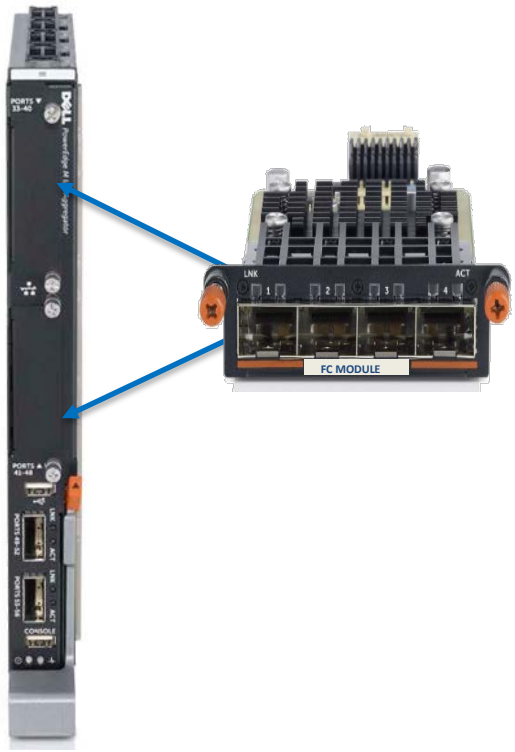
Deploying Dell Networking MXL and PowerEdge M IO Aggregator with the FC FlexIO Module in a Cisco MDS Environment

Dell Networking Solutions Engineering
January 2014



A Dell Deployment/Configuration Guide

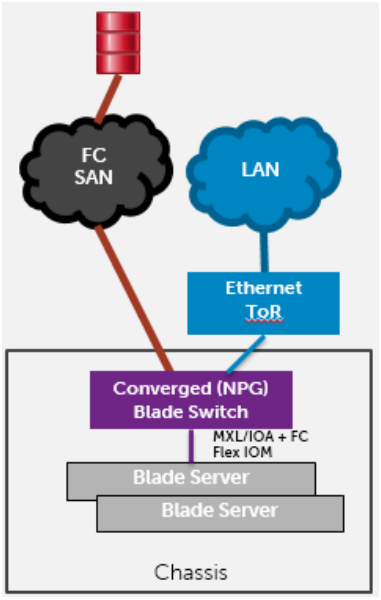
Revisions

Date	Description	Authors
January 2014	Initial release	Neal Beard, Humair Ahmed

Table of contents

1 Executive Summary

In the “Dell Networking FC Flex IO: Infrastructure & Network Convergence w/ FCoE” whitepaper, it is demonstrated and explained how to move from a traditional non-converged LAN/SAN network to a converged LAN/SAN infrastructure and network, and how the Dell FC Flex IO with Dell MXL/IOA is an ideal solution for this transition. The many benefits incorporated by the FC Flex IO module offer not only a converged networking solution, but also a converged infrastructure where the FCoE-FC bridging functionality is implemented directly within the blade switch.



The Dell FC Flex IO module transforms a Dell MXL/IOA blade switch from an Ethernet-only switch to a converged switch capable of bridging between Ethernet and Fibre Channel. The Dell MXL/IOA blade switch with the FC Flex IO module moves the convergence layer from a typical dedicated Top of Rack (ToR) switch down to the blade switch via an extensible IO module, providing the benefits of IO consolidation such as less infrastructure hardware, less maintenance, and considerable cost savings. With its unique modular design the FC Flex IO module allows end users to migrate to a converged solution and introduce FCoE-FC functionality to the MXL/IOA blade switch at their own pace without replacing the entire switch. This benefit is unmatched in the industry. In this deployment guide we cover detailed Dell FC Flex IO module topology and configuration examples.

2 Dell PowerEdge M1000e Overview

The PowerEdge M1000e Modular Server Enclosure solution supports up to (32) server modules, and (6) network IO modules. The M1000e contains a high performance and high availability passive midplane that connects server modules to the infrastructure components, power supplies, fans, integrated KVM and Chassis Management Controllers (CMC). The PowerEdge M1000e uses redundant and hot-pluggable components throughout to provide maximum uptime. The chassis has the ability to house 6 x network IO modules allowing for a greater diversity of roles for all of the enclosed blade servers.

The six (6) network IO slots in the back of the chassis are classified as 3 separate fabrics, each fabric contains 2 slots (A1/A2, B1/B2, and C1/C2), these fabric IO slots relate to the ports found on the server side network adapters. The IO modules can be used independently of each other, and each network IO module must contain the same technology. For example, Fabric A is hardwired to the 2 network adapters on the blade server mainboards, which means the IO modules in Fabric A must support Ethernet. Fabrics B and C can be used for Ethernet, Fibre Channel, or InfiniBand. Figure 2 below exemplifies the IO mappings between the server side dual/quad port networking adapters and the IO modules.

Note: The networking adapters in Fabric A have also been described as LOMs (LAN on Motherboards), and bNDCs (blade Network Daughter Cards). All of these terms describe the same device: A network adaptor that performs Ethernet/iSCSI/FCoE tasks on behalf of the Server and its operating system.

Figure 1 M1000e Front and Back view

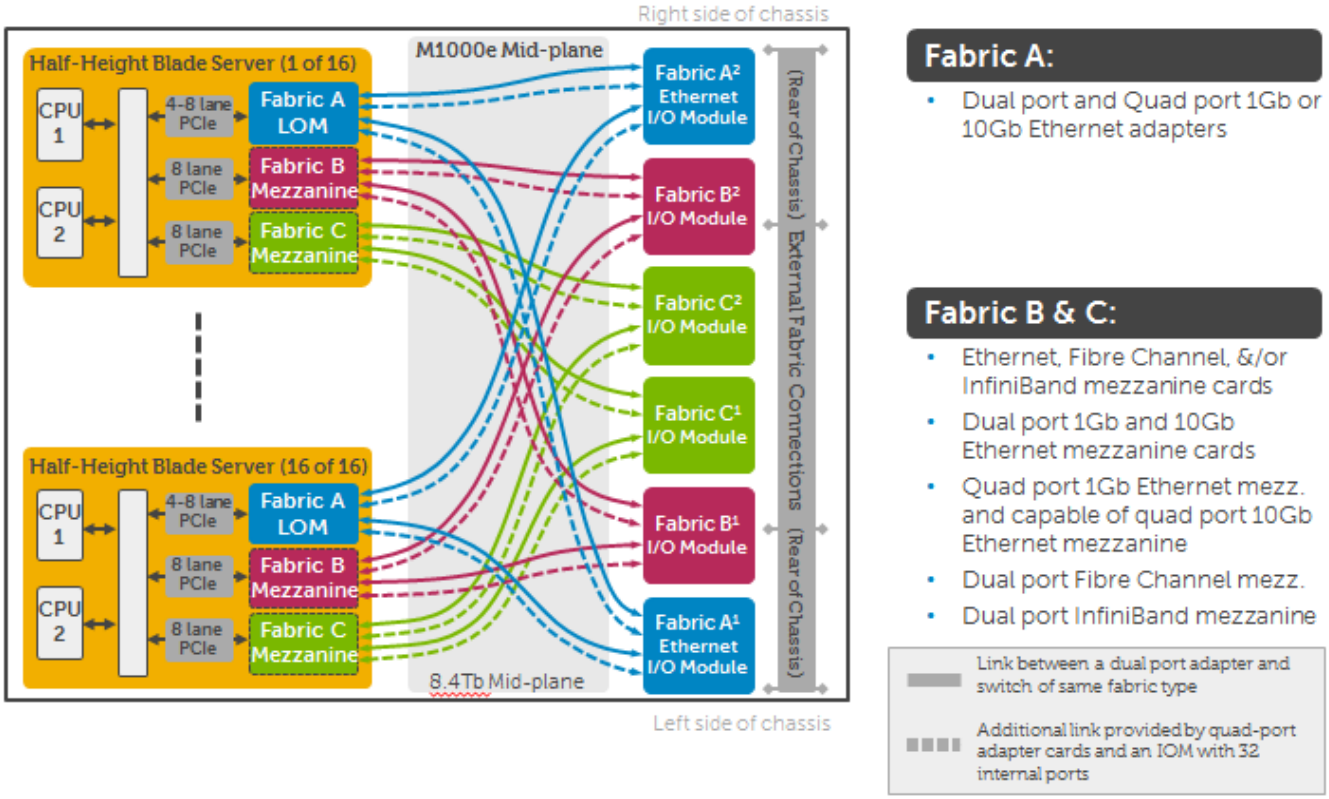


Figure 2 M1000e Midplane Dual/Quad Port Network Adaptor IO Mappings

2.1 Flex IO Expansion Modules (External Ports)

The Dell MXL/IOA blade switches support a combination of Flex IO Modules. The four (4) different types of Flex IO expansion modules are:

- 4-port 8Gb Fiber Channel FlexIO module
- 4-port 10Gbase-T FlexIO module
- 4-port 10Gb SFP+ FlexIO module
- 2-port 40Gb QSFP+ FlexIO module

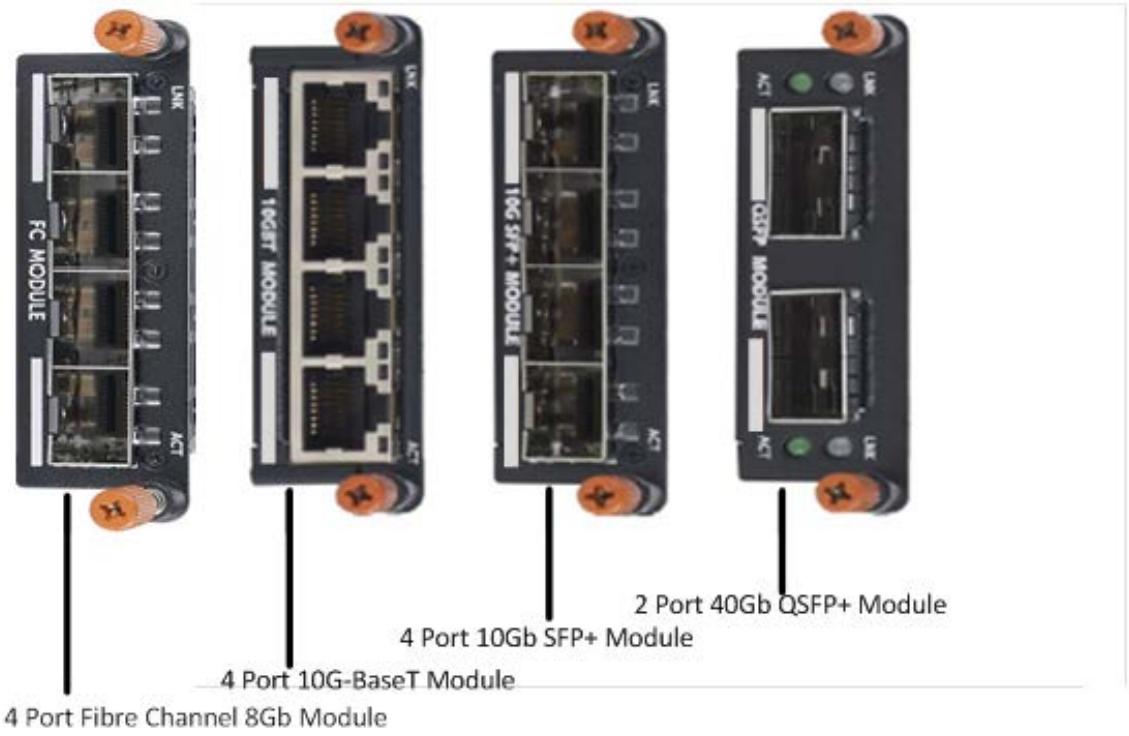


Figure 3 FlexIO expansion modules

Notes:

1. The 4 Port Fiber Channel 8Gb module can only be used with the release of 9.3 firmware and later for the Dell MXL/IOA modular switches.
2. Each MXL/IOA only supports one 10Gbase-T module.

2.2 Dell FC Flex IO and Convergence on Dell MXL/IOA Overview

The Dell FC FlexIO module allows the standard FCoE convergence model to move from a typical dedicated Top of Rack (ToR) switch with an FCoE Forwarder (FCF) down to an MXL or IOA IO module in an M1000e chassis. The FC FLexIO module utilizes NPIV Proxy Gateway (NPG) technology to allow the breakout of LAN and FC traffic at the M1000e chassis rather than at the ToR switch. The FC FlexIO module is the FCF (FCoE Forwader) that extracts the FC frame from the Ethernet frame generated by the CNA (Converged Network Adaptor) and sends it to the Fibre Channel switch. NPG technology in the FC FlexIO module performs a proxy fabric login (FLOGI) on behalf of the CNA's in the M1000e blade servers. Fibre Channel fabric services are maintained by the FC switch to which the FC FlexIO module is attached. The features that the FC FlexIO module provides are:

1. Management of the FLOGI and FDISC conversion process
2. Management of the PFC (Priority Flow Control)
3. Management of ETS (Enhanced Transmission Selection)
4. Management of the FIP keep-alives

From the perspective of the network adapter on the server, the FlexIO looks like an FCoE forwarder (FCF) and is performing FIP(FCoE initialization Protocol and DCBx(Data Center Bridging Exchange) provisioing. The NPG technology makes the FC FlexIO module appear to be an FC device or N port.

2.3 Component Information

In this table are the components used at the time this document was created, along with their FW versions

	Component	Version
Chassis / Server	M1000e Chassis Management Controller	4.50
	Dell PowerEdge M IO Aggregator	9.3(0.0)
	Dell Networking MXL	9.3(0.0)
	Dell PowerEdge M820 Blade Server(s)	1.7.3 (BIOS)
	Lifecycle Controller	1.1.5.165
	Broadcom 10Gb 2P 57810-k Mezzanine Card	7.6.15 (Firmware)
	QLogic 10Gb 2P QME8262-k Mezzanine Card	1.12.61 (Firmware)
	Intel 10Gb 2P X520-k blade Network Daughter Card (bNDC)	14.5.8 (Firmware)
Storage Equipment	EMC_3U VNX 5300	05.32.000.5.008
Cisco Equipment	Cisco MDS 9148 (system and kickstart)	6.2(3)
Cables	SFP+ Optical Transceivers (SR or LR) with Fiber Cables	5 Meter Cable

Table 1 Component Information

3 Converged Network Solution - Dell PowerEdge M1000e, EMC VNX 5300 storage array, Cisco MDS 9148 FC Switch, and Dell MXL/IOA w/ FC Flex IO module

This converged network scenario will concentrate on creating an easy to follow step by step guide that walks a network engineer through provisioning the various components of a Converged Network Solution. The configuration steps that will be used to exemplify the solution are:

1. Server side network adaptor configuration
2. Setup of the MXL and IOA
3. Configuration of the Cisco MDS 9148 Fibre Channel switch
4. Provisioning of the LUNS on the EMC VNX 5300 storage array

3.1 Converged Network Topology

The Figure 4 diagram below shows the components that are being used to create this particular converged network solution.

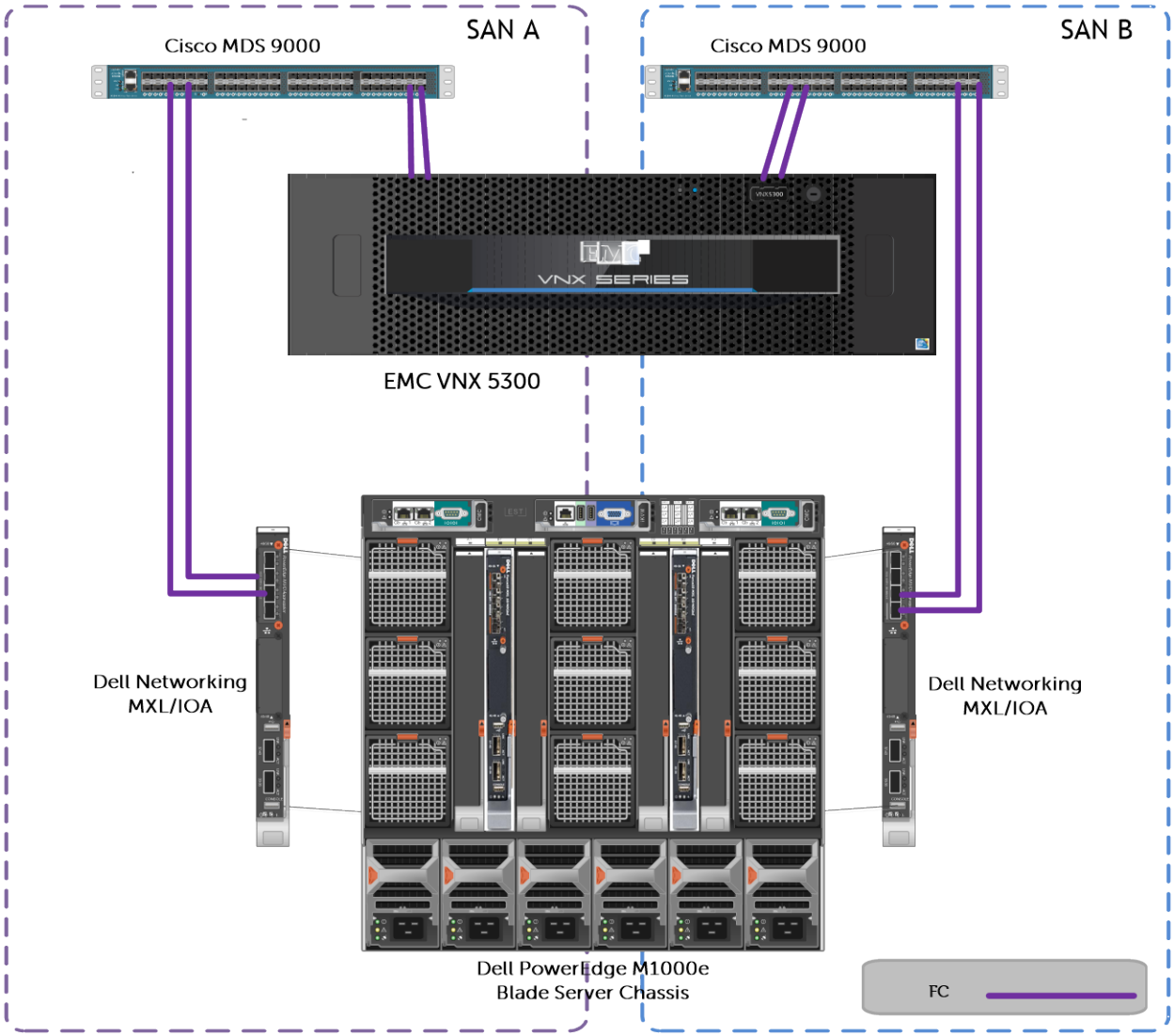


Figure 4 Converged Network Configuration with MDS FC Switch

3.2 M1000e Internal Port Mappings

A main item of focus is understanding port mappings between the server's network adaptor and the MXL/IOA's 32 internal ports. Referring to Figure 5, it can be seen that the MXL/IOA IO modules can be provisioned for up to 32 internal connections depending on whether the 16 slots in the M1000e are filled with half height, or full height blades, which would only use ports 1-16. If quarter height blades or quad port network adapters are used, then ports 17-32 would be available for configuration as well. External port numbering on the IO modules is dependent on whether the (2) integrated 40Gb Ethernet ports are in 4 x 10Gb breakout mode and/or which FlexIO modules are in the two external IO bays.

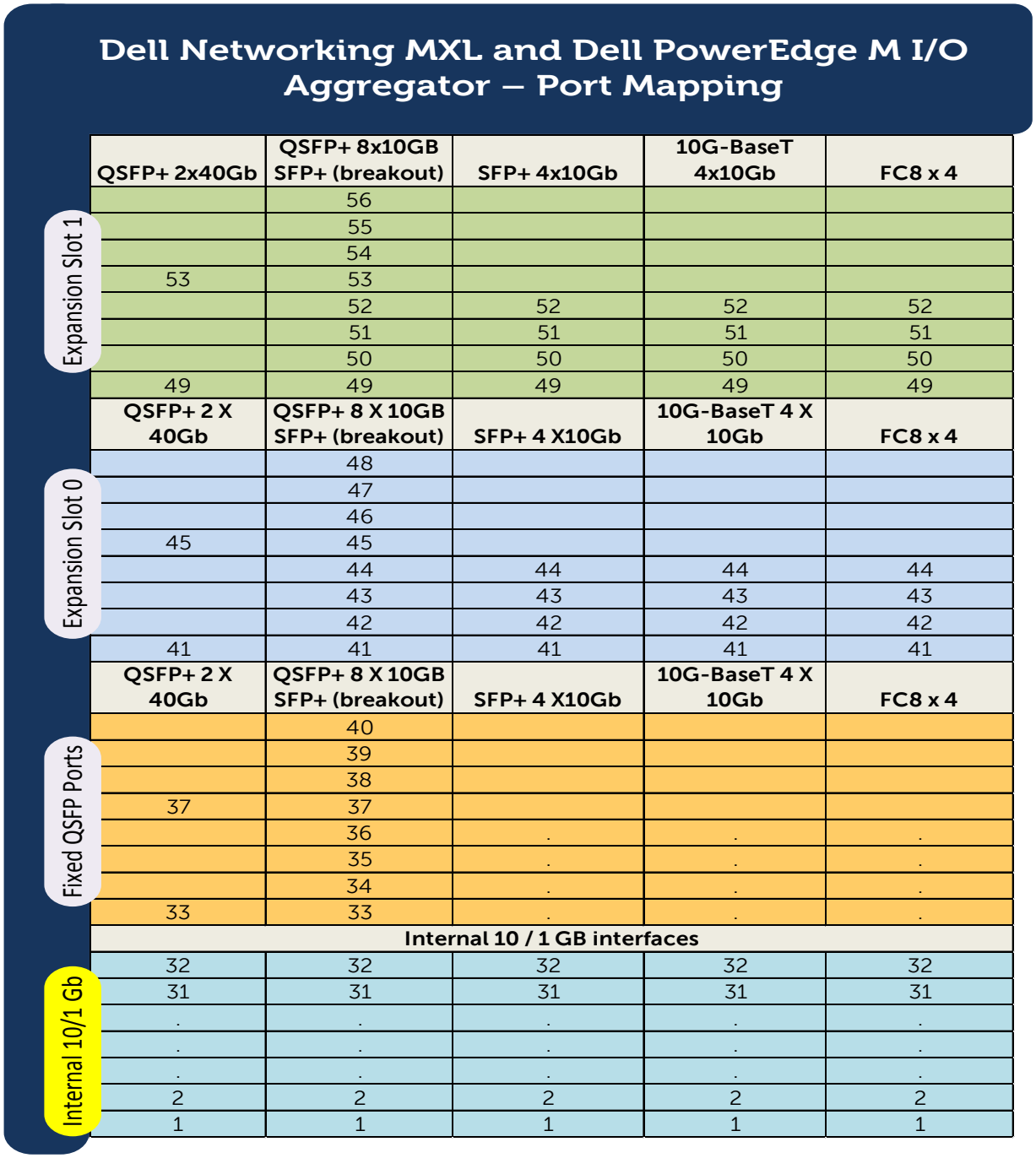


Figure 5 Port mapping for Dell Networking MXL/IOA expansion modules

3.3 M1000e FlexAddress enablement

When working with network adapters that provide a MAC address as well as a World Wide Port and Node name, it's important to understand where these addresses originate. The M1000e chassis has a feature called Flexaddressing. Flexaddressing is a pool of 208 MAC's and 64 WWN contained on an SD card that

can either be added at the Dell factory to the M1000e's chassis management controller (CMC) or to an existing M1000e CMC at a datacenter via a customer kit. Referring to Figure 7 below, enabling Flexaddressing is facilitated by:

1. Enter CMC's IP address
2. Click on Server Overview
3. Click on Setup
4. Click on FlexAddress
5. Choose to enable FlexAddressing to be either Fabric based or Slot based

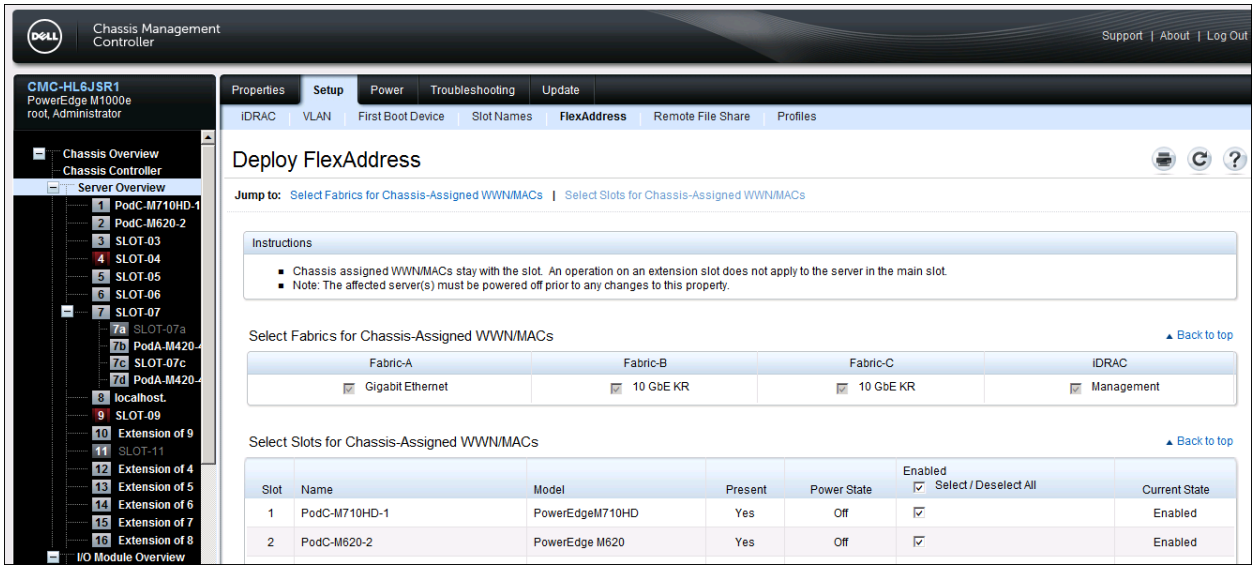


Figure 6 FlexAddress Enablement

Note: When the FlexAddress feature is deployed for the first time on a given server, it requires at least one power-down and power-up sequence of the server for the FlexAddress feature to take effect. Certain network configurations may require refreshing of network tables, such as ARP tables on IP switches and routers in order for new MAC/WWNs to be logged in.

To check that FlexAddressing is enabled:

1. Click on Server Overview
2. Click on Properties
3. Click on WWN/MAC
4. WWN/MAC Addresses section – FlexAddress: Enabled should be visible. Chassis-Assigned should have check marks next to each address.

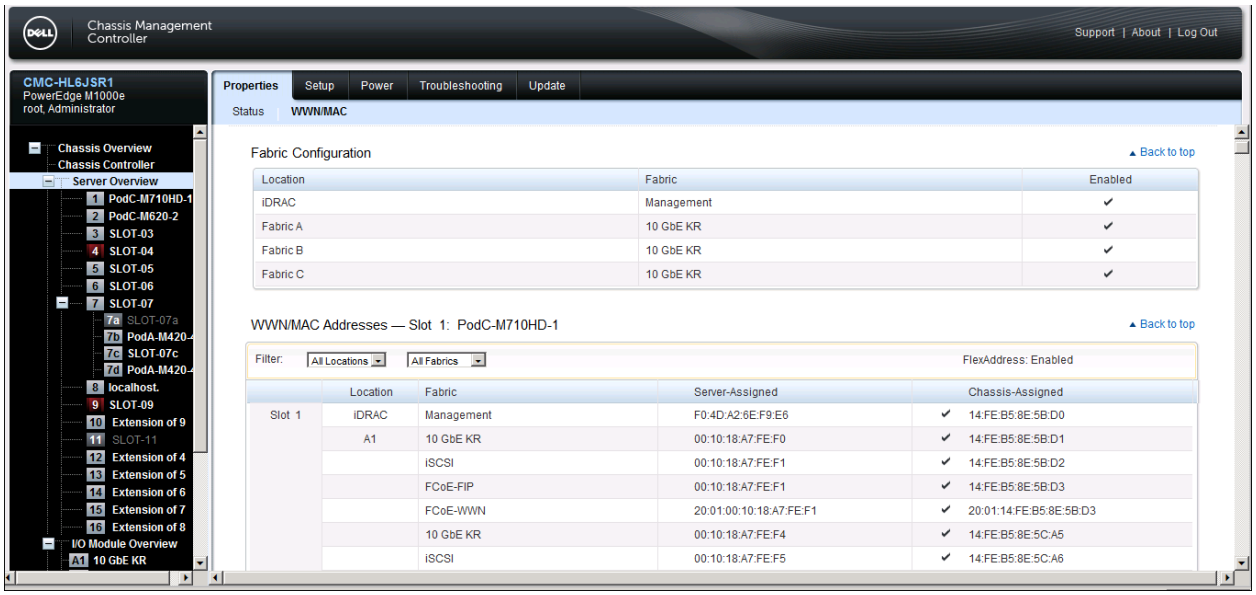


Figure 7 FlexAddress Enablement Confirmation

Once FlexAddress has been confirmed as enabled, it's a simple matter of matching up the MAC or WWN of the network adaptor with the Server Slot and the Fabric that it is in.

3.4 Network Adaptor Provisioning

The next configuration step is setting up the network adapters for FCoE traffic. The network adapters that will be utilized are:

1. Broadcom 57810-K Dual Port 10Gb network adaptor
2. Qlogic QME8262-K Dual Port 10Gb network adaptor
3. Intel X520-K Dual Port 10Gb network adaptor

3.4.1 Broadcom 57810-K network adaptor FCoE configuration steps

To configure the Broadcom BCM57810 network adapter the following information is required (if FCoE boot from SAN is desired). Storage Processor WWPN and Boot LUN (WWPN 50:06:01:6F:3E:E0:18:70 (EMC VNX 5300) and Boot LUN = 0

1. Power On the server.
2. Press **Ctrl-s** when the Broadcom adapters Boot Code appears on screen.

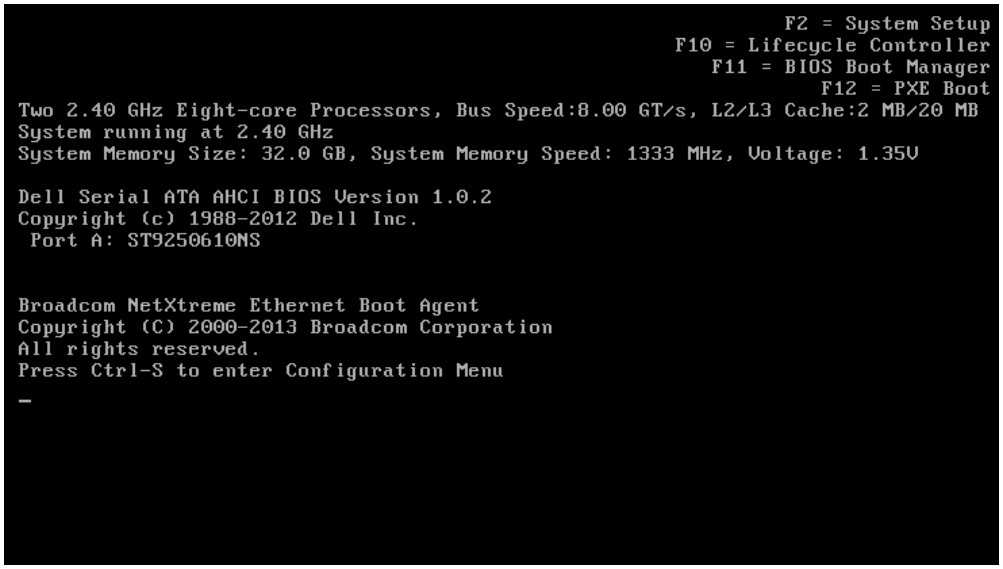


Figure 8 Ctrl-S screen

3. Highlight the BCM57810 adaptor that will be used for FCoE and press the **enter** key

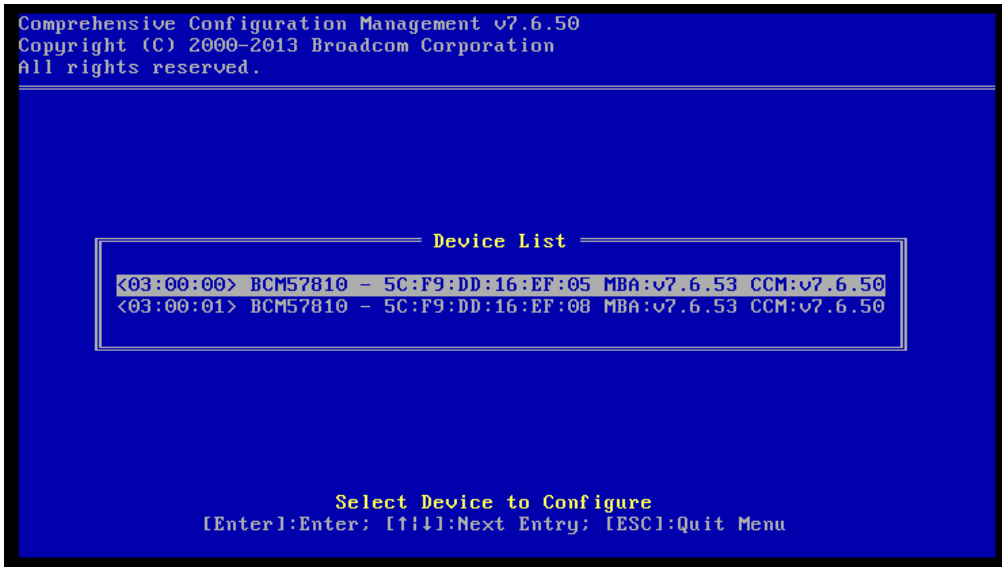


Figure 9 Adaptor Selection

4. You are now in the “Main Menu” to configure the Broadcom adapter.
5. From the Main Menu choose Device Hardware Configuration and press the **Enter** key.

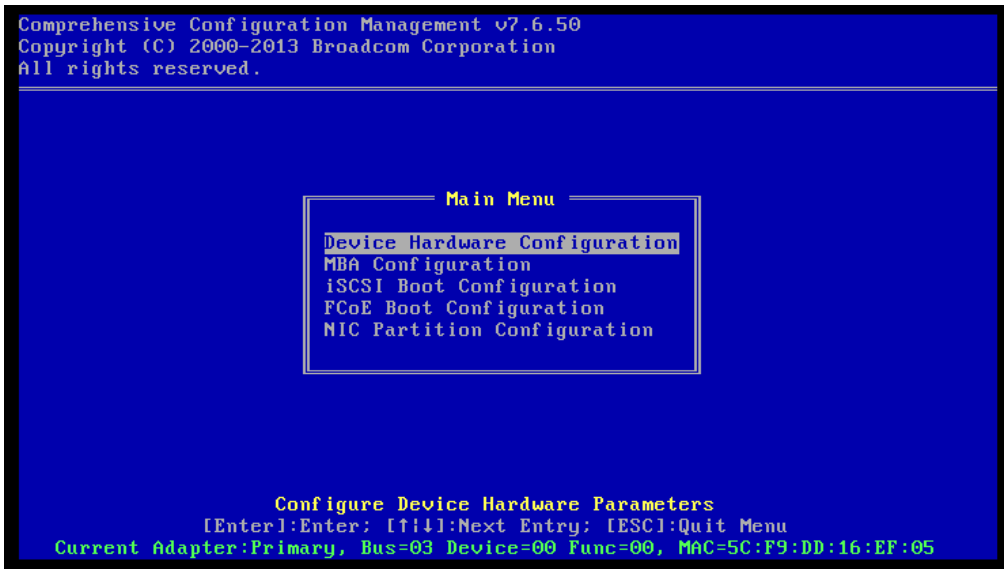


Figure 10 Device Hardware Configuration

6. Change **Multi-Function Mode** to **NPAR** if desired.

Note: NPAR (NIC Partitioning) is not required for FCoE.

7. Change **DCB Protocol** to **Enabled** (disabled by default).
8. Change **SR-IOV** to **Enabled** if desired. Note: SR-IOV is not required for FCoE

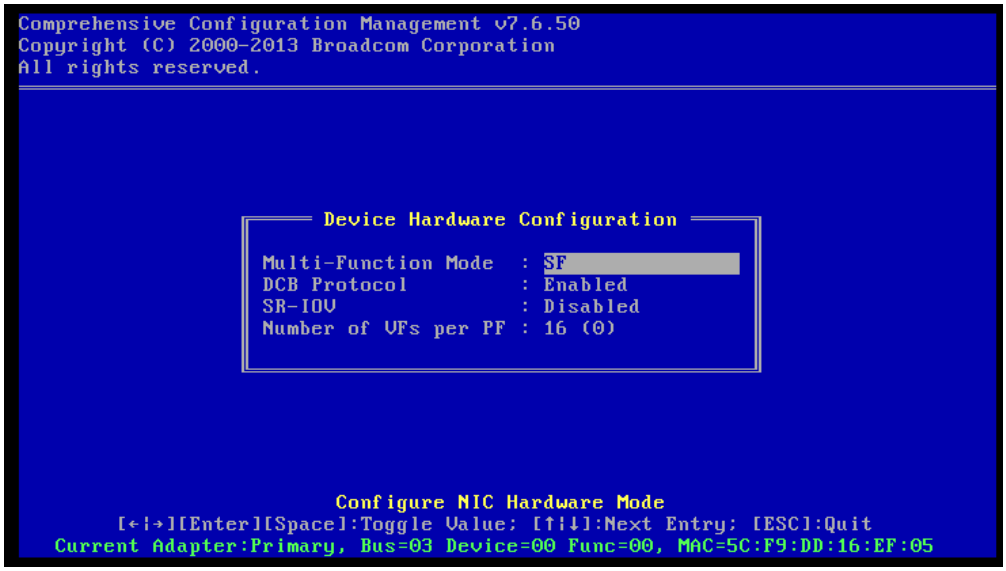


Figure 11 Device Hardware Configuration Details

9. Press the **Escape** key to return to the Main Menu.
10. From the **Main Menu** choose **MBA Configuration** and press the **Enter** key.

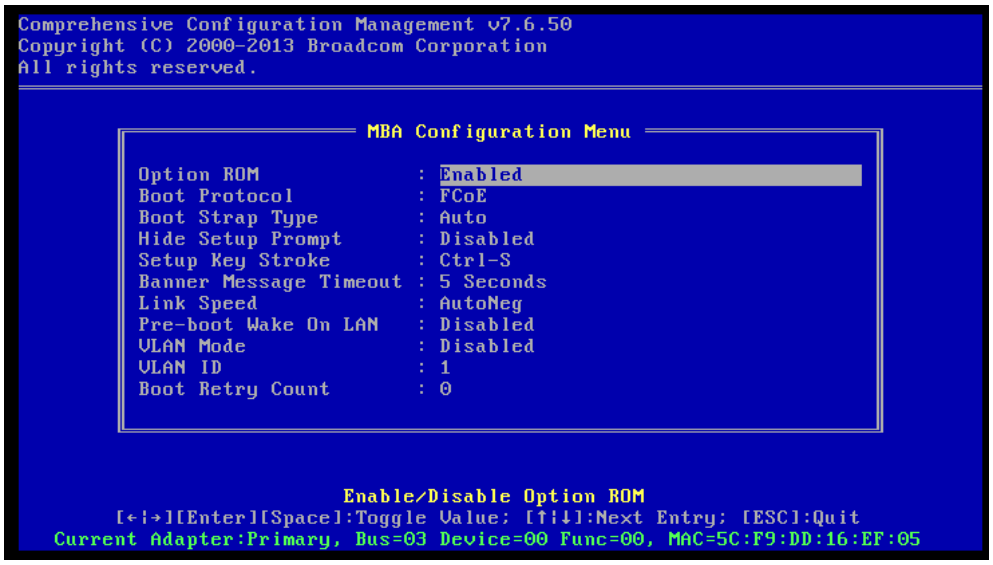


Figure 12 MBA Configuration

11. Option ROM should be Enabled
12. **Boot protocol** should be set to **FCoE**
13. All other settings can be left at default
    - Boot Strap Type = Auto
    - Hide Setup Prompt = Disabled
    - Setup Key Stroke = Ctrl-S
    - Banner Message Timeout = 5 Seconds
    - Link Speed = AutoNeg
    - Pre-boot Wake On LAN = Disabled
    - VLAN Node = Disabled
    - VLAN ID = 1
    - Boot Retry Count = 0
14. Press the **Escape** key to exit back to the Main Menu
15. From the **Main Menu** choose **FCoE Boot Configuration** and press the **Enter** key.

16. From the FCoE Boot Main Menu, highlight General Parameters and press the Enter key

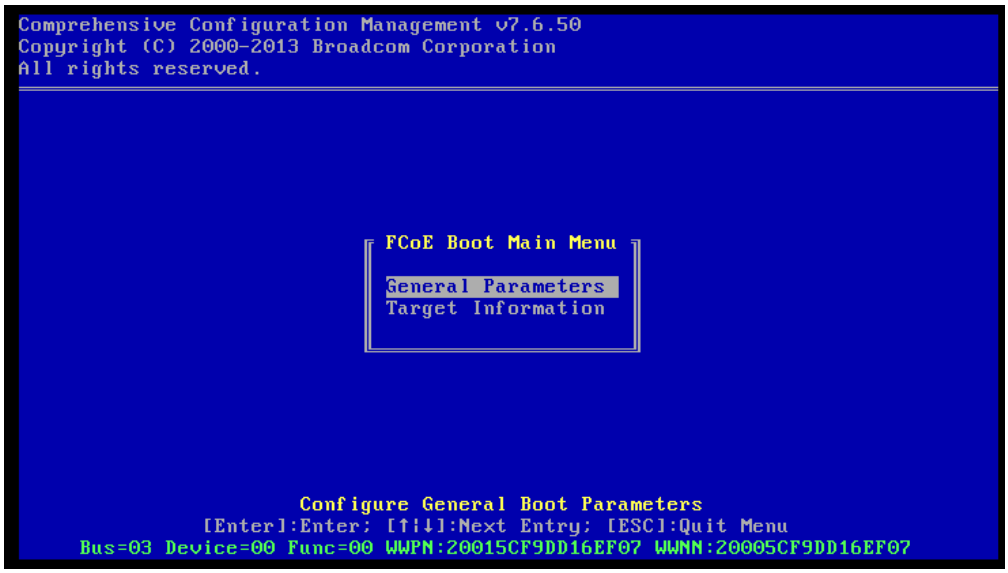


Figure 13 FCoE Boot Main Menu

17. Set the Boot to FCoE Target to One Time Disabled

Note: Setting one time disabled allows the CNA to bypass booting to the LUN (which has no OS installed) and allows the cd rom to be the boot device for loading the OS. Please understand that this setting will need to be re-enabled after each reboot if a further use of the CD-ROM is needed.

18. Set Target as First HDD to Enabled.

19. All other settings can be left at default
    - Link Up Delay Time = 0
    - LUN Busy Retry Count = 0
    - Fabric Discovery Retry = 4
    - HBA Boot Mode = Enabled

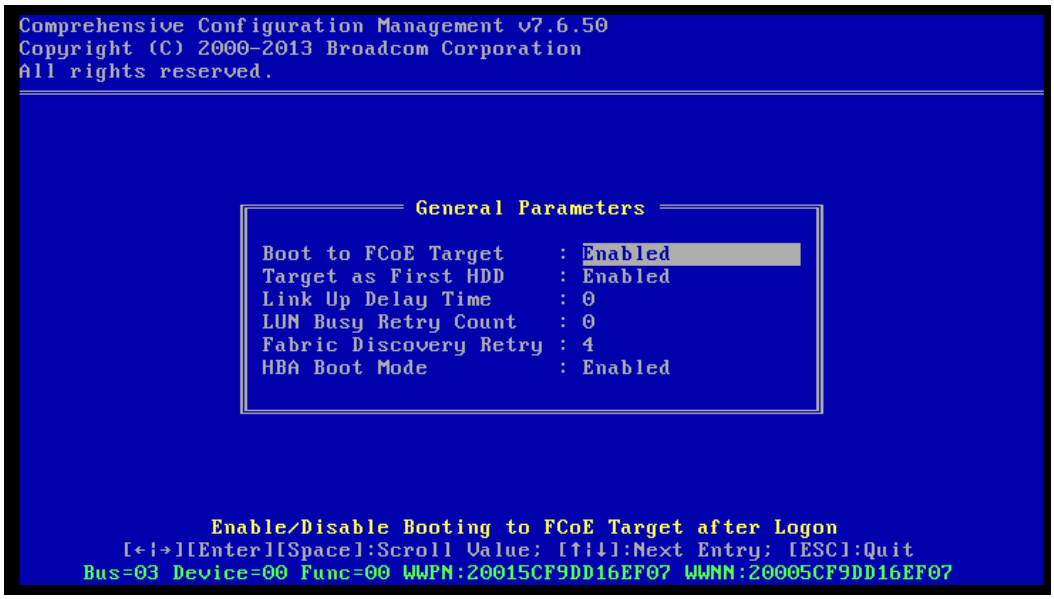


Figure 14 FCoE Boot General Parameters

20. Press the **Escape** key to exit back to the FCoE Boot Main Menu
21. From the **FCoE Boot Main Menu**, highlight **Target Information** and press the Enter key
22. Highlight **No. 1 Target** and press the **enter** key

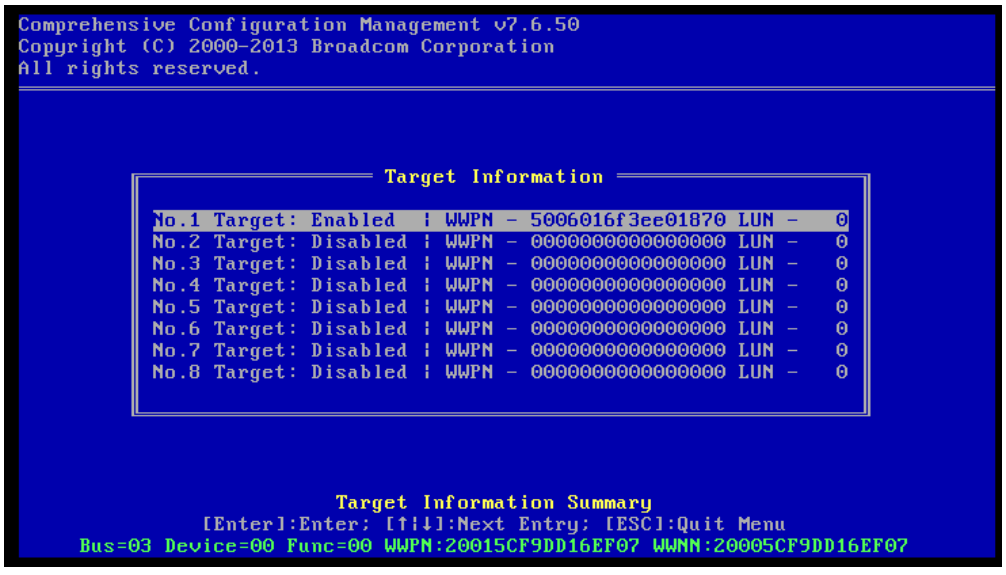


Figure 15 FCoE Boot Target Information

23. Change Connect to Enabled
24. Highlight **WWPN** and press the **enter** key to enter your Storage Processor WWPN information (500601693EE01870 was the WWPN used in this lab configuration).
25. Set **Boot LUN** field to **0** (default for most arrays),

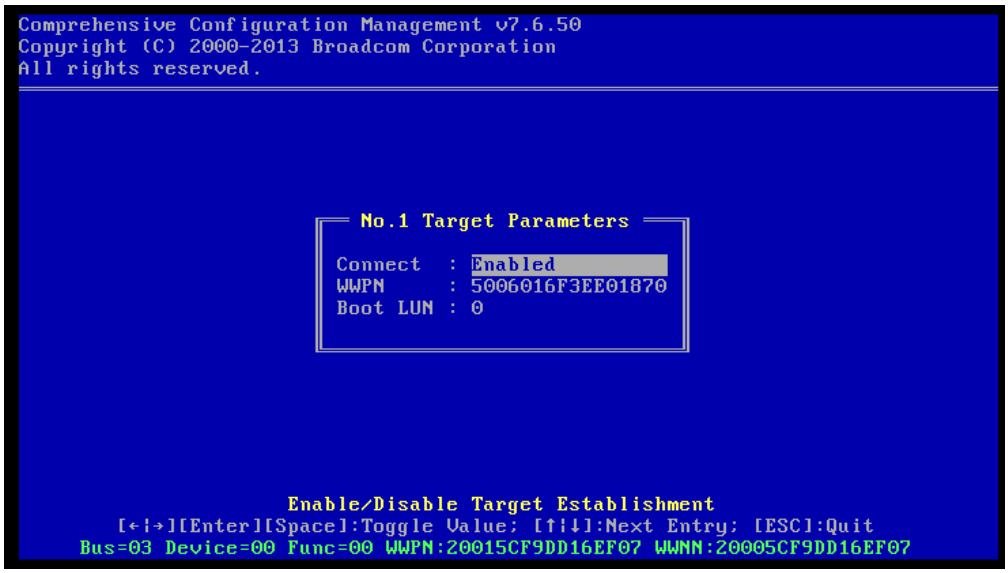


Figure 16 FCoE Target Parameters

26. Press the escape key 2 times to exit back to the Main Menu

Note: The best way to determine the WWPN of Storage Processor is to look at the Name Server or Flogi database on the FC switch or top-of-rack FCF switch.

27. From the **Main Menu** choose **NIC Partition Configuration** and press the **Enter** key if NPAR was enabled earlier during **Device Hardware Configuration**.
28. Flow Control= Auto
29. Select the first partition **PF#0** and press the **Enter** key
30. Ethernet protocol should be **Enabled**
31. iSCSI Offload Protocol should be **Disabled**
32. FCoE Offload Protocol should be **Enabled**
33. All other settings can be left at the default settings
34. Press the **Escape** key to exit back to the NIC Partition Configuration Menu
35. Enter each of the other partitions (I.E PG#2, PF#4 and PF#6) and disable the protocols (Ethernet, iSCSI, and FCoE) for this initial setup

Note: By disabling the other protocols this will ensure that, only one path will be presented to the Array. It will also ensure easier trouble shooting. However for the other partitions to function they must be reconfigured later for Ethernet or applicable protocol.

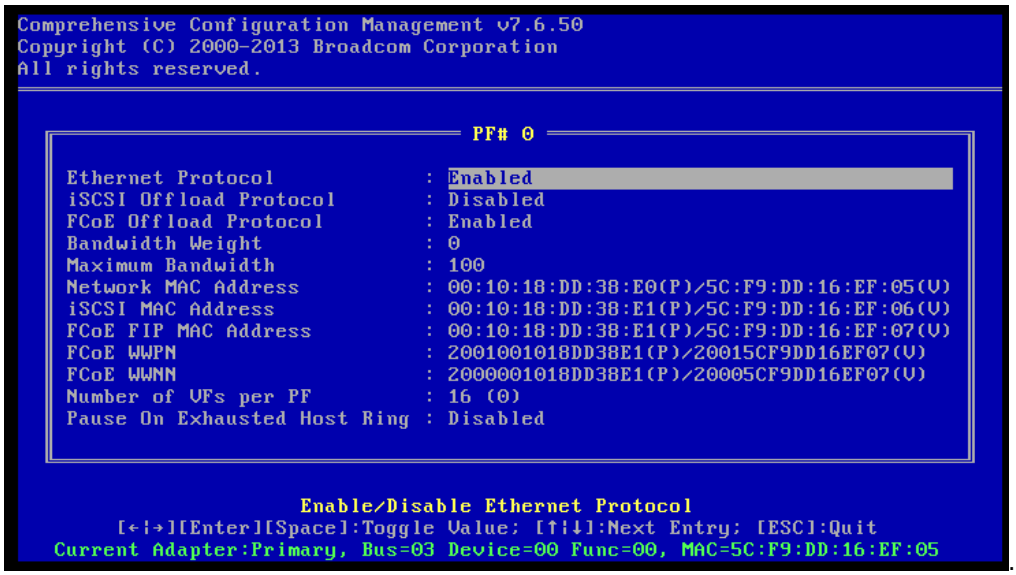


Figure 17 Partition 0 Configuration

36. Press the Escape key until you get back to the Exit Configuration Screen and select Exit and Save Configuration.
37. The Broadcom menu should reappear for any further configuration tasks
38. If no other menu configuration tasks are needed press escape and the system will reboot

3.4.2 Qlogic QME8262 network adapter FCoE configuration steps

To configure the Qlogic QME8262 network adapter the following information is required (if FCoE boot from SAN is desired).

- Storage Processor target WWPN = 50:06:01:6F:3E:E0:18:70 (EMC VNX 5300)
- Boot LUN number 0.

In this example the network adaptor will be configured through the **Life Cycle Controller** pages.

1. Power On the server
2. When you see the Dell Logo press **F10** to enter the **Life Cycle Controller**

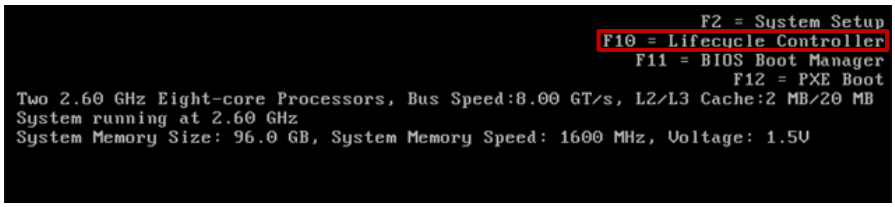


Figure 18 Lifecycle Controller F10 screen

3. On the Left side of the screen, highlight the **System Setup** then select **Advanced Hardware Configuration**.

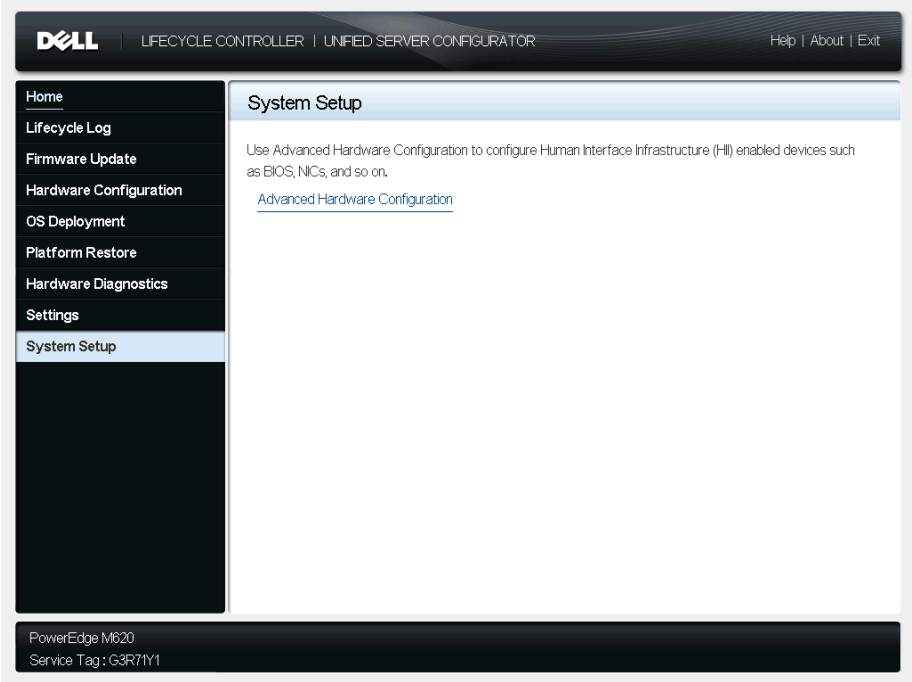


Figure 19 Lifecycle Controller System Setup

4. On the System Setup Main Menu select **Device Settings**

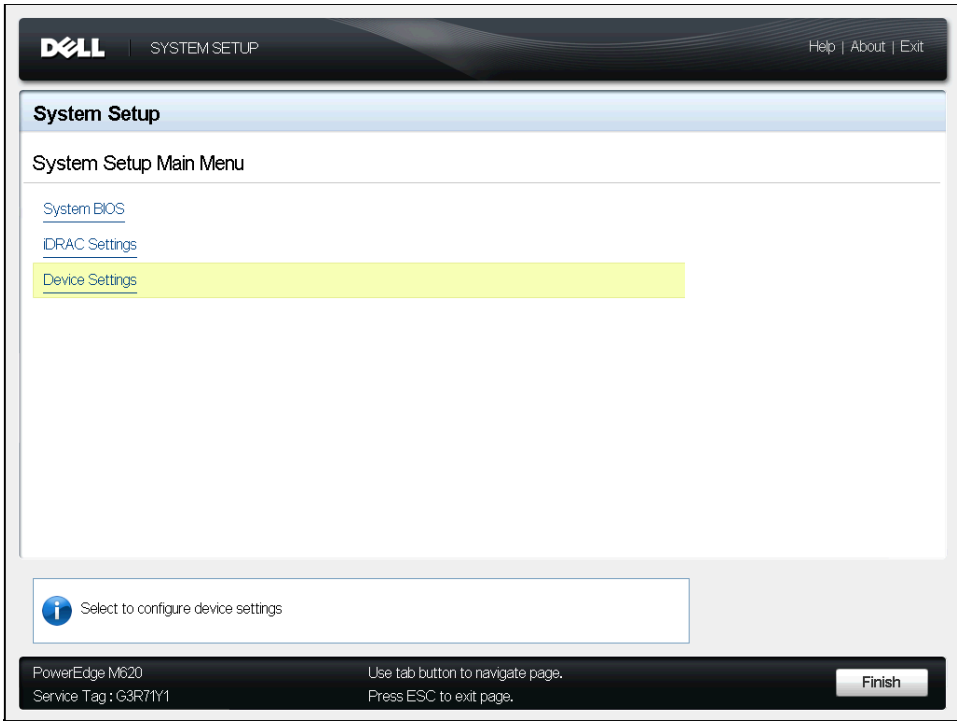


Figure 20 Device Settings Screen

5. From the Device Setting Page select the FCoE QME8262 adaptor. In this configuration, select the **NIC in Mezzanine 2B Port 1** Device.

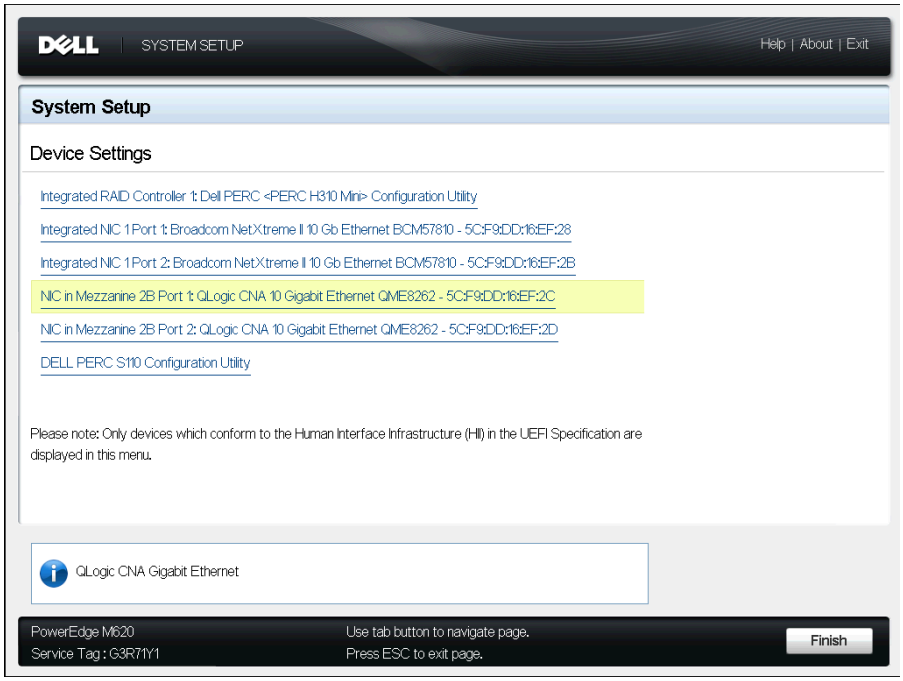


Figure 21 Mezzanine 2B Port 1 Screen

6. From the Main Configuration Page select **NIC Partitioning Configuration**.

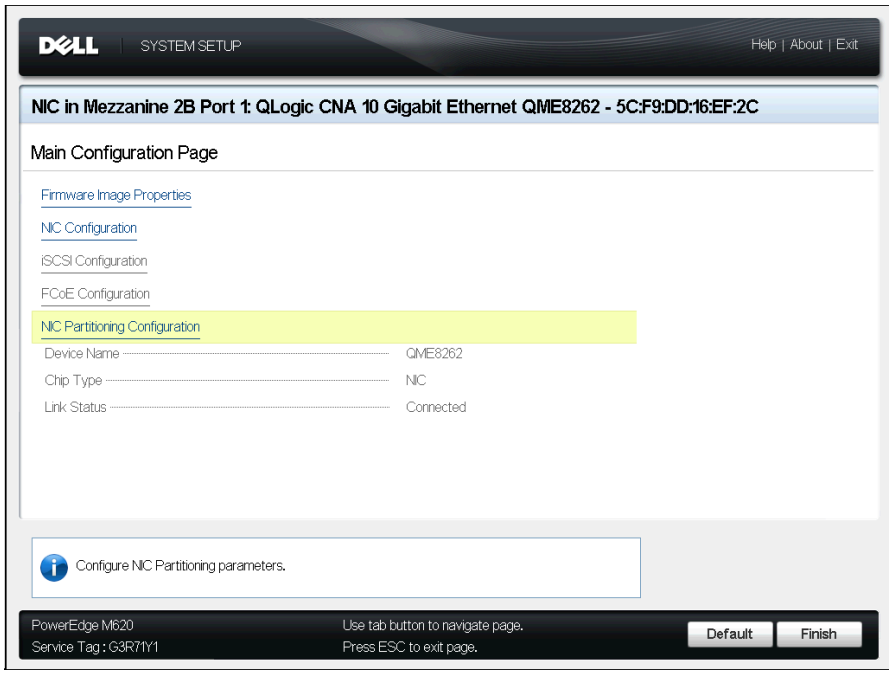


Figure 22 Main NIC Partition Configuration Menu

7. In the NIC Partitioning Configuration page select **Partition 4 Configuration**

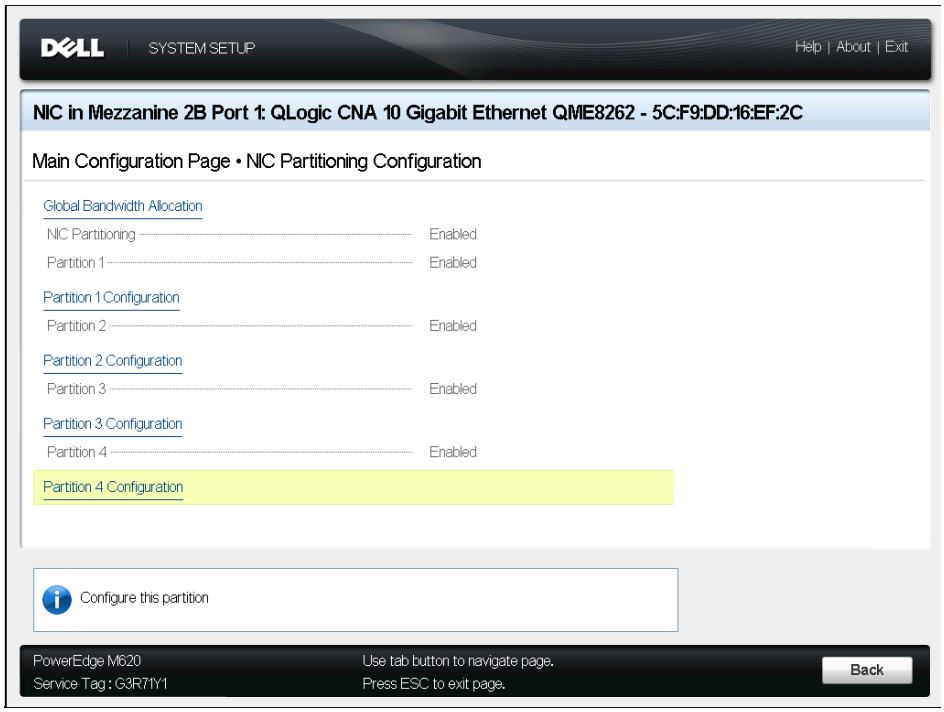


Figure 23 NIC Partition 4 Configuration Screen

8. In the Partition 4 Configuration, Enable FCoE Mode.

Figure 24 Partition 4 Configuration Screen

9. Press the **Back** Button two times to get you back to the Main Page
10. From the Main Page you can now select **FCoE Configuration**

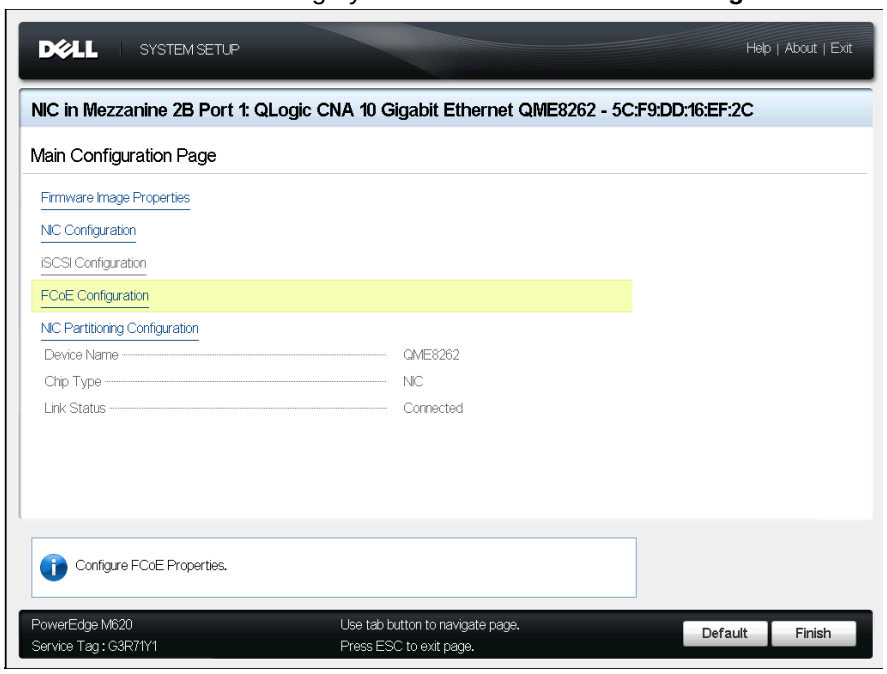


Figure 25 FCoE Configuration main menu

11. In the FCoE Configuration Page make sure Connect is Enabled
12. Enter in your Boot LUN (Typically LUN 0)
13. Enter in your **World Wide Port Name Target** (This configuration uses 50:06:01:6F:3E:E0:18:70)

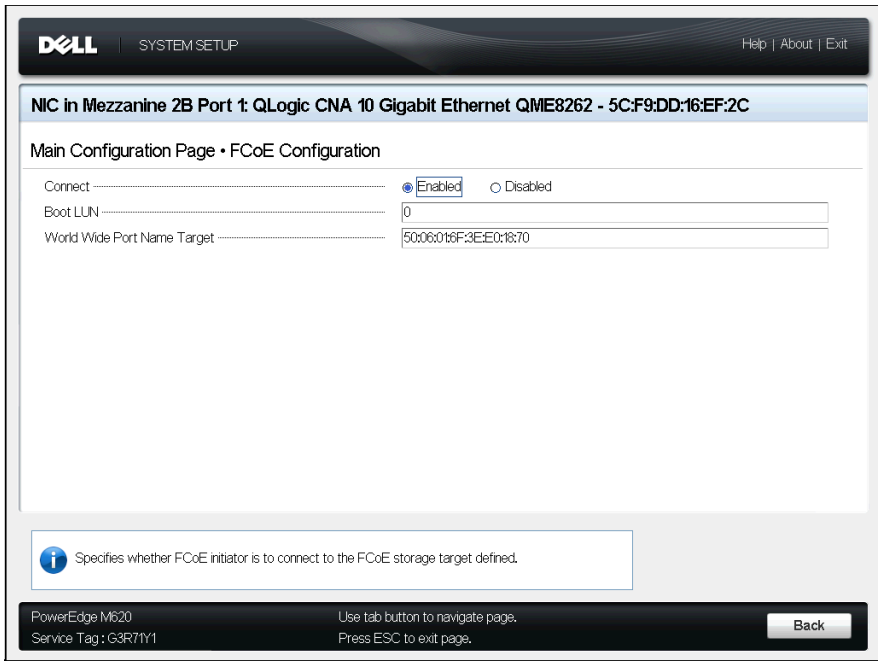


Figure 26 FCoE Configuration Screen

14. Press the **Back** button to finish the configuration and save.
15. Finish and exit until the server restarts.

3.4.3 Intel X520 network adaptor FCoE configuration steps

1. The Intel X520 network adaptor will be initially configured from the **F2 System Setup** program.

Figure 27 F2 System Setup

2. Once F2 is invoked the next step is to double click **Device Settings**.

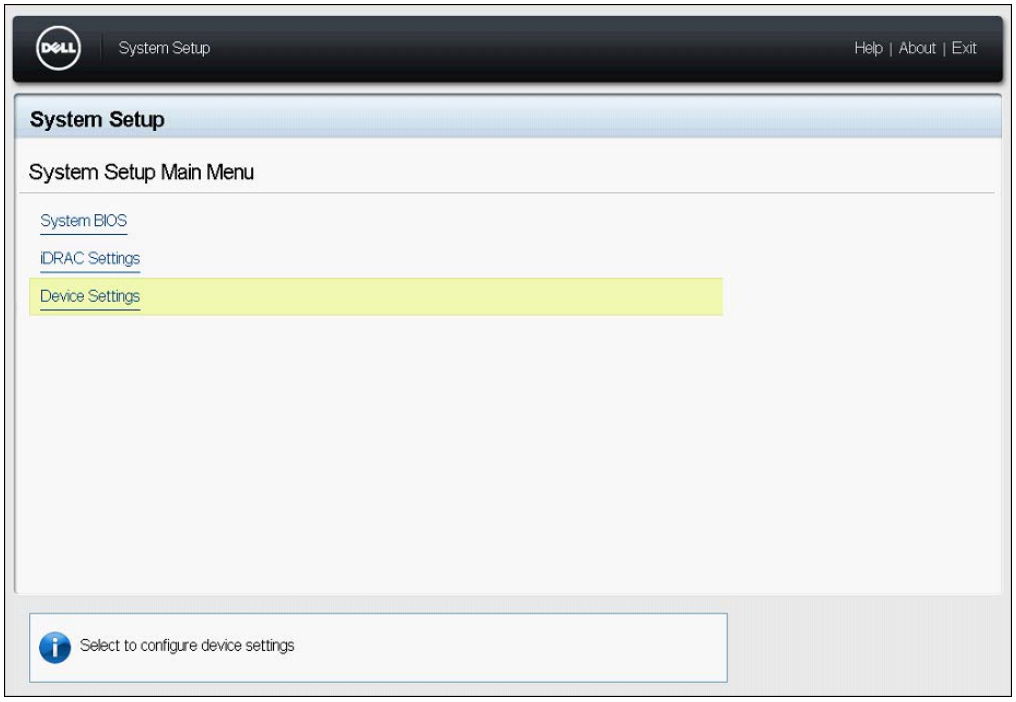


Figure 28 Device Settings main menu

3. In the **Device Settings** initial screen multiple adapters can be seen. The adaptor to be configured is the **Integrated NIC 1Port 1:Intel Ethernet 10G 2P X520-k bNDC – 14:FE:B5:8E:5B:F8**. Port 1 and Port 2 will need to be setup exactly the same for this deployment guide.

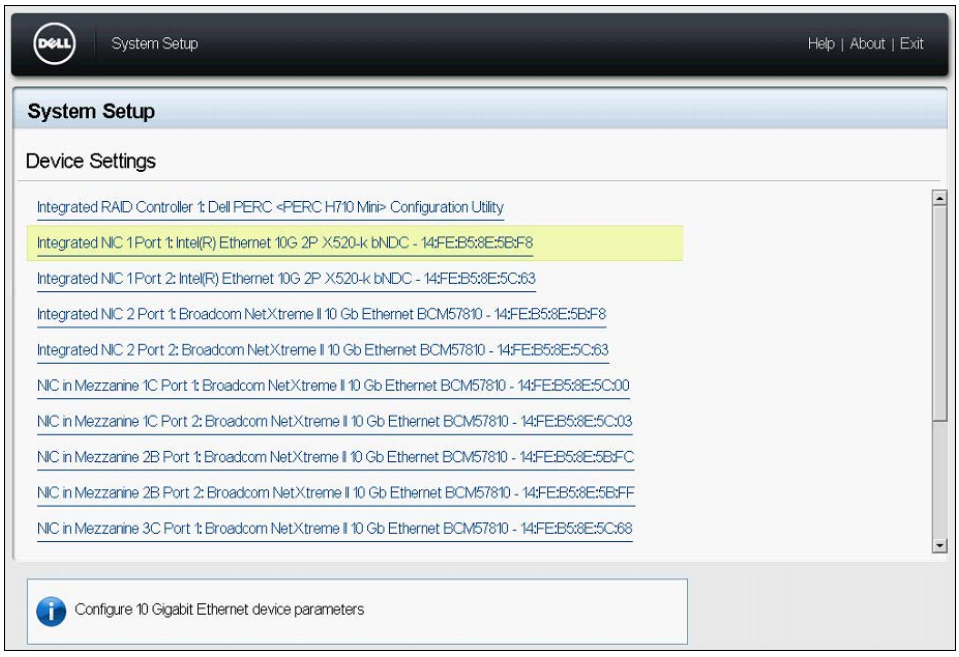


Figure 29 Intel X520 bNDC Port 1

4. Highlight the **NIC Configuration** menu and hit enter.

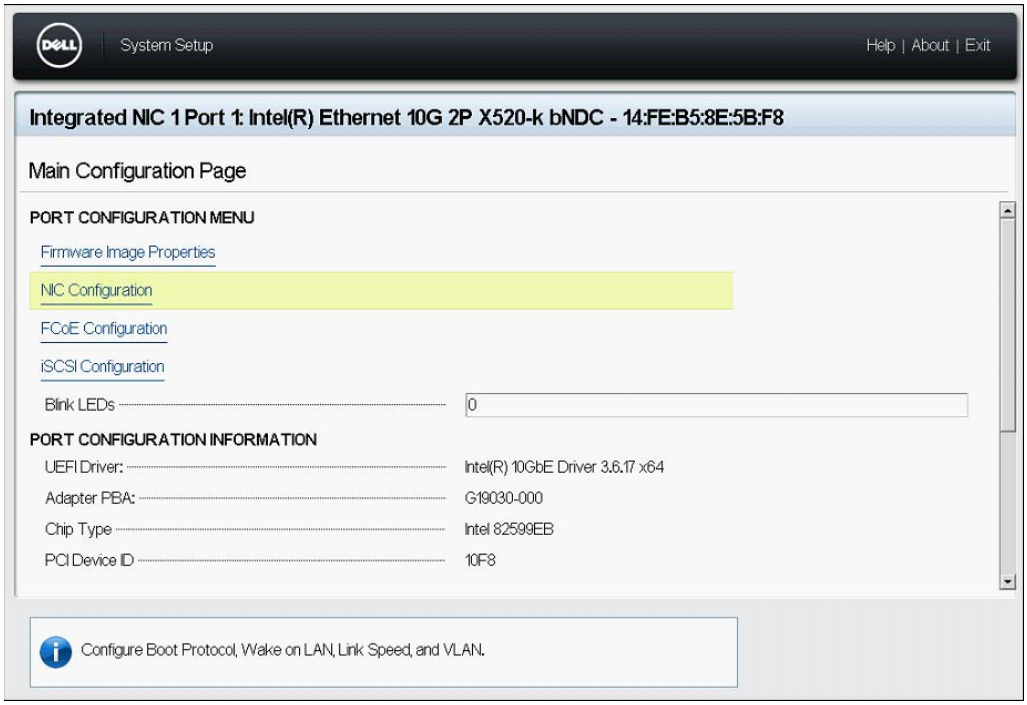


Figure 30 NIC Configuration main menu

5. Highlight the **Legacy Boot Protocol** and choose **FCoE.**

6. Press **Back** to go to the main menu.

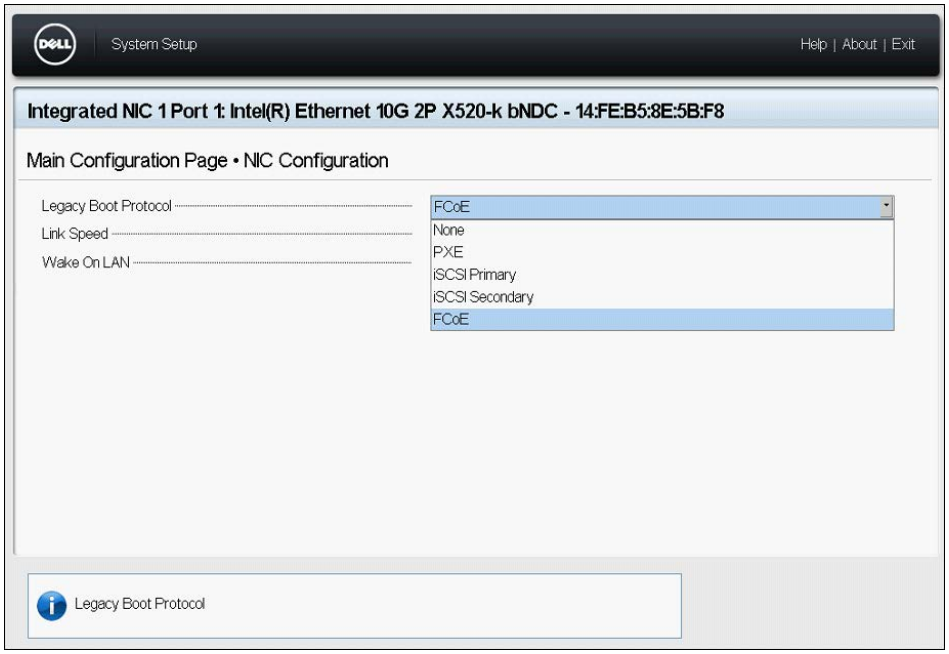


Figure 31 NIC Configuration FCoE Protocol

7. Highlight **FCoE configuration** and hit enter

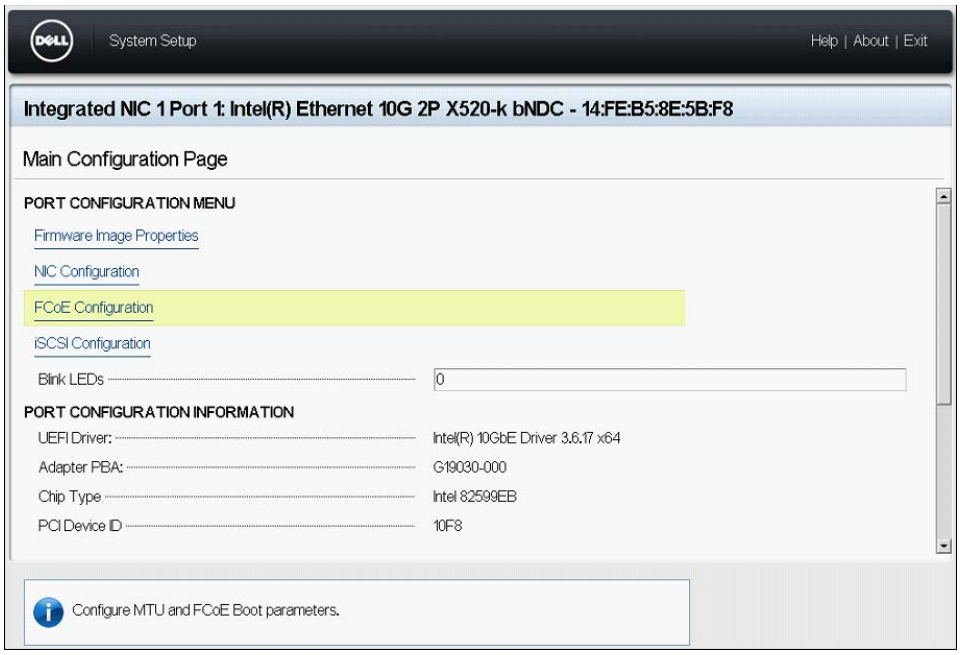


Figure 32 FCoE Configuration main menu

8. Highlight **Enabled** on the Connect menu and hit enter. The other fields would only need to be filled in if a remote Boot from SAN configuration was needed.
9. Press **Back** and hit enter to go to the main menu.

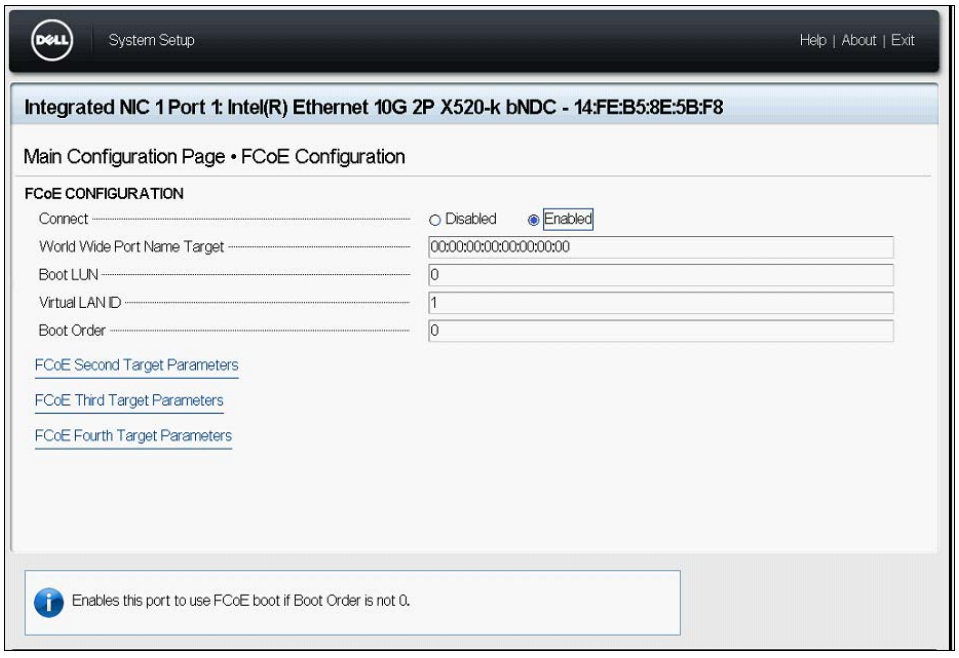


Figure 33 FCoE Configuration enablement

10. Click **Finish** in the lower right of the screen and click on **Yes** to save changes.
11. Click on **Finish** twice to exit the System Setup and click on **Yes** to reboot.

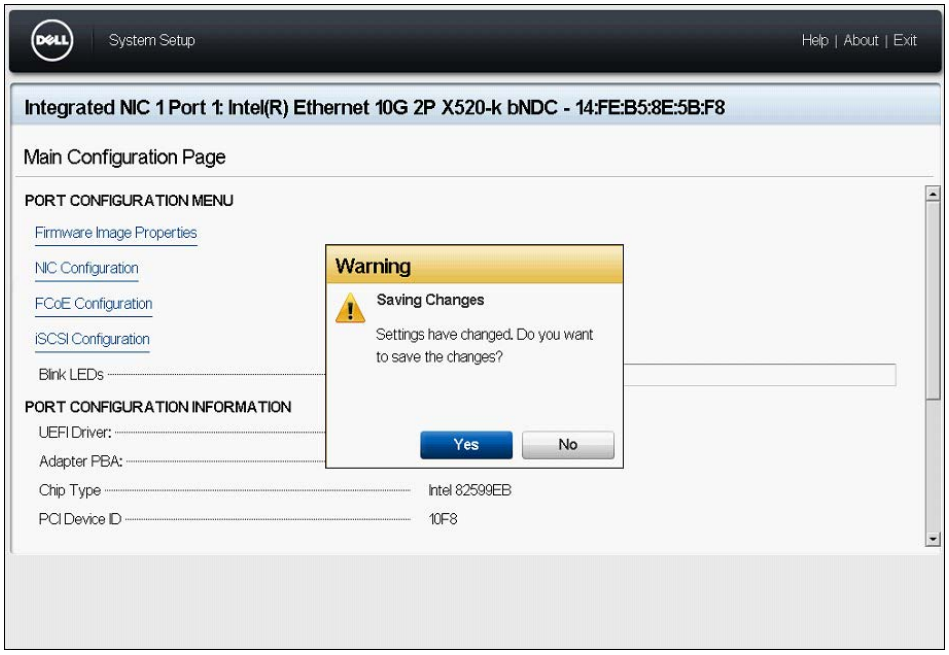


Figure 34 Intel X520 Port 1 change confirmation

4 MXL/IOA FCoE Configuration

The next task will be to configure the MXL and IOA IO modules.

When configuring the IO modules, an understanding of the slot numbers that the blade servers reside in and their relationship to the internal or server facing ports of the MXL/IOA is paramount to understanding what internal port to configure for FC traffic.

4.1 MXL FCoE Configuration

The M820 server that is being provisioned for this document resides in slot 4 and in slot 12 since it is a full height blade. This configuration effort will concentrate on internal port 0/4 of the MXL/IOA.

The first feature to enable is the MXL's FC capabilities.

```
mxl_b2_190#
mxl_b2_190#config
mxl_b2_190(conf)#feature fc
mxl_b2_190(conf)#
```

Figure 35 MXL Feature FC enablement

- Command feature fc. This command allows the MXL to recognize the FC FlexIO module and provision all internal support process's.

Next we'll create a DCB map named test2.

```
mxl_b2_190(conf)#
mxl_b2_190(conf)#dcb-map test2
```

Figure 36 MXL DCB map creation

1. Command dcb-map test2. This command enters the configuration mode of the dcb-map as well as setting the name of this dcb map configuration to test2.

The next commands to run are **priority-group**, and **priority-pgid**

```
mxl_b2_190(conf-dcbmap-test2)#
mxl_b2_190(conf-dcbmap-test2)#priority-group 0 bandwidth 50 pfc off
mxl_b2_190(conf-dcbmap-test2)#priority-group 1 bandwidth 50 pfc on
mxl_b2_190(conf-dcbmap-test2)#priority-pgid 0 0 0 1 0 0 0 0
```

Figure 37 DCB map configuration overview

1. Command **priority-group 0 bandwidth 50 pfc off**. Priority group 0 is for LAN traffic. LAN traffic will use the standard PAUSE frame for flow control, hence Priority flow control (PFC) is disabled.
2. Command **priority-group 1 bandwidth 50 pfc on**. Priority group 1 is lossless for Storage traffic. Storage traffic has PFC turned on, which means that once the receive que on either the Target or Initiator (for Writes or Reads) has reached its buffer full high-water mark, a priority pause frame will be sent to the transmitting port. The transmitting port will stop sending data, the receiving port will send all the storage data in its que for Class of Service (COS) 3. Once the receive que has sent all its data, it will send a PAUSE 0 frame to notify the transmitting port that it is to resume.
3. Both Priority groups have Enhanced Transmission Selection (ETS) set to 50%. ETS allows bandwidth to be limited for both LAN and Storage protocols to no more than 50% during congestion events.
4. Priority group 1 has been assigned to COS 3 with the command **priority-pgid**, while priority group 0 has been assigned COS 0-2, and 4-7.

After creating the DCBx parameters above, the next step is to create an FCoE vlan

The FCoE VLAN can be any non-conflicting numerical designation up to the limit of VLAN 4094. The example below is using **vlan 1001**.

```
mxl_b2_190(conf)#int vlan 1001
mxl_b2_190(conf-if-vl-1001)#no ip address
mxl_b2_190(conf-if-vl-1001)#no shutdown
mxl_b2_190(conf-if-vl-1001)#
```

Figure 38 FCoE VLAN creation

1. Command **int vlan 1001**. This command enters the configuration mode for setting a vlan numerical designation.
2. Command **no ip address**. Since this vlan is for storage traffic no ip address is needed.
3. Command **no shutdown**. This command activates the vlan for storage traffic.

Note: For troubleshooting purposes, in a highly available multiple FC fabric environment, each FCoE VLAN should be unique.

Once the FCoE VLAN has been created, the next step is to enter the FCoE map parameters.

```
mxl_b2_190(conf)#
mxl_b2_190(conf)#fcoe-map fc2
mxl_b2_190(conf-fcoe-fc2)#fc-map 0efc01
mxl_b2_190(conf-fcoe-fc2)#fabric-id 1001 vlan 1001
mxl_b2_190(conf-fcoe-fc2)#
```

Figure 39 FCoE map creation

1. Command fcoe-map fc2. This command enters the fcoe-map configuration mode as well as setting the name (in this configuration instance fc2 was used)
2. Command fc-map 0efc01. This is a unique 24 bit MAC address pre-fix that is used to generate a 48 bit Fabric Provided MAC Address (FPMA). The second 24 bits of the FPMA are provided by the Fibre Channel switch and it's called a Port ID (PID) or FC_ID. The PID is based on what port

the FC FlexIO module is connected to as well as what domain ID the switch has currently. For example if the FC FlexIO module is connected to port 1 and the Fibre Channel switch has a domain ID of 2 then the 24 bit Port ID address would be 02:01:00. This means that the FPMA address would be 0e:fc:01:02:01:00.

3. Command fabric-id 1001 vlan 1001. The fabric-id and the FCoE VLAN must match since the fabric-id is what is used to bind the VLAN to the FCoE map. The fabric-id is also used to identify the SAN fabric that the FC traffic is to be forwarded to.

Note: For troubleshooting purposes, in a highly available multiple FC fabric environment, each FCoE map should be unique ie. 0efc01, 0efc02, etc. This ensures that no FPMA 48 bit address could be the same between FC fabric's.

The DCB map, FCoE VLAN, and FCoE map have been created. These values will now be applied to internal port 0/4.

```
mxl_b2_190(conf)#
mxl_b2_190(conf)#int te 0/4
mxl_b2_190(conf-if-te-0/4)#no ip address
mxl_b2_190(conf-if-te-0/4)#mtu 12000
mxl_b2_190(conf-if-te-0/4)#portmode hybrid
mxl_b2_190(conf-if-te-0/4)#switchport
mxl_b2_190(conf-if-te-0/4)#fcoe-map fc2
mxl_b2_190(conf-if-te-0/4)#dcb-map test2
mxl_b2_190(conf-if-te-0/4)#no shutdown
mxl_b2_190(conf-if-te-0/4)#
```

Figure 40 MXL Internal Port 0/4 FCoE provisioning

Internal port 0/4's configuration commands are:

1. Command **int te 0/4**. This command enters internal port 0/4's configuration mode.
2. Command **no ip address**. Storage traffic does not require an IP address to function.
3. Command **mtu 12000**. This is the maximum size an Ethernet frame can be. FCoE requires a mini-jumbo frame of 2500 in order to encapsulate Fibre Channels frame size of 2148. Typical MTU sizes seen on various FCoE configurations are, 2500, 9260, or 12000. The main goal of a 2500 byte mini-jumbo frame is to avoid fragmentation of the FC frame. With fragmentation comes disassembly and reassembly overhead which always incurs latency, as well the potential for incorrect CRC re-calculation, and header re-creation.
4. Command **portmode hybrid**. This command ensures internal port 0/4 can transport tagged and untagged frames.
5. Command **switchport**. This command puts the port in a layer 2 only mode. Since FCoE does not contain an IP layer and operates directly above Ethernet in the network protocol stack it is not routable.
6. Command **fcoe-map fc2**. This command applies the 24 bit fc map, the fabric-id, and storage VLAN to port 0/4
7. Command **dcb-map test2**. This command applies a template that assigns PFC, ETS, and COS values to port 0/4
8. Command **no shutdown**. This last command allows the port to become active.

Now that the MXL's internal port 0/4 has been configured, the next step is to apply a FCoE map to an external fibre channel port on the FC FlexIO module. This example will use the second external port of the FC FlexIO module in IO bay one. Referring back to the Dell Networking MXL/IOA - Port Mapping table, it can be seen that the port is numbered 0/42.

```
mxl_b2_190#
mxl_b2_190#config
mxl_b2_190(conf)#int fibreChannel 0/42
mxl_b2_190(conf-if-fc-0/42)#fabric fc2
mxl_b2_190(conf-if-fc-0/42)#no shutdown
mxl_b2_190(conf-if-fc-0/42)#
```

Figure 41 MXL External Port 0/42 configuration

External Port 0/42's settings are:

- Command **int fibreChannel 0/42**. This command enters fibreChannel 0/42's configuration mode.
- Command **fabric fc2**. This applies a template for FCoE and FC parameters
- Command **no shutdown**. This enables the port

4.2 IOA FCoE Configuration

Prior to continuing on with the configuration of the Cisco MDS 9148, the IOA will be configured for the FC FlexIO module. Since the IOA is considered a zero-touch self-configuring layer 2 switching module, The IOA does not require the configuration steps which the MXL needed. The show commands below will allow the health of the automatic configuration to be confirmed.

1. The first command is **show qos dcb-map SAN_DCB_MAP**

```
IOA_A2_190#show qos dcb-map SAN_DCB_MAP
----------------------

State   :Complete
PfcMode:ON
--------------------
PG:0 TSA:ETS  BW:30  PFC:OFF
Priorities:0 1 2 5 6 7


PG:1 TSA:ETS  BW:30  PFC:OFF
Priorities:4


PG:2 TSA:ETS  BW:40  PFC:ON
Priorities:3


IOA A2 190#
```

Figure 42 IOA DCB-Map command

The command **show qos dcb-map SAN_DCB_MAP** shows that there are:

- Three (3) default priority groups created.
- Priority groups 0 and 1 have Priority Flow Control (PFC) turned off
- Priority group 2 has it turned on.
- Enhanced Transmission Selection (ETS) bandwidth allocation is 40% for priority group 2 with a Class of Service (COS) of 3, and 30% for priority groups 0 and 1 with their respective Classes of service.

2. The next command is **show fcoe-map**

```
IOA_A2_190#show fcoe-map

Fabric Name          SAN_FABRIC
Fabric Id            1002
Vlan Id              1002
Vlan priority        3
FC-MAP               0efc00
FKA-ADV-Period       8
Fcf Priority         128
Config-State         ACTIVE
Oper-State           UP
Members
Fc 0/41 Fc 0/42 Fc 0/43 Fc 0/44 Fc 0/45 Fc 0/46 Fc 0/47 Fc 0/48
Te 0/4 Te 0/12
IOA A2 190#
```

Figure 43 IOA Show FCoE-Map command

The command **show fcoe-map** confirms:

- The state of the FCoE session is active and up
- The FC-MAP is 0efc00
- The FCoE VLAN is 1002

3. Another useful command is **show config** on internal port 4.

```
IOA_A2_190#config
IOA_A2_190(conf)#interface tengigabitethernet 0/4
IOA_A2_190(conf-if-te-0/4)#show config
!
interface TenGigabitEthernet 0/4
 mtu 12000
 portmode hybrid
 switchport
 auto vlan
 dcb-map SAN_DCB_MAP
 fcoe-map SAN_FABRIC
!
 protocol lldp
  advertise management-tlv management-address system-name
 no shutdown
IOA A2 190(conf-if-te-0/4)#
```

Figure 44 IOA Internal port 0/4 show config

While it was already known what the name of the dcb and fcoe maps were. A key take away is that there was no need to configure anything on internal port 0/4.

4. The final command is **show npiv devices**

```
IOA_A2_190#show npiv devices

ENode[0]:
ENode MAC    :   14:fe:b5:8e:5e:c9
ENode Intf   :   Te 0/4
FCF MAC      :   d0:67:e5:ac:ac:d4
Fabric Intf  :   Fc 0/42
FCOE Vlan    :   1002
Fabric Map   :   SAN_FABRIC
ENode WWPN   :   20:01:14:fe:b5:8e:5e:c9
ENode WWNN   :   20:00:14:fe:b5:8e:5e:c9
FCoE MAC     :   0e:fc:00:eb:02:01
FC-ID        :   eb:02:01
LoginMethod  :   FLOGI
Secs         :   2131
Status       :   LOGGED_IN
IOA_A2_190#
```

Figure 45 IOA Show NPIV Devices command

The command **show npiv devices** confirms what we have seen from the previous show commands:

- The FC-ID
- The FCoE MAC address
- The ENode/network adaptor interface
- The Fabric/External FC interface
- The Status is LOGGED_IN

5 Cisco MDS 9148 FC Switch Configuration

The configuration of the MXL and IOA is complete. The next step is to login to the Cisco MDS 9148 Fibre Channel switch and configure it to accept the logins from the FC FlexIO module's port 0/42, and the EMC VNX 5300.

1. The initial configuration step will be to issue a **no shut** on the MDS 9148 ports 15 and 16 connected to the FC FlexIO module and the EMC VNX 5300 storage array respectively.

```
MDS9148# config
Enter configuration commands, one per line.  End with CNTL/Z.
MDS9148(config)# int fc1/15
MDS9148(config-if)# no shut
MDS9148(config-if)# int fc1/16
MDS9148(config-if)# no shut
MDS9148(config-if)#
```

Figure 46 Cisco MDS 9148 No Shut command

The Cisco MDS9148 ports have been enabled:

- Port 15 is connected to the FC FlexIO module. It could also be considered as a host or initiator port.
- Port 16 is connected to the EMC VNX 5300 array. It could also be considered the target port.

2. The next step will be to enable the feature NPIV with the command **feature npiv**.

```
MDS9148# config
Enter configuration commands, one per line.  End with CNTL/Z.
MDS9148(config)# feature npiv
MDS9148(config)#
```

Figure 47 Cisco MDS9148 Feature NPIV command

NPIV stands for Node Port Identification Virtualization. This N-Port ID Virtualization feature allows multiple logins per physical connection. This means that for example, if a physical host has multiple virtual machines residing on it, the FC switch can give each virtual machine its own:

- 64-bit World Wide Node Name (WWNN)
- 64- bit World Wide Port Name (WWPN)
- 24 bit Port ID(PID)/FC-ID. If a PID has a non-zero byte in the last octet of its address, this means that it was a network adaptor or virtual machine behind the FC FlexIO module that logged into the FC switch: For example if the PID is 02:01:01, this would indicate that (1) network adaptor had successfully logged into the FC switch. The original PID was 02:01:00, which points towards the FC FlexIO module logging in to the FC switch. Any other value in the last byte would have to be an NPIV device such as a network adaptor or a hypervisor based virtual machine.

3. Now that NPIV has been enabled, VSAN 2227 can be created and interfaces 15 and 16 can be added to it

```
MDS9148(config)# vsan database
MDS9148(config-vsan-db)# vsan 2227
MDS9148(config-vsan-db)# vsan 2227 interface fc1/15
Traffic on fc1/15 may be impacted. Do you want to continue? (y/n)  [n] y
MDS9148(config-vsan-db)# vsan 2227 interface fc1/16
Traffic on fc1/16 may be impacted. Do you want to continue? (y/n)  [n] y
MDS9148(config-vsan-db)#
```

Figure 48 Cisco MDS9148 VSAN creation

VSAN creation:

- Command **vsan database**. This command enters the VSAN database for creation of VSAN's
- Command **vsan 2227**. This command creates a VSAN database instance
- Command **vsan 2227 interface fc1/15**. This command adds port 15 to VSAN 2227.
- Command **vsan 2227 interface fc1/16**. This command adds port 16 to VSAN 2227

VSAN's allow:

- Traffic isolation between initiator/target pairs
- Separate configuration settings
- Production on one VSAN and testing in another VSAN, while still utilizing the same physical switch

Note: Since ports 15 and 16 can reside in additional VSANs, the warning about traffic disruption will appear. This is a normal message and will occur in production environments when multiple hosts and targets are shared across multiple VSANs. Normal data center requirements would invoke an outage or maintenance window to ensure IO disruption of critical business applications is kept to a minimum.

Once the VSAN is created, it needs to be placed inside a zone.

4. The zone is given a name, then bound to a VSAN.

```
MDS9148# config
Enter configuration commands, one per line.  End with CNTL/Z.
MDS9148(config)# zone name zone2228 vsan 2227
MDS9148(config-zone)# member interface fc1/15
MDS9148(config-zone)# member interface fc1/16
MDS9148(config-zone)#
```

Figure 49 Cisco MDS9148 Zone creation

Zone creation:

- Command **zone name zone2228 vsan 2227**. This command creates zone2228 and adds VSAN 2227 to it.
- Command **member interfaces fc1/15**. This commands adds member interface fc1/15 to zone2228. Even though interface fc1/15 is existing inside of VSAN 2227, the interface must be explicitly added to the zone
- Command **member interfaces fc1/16**. This commands adds member interface fc1/16 to zone2228. Even though interface fc1/16 is existing inside of VSAN 2227, the interface must be explicitly added to the zone

5. Now that the zone is created and the FC ports are assigned. A zoneset must be created as an overall "container" for the zones, the vsans, and the FC ports.

```
MDS9148# config
Enter configuration commands, one per line.  End with CNTL/Z.
MDS9148(config)# zoneset name set2229 vsan 2227
MDS9148(config-zoneset)# member zone2228
MDS9148(config-zoneset)# zoneset activate name set2229 vsan 2227
Zoneset activation initiated. check zone status
MDS9148(config)#
```

Figure 50 Cisco MDS9148 Zoneset creation

Zoneset creation:

- Command **zoneset name set2229 vsan 2227**. This command creates the zoneset container set2229 and links VSAN 2227 to it.
- Command **member zone2228**. This command adds zone2228 to zoneset container set2229.
- Command **zoneset activate name set2229 vsan 2227**. This command activates the zone and allows the initiator and target to login to the switch.

6. Once the zoneset is activated, there are a number of commands on the Cisco MDS9148 FC switch that can be invoked to check for proper connectivity.

```
MDS9148(config)# show zoneset active vsan 2227
zoneset name set2229 vsan 2227
  zone name zone2228 vsan 2227
  * fcid 0xeb0001 [interface fc1/15 swwn 20:00:54:7f:ee:50:45:e8]
  * fcid 0xeb0000 [interface fc1/15 swwn 20:00:54:7f:ee:50:45:e8]
  * fcid 0xeb0100 [interface fc1/16 swwn 20:00:54:7f:ee:50:45:e8]

  zone name $default_zone$ vsan 2227
MDS9148(config)#
```

Figure 51 Cisco MDS9148 Show Zoneset Active VSAN 2227 command

- Command **show zoneset active vsan 2227**. This command can be used to confirm that Zone2228, along with vsan 2227 and its member FC ports are activated and logged in
- FC port 15 has two fcid's.
    a. **fcid 0xeb0001**. This fcid represents the network adaptor behind the FC FlexIO module since the last 8 bits of the 24 bit fcid is a non-zero value
    b. **fcid 0xeb0000**. This fcid represents the FC FlexIO module since the last 8 bits of the 24 bit fcid is a zero value
- FC port 16 has one fcid.
    a. **fcid 0xeb0100**. This fcid represents the EMC VNX 5300. Port fc1/16 is what the storage array is connected to, but, further confirmation can be realized with the **show fcns database** command below.

7. This next command can be used to confirm if the FC ports have come online and registered with the FC name server database.

```
MDS9148# show fcns database vsan 2227

VSAN 2227:
--------------------------------------------------------------------------
FCID         TYPE  PWWN                    (VENDOR)         FC4-TYPE:FEATURE
--------------------------------------------------------------------------
0xeb0000     N     20:2a:00:1e:c9:f1:05:42                   
0xeb0001     N     20:01:14:fe:b5:8e:5e:c9                   scsi-fcp:init
0xeb0100     N     50:06:01:68:3e:e0:18:70 (Clariion)        scsi-fcp

Total number of entries = 3
MDS9148#
```

Figure 52 Cisco MDS9148 Show FCNS Database VSAN 2227 command

- Command **show fcns database vsan 2227**. This command confirms that the devices have logged in successfully.
- The FC FlexIO module has logged in. The PID/FCID 0xeb0000 confirms that it is a full fabric device logging in due to the last byte being 00.
- The network adaptor has logged in. The PID/FCID 0xeb0001 confirms that an NPIV device has logged in due to the last byte being 01
- The EMC VNX 5300 storage array has logged in. The PID/FCID 0xeb0100 confirms that it is a full fabric device logging in due to the last byte being 00. The secondary indication is due to the Vendor field being filled in with the name "Clariion". This name is unique only to EMC.

Note: FCNS stands for Fibre Channel Name Server. A Fibre Channel Name Server service is a central repository of device attributes. These attributes are available to any device on the local Fibre Channel fabric. Typical usage is for an Initiator/Host to query the Name Server for any storage arrays in the Fibre Channel fabric.

6 EMC VNX 5300 Storage Array initiator confirmation

The final configuration step is to provision the EMC VNX 5300 storage array.

1. Confirm EMC storage can see the initiators.
2. Create new Host and Register initiator with storage device, typically done same time on first initiator.
3. Create required number of LUNs
4. Next Create storage group, this will be listed under the Host Menu
5. Add Host and LUNs to new storage group.
   a. Confirm EMC storage sees the initiators
   b. From tab selection Menu , select “System List”, EMC VNX 5300 link
   c. A set of tab selections will be presented, select Host tab > initiators
   d. Look for initiator in list, it should show as not registered. Select “Not Registered” from Connection status to help filter initiators. It is also possible to use the last four digits of the WWPN of the CNA as a filter.

Note: The order of the following steps is just one suggested manner of implementation, there are several other methods for achieving the same end configuration,

6. Create New Host and register initiator.
   a. Select the row with desired initiator
   b. On bottom left corner of screen, select Register
   c. A New window will show up, use the following suggested settings
      i. Initiator type= Clarion VNX (for Windows/VMware)
      ii. Fail over Mode=Active-Active (default)
      iii. Select New Host
      iv. Enter host name , it may be preferable to include Rack_Server_Fabric_SAN info (just to help with later identification)
      v. IP address is needed even if not using host agent or PowerPath, IP address can be a fake address in order to create/register the initiator.
      vi. Select OK
   d. There will be several prompts to continue/ confirm your request.
   e. At this point the final success box will appear, select OK when prompted
   f. Continue registering the other initiators if multiple paths are available.
   g. Select “existing host” (the one just created)
      i. Browse Host…
      ii. In the filter box, host name can be typed in partially to make the search easier if there are several hosts created.
      iii. Select the host desired and select OK
   h. Create desired number of LUNs.
      i. From top menu tabs, select Storage>LUNs
      ii. At the bottom left select “create” to open the Create LUNs dialog box and use the following suggested settings
         1. Storage type: Leave at Pool
         2. RAID type: Select default this example uses RAID 6
         3. Select Storage Pool for new LUN, this example uses “Pool 0”

4. LUN Properties
    a. User Capacity: select size, for ESX host at least 20GB minimum
    b. LUN ID : allow system to assign, no selection required
5. Number of LUNS to Create: Enter a number based on the infrastructure storage needs. (Note LUNs assigned to SPA and SPB in round robin) So the storage should be configured with access to SPA and SPB.
6. LUN Name: leave at default, which automatically assigns LUN IDS.
7. Select “Apply”

7. Create storage group (listed under the Host Menu).
    a. From top menu tabs, select Host > Storage Groups
    b. Mid to Bottom left , select “Create”
        i. Create Storage box is presented
            1. Enter Storage Name (Storage admins will typically have protocols for this naming convention)
            2. Select OK, (this will show success then step through to add LUNs and Host)

8. Add Host and LUNs to new storage group
    a. From LUNs tab, Select LUNs to add, note SPA and SPB should show the applicable LUNs
    b. Select “add”, LUN moved to selected LUNs
    c. Select “Hosts” tab
        i. Select Host from Available Hosts
        ii. Select right pointing arrow to move Host to “Hosts to be Connected”
        iii. Select “OK”

6.1 Two ways to confirm the Host LUN ID info

- This can be viewed through the storage group and note the Host LUN ID
    a. From Storage group list: Host >Storage group
    b. Select Storage group desired
    c. In the Details window, bottom of web page, select LUNs tab
    d. View the LUNs assigned to the storage group
    e. The far right column will list the “Host LUN ID”
- From Host > Host List
    a. Select Host from List
    b. In the Details pane, far right column will show Host LUN ID.

The Cisco MDS FC switch is the device providing fabric services

```
MDS9148# show flogi database
--------------------------------------------------------------------------------
--------
INTERFACE        VSAN    FCID           PORT NAME               NODE
NAME
--------------------------------------------------------------------------------
--------
fc1/1            2     0x850000  20:1d:54:7f:ee:56:55:40
20:02:54:7f:ee:56:55:41
fc1/1            2     0x85001d  20:01:e0:db:55:1c:82:91
```

Figure 53 “show flogi database” from the Cisco MDS 9148

This command shows that the EMC VNX 5300 storage array performed a successful fabric login .

7 Operating System Confirmation of FC FlexIO LUN Presentation

Windows Server 2012 is the operating system that will be used to confirm that all devices are properly configured to present a LUN in Disk Management.

1. In order to access Disk Management in Windows 2012, the procedure as can be seen in Figure 54 is to first click on the Server Management icon in the lower left hand corner of the screen.
2. Then click on Tools in the upper right hand corner of the Server Management applet.
3. Then click on Computer Management.

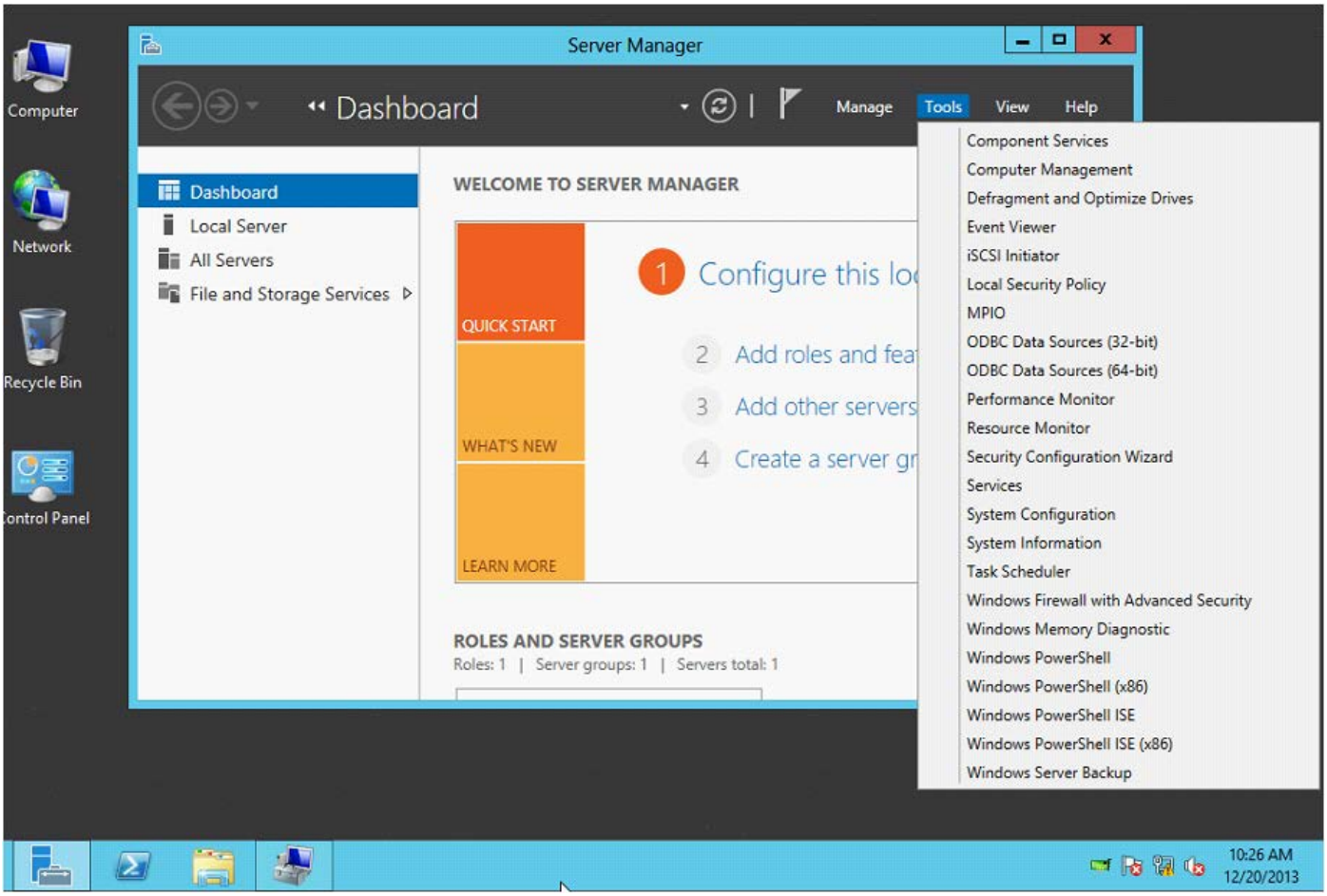


Figure 54 Server Manager Tools Menu

4. Once the Computer Management applet has launched, click on Disk Management and the new remote LUN should be seen as a Basic Disk as can be seen in Figure 55.
5. This new LUN may need to be brought online from a default offline state by right clicking on the new LUN and choosing “Online”.
6. If the new LUN does not appear, right click on Disk Management and choose “Rescan Disk”.

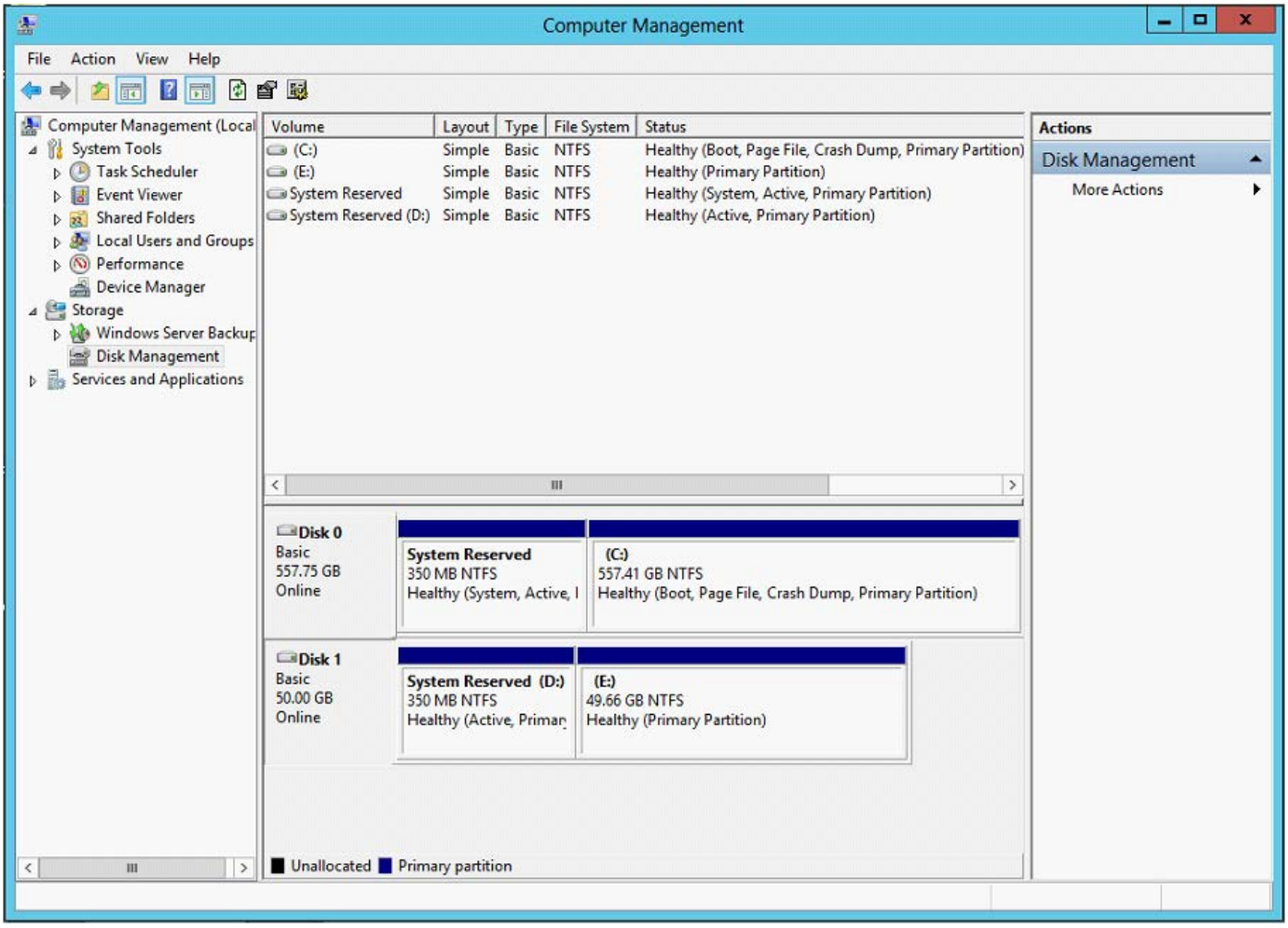


Figure 55 Disk Management

7. Once the new LUN has been discovered and brought online, right click next to Basic and choose properties. As can be seen in Figure 56, DGC VRAID Multi-Path Disk Device is the identification information sent from the EMC VNX 5300 storage array.

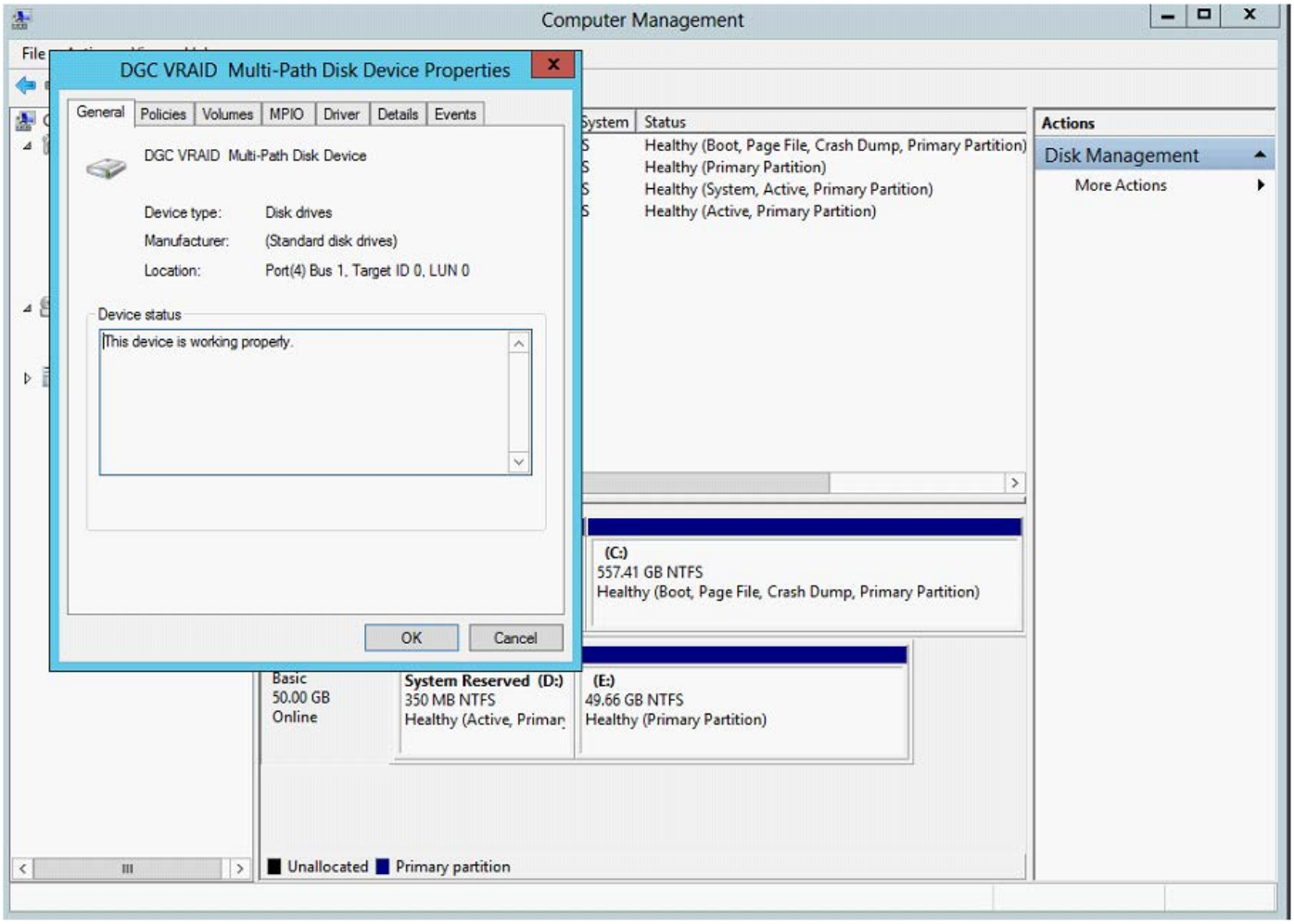


Figure 56 Remote LUN Properties

8 Basic Terminology

CLI

Command line interface (CLI). This is the text-based telnet, SSH, or serial type interface that are used for entering commands into the Dell Networking MXL model switch. At release, the MXL is only configurable via CLI.

CMC

Chassis management controller (CMC) is the module, which controls the Dell PowerEdge™ M1000e blade server chassis. Through this controller a telnet, SSH or serial based connection can be used to manage the MXL switch.

ETS

Enhanced Transmission Selection (ETS) is defined in the IEEE 802.1Qaz standard (IEEE, 2011). ETS supports allocation of bandwidth amongst traffic classes. It then allows for sharing of bandwidth when a particular traffic class does not fully utilize the allocated bandwidth. The management of the bandwidth allocations is done with bandwidth-allocation priorities, which coexist with strict priorities (IEEE, 2011).

FIP-Snooping

With FIP-Snooping enabled on the Dell Networking MXL model switch, FIP logins, solicitations, and advertisements are monitored. In this monitoring or snooping process, the switch gathers information pertaining to the ENode and FCF addresses. With this information, the switch will then place filters that only allow access to ENode devices that have logged-in successfully. This enables the FCoE VLAN to deny all other traffic except this lossless FCoE storage traffic.

The filtering process also secures the end-to-end path between the ENode device and the FCF. The ENode will only be able to talk with the FCF in which it has logged into.

FIP-Snooping bridge (FSB)

With a switch configured to performing FIP-Snooping the industry, term for this switch is FSB or FIP-Snooping Bridge. It is performing FIP-Snooping as described in the previous term.

FCF

FCoE forwarders (FCFs) act as an Ethernet and FC switch combined. All typical termination functions that would occur on a FC switch occur on the FCF. FCF's give VF_Ports and VE_Ports for their virtual FC interfaces.

IOM

IO module (IOM) refers to the modules on the rear of the Dell PowerEdge M1000e chassis that will receive and transmit IO (Ethernet, FC, InfiniBand, etc.) from the blade servers. The Dell Networking MXL switch is as an IOM for the M1000e blade server chassis.

MAC Address

Media Access Control Address (MAC Address) is a layer-2 node identifier. In Ethernet bridging, MAC addresses are used for source and destination identification. They can also be used as system identifiers since vendor-assigned (or burned-in) MAC addresses are globally unique. An Ethernet MAC address is 48 bits long and generally written in groupings of two hexadecimal digits often separated by colons or

hyphens like this: 00:1e:c9:00:cb:01. But, are sometimes written in groupings of four hexadecimal digits separated by periods like this: 001e.c900.cb01

NPG
NPIV Proxy Gateway (NPG) is a Fibre Channel over Ethernet (FCoE) – Fibre Channel gateway that connects Converged Network Adapters (CNA's) in Dells M-Series blade servers to an FC switch.

NPIV
N-port identifier virtualization, which enables multiple N-port fabric logins at the same time on the same physical FC link (Cisco Systems, Inc., 2011) .This term, is in reference to the Cisco Nexus 5000 series switches implementation of NPIV. NPIV must be enabled to share multiple logins across a single port/link or a port-channel/multiple-line link.

NPV
N-port Virtualizer is a FC aggregation method, which passes traffic through to end devices, while eliminating the need to use a domain ID for this device (Cisco Systems, Inc., 2011). This term is also in reference to configuration settings on the Cisco Nexus 5000 series switches.

PFC
Priority Flow Control or Per-Priority Pause is defined in the IEEE 802.1Qbb standard. PFC is flow control based on priority settings and adds additional information to the standard pause frame. The additional fields added to the pause frame allow devices to pause traffic on a specific priority instead of pausing all traffic. (IEEE, 2009) Pause frames will be initiated by the FCF in most cases when its receive buffers are starting to reach a congested point. With PFC, traffic is paused instead of dropped and retransmitted. This provides the lossless network behavior necessary for FC packets to be encapsulated and passed along the Ethernet paths.

ToR
Top of Rack is a term for a switch that is actually positioned at the top of a server rack in a data center.

VLAN
Virtual Local Area Network (VLAN) is a single layer-2 network (also called a broadcast domain, as broadcast traffic does not escape a VLAN on its own). Multiple VLANs can be passed between switches using switchport trunk interfaces. When passed across trunk links, frames in a VLAN are prefixed with the number of the VLAN that they belong to—a twelve-bit value that allows just over 4000 differently numbered VLANs.

VSAN
Virtual SAN is a logical partitioning of physical connections to provide for fabric or SAN separation. VSAN is a term that is particular to the Cisco Nexus series switches.