

Technical White Paper

Booting VMware ESXi from FCoE

Abstract

FCoE is a network technology which transports Fibre Channel traffic over Ethernet network to make use of 10 Gbit speed and reduce the operational costs. This technical white paper gives an overview of FCoE, software and hardware requirements to boot VMware ESXi from an FCoE LUN. In this white paper, we focus on a hardware CNA initiator, a switch and a storage array to make it more specific, however majority of the flow would be same across storage arrays and initiators except the switch specific commands. This white paper also provides useful ESXi commands which provides the initiator with target details and the statistics of FCoE LUN.

November 2019

Revisions

Date	Description
November 2019	Initial release

Acknowledgements

This paper was produced by the following:

Author: Nikhil Kumar Chaturvedi

Support: Ramya D R, Technical Writer, IDD team

The information in this publication is provided "as is." Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

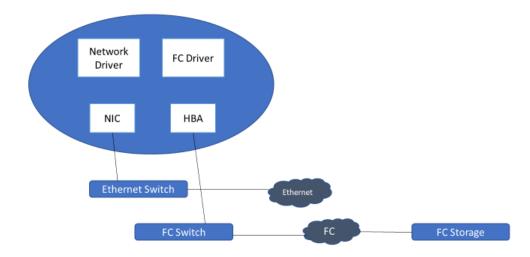
Copyright © <11/20/2019> Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

Table of contents

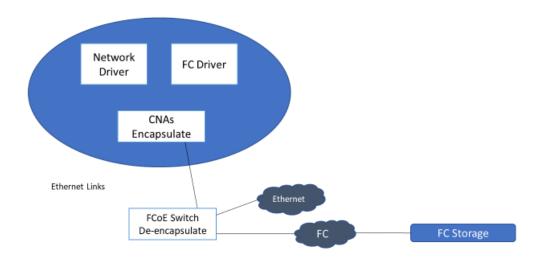
Revisions			2		
Acknowledgements					
Table of contents					
Executive summary					
1	Introduction				
•	1.1	Audience and Scope			
	1.2	Hardware and Software Requirements			
	1.3	Storage configuration			
	1.4	Enabling FCoE in PowerEdge servers	7		
2	ESXi FCoE commands		14		
	2.1	FCoE capable CNA devices	14		
	2.2	LUN pathing information	15		
	2.3	FCoE adapter attributes	15		
	2.4	FCoE adapter statistics	16		
3	Logging into FCoE during system boot				
4	Troubleshooting1				
5	Summary				
6	References 21				

Executive summary

Traditional SCSI was designed to transport Storage Area Network (SAN) traffic between devices in a fabric with shorter distance. Data transport was quick and reliable, but over a short distance. Ethernet was developed to overcome this problem by transporting data over long distances. But the dependency on transmission mechanics such as TCP and UDP brings in latency (especially during TCP retransmission).



FCoE combines the best of both worlds by transporting Fibre Channel SAN traffic encapsulated over Ethernet networks which also cuts down the overall capital and operational cost. With FCoE, the native FC protocol is preserved yet leveraging the use of SCSI directly over Ethernet unlike iSCSI which sends SCSI over TCP/IP which can introduce latency over the network.



VMware ESXi supports booting from FCoE capable network adapters in Dell EMC PowerEdge servers. When ESXi is installed from an FCoE LUN, the host requires a dedicated CNA with FCoE capabilities.

1 Introduction

This document is intended to help the user with configuring Dell EMC PowerEdge server and Dell EMC supported FCoE CNA adapter to boot from SAN.

1.1 Audience and Scope

This whitepaper is intended for IT administrators planning to boot ESXi over FCoE LUN with DELL EMC PowerEdge servers utilizing Dell S5000 switch which offers converged network capabilities and Dell Compellent SC8000 storage center controller which provides failover capabilities and support datacenter requirements. The whitepaper covers switch and storage configurations to set up ESXi installation over FCoE LUN exposed to the fabric.

1.2 Hardware and Software Requirements

To boot VMware ESXi from an FCoE LUN, following components are required:

- A Converged network adapter (CNA)
 For more information about Dell EMC supported FCoE CNAs, see <u>VMware Compatibility Guide</u>.
- Ethernet link
- FCoE capable switch
- Storage array with FCoE IO card(s)



The CNA adapter can be based on hardware or software (capable of running FCoE software stack). Hardware FCoE adapters contain network and fibre channel functionalities on the same card. When the host is installed with the hardware FCoE adapter, the hosts can use both the CNA components. The networking components appear as a standard network adapter (vmnic), and the Fibre channel component appears as an FCoE adapter (vmhba) in the vSphere client.

A software based FCoE adapter uses software code to perform FCoE processing. The software adapter must be added before using the software based FCoE adapter as a NIC supporting partial FCoE offload. To add the software adapter, navigate to **Datacenter Host Configure Storage Adapters Add Software FCoE Adapter**. Select the appropriate **Physical network adapter** and **VLAN ID**, if any.

The Ethernet link must be an extension to existing Ethernet and provide similar lossless behavior like Fibre channel links, where frames are not lost during congestion, as FCoE enables FC over existing Ethernet. An FCoE switch is an L2 switch capable of transporting FCoE frames over Ethernet. This is also termed as data center bridging (DCB) switch supports both Ethernet and native FC traffic on its interfaces.

The FCoE capable switch has the capabilities to converge LAN and SAN traffic over a single 10GbE connection to optimize enterprise-scale data center for maximum flexibility and scalability.

The testbed setup used in the above example to explain the FCoE boot for VMware ESXi comprises of the following hardware:

- Dell EMC PowerEdge R630 server
- QLogic BCM57xxx series CNA adapter

- Dell S5000 Ethernet switch
- Dell Compellent SC8000 Storage Center Controller

For more information about the FCoE boot from SAN, BIOS, and UEFI perspective, see VMware Knowledge
Base Article 2148531. In this example testbed, we use Qlogic BCM 57xxx series adapter which is an FCoE hardware offload capable CNA and it uses qfle3f driver module for storage adapter and qfle3 driver module for network processing FCoE frames. For more information about the CNA adapter, see VMware Compatibility Guide.

1.3 Storage configuration

To proceed with storage configuration, complete the following steps:

- 1. On the Storage Center Login page, enter your username and password, and then click Login.
- 2. Read the license agreement and on accepting, click **Accept**.
- On the Load license page, browse and select the license file and then click Load License.
 Note:

The serial number of the license file must match with the serial number of the controller.

The license submission must be completed successfully complete.

4. Click Continue.

The Create Disk Folder page appears.

Note: Use this to assign disks to a folder and to create a single pool of storage for volumes.

- 5. On the **Create Disk Folder** page, select the required disk folder, and then click **Continue**. **Note:** By default, all the disks are selected.
- 6. If required, rename the default folder name.
- 7. To configure Disk Folder, click **Advanced**.
 - a. Configure Tier Redundancy for each Tier, by selecting **Single Redundant** or **Dual Redundant** from the drop-down menu.
 - b. From the **Datapage Size** drop-down menu, select the required Datapage size.
 - c. Select Prepare Disk Folder for redundant storage, and then click Continue.
- 8. Click Create Now, and then click OK.
- 9. On the Add Controller page, for Single-controller storage center, click Continue Setup.
 - a. For Dual- controller storage center, click **Add Controller** to add the second controller.
 - b. In Controller ID, enter the HSN value.
 - **Note:** If the HSN for Controller 2 is included in the license file, then the value in Controller ID is auto-populated, and the value cannot be changed. If the HSN for Controller 2 is not included in the license file, you must enter the HSN value.
 - c. Under the Ether 0 interface, enter the values in IP Address, Net Mask, and Gateway.
 - d. Under the Ether 1 interface, enter the values in IP Address, Net Mask, and Gateway.
 - e. In **Primary DNS Server**, enter the IP Address of the primary DNS server.
 - f. If there is a secondary DNS server, then you must enter the IP address in **Secondary DNS** server.
- 10. Click Continue.
- 11. Click Join Now.
- 12. On the Time Settings page, set the system time for the Storage Center, and then click Continue.

- 13. On the **System Setup** page, enter the system name and management IP address for the Storage Center, and then click **Continue**.
 - The wizard prompts you to enable or disable the read and write cache.
- 14. Select Enable Read Cache and Enable Write Cache, and then Click Continue.
- 15. **Configure SMTP** page, to enable alert message emails to be sent to users who have specified a recipient address in their contact properties. Click **Continue**.
- 16. Use the **Update setup** page to configure how Storage Center handles software updates. Click **Continue**.
- 17. On the **User Setup** page, specify session timeout and email addresses for the Admin account. Click **Continue**.
- 18. On the **Configure SMTP** page, enter the IP address or a domain name of the SMTP email server. **Note**: To configure SMTP later, click **Skip SMTP Configuration**. SMTP settings can be configured later in the **System Manager**.
- 19. In the Sender E-mail Address field, enter the email address of the sender.
- 20. On the **Update Setup p**age, select an update option from the drop-down menu, and then click **Continue**.
- 21. On the User Setup page, for Session Timeout, select the time from the drop-down menu.
- 22. In the **Email**, **Email 2**, and **Email 3** fields, enter email addresses to which the Storage Center has to send system alerts.
- 23. Click Continue.
- 24. On the **Configure IO Cards** page, enter the values in the **IP Address**, **Subnet Mask**, and **Gateway** fields.
- 25. Click Continue.

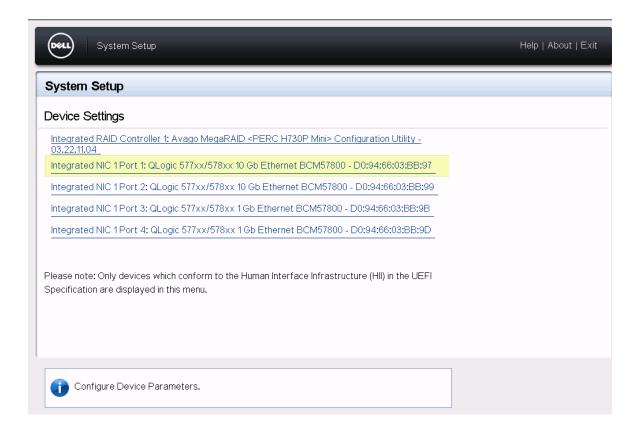
of the port(s).

- 26. On the Configure Ports page, click Configure Local Ports.
 The wizard lists the IO cards installed in the storage center (FC, iSCSi and SAS) along with the status
- 27. Configure Purpose, Fault Domain and User Alias (optional) based on your requirement.
 - a. For Single Controller Storage Center, by default, fault domain is created for each FC port.
 - b. For Dual-Controller Storage Center, create a fault domain for each pair of redundant FC ports.
- 28. Click Continue.
- 29. On the Generate SSL Certificate page, click Import.
- 30. Browse to the location of the public key (*.pem) file, and then select the file.
 Note: If you do not have a certificate containing the Storage Center host name or IP address, generate a certificate by clicking Generate Now.
- 31. Click Next.
- 32. To import the certificate, click Save.

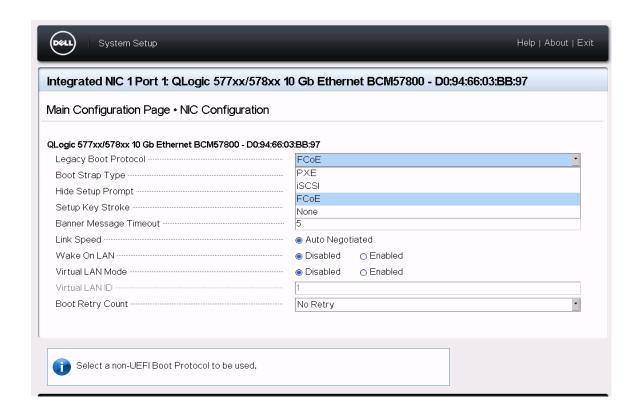
1.4 Enabling FCoE in PowerEdge servers

To enable FCoE, complete the following steps:

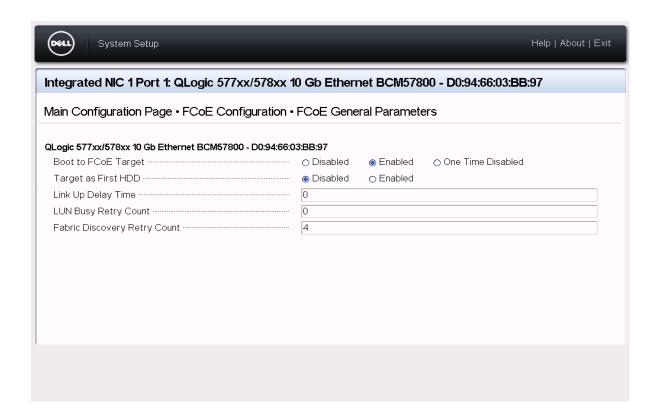
- 1. On a supported Dell EMC PowerEdge server, navigate to **Device Settings** in the BIOS setup, and then select the CNA adapter port that you want to boot.
- 2. To enable logical interfaces partitioned from a single interface, select **enable**.



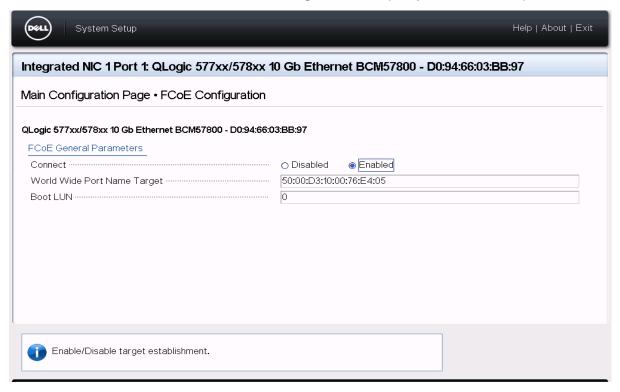
- 3. In the Main Configuration Page, select **NIC Configuration**.
- 4. From the Legacy Boot Protocol drop-down menu, select FCoE.



- 5. Select FCoE Configuration, and then click FCoE General Parameters.
- 6. To enable Boot to FCoE target, in FCoE General Parameters, select **Enabled**. Following are the options available in the FCoE General Parameters:
 - Boot to FCoE Target: Enable/Disable booting to FCoE target after logging in.
 - Target as First HDD: Target appears as a first Hard disk drive (HDD) in the system.
 - Link Up Delay Time: The time (in seconds) of the FCoE initiator that waits after an Ethernet link is established before sending any data over the network. By default, it is set to 0.
 - **LUN busy retry count**: The number of connections retries that the FCoE initiator attempts when the FCoE target LUN is busy. By default, it is set to 0.
 - **Fabric Discovery Retry Count:** The retry count for FCoE fabric discovery. By default, it is set to 4.



7. To enable Connect, select **Enable**, and then enter the World-Wide Port Name (WWPN) of the target device in the **World Wide Port Name Target** field and specify Boot LUN as required.



Note: To get the target WWPN (Controller Port), log in to the storage array and navigate to Hardware-Controllers-IO Ports-Fibre Channel.

Note: Boot LUN is the first FCoE storage target's LUN that is used by the FCoE initiator as the system boot device. LUN ID 0 is reserved for Boot LUNs.

8. The host or initiator WWPN is listed in the Main configuration page.



After configuring the BIOS settings, the next step is to configure Zoning to control the Fibre channel ports that communicates with each other. This configuration prevents the zone from discovering unauthorized target devices and reduces traffic, which makes the Fibre channel network more reliable and stable.

Separate logical group of ports are configured to address the separate set of connectivity requirements. For example, zone A acts as server 1 to communicate with the storage system and zone B acts as server 2 to communicate with the storage system and so on.

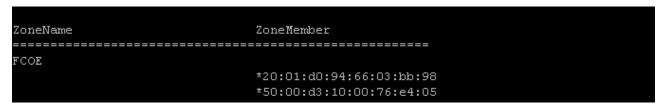
The following are the types of zoning:

- **Soft zoning:** Soft zoning relies on the WWN of the node in the fabric. That means, a host WWN is placed in a zone without concerning about the ports they are connected in the FC switch.
- Hard zoning: Hard zoning is implemented in hardware. A physical switch port is placed for zone
 membership. This method physically blocks access to a zone from any devices which resides outside
 of the zone.

In the below example, Soft zoning is performed.

Note: The zoning configuration varies by switch vendor or model.

The below screenshot shows an active zone in a Dell S5000 Switch. The output shown is generated by executing show fc zoneset active command in the Switch command-line interface.

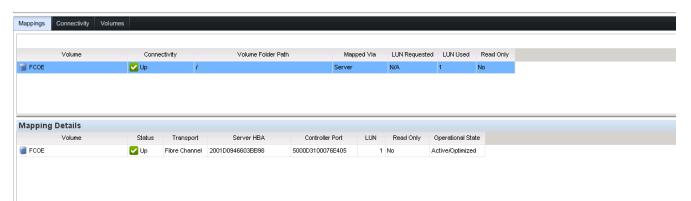


For more information, see the PowerSwitch S5000 documents.

Following are the two ways to access the storage controller:

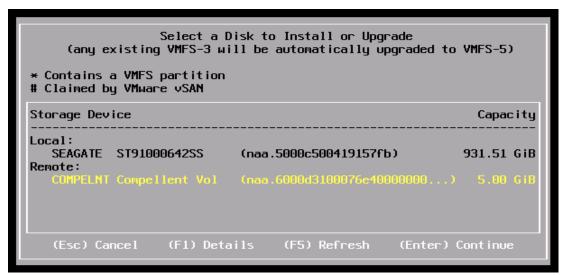
- Use a web browser with network access and connect to the eth0 IP address of the controller.
- <u>Download</u> Storage Manager Client (Windows/Linux) to manage and monitor Storage center (s).

The following screenshot displays a volume mapped to an FCoE initiator in a Dell Compellent SC8000 storage array.



For more information, see the <u>Dell Storage SC8000</u> documents.

Boot PowerEdge server with ESXi ISO build mounted. The configured FCoE LUN is listed as one of the storage devices as shown:



To get the device driver, firmware details, and additional information in the test setup, see <u>VMware</u> <u>Compatibility Guide</u>.

2 ESXi FCoE commands

This section lists the commands used for FCoE.

2.1 FCoE capable CNA devices

The following screenshot captured with the command that lists the FCoE capable CNA devices:

```
[root@he-dhcp-100-98-13-204:~] esxcli fcoe adapter list
vmhba64
Source MAC: d0:94:66:03:bb:97
FCF MAC: 5c:f9:dd:ef:aa:83
VNPort MAC: 0e:fc:02:01:30:00
Physical NIC: vmnic0
User Priority: 3
VLAN id: 1002
VN2VN Mode Enabled: false
```

The following are list of commands and its description mentioned in the above screenshot:

- vmhba64: Name of the storage adapter. This refers to the physical adapter on the host.
- Source MAC: The unique hardware address i.e. physical address of the source node.
- **FCF MAC**: The unique address assigned by Fibre channel forwarder (FCF) during FLOGI (Fibre channel login) or FDISC (Fabric discovery) process.
- VNPort MAC: A MAC address that uniquely identifies a CNA port (VN_port in FCoE terminology)
- Physical NIC: A virtual machine network interface card, also a physical adapter on the ESXi host.
- User Priority: Priority to use for FCoE traffic.
- VLAN id: To identify VLAN.
- VN2VN mode: Allows FCoE network to work without any FCoE switches (or FCFs). The default setting is false.

2.2 LUN pathing information

The following screenshot captured with the commands to obtain LUN pathing information:

```
TrootBendersDepthOn-Ministration of secular storage decay path last concession of the control of
```

2.3 FCoE adapter attributes

The following screenshot captured with the commands to list adapter attributes:

```
[root@he-dhcp-100-98-13-204:~] esxcli storage san fcoe list
   Adapter: vmhba64
   Port ID: 013A00
  Node Name: 20:00:d0:94:66:03:bb:98
  Port Name: 20:01:d0:94:66:03:bb:98
   Speed: O Gbps
  Port Type: NLPort
  Port State: ONLINE
  Controller MAC: d0:94:66:03:bb:98
  FCF MAC: 5c:f9:dd:ef:aa:8d
  VN Port MAC: 0e:fc:02:01:3a:00
  VLAN ID: 1002
  Vmnic Name: vmnic0
  Model Description: Broadcom NetXtreme II 57800 FCoE Driver
  Hardware Version:
  OptionROM Version:
  Firmware Version:
  Driver Name: qfle3f
  DriverVersion: 1.0.63.0
 root@he-dhcp-100-98-13-204:~]
```

FCoE adapter statistics 2.4

```
The following screenshot captured with the command to obtain adapter statistics: [root@he-dhcp-100-98-13-204:~] esxcli storage san fcoe stats get
   Adapter: vmhba64
   Tx Frames: 12388563
   Rx Frames: 15172830
   Error Frames: 0
  Dumped Frames: 0
  Link Failure Count: 0
  Loss of Signal Count: 0
   Invalid Tx Word Count: 0
   Invalid CRC Count: 0
   VLink Failure Count: 0
  Miss Disc Adv Count: 0
   Input Requests: 0
  Output Requests: 0
   Control Requests: 0
 root@he-dhcp-100-98-13-204:~]
```

3 Logging into FCoE during system boot

The screenshot below shows the DCBX startup and initiator login to the target to enable OS boot from the FCoE LUN exposed over the network:

```
Copyright (C) 2017 QLogic Corporation
FCoE Boot v7.15.0
Starting DCBX process with interface (D0:94:66:03:BB:98) ... Succeeded
Discovering FC Fabric with interface (D0:94:66:03:BB:98) ... Succeeded
World Wide Node Name : 20:00:D0:94:66:03:BB:98
World Wide Port Name : 20:01:D0:94:66:03:BB:98
Fabric Name
                     : 10:00:5C:F9:DD:EF:AA:80
Fabric name
FCF MAC Address
                     : 5C:F9:DD:EF:AA:83
                     : 0E:FC:02:01:30:00
ULAN ID
                     : 1002
Fabric Login via interface (D0:94:66:03:BB:98) ... Succeeded
Login to target [5000D3100076E405:013600:LUN=000] ... Succeeded
FC Target Drive: COMPELNT Compellent Vol
                                            (Rev: 0703)
Press <Ctrl-D> within 4s to stop booting from the target ... _
```

Data center bridging exchange (DCBX) is a protocol which is an extension to the Link layer discovery protocol (LLDP). DCBX cannot run if LLDP is disabled. DCBX enables the devices connected in a fabric to discover and exchange configuration information.

Note that the DCBX capabilities are supported by the FCoE capable switch.

In this case, DCBX exchanges FCoE information by default as the interface carries FCoE traffic.

During FC fabric discovery, The FC switch discover the WWPN of a device or host and assign a port address to the device. This port address is unique and routable over the network and valid only till the device (host or target) is logged on.

After the DCBx setup and FCF discovery, the initiator successfully logs in to the target, i.e. the FCoE disk is ready for read/write operations.

The following screenshot captured with the commands that lists the FCoE software packages installed after

ESXi completes to boot to DCUI.

```
[root@he-dhcp-100-98-13-204:~] esxcli software vib list | grep -i fcoe
                                                                                               2019-07-01
brcmfcoe
                              11.4.1078.5-11vmw.670.1.28.10302608
                                                                    VNW
                                                                             VMwareCertified
                                                                             VMwareCertified
                                                                                               2019-07-01
net-fcoe
                              1.0.29.9.3-7vmw.670.0.0.8169922
                                                                    VNW
net-libfcoe-92
                              1.0.24.9.4-8vmw.670.0.0.8169922
                                                                    VMW
                                                                             VMwareCertified 2019-07-01
shim-libfcoe-9-2-1-0
                              6.7.0-0.0.8169922
                                                                    VMW
                                                                             VMwareCertified 2019-07-01
                              6.7.0-0.0.8169922
shim-libfcoe-9-2-2-0
                                                                    VMW
                                                                             VMwareCertified
                                                                                               2019-07-01
                                                                    VMW
                                                                                               2019-07-01
/mkfcoe
                              1.0.0.1-1vmw.670.1.28.10302608
                                                                             VMwareCertified
```

The following screenshot captured with the commands FCoE driver loaded after ESXi completes boot to DCUI:

```
[root@he-dhcp-100-98-13-204:~] esxcli system module list | grep -i qfle3f
qfle3f true true
[root@he-dhcp-100-98-13-204:~]
```

```
[root@he-dhcp-100-98-13-204:~] vmkload_mod -l | grep -i qfle3f
qfle3f 2 708
[root@he-dhcp-100-98-13-204:~]
```

```
[root@he-dhcp-100-98-13-204:~] esxcli system module get -m qfle3f
Module: qfle3f
Module File: /usr/lib/vmware/vmkmod/qfle3f
License: ThirdParty:QLogic_Proprietary
Version: 1.0.63.0-10EM.670.0.0.8169922
Build Type: release
Provided Namespaces:
Required Namespaces: com.vmware.vmkapi@v2_5_0_0
Containing VIB: qfle3f
VIB Acceptance Level: certified
```

4 Troubleshooting

Switches are located between the hosts and storage devices and have visibility to both sides of the storage network. When there is an issue in the fabric, observe if the issue persisting on the host or storage or the switch and then perform a detailed diagnosis.

Verify the following issues:

Host issues:

- Missing device(s)
- CNA firmware
- Incorrect device driver installation and device configuration
- Interrupted or damaged links between host and switch

Storage issues:

- Physical connectivity between switch and storage
- Incorrect storage configuration

Switch issues:

- Missing or interrupted links
- Damaged cables
- Incorrect switch and zoning configuration

5 Summary

This paper explains about configuring Dell EMC PowerEdge server, Dell S5000 switch and Dell Compellent SC8000 Storage Center to VMware ESXi from FCoE LUN exposed to SAN.

6 References

Booting ESXi with Software FCoE