute, Network, and Storage resources such as the Dell PowerEdge R620 with Broadcom 57810, Dell Networking S4810, and Dell EqualLogic PS series storage.

1.1 Objectives

The primary objectives of the tests conducted for this paper were:

- Develop a reference architecture and a set of best practices for a server virtualization solution using VMware vSphere Enterprise and connected to a DCB enabled converged network supporting the communications for both the LAN and SAN infrastructure
- Develop step by step guide for deployment of the reference architecture

The test infrastructure used for the tests included:

- VMware vSphere 5.1 hypervisor
- Dell PowerEdge R620 Rack Servers with Broadcom® 57810 Converged Network Adapters (CNA)
- Dell Networking S4810 switches
- Dell EqualLogic PS6110 iSCSI storage arrays

1.2 Audience

This paper is intended for solution architects, storage network engineers, system administrators, and IT managers who need to understand how to design and implement VMware vSphere virtualization technology with DCB and Dell EqualLogic iSCSI Storage solutions. It is expected that the reader has a working knowledge of basic layer 2 networking, basic VMware ESXi configuration, and Dell EqualLogic iSCSI SAN operation.

1.3 Terminology

The following list defines key terms used throughout this paper.

Converged Network Adapter (CNA): An adapter that supports more than one type of networking function that has traditionally been separated.

Data Center Bridging (DCB): DCB is a set of IEEE standards that help ensure performance and delivery of Ethernet traffic throughout the network topology. These standards are:



- **Congestion Notification (CN; IEEE 802.1Qau):** Enables DCB switches to identify primary bottlenecks and take preventative action to ensure that these primary points of congestion do not spread to other parts of the network infrastructure.
- **Datacenter Bridging Capability Exchange (DCBx):** The glue that binds all of the standards by allowing networking components to understand and communicate the settings required to operate within the DCB environment.
- **Enhanced Transmission Selection (ETS; IEEE 802.1Qaz):** Provides administrators the ability to group multiple classes of service together and then define a guaranteed minimum bandwidth allocation from the shared network connection.
- **Priority-based Flow Control (PFC; IEEE 802.1Qbb):** Expands the function of the standard class of service structure of Ethernet to provide a mechanism to allow for lossless classes of service since a non-lossless class cannot be paused.

Host Bus Adapter (HBA): Dedicated adapter to connect a host system to a storage network.

iSCSI Qualified Name (IQN): Provides a unique identification string for storage targets and initiators documented in RFC 3720 section 3.2.6.

Link Layer Discovery Protocol (LLDP): IEEE standard 802.1AB, used for informational exchange about identity and abilities between networking devices.

Multipath Extension Module (MEM): Dell EqualLogic specific software for iSCSI hosts, provides optimization for EqualLogic Arrays to the default host MPIO scheme.

NIC Partitioning (NPAR): Ability to separate, or partition, one physical adapter port into multiple simulated adapter port instances within the host system.

VMKernel Port (VMK): Port created and used by the host hypervisor, not visible to virtual machines.

vSphere Command Line Interface (vCLI): Installation package provided from VMware for Linux® and/or Microsoft® Windows® providing command line interface management access.

vSphere Management Assistant (vMA): Prepackaged virtual machine from VMware providing command line interface management access.



2 Solution components overview

The following section presents an overview of the major technology used to develop the reference architecture shown in this paper.

2.1 Data Center Bridging overview

Many enterprises run dedicated networks to separate and protect iSCSI traffic from the unpredictable nature of regular LAN, but with the introduction of DCB, both of these traffic types can be converged onto the same Ethernet infrastructure in a controlled manner. Many detailed white papers about how and why DCB provides benefits to the Ethernet stack are available in Appendix C, Additional resources. This section provides a brief overview of the DCB standards as well as the benefits of DCB to iSCSI storage traffic.

2.1.1 Data Center Bridging standards

DCB is made up of several IEEE standards, and each of these standards works together to provide a lossless Ethernet environment.

Priority-based Flow Control (PFC) - Priority-based Flow Control (PFC) is an enhancement of the original Ethernet Flow Control concept that was implemented with MAC PAUSE. The MAC PAUSE feature allows for the receiver to request the sender to stop (or pause) sending Ethernet frames, this request applied to all traffic regardless of traffic classes or priorities. PFC works in a similar, but improved, fashion. PFC uses classes of traffic to apply more control to the process. For example, PFC can request a single class of traffic be paused for a specified time period and allow other classes to continue.

Enhanced Transmission Selection (ETS) - Enhanced Transmission Selection (ETS) provides the perfect framework for PFC to operate within. ETS provides a means of allocating bandwidth to a specified traffic class, and a traffic class can contain one or more Class of Services. Each traffic class is allocated a portion of bandwidth defined as a percentage; all allocations must add up to 100%. This allocated portion of bandwidth is not a maximum allocation, because traffic classes may use more than allocated if there is unutilized bandwidth available. Instead, the allocation represents the minimum amount of bandwidth available to the traffic class.

Congestion Notification (CN) – Congestion Notification (CN) is not currently implemented widely. It aims to provide a mechanism for managing congestion throughout a DCB fabric. Due to the slow adoption and implementation, CN is not discussed in this paper.

Data Center Bridging Exchange (DCBx) – DCBx is an extension to the familiar LLDP IEEE standard. DCBx uses the LLDP framework to advertise capabilities of the devices through Type-Length-Values (TLVs) frames. These TLVs include a PFC TLV, ETS TLV, and an iSCSI Application TLV.

iSCSI Application TLV – The iSCSI Application TLV is required for DCB operation in EqualLogic storage environment.

Note: Not all devices claiming DCB support will support all of the DCB standards. Support of each DCB standard needs to be verified.



2.1.2 Data Center Bridging benefits

One of the biggest benefits of using DCB is the increase in network control that it gives to the network architect when distributing Network bandwidth and in managing its use. It is now possible to take a 10Gb, 40Gb, or larger Ethernet pipe and allocate it to resources as needed. No longer is it required to dedicate separate hardware for LAN and separate hardware for SAN.

A second and equally important benefit to DCB is the introduction of lossless Ethernet. Ethernet was originally designed without any form of flow control. Management of that was left to the upper layers of the OSI stack, such as TCP. However, as the demands on the network have grown, so have the expectations. It is now important to make the network as efficient as possible, and one way of doing this is with Lossless Ethernet. DCB, with its PFC mechanism, is able to reduce network retransmissions to close to zero percent, thereby eliminating wasted network bandwidth on retransmitted data.

2.2 Server virtualization overview

Server virtualization can be accomplished many different ways. In this reference architecture, VMware vSphere® is used to provide the server virtualization piece.

2.2.1 vSphere vCenter

VMware vSphere vCenter® provides centralized management of the server virtualization environment while greatly extending the platform abilities. If managing more than one ESXi host, it is highly recommended to use vSphere vCenter as the single point of management.

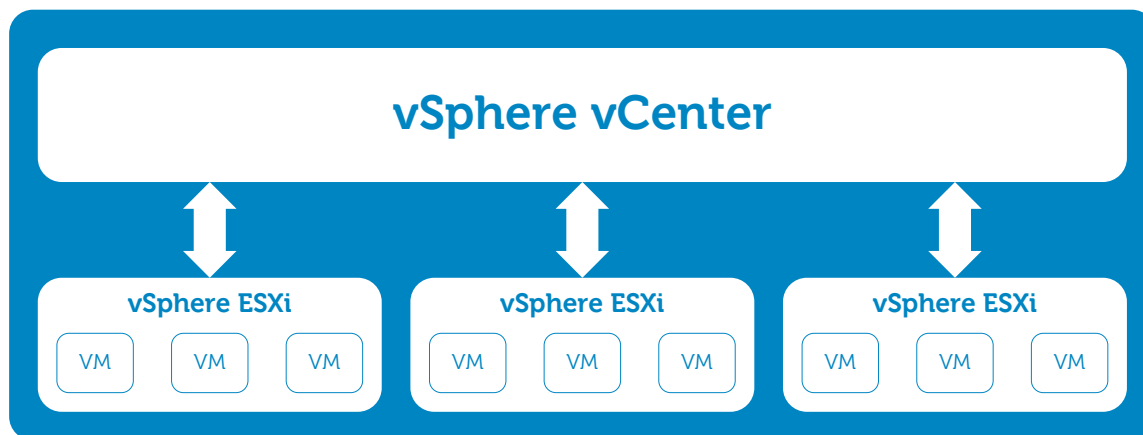


Figure 1 vSphere centralized management

2.2.2 vSphere Client

The vSphere Client provides a GUI to manage the vSphere environment that can connect to individual ESXi hosts or to a vCenter server to provide centralized management.

2.2.3 vSphere vMA / vCLI

When command line access is needed, VMware provides two options to manage the environment: vSphere Management Assistant (vMA) and vSphere Command Line Interface (vCLI).

The vCLI is used in this paper during the installation of the EqualLogic Multipath Extension Module (MEM) and during the creation of custom named vSwitches.

The more information and downloads for vMA and vCLI are available here:

- VMware vSphere Management Assistant documentation and download
<http://www.vmware.com/support/developer/vima/>
- VMware vSphere Command Line Interface documentation and download
<http://www.vmware.com/support/developer/vcli/>

2.2.4 ESXi Hypervisor 5.1

The VMware vSphere hypervisor layer is provided by VMware ESXi. ESXi provides a small footprint bare-metal hypervisor.

VMware currently does not natively support DCB for iSCSI. NIC Partitioning (described in section 2.3) provides a complete solution by presenting virtual network adapters to ESXi.

2.3 NIC Partitioning (NPAR) overview

NIC Partitioning begins by configuring a single 10 Gb port to represent up to four separate functions or partitions. Each partition appears to the virtualization hypervisor (or operating system) as a discrete NIC with its own driver software functioning independently. Dedicating one of the independent partitions to iSCSI traffic allows the Broadcom CNA to manage the DCB parameters and provide lossless iSCSI traffic without VMware support. In this white paper, one of the partitions is dedicated from each physical port, Function 0 and Function 1, as shown in the graphic below.

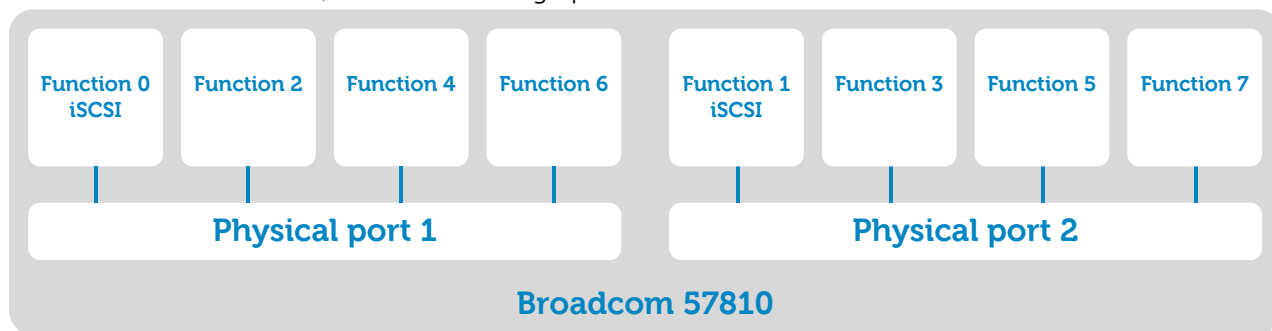


Figure 2 NIC Partition with Broadcom 57810

2.3.1 iSCSI initiator types

There are three categories of iSCSI initiators supported by VMware vSphere, including the built in software initiator, Independent, and Dependent Hardware Initiators, as shown in the figure below. This section briefly discusses each type and its suitability in this reference architecture.

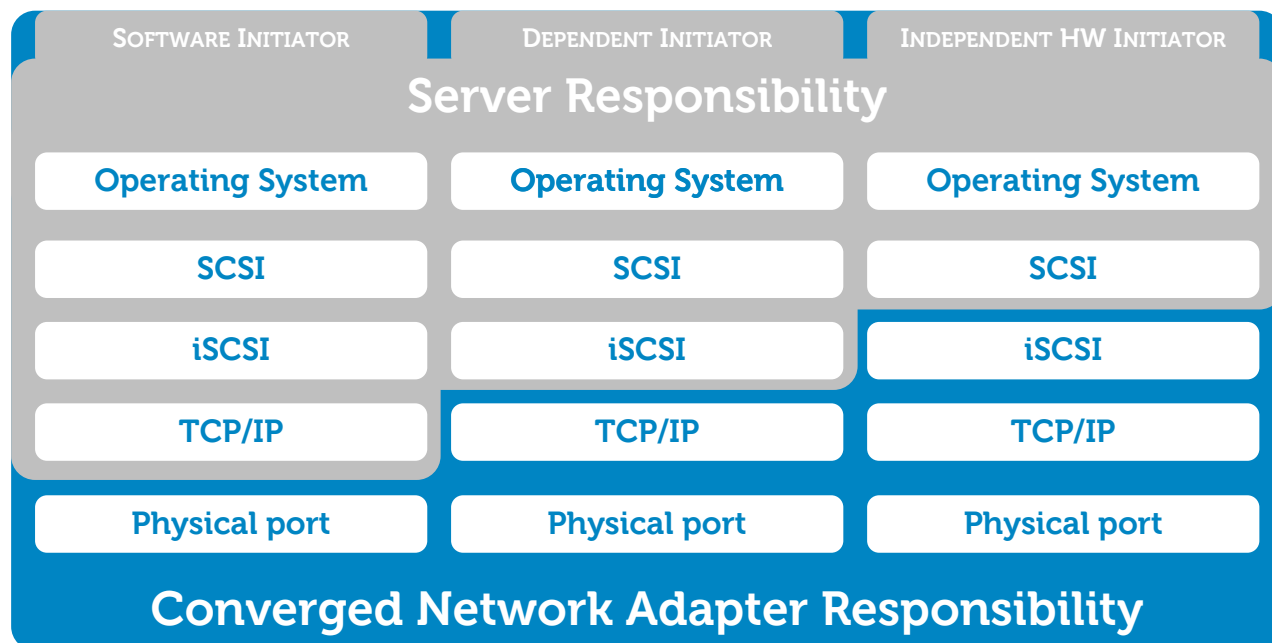


Figure 3 iSCSI initiator types

Software iSCSI Initiator

VMware vSphere provides, in the hypervisor, a built-in software iSCSI initiator capable of using any Ethernet adapter with IP connectivity to the iSCSI storage for iSCSI connectivity. However, VMware vSphere does not support DCB for iSCSI connectivity making this option unsuitable for an end-to-end DCB architecture.

Independent Hardware iSCSI Adapter

There are various manufacturers who make Independent Hardware iSCSI initiators. These adapters implement their own networking and iSCSI stacks and present to the hypervisor in the same way as a dedicated HBA. This adapter type is not supported by Dell EqualLogic MEM. They can be used with other MPIO policies (such as round robin), however, this may not provide optimal performance compared to a system with MEM installed and configured.

Dependent Hardware iSCSI Adapter

The Dependent Hardware iSCSI Adapter provides a hybrid approach. This adapter type depends on the vSphere networking and iSCSI configuration but provides iSCSI offload capabilities with supplemental networking support such as DCB. This form of adapter is also supported with Dell EqualLogic MEM. The Broadcom 57810, used in this architecture, is an example of this type of adapter.

2.3.2 EqualLogic MEM

The Dell EqualLogic MEM provides an optimized VMware Path Selection Policy for Dell EqualLogic storage. This provides an advantage over normal VMware PSP options that provide basic connectivity. The EqualLogic MEM has been developed with the unique architecture of EqualLogic Storage in mind to provide optimized path selection, iSCSI connection management, and load balancing.

As an additional benefit, EqualLogic MEM provides a configuration script that walks through the setup process step-by-step in interactive mode or can provide single command installation in unattended mode. Here is what the setup script configures (with provided input):

- Create iSCSI dedicated vSwitch
- Set vSwitch MTU to 9000
- Create iSCSI Portgroups
- Assign iSCSI IP Addresses
- Add vmnic uplinks to vSwitch
- Bind vmnics to vSwitch uplinks correctly
- Run host storage system refresh
- Add Group IP address to storage adapters
- Rescan all HBAs

Configuration is covered in more detail in section 3.3.3 EqualLogic MEM installation and .



3 Test configuration

The following sections have detailed configuration steps for each of the Network, Storage, and Compute resources.

3.1 Network configuration

This section covers configuring the Dell Networking S4810 for the reference architecture described in this white paper. Notice in the

Network flow diagram, 0, how traffic moves from the Virtual machine to the iSCSI storage below.



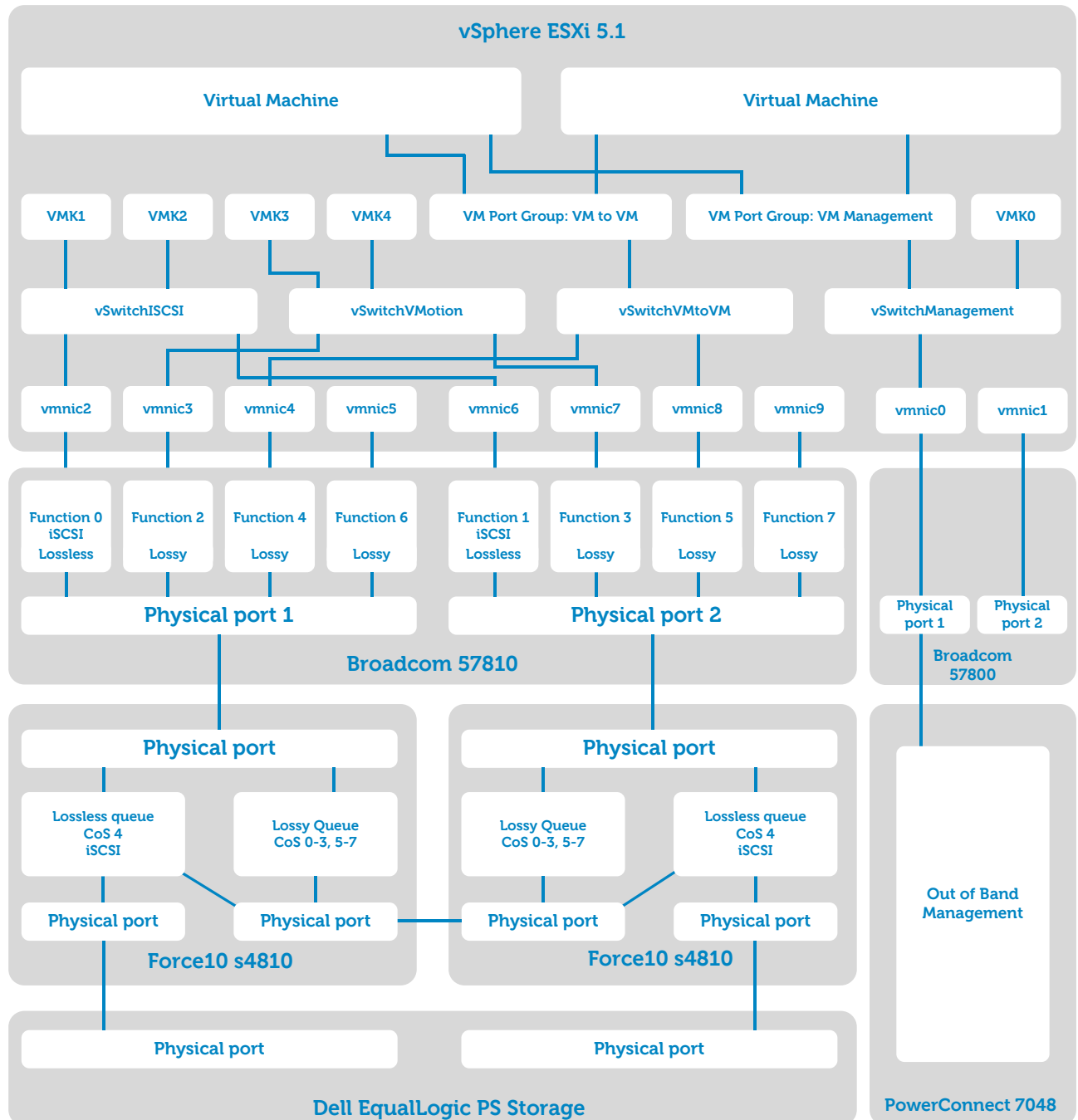


Figure 4 Network flow diagram

3.1.1 Switch DCB configuration

This section shows how to disable traditional flowcontrol, enable DCB (requires switch reload), and configure the initial requirements for DCB including optimizing switch buffering into two queues to accommodate the lossy and lossless traffic classes.

```

sw122(conf)#interface range tengigabitethernet 0/0 - 47
sw122(conf-if-range-te-0/0-47)#no flowcontrol rx on tx off
sw122(conf-if-range-te-0/0-47)#mtu 12000
sw122(conf)#interface range fortyGigE 0/48 - 60
sw122(conf-if-range-fo-0/48-52)#no flowcontrol rx on tx off
sw122(conf-if-range-fo-0/48-52)#mtu 12000
sw122(conf)#dcb-input pfc
sw122(conf-dcb-in)#pfc priority 4
sw122(conf-dcb-in)#exit
sw122(conf)#priority-group iSCSI
sw122(conf-pg)#priority-list 4
sw122(conf-pg)#set-pgid 1
sw122(conf-pg)#exit
sw122(conf)#priority-group OTHER
sw122(conf-pg)#priority-list 0-3,5-7
sw122(conf-pg)#set-pgid 2
sw122(conf-pg)#exit
sw122(conf)#dcb-output ets
sw122(conf-dcb-out)#priority-group iSCSI qos-policy iSCSI
sw122(conf-dcb-out)#priority-group OTHER qos-policy OTHER
sw122(conf-dcb-out)#exit
sw122 (conf)#service-class dynamic dot1p
sw122(conf)#qos-policy-output iSCSI ets
sw122(conf-qos-policy-out-ets)#bandwidth-percentage 50
sw122(conf-qos-policy-out-ets)#exit
sw122(conf)#qos-policy-output OTHER ets
sw122(conf-qos-policy-out-ets)#bandwidth-percentage 50

```




```
sw122(conf-qos-policy-out-ets)#exit

sw122(conf)#dcb enable

sw122(conf)#dcb stack-unit all pfc-buffering pfc-ports 56 pfc-queues 2
```

Note: For the complete configuration guide and most up to date configuration steps please refer to the Force10™ s4810 Switch Configuration Guide at <http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20220824/download.aspx>

3.1.2 Switch interconnect and VLAN configuration

This section shows how to create the switch to switch interconnect using two of the 40GbE ports provided on the Dell Networking S4810. It also shows how to create VLAN 100, 200, and 400 used for iSCSI, VM to VM, and vMotion® traffic respectively as shown in 0

Network flow diagram.

```
sw122(conf)#interface Port-channel 1

sw122(conf-if-po-1)# no ip address

sw122(conf-if-po-1)# mtu 12000

sw122(conf-if-po-1)# switchport

sw122(conf-if-po-1)# channel-member fortyGigE 0/48,52

sw122(conf-if-po-1)# no shutdown

sw122(conf-if-po-1)#exit

sw122(conf)#interface Vlan 100

sw122(conf-if-vl-100)# name iSCSI

sw122(conf-if-vl-100)# no ip address

sw122(conf-if-vl-100)# mtu 12000

sw122(conf-if-vl-100)# tagged TenGigabitEthernet 0/0-7

sw122(conf-if-vl-100)# tagged Port-channel 1

sw122(conf-if-vl-100)# no shutdown

sw122(conf-if-vl-100)#interface Vlan 200

sw122(conf-if-vl-200)# name vm2vm

sw122(conf-if-vl-200)# no ip address
```



```

sw122(conf-if-vl-200)# mtu 12000

sw122(conf-if-vl-200)# tagged TenGigabitEthernet 0/0-3

sw122(conf-if-vl-200)# tagged Port-channel 1

sw122(conf-if-vl-200)# no shutdown

sw122(conf-if-vl-200)#interface Vlan 400

sw122(conf-if-vl-400)# name vMotion

sw122(conf-if-vl-400)# no ip address

sw122(conf-if-vl-400)# mtu 12000

sw122(conf-if-vl-400)# tagged TenGigabitEthernet 0/0-3

sw122(conf-if-vl-400)# tagged Port-channel 1

sw122(conf-if-vl-400)# no shutdown

sw122(conf-if-vl-400)#exit

```

3.1.3 Verifying switch DCB configuration

There are numerous commands available on the s4810 to verify configuration. The most basic and most obvious is `show running-config`. In addition, here are some DCB show commands to see more details.

`show interfaces dcbx` - This command produces a lot of output, so look for "DCBX Operational Status is Enabled". Note however that only ports connected to EqualLogic Active controller ports will display "Enabled", and ports connected to EqualLogic Passive Controller ports will display "Disabled" until becoming active.

`show interfaces pfc summary` - This command also produces a large amount of output, and several things to note have been highlighted in bold below.

```

Interface TenGigabitEthernet 0/0

    Admin mode is on

    Admin is enabled, Priority list is 4

    Remote is enabled, Priority list is 4

    Remote Willing Status is enabled

    Local is enabled, Priority list is 4

    Oper status is init

```



PFC DCBX Oper status is Up

State Machine Type is Feature

TLV Tx Status is enabled

PFC Link Delay 45556 pause quantams

Application Priority TLV Parameters :

FCOE TLV Tx Status is disabled

ISCSI TLV Tx Status is enabled

Local FCOE PriorityMap is 0x0

Local ISCSI PriorityMap is 0x10

Remote ISCSI PriorityMap is 0x10

show interfaces ets summary – This command shows the traffic class groups, the priorities mapped to them, and the ETS settings allocated to each as well as confirming ETS DCBx is up and running. Note here that both the local switch port DCB settings are shown as well as what the switch has received from the remote port as its configured DCB settings, remote and local information should match.

Interface TenGigabitEthernet 0/0

Max Supported TC Groups is 4

Number of Traffic Classes is 8

Admin mode is on

Admin Parameters :

Admin is enabled

TC-grp	Priority#	Bandwidth	TSA

0		-	-
1	4	50 %	ETS
2	0,1,2,3,5,6,7	50 %	ETS
3		-	-
4		-	-
5		-	-
6		-	-



7	-	-
15	-	-

Remote Parameters :

Remote is enabled.

TC-grp	Priority#	Bandwidth	TSA

0		-	-
1	4	50 %	ETS
2	0,1,2,3,5,6,7	50 %	ETS
3		-	-
4		-	-
5		-	-
6		-	-
7		-	-
15		-	-

Remote Willing Status is enabled

Local Parameters :

Local is enabled

TC-grp	Priority#	Bandwidth	TSA

0		-	-
1	4	50 %	ETS
2	0,1,2,3,5,6,7	50 %	ETS
3		-	-
4		-	-
5		-	-
6		-	-
7		-	-
15		-	-

Oper status is init

ETS DCBX Oper status is Up



State Machine Type is Feature

Conf TLV Tx Status is enabled

3.2 Storage configuration

This reference architecture assumes the reader is familiar with Dell EqualLogic storage configuration of volumes, so this section covers only the unique configuration of DCB and multiple IQN access to a single volume.

3.2.1 Data Center Bridging configuration

Since firmware version 5.1.1, Dell EqualLogic supports DCB. The figure below shows the GUI for configuring EqualLogic to support DCB.

1. In Group manager, select **Group Configuration** on the left.
2. Under the Advanced tab, **Enable DCB** should be checked by default.
3. Change the **VLAN ID** to match the iSCSI VLAN defined on the switch. In this example we set this parameter to VLAN 100.
4. Save the setting by clicking the disk icon on the top right. DCB is now configured.

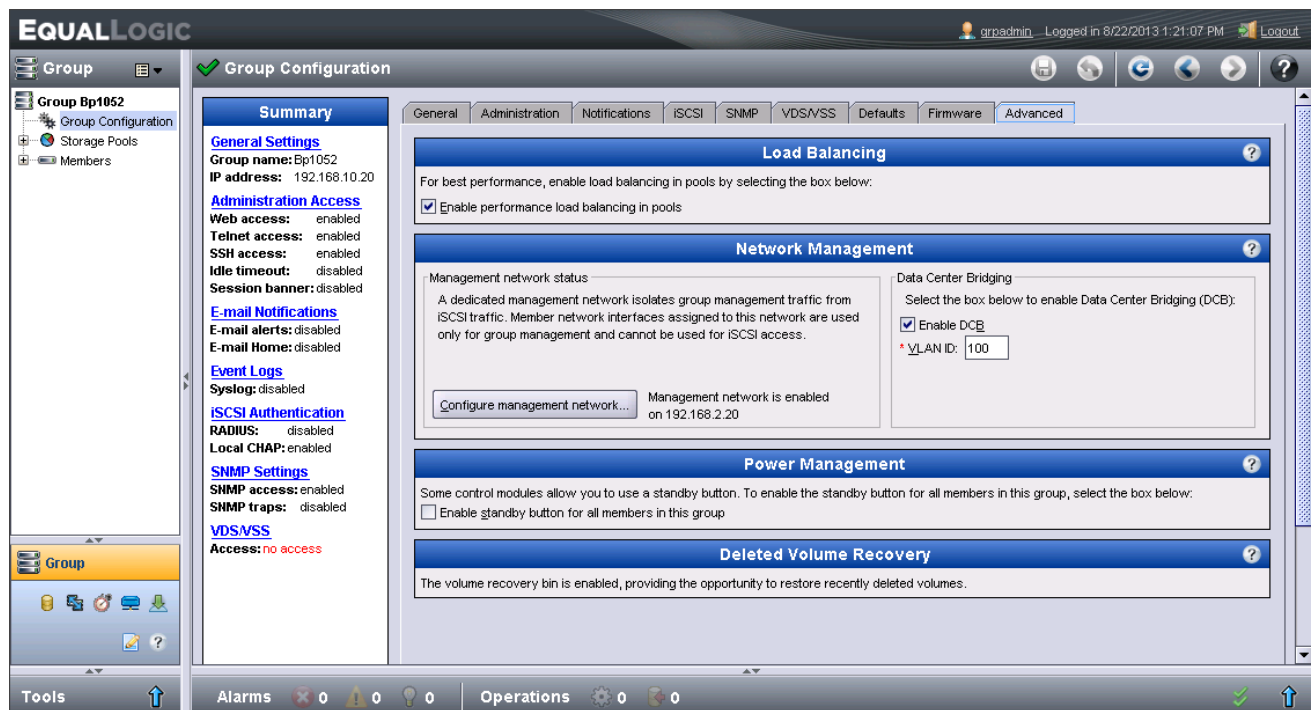


Figure 5 EqualLogic DCB VLAN ID

Dell EqualLogic iSCSI arrays are always in a DCB willing state to accept DCBX parameters from the switch, there is no way to manually change the DCB settings in the array.

3.2.2 Volume access configuration

By default EqualLogic restricts access to each volume to a single IQN. This behavior can be modified on a per volume basis. To enable multiple initiators to access a single volume, select a volume from the list of volumes and click **Modify settings** under the Activities column as shown in the figure below.

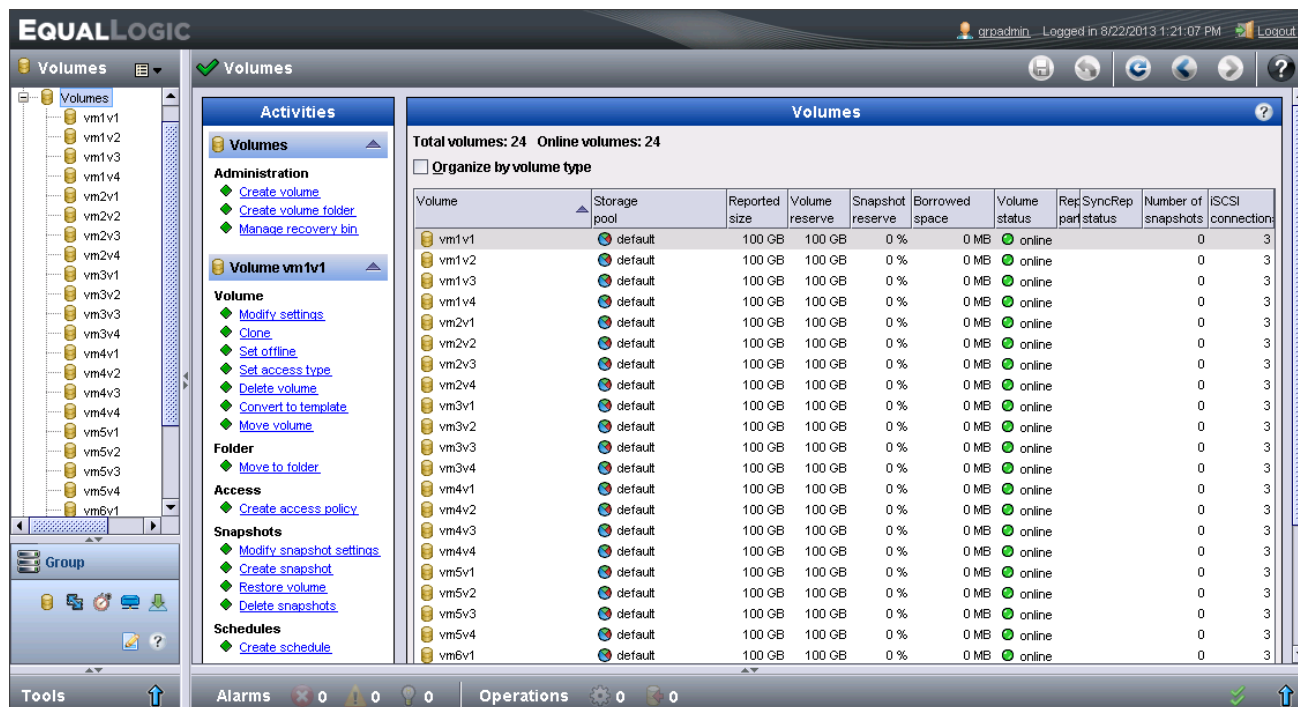


Figure 6 EqualLogic Volume - Modify settings option

In the Modify volume settings window, select the **Advanced** tab, then select **Allow simultaneous connections from initiators with different IQNs** as shown in the figure below. This is required to allow multiple ESXi hosts, so that multiple IQNs can simultaneously access the same volumes.

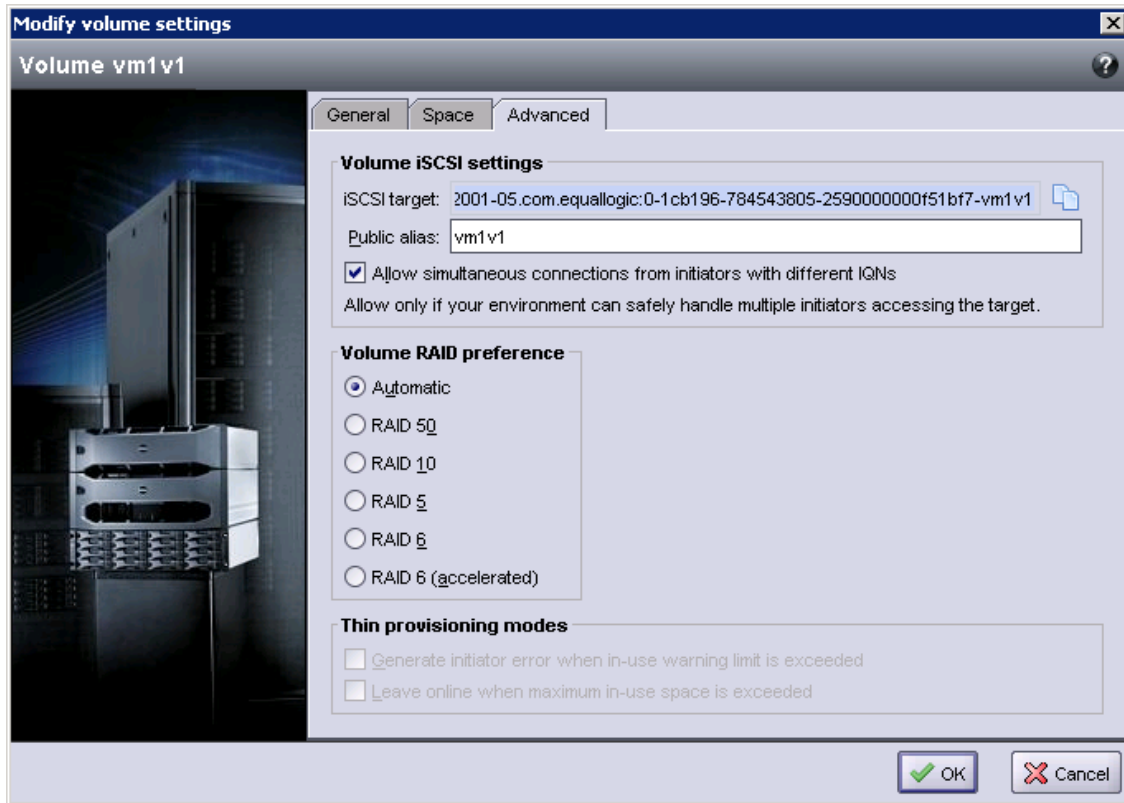


Figure 7 Modify volume settings – Advanced tab

3.3 ESXi Host configuration

This section shows the steps for configuring NPAR, iSCSI offload, and the EqualLogic MEM.

3.3.1 Configure NPAR in Dell System Setup

The key to using Broadcom's NPAR technology to provide support for lossless Ethernet for iSCSI traffic is enabling iSCSI Offload Engine on the Broadcom partition used for iSCSI traffic. This can be done through the Dell System Setup utility on Dell 12G servers, such as the R620 shown here. To access the utility, power on (or reboot if currently powered on) the server and press F2 during the boot process. This opens the System setup main menu. Select **Device Settings**. This menu presents a list of devices that can be configured by the system setup utility. For this example, select Port 1 Broadcom NetXtreme® II 10Gb Ethernet BCM57810 (this will need to be repeated for Port 2).

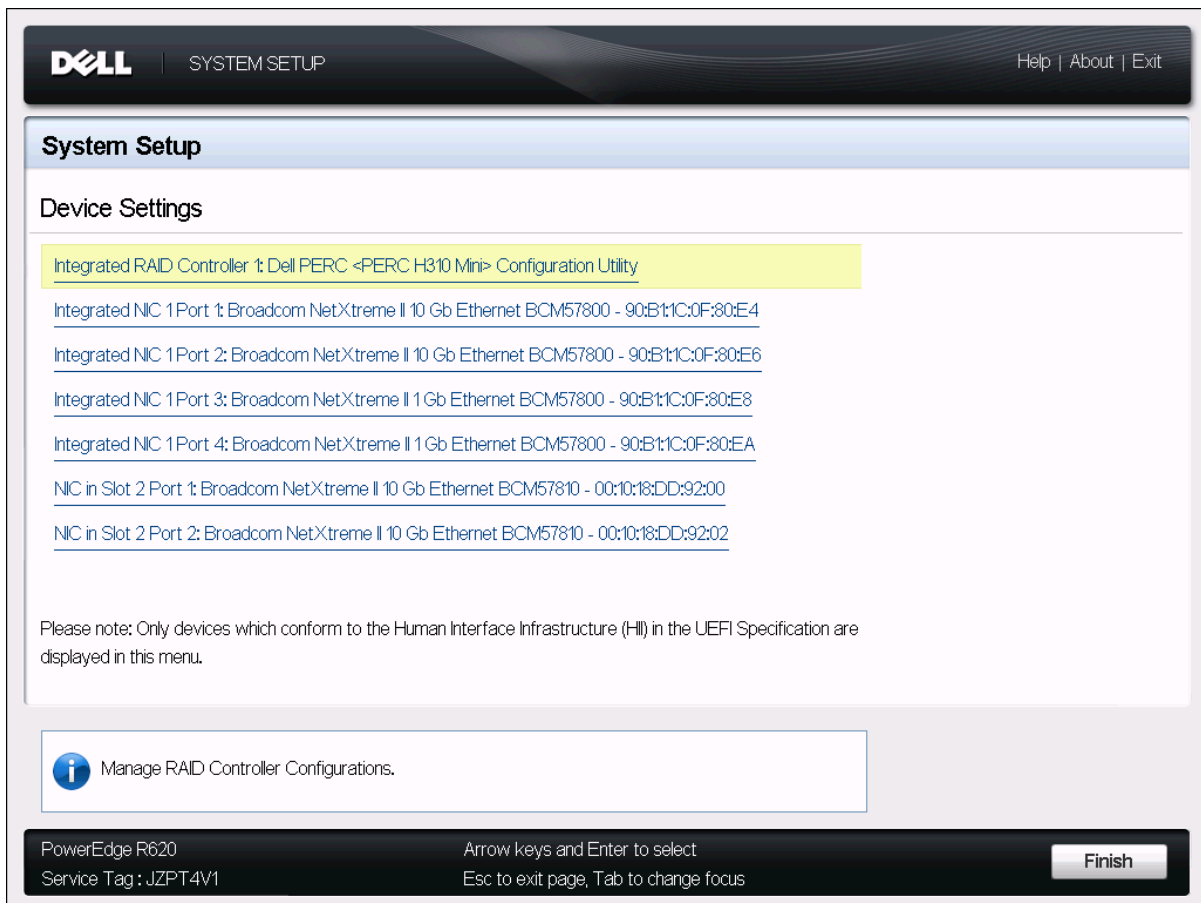


Figure 8 Dell System Setup – Device Settings Menu

Select the **Device Configuration Menu** and ensure DCB protocol is enabled as shown in Figure 9. Select **Back**.

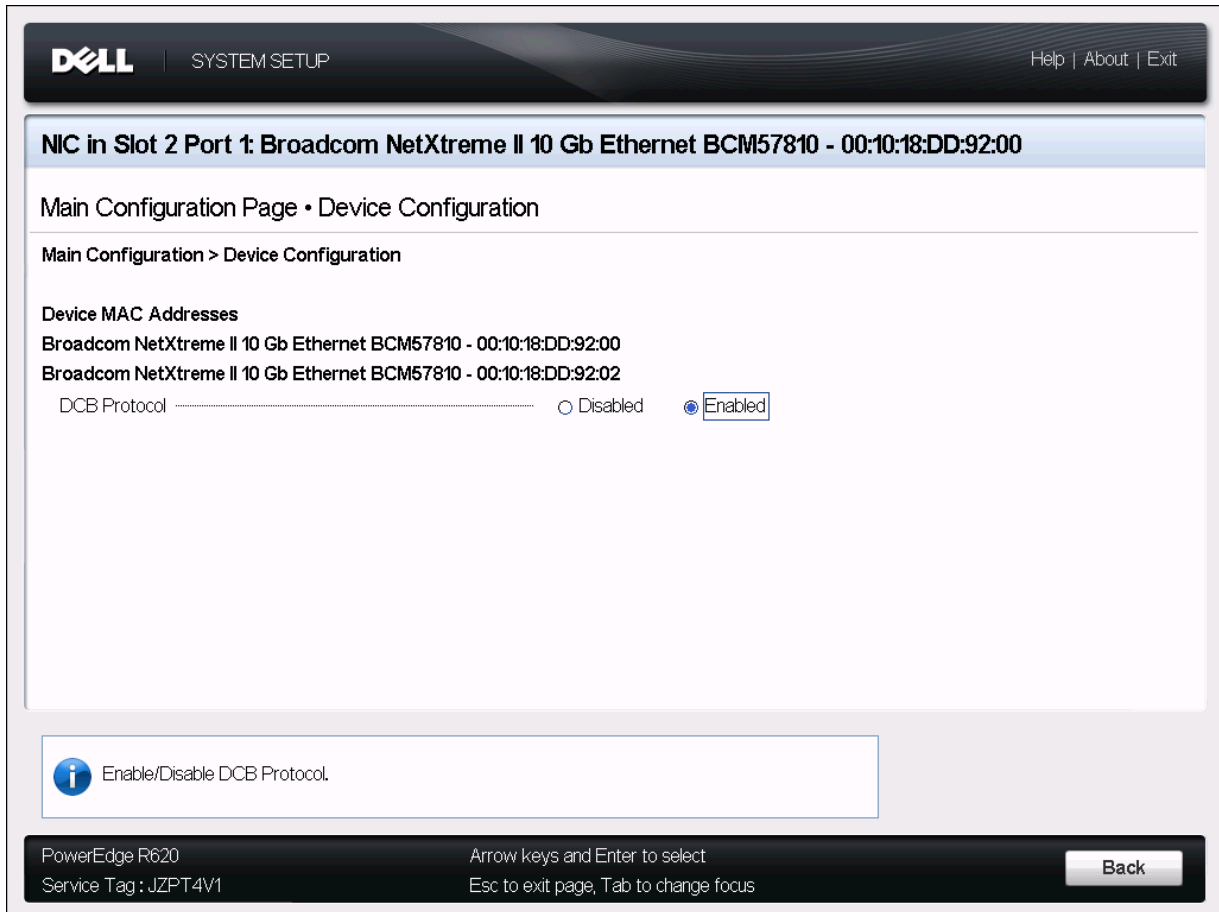


Figure 9 Dell System Setup – Device Configuration – DCB Enabled

As shown in Figure 10, change virtualization mode to NPAR, which presents a new menu item named NIC Partitioning Configuration Menu.

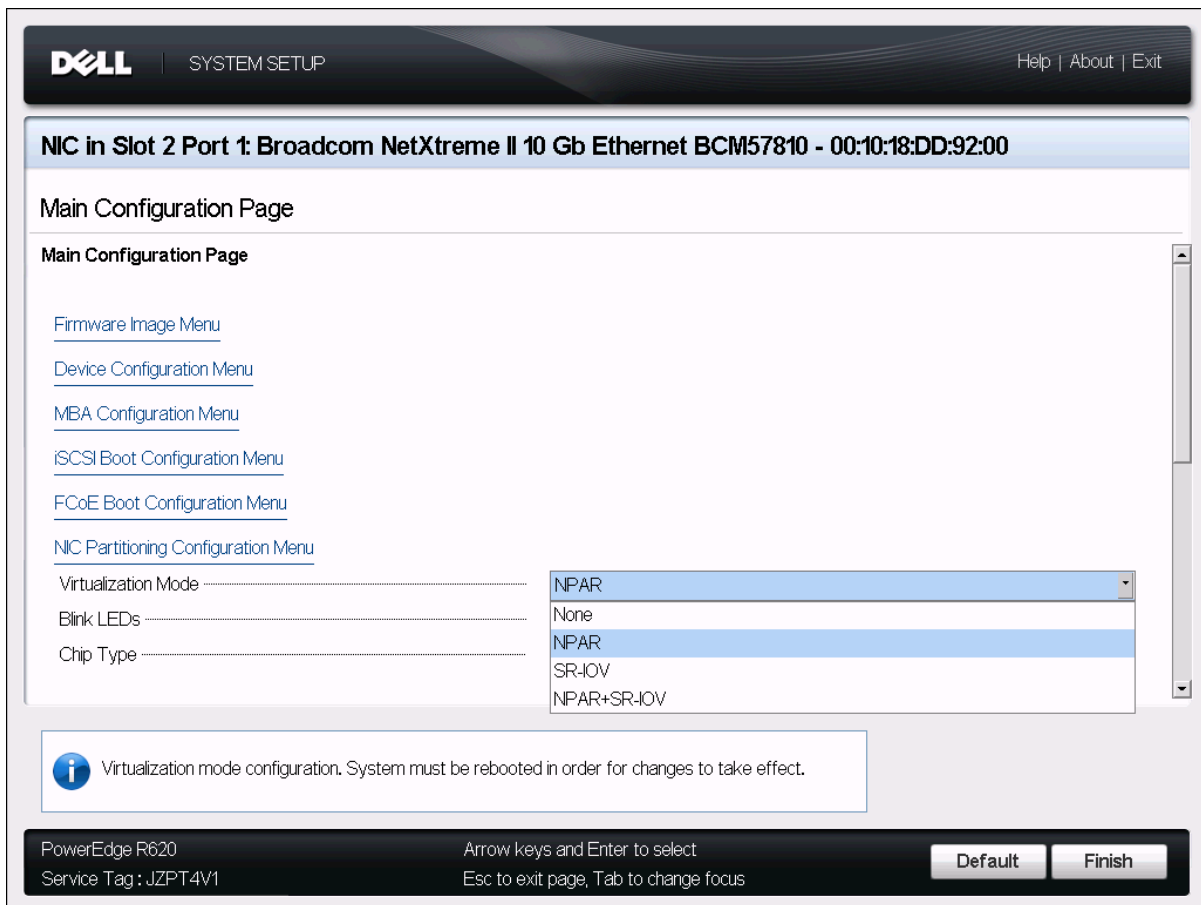


Figure 10 Dell System Setup – Device Configuration – Virtualization Mode

Select **NIC Partitioning Configuration Menu**, then select the Partition to be used for iSCSI traffic. In that partition menu, enable **iSCSI Offload Protocol** as shown below.

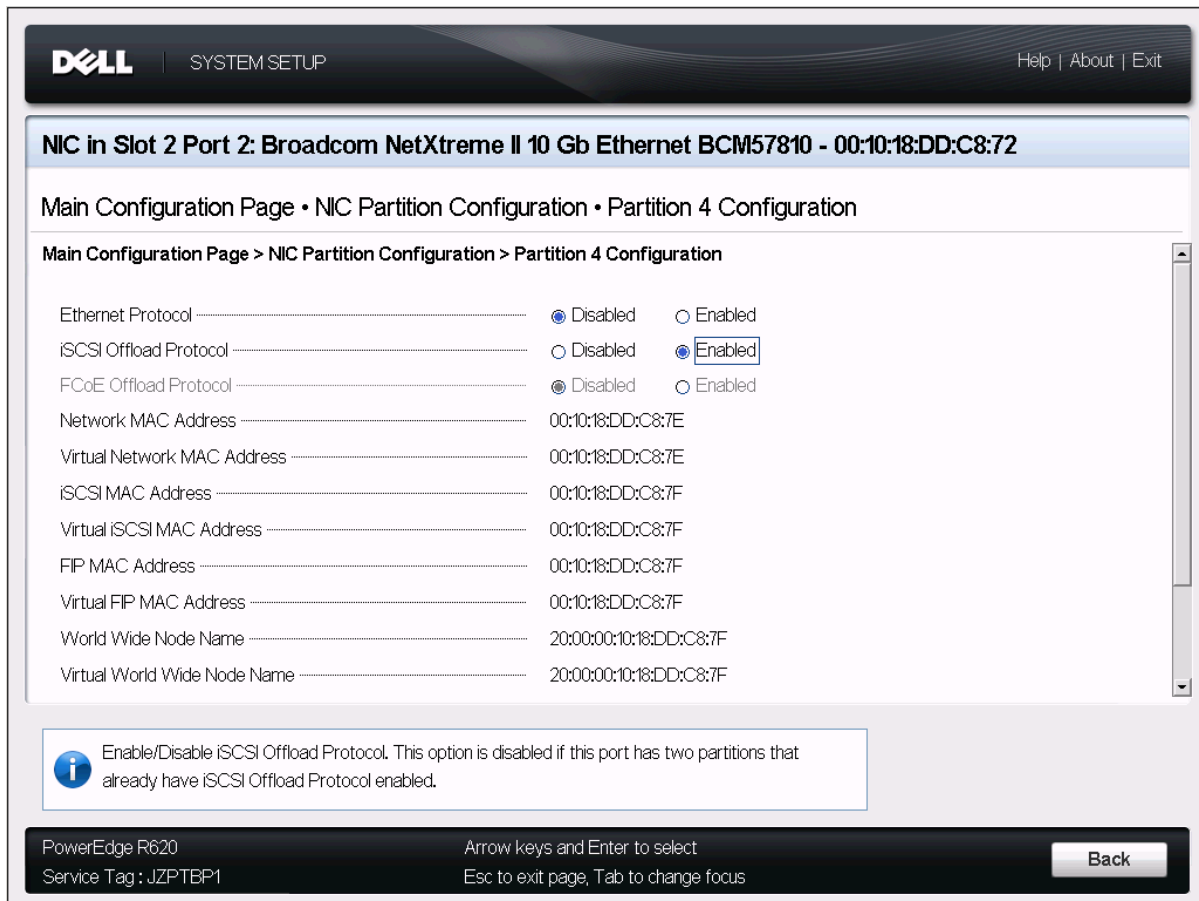


Figure 11 Dell System Setup – Partition Configuration - Enable iSCSI Offload Protocol

Make sure to note the Network MAC Address of the Partition with iSCSI Offload enabled, because this will ensure the correct vmnic is used for iSCSI traffic when configuring the vmnic to vSwitch setup. While there are many MAC addresses listed for the various features and functions of each partition, only the one identified as “Network MAC Address” will be visible in the vSphere Client.

3.3.2 Server virtualization configuration

If not completed yet, install vSphere vCenter and vSphere ESXi 5.1 now. For more instructions, see VMware installation documentation at <http://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>.

vSwitch Configuration

The following diagram shows the VMware vSphere vSwitch configuration for this white paper. Notice the VLAN ID defined for each VMKernel Port connected to vSwitchiSCSI. The VLAN assigned, VLAN 100 in this paper, is the same as the VLAN defined on the switch and the storage arrays for iSCSI traffic.

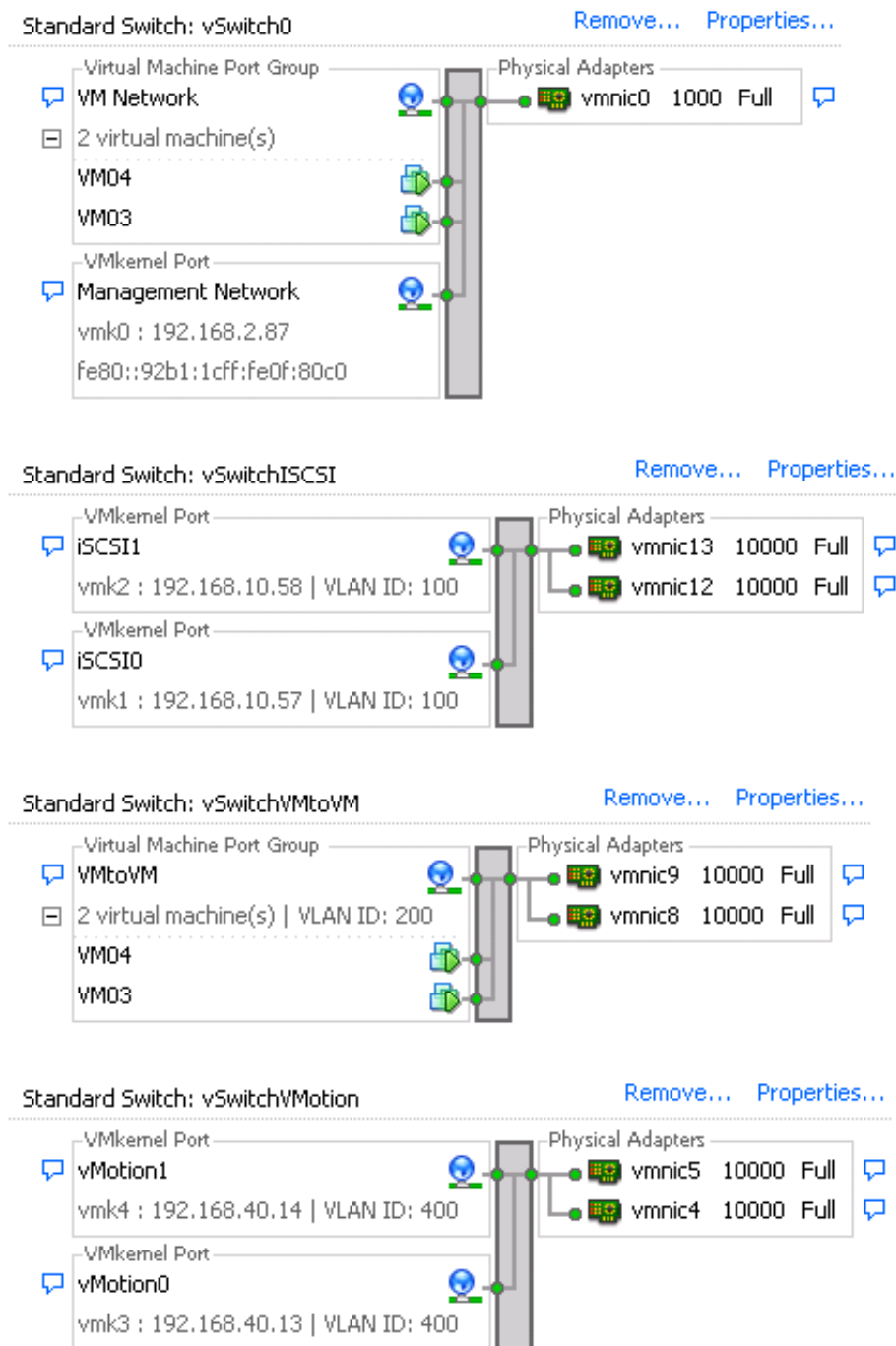


Figure 12 VMware Virtual Switch configuration

Note: vSwitches created through the vSphere Client GUI are automatically named vSwitch0, vSwitch1, etc. To create custom vSwitch names use the vCLI or vMA with the command `esxcfg-vswitch -a [vSwitchName]`

3.3.3 EqualLogic MEM installation and configuration

The following sections cover the installation and configuration of EqualLogic MEM for VMware.

3.3.3.1 EqualLogic MEM Installation

The Dell EqualLogic MEM can be installed through the vCLI/vMA or through the vSphere Update Manager. This document uses the vCLI installation from a remote windows host. During the installation, the following information is required:

- IP Address of vSphere ESXi server
- Login credentials for vSphere ESXi server
- EqualLogic MEM bundle that is available from the EqualLogic support site

Included in the download from the EqualLogic support site is a setup script to aid in the installation. The script requires the server IP Address, login credentials, and bundle name. For detailed installation instructions for EqualLogic MEM please see the Installation Guide found on <http://support.equallogic.com> (Requires a login).

```
C:\Program Files (x86)\VMware\VMware vSphere CLI\bin\EqualLogic-ESX-Multipathing-Module>setup.pl --install --server=192.168.2.86 --username=root --password=password --bundle=dell-eql-mem-esx5-1.1.2.292203.zip --reboot
```

3.3.3.2 EqualLogic MEM configuration

The Dell EqualLogic MEM can be configured two ways. The most intuitive way is through the interactive mode. To run in interactive mode, run `setup.pl --configure --server=[ESXi IP Address]` This information is helpful to have when completing the configuration:

- ESXi Server management IP Address
- ESXi login credentials
- Name of vSwitch to be created
- MTU size
- The vmnics to be used - These should correspond to the NPAR partition with iSOE enabled.
- IP Addresses to be assigned for the VMKernel interfaces
- EqualLogic Group IP Address

The second method to configure the EqualLogic MEM is through the unattended mode. In this mode, all the required information is given as arguments to the MEM configuration command. For more information, see the Installation Guide found on the EqualLogic Support site.



3.3.3.3 Datastore creation

Before storage can be used in VMware, it must be placed in a Datastore container to allow virtual machine access. This section provides a brief overview on creating the Datastore and selecting the correct volume when multiple exist.

1. Select a host in Hosts and clusters view.
2. Select the **Configuration** tab.
3. Select storage and then select **Add Storage**. This presents a disk selection wizard.
4. Select **disk/LUN** and select **Next**. This brings up a disk selection screen as shown below.

The simplest method to map the available disks to the correct iSCSI volume on EqualLogic is to extend the PATH ID column. The last segment of the path ID corresponds to the volume name given during volume creation on EqualLogic.

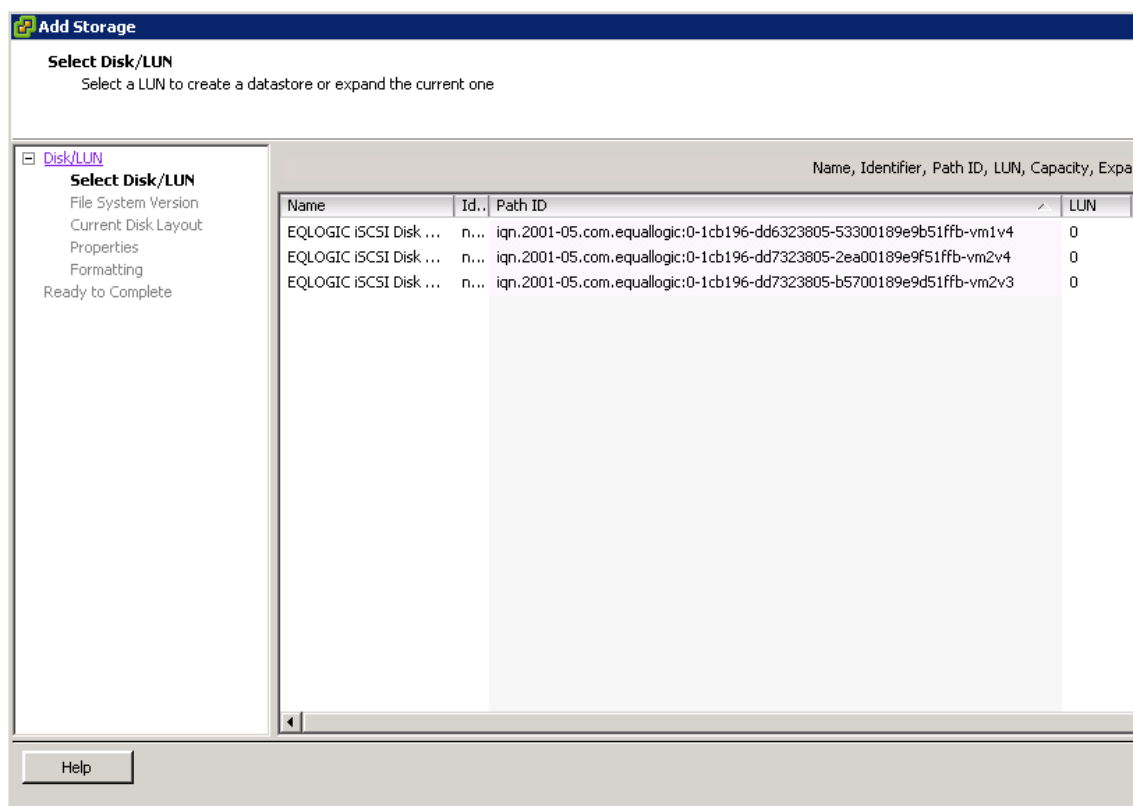


Figure 13 VMware – Add Storage – Disk Selection – Path ID

3.3.3.4 VMware Path Selection Policy (PSP)

The EqualLogic MEM configuration correctly modifies the VMware PSP to reflect the optimized DELL_PSP_EQL_ROUTED for EqualLogic storage. To manually verify this change, select a host in the vSphere client management application, select the **Configuration** tab, then select **Storage**, then select a Datastore and view the Path Selection field in the Datastore Details pane.

View: Datastores Devices


Datastores

Identification	Device	Drive Type	Capacity	Free	Type	Last Update
VM01_Disk01	EQLOGIC iSCSI Disk (naa.6019...	Non-SSD	99.75 GB	8.80 GB	VMFSS	8/28/2013 10:39:22 AM
VM01_Disk02	EQLOGIC iSCSI Disk (naa.6019...	Non-SSD	99.75 GB	8.80 GB	VMFSS	8/28/2013 10:39:22 AM
VM01_Disk03	EQLOGIC iSCSI Disk (naa.6019...	Non-SSD	99.75 GB	8.80 GB	VMFSS	8/28/2013 10:39:22 AM
VM01_Disk04	EQLOGIC iSCSI Disk (naa.6019...	Non-SSD	99.75 GB	8.80 GB	VMFSS	8/28/2013 10:39:22 AM
VM02_Disk01	EQLOGIC iSCSI Disk (naa.6019...	Non-SSD	99.75 GB	8.80 GB	VMFSS	8/28/2013 10:39:22 AM

Datastore Details

VM01_Disk01
 Location: /vmfs/volumes/51dc0f98-4e516f44-ebb3-90b11c0f8048
 Hardware Acceleration: Supported
[Refresh Storage Capabilities](#)
 System Storage Capability: N/A
 User-defined Storage Capability: N/A

99.75 GB Capacity
 90.95 GB Used
 8.80 GB Free



Path Selection	Properties	Extents	Storage I/O Control
DELL_PSP_EQL_R...	Volume Label: VM01_Disk01 Datastore Name: VM01_Disk01	EQLOGIC iSCSI Disk (naa.6... Total Formatted Capacity 99.75 GB	Disabled

Paths
 Total: 3
 Broken: 0

Formatting
 File System: VMFS 5.58

Figure 14 VMware – Datastore Properties – Dell Path Selection Properties

At this point the Datastore may be used to hold virtual machines files, secondary VMDKs, and others.

4 Test results and analysis

This section outlines the parameters examined during each test run through the DCB network topology described in this white paper. These tests were run to drive both iSCSI Ethernet traffic and LAN Ethernet traffic to verify the configuration DCB in the network.

4.1 Ethernet retransmissions

Retransmitted Ethernet frames are Ethernet frames that have to be sent again because the original frame did not arrive at the destination – either they did not arrive as expected (for example, corruption or CRC error), did not arrive at all (dropped, usually due to congestion in the network), or were not acknowledged as being received (missing ACK). These frames are vital to ensure all data that needs to be transmitted from the sender is received at the other end, but they waste bandwidth with data that is repeated. DCB provides a method to try and prevent transmitted data from being dropped due to congestion. With retransmitted frames due to switch congestion close to eliminated, the efficiency of the network is greatly improved, even with some retransmission caused by things like CRC errors. Figure 15 shows the iSCSI retransmissions as measured at the storage arrays.

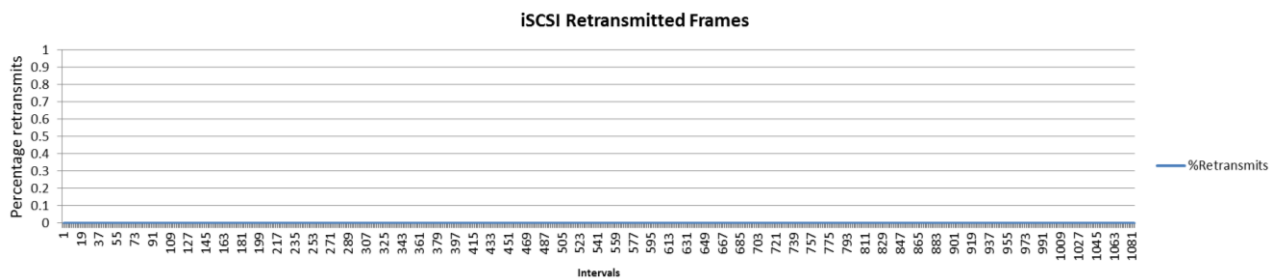


Figure 15 iSCSI retransmitted frames

No retransmitted frames were recorded for the lossless Ethernet class. The results were similar for every run.

4.2 Response time

Observing the disk response time gives insight into the type of performance that can be expected during different types of IO load. The chart below shows the response time of each iSCSI disk measured on the six virtual machines averaged over a test run. The changes in response time patterns reflect the changes in IO workload during the test.

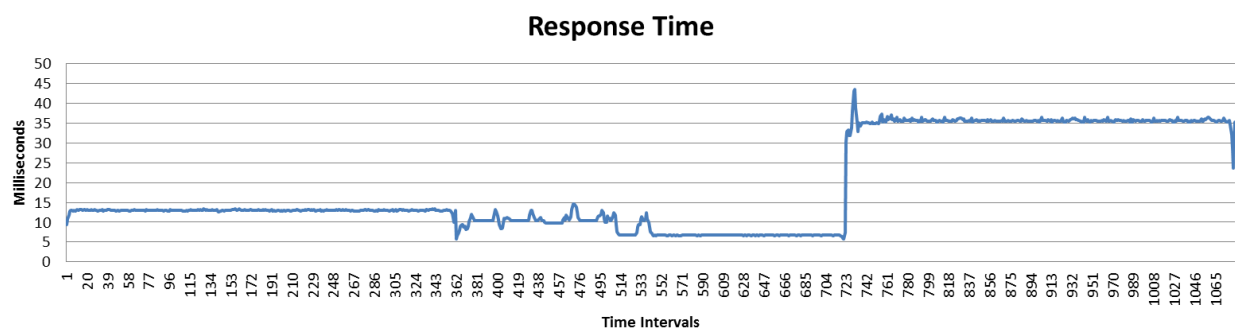


Figure 16 Response time

Note: For more information on the workloads run, see Appendix B Workload .

5 Best practice recommendations

The following section lists the best practices that should be used throughout the solution stack. Use this list as a check list when designing and deploying this reference architecture.

- ESXi host segment
 - Configure NPAR through Dell System Setup utility.
 - Configure NPAR to use DCB on a separate partition dedicated for iSCSI traffic with iSCSI offload enabled (iSOE).
 - Configure NPAR to use one or more of the remaining partitions for LAN traffic.
 - Install and configure the Dell EqualLogic Multipath Extension module.
 - Use the setup script provided with EqualLogic MEM.
 - Configure VLAN for DCB use on the VMKernel ports belonging to the iSCSI vSwitch.
- Network segment
 - Use DCB to provide lossless Ethernet for iSCSI Storage connectivity.
 - Configure dedicated VLAN for iSCSI traffic.
 - Disable traditional flowcontrol.
 - Configure PFC priority and Priority groups.
 - Setup ETS parameters to reflect desired network partitions.
 - Configure Jumbo Frame support end-to-end to accommodate larger Ethernet Frames.
- Storage segment
 - Ensure iSCSI logins are using Jumbo Frames.
 - To verify login to Group Manager, select Monitor and verify recent logins are with Jumbo Frames, not Standard Frames.
 - Configure VLAN for DCB use in Group Manager.
- Verification: Verify every section of the setup to ensure it is configured as designed.
 - DCB requires all components in the topology to participate to provide lossless Ethernet.
 - Virtual switches currently do not support DCB, so this prohibits the use of Guest initiators within the Virtual machine for iSCSI connectivity.
 - DCB configuration parameters can be checked as shown in section 3.1.3 Verifying switch DCB configuration.



6 Conclusion

This reference architecture is customizable depending on the exact challenges and requirements presented. The solution documented here represents a basic converged server virtualized environment when using the Dell Networking S4810 with Dell EqualLogic iSCSI storage in a converged LAN and SAN solution leveraging industry standard Data Center Bridging features.

It shows the configuration of each portion of the solution stack, starting with configuring the basic switch management, VLAN setup, and DCB in the network stack. Then the configuration continues with the EqualLogic storage. Finally, the configuration ends with NPAR, VMware vSwitches, and EqualLogic MEM for VMware.



A Solution configuration details

Table 1 Compute segment components

Component	Description
Server Hardware	PowerEdge R620
Server Bios	1.4.8
Server Virtualization	VMware vSphere ESXi 5.1
VM Operating System	Windows 2008 R2
Converged Network Adapter	Broadcom NetXtreme II BCM57810 (Dual port 10Gb) Family Firmware Version 7.4.8

Table 2 Network segment components

Component	Description
Model	Dell Networking S4810
Firmware	8.3.12.0

Table 3 Storage segment components

Component	Description
Storage Array	Dell EqualLogic PS6110
Firmware	V6.0.5



B Workload characterizations

Multiple tools were used for load generation. To provide storage traffic to the disks, VDBench was used for its simplicity and script ability.

```
sd=A-a,lun=\\.\PhysicalDrive1
```

```
sd=A-b,lun=\\.\PhysicalDrive2
```

```
sd=A-c,lun=\\.\PhysicalDrive3
```

```
sd=A-d,lun=\\.\PhysicalDrive4
```

```
sd=B-a,lun=\\.\PhysicalDrive1,range=(30m,60m)
```

```
sd=B-b,lun=\\.\PhysicalDrive2,range=(30m,60m)
```

```
sd=B-c,lun=\\.\PhysicalDrive3,range=(30m,60m)
```

```
sd=B-d,lun=\\.\PhysicalDrive4,range=(30m,60m)
```

```
wd=wd1,sd=A-*,seekpct=100,rdpct=67,xfersize=8k,iorate=9999999,priority=1
```

```
wd=wd2,sd=B-*,seekpct=0,rdpct=100,xfersize=256k,iorate=9999999,priority=1
```

```
wd=wd3,sd=B-*,seekpct=0,rdpct=0,xfersize=256k,iorate=9999999,priority=1
```

```
rd=rd1,wd=wd1,elapsed=10800,interval=30,forthreads=20
```

```
rd=rd2,wd=wd2,elapsed=10800,interval=30,forthreads=5
```

```
rd=rd3,wd=wd3,elapsed=10800,interval=30,forthreads=5
```



C Additional resources

Support.dell.com is focused on meeting your needs with proven services and support.

DellTechCenter.com is an IT Community where you can connect with Dell Customers and Dell employees for the purpose of sharing knowledge, best practices, and information about Dell products and installations.

Referenced or recommended Dell publications:

- Dell EqualLogic Configuration Guide:
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19852516/download.aspx>
- Dell Force10 S4810 Switch Configuration Guide
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20220824/download.aspx>
- Dell EqualLogic DCB Configuration Best Practices
<http://en.community.dell.com/techcenter/storage/w/wiki/4355.configuring-dcb-with-equallogic-sans.aspx>
- Configuring and Installing the EqualLogic Multipathing Extension Module for VMware vSphere 5.1, 5.0 and 4.1 and PS Series SANs
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19991633.aspx>
- Data Center Bridging: Standards, Behavioral Requirements, and Configuration Guidelines
<http://en.community.dell.com/techcenter/storage/w/wiki/4396.data-center-bridging-standards-behavioral-requirements-and-configuration-guidelines-by-sis.aspx>
- Configuring DCB with EqualLogic SANs
<http://en.community.dell.com/techcenter/storage/w/wiki/4355.configuring-dcb-with-equallogic-sans.aspx>
- Creating a DCB Compliant EqualLogic iSCSI SAN with Mixed Traffic
<http://en.community.dell.com/techcenter/storage/w/wiki/2634.creating-a-dcb-compliant-equallogic-iscsi-san-with-mixed-traffic-by-sis.aspx>

Referenced or recommended Broadcom publications:

- Broadcom Switch Independent Partitioning Setup Guide
<http://www.broadcom.com/collateral/swum/2CS57712-SWUM103-R.pdf>

Referenced or recommended VMware publications:

- VMware vSphere 5 Documentation Center - Configuring Dependent Hardware iSCSI initiator
http://pubs.vmware.com/vsphere-50/topic/com.vmware.vsphere.storage.doc_50/GUID-F675B2EF-3B6A-4107-83D5-49262DEA1F4D.html
- VMware vSphere Management Assistant documentation and download
<http://www.vmware.com/support/developer/vima/>
- VMware vSphere Command Line Interface documentation and download
<http://www.vmware.com/support/developer/vcli/>



- Multiple-NIC vMotion in vSphere 5 (KB: 2007467)
http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2007467
- Creating vSwitches with custom names (KB: 1020757)
http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1020757

