Dell EMC PowerEdge MX Networking Architecture Guide

Abstract
This document provides an overview of the architecture, features, and functionality of the PowerEdge MX networking infrastructure.

January 2019

Revisions

Date	Description
January 2019	Minor edits in sections 1.2.5 and 2.2
November 2018	Added Appendix A and B, minor edits for clarity, and updated images
September 2018	Initial release

DELLEMC

Dell believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

Table of contents

DELLEMC

DELLEMC

1 Dell EMC PowerEdge MX platform overview

The Dell EMC PowerEdge MX, a unified, high-performance data center infrastructure, provides the agility, resiliency, and efficiency to optimize a wide variety of traditional and new, emerging data center workloads and applications. With its kinetic architecture and agile management, PowerEdge MX dynamically configures compute, storage, and fabric, increases team effectiveness and accelerates operations. Its responsive design and delivers the innovation and longevity customers of all sizes need for their IT and digital business transformations.

This section contains information about the hardware and options available in the PowerEdge MX7000. The section is divided into two parts:

- The front of the MX7000 chassis, containing compute and storage sleds
- The back of the MX7000 chassis, containing networking, storage, and management components
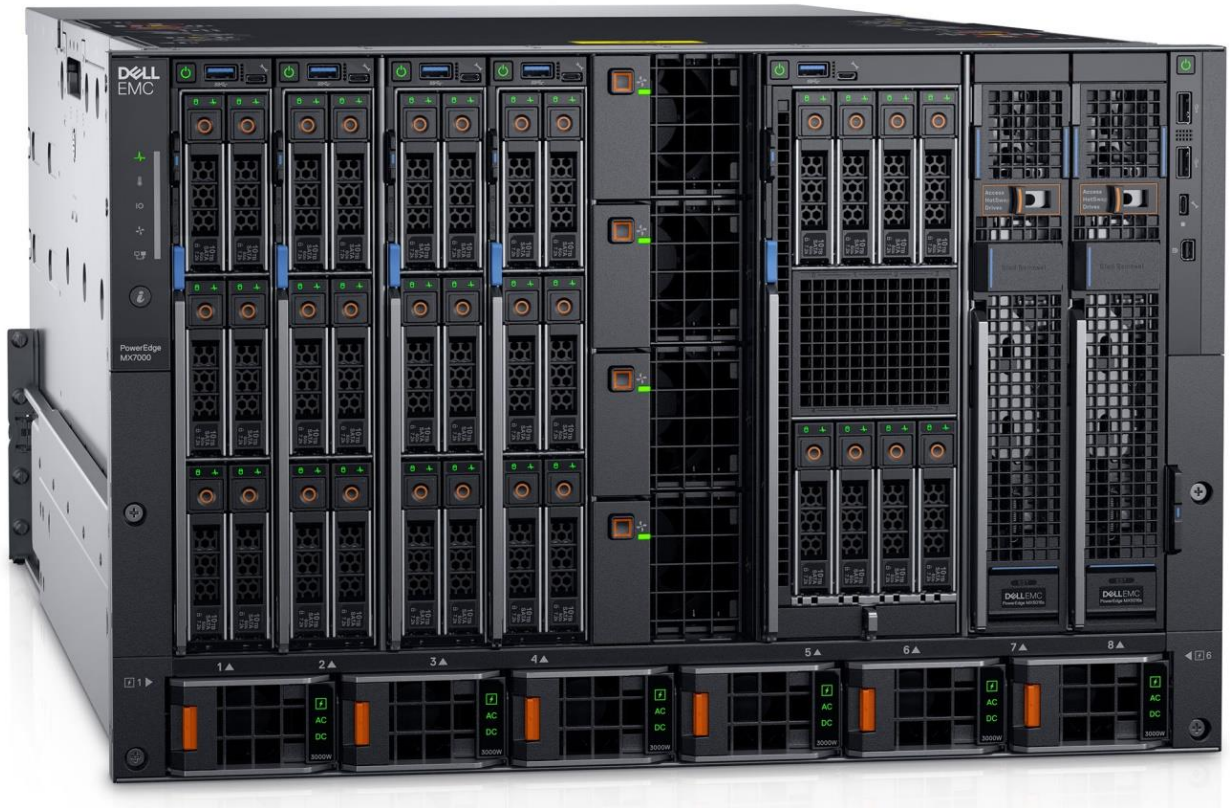
Figure 1   Dell EMC PowerEdge MX7000 chassis

1.1 Dell EMC PowerEdge MX7000 (front)

Figure 2 shows the front view of the Dell EMC PowerEdge MX7000 chassis. The left side of the chassis can have one of three control panel options:

- LED status light panel
- Touch screen LCD panel
- Touch screen LCD panel equipped with Dell EMC PowerEdge iDRAC Quick Sync 2

The bottom of Figure 2 shows six hot-pluggable, redundant, 3,000-watt power supplies. Above the power supplies eight slots are available. In the example below, the slots contain:

- Four Dell EMC PowerEdge MX740c sleds in slots one through four
- One Dell EMC PowerEdge MX840C sled in slots five and six
- Two Dell EMC PowerEdge MX5016s sleds in slots seven and eight
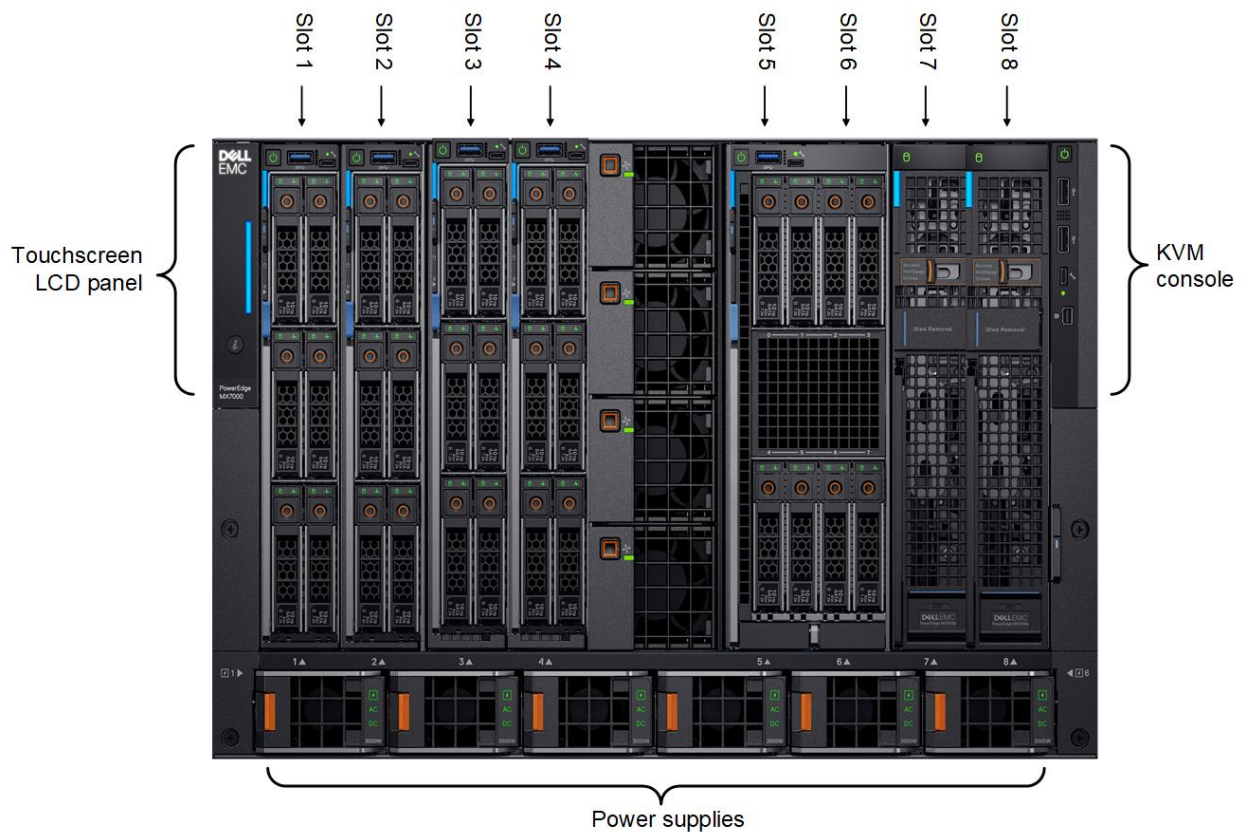
Figure 2   Dell EMC PowerEdge MX7000 – front

1.1.1 Dell EMC PowerEdge MX740c compute sled

The Dell EMC PowerEdge MX740c is a two-socket, full-height, single-width compute sled that offers impressive performance and scalability. It is ideal for dense virtualization environments and can serve as a foundation for collaborative workloads. The MX7000 chassis supports up to eight MX740c sleds.

Key features of the MX740c include:

- Single-width slot design
- Two CPU sockets
- 24 DIMM slots of DDR4 memory
- Boot options include BOSS-S1 or Internal Dual SD Modules (IDSDM)
- Up to six SAS/SATA SSD/HDD and NVMe PCIe SSDs
- Two PCIe mezzanine card slots for connecting to network Fabric A and B
- One PCIe mini-mezzanine card slots for connecting to storage Fabric C
- iDRAC9 with Lifecycle Controller

Figure 3   Dell EMC PowerEdge MX740c sled with six 2.5-inch SAS drives

DELLEMC

1.1.2   Dell EMC PowerEdge MX840c compute sled

The Dell EMC PowerEdge MX840c is a powerful four-socket, full-height, double-width server that features dense compute, exceptionally large memory capacity, and a highly expandable storage subsystem. It is the ultimate scale-up server that excels at running a wide range of database applications, substantial virtualization, and software-defined storage environments. The MX7000 chassis supports up to four MX840c compute sleds.

Key features of the MX840c include:

- Dual-width slot design
- Four CPU sockets
- 48 DIMM slots of DDR4 memory
- Boot options include BOSS-S1 or IDSDM
- Up to eight SAS/SATA SSD/HDD and NVMe PCIe SSDs
- Four PCIe mezzanine card slots for connecting to network Fabric A and B
- Two PCIe mini-mezzanine card slots for connecting to storage Fabric C
- iDRAC9 with Lifecycle Controller

Figure 4   Dell EMC PowerEdge MX840c sled with eight 2.5-inch SAS drives

1.1.3 Dell EMC PowerEdge MX5016s storage sled

The PowerEdge MX5016s storage sled delivers scale-out, shared storage within the PowerEdge MX architecture. The MX5016s provides customizable 12 GB/s direct-attached SAS storage with up to 16 SAS HDDs/SSDs. Both the MX740c and the MX840c compute sleds can share drives with the MX5016s using the PowerEdge MX5000s SAS module. Internal server drives may be combined with up to seven MX5016s sleds in one chassis for extensive scalability. The MX7000 chassis supports up to seven MX5016s storage sleds.

Figure 5   Dell EMC PowerEdge MX5016s sled with the drive bay extended

1.2 Dell EMC PowerEdge MX7000 (rear)

The MX7000 includes three I/O fabrics. Fabrics A and B are for Ethernet and future I/O module connectivity, and Fabric C is for SAS and Fibre Channel (FC) connectivity. Each fabric provides two slots to provide redundancy. Figure 6 shows the rear of the PowerEdge MX7000 chassis. From top to bottom the chassis is configured with:

- One Dell EMC Networking MX9116n Fabric Switching Engine (FSE) installed in fabric slot A1
- One Dell EMC Networking MX7116n Fabric Expander Module (FEM) installed in fabric slot A2
- Two Dell EMC Networking MX5108n Ethernet switches installed in fabric slots B1 and B2
- Two Dell EMC Networking MXG610s Fibre Channel switches installed in fabric slots C1 and C2
- Two Dell EMC PowerEdge MX9002m modules that are installed in management slots MM1 and MM2
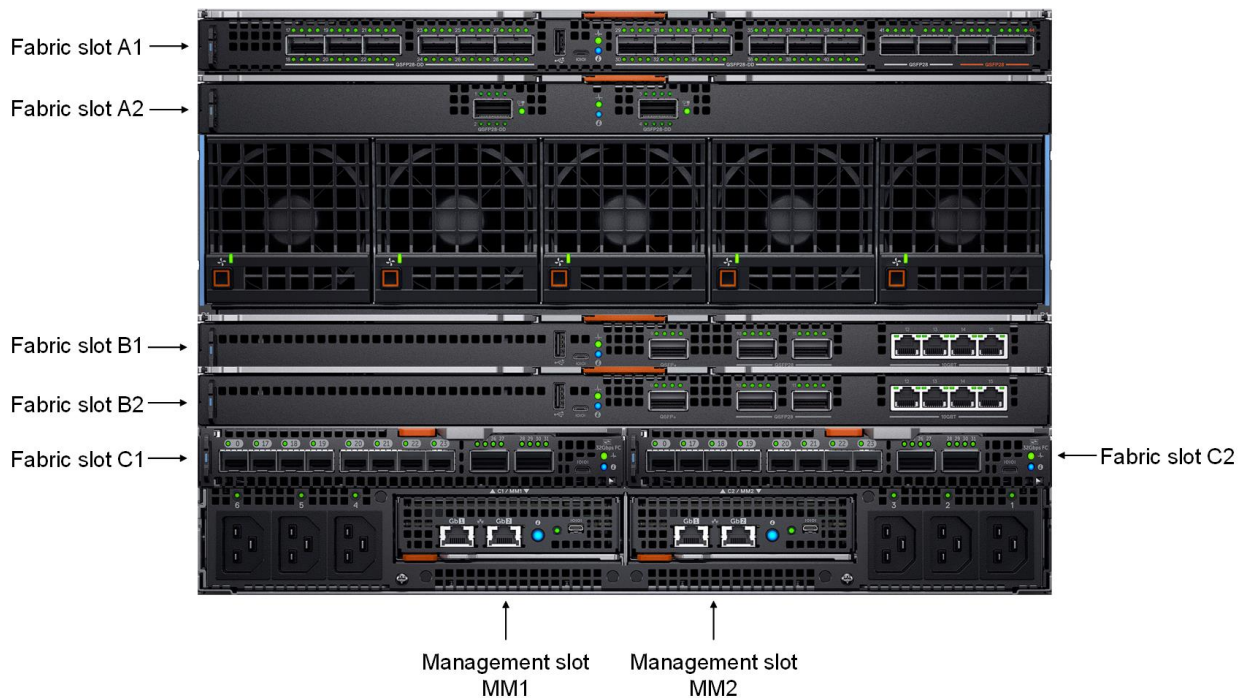
Figure 6   Dell EMC PowerEdge MX7000 – rear

1.2.1 Dell EMC PowerEdge MX9002m management module

The Dell EMC PowerEdge MX9002m management module controls the overall chassis power, cooling, and hosts the OpenManage Enterprise - Modular (OME-M) console. Two external 1G-BaseT Ethernet ports are provided to enable management connectivity and to connect more MX7000 chassis into a single logical chassis. The MX7000 chassis supports two MX9002m modules for redundancy. Figure 7 shows a single MX9002m module and its components.
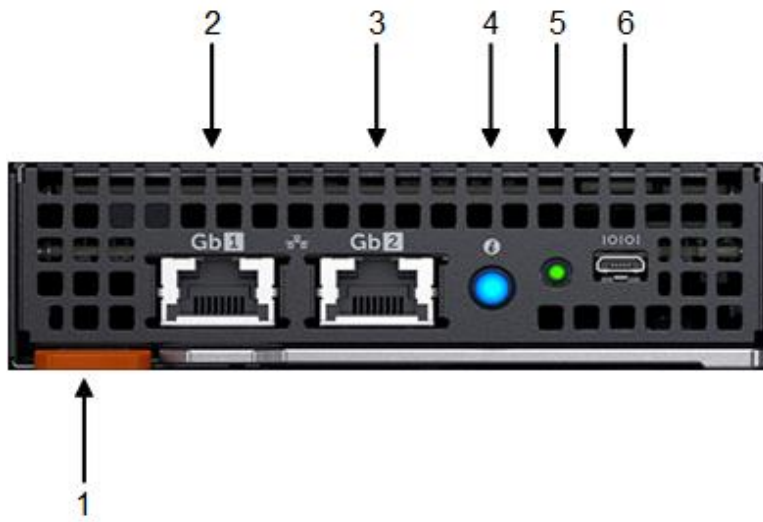
Figure 7   Dell EMC PowerEdge MX9002m module

The following MX9002m module components are labeled Figure 7:

1. Handle release
2. Gigabit Ethernet port 1
3. Gigabit Ethernet port 2
4. ID button and health status LED
5. Power status LED
6. Micro-B USB port

1.2.2   Dell EMC Networking MX9116n Fabric Switching Engine

The Dell EMC Networking MX9116n Fabric Switching Engine (FSE) is a scalable, high-performance, low latency 25 GbE switch purpose-built for the PowerEdge MX platform. The MX9116n FSE provides enhanced capabilities and cost-effectiveness for enterprise, mid-market, Tier 2 cloud, and NFV service providers with demanding compute and storage traffic environments.

In addition to 16 internal 25GbE ports, the MX9116n FSE provides:

- Two 100GbE QSFP28 ports
- Two 100GbE QSFP28 unified ports
- Twelve 2x100GbE QSFP28-Double Density (DD) ports

The QSFP28 ports can be used for Ethernet uplink connectivity (see Section 3.1 for more information). The unified ports can be used for Ethernet uplink connectivity and supporting eight 32Gb Fibre Channel ports for SAN connectivity supporting both NPIV Proxy Gateway (NPG) and direct attach FC capabilities.

The QSFP28-DD ports provide fabric expansion connections for up to nine more MX7000 chassis using the MX7116n Fabric Expander Module. The QSFP28-DD ports also provide capacity for extra uplinks, VLTi links, and connections to rack servers at 10GbE or 25GbE using breakout cables. The MX7000 chassis supports up to four MX9116n FSEs in Fabric A and/or B. See Appendix C.4 for more information.

Figure 8   Dell EMC Networking MX9116n FSE

The following MX9116n FSE components are labeled in Figure 8:

1. Express service tag
2. Storage USB port
3. Micro-B USB console port
4. Power and indicator LEDs
5. Handle release
6. Two QSFP28 ports
7. Two QSFP28 unified ports
8. 12 QSFP28-DD ports

DELLEMC

1.2.3 Dell EMC Networking MX7116n Fabric Expander Module

The Dell EMC Networking MX7116n Fabric Expander Module (FEM) acts as an Ethernet repeater, taking signals from an attached compute sled and repeating them to the associated lane on the external QSFP28-DD connector. The MX7116n FEM provides two QSFP28-DD interfaces, each providing up to eight 25 GbE connections to the chassis.

There is no operating system or switching ASIC on the MX7116n FEM, so it never requires an upgrade. There is also no management or user interface, making the MX7116n FEM maintenance-free. The MX7000 chassis supports up to four MX7116n FEMs in Fabric A and/or B. See Appendix C.4 for more information.
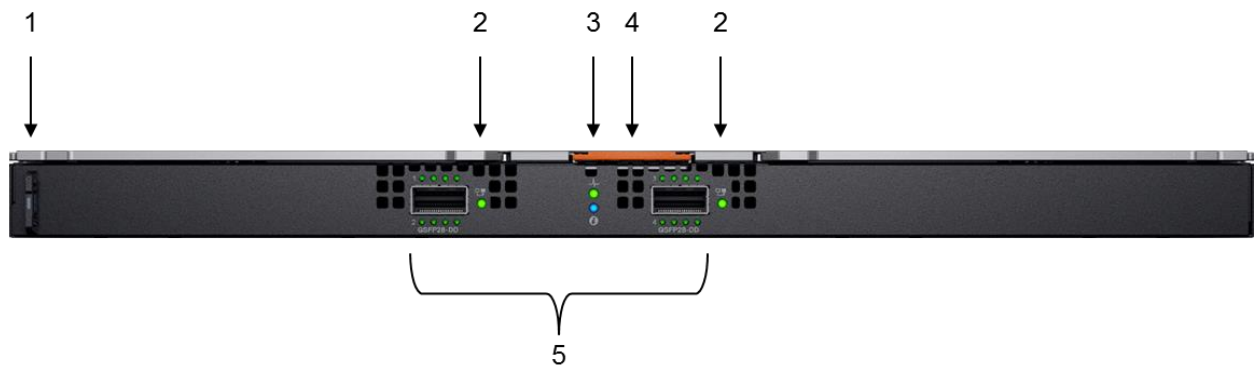
Figure 9   Dell EMC Networking MX7116n FEM
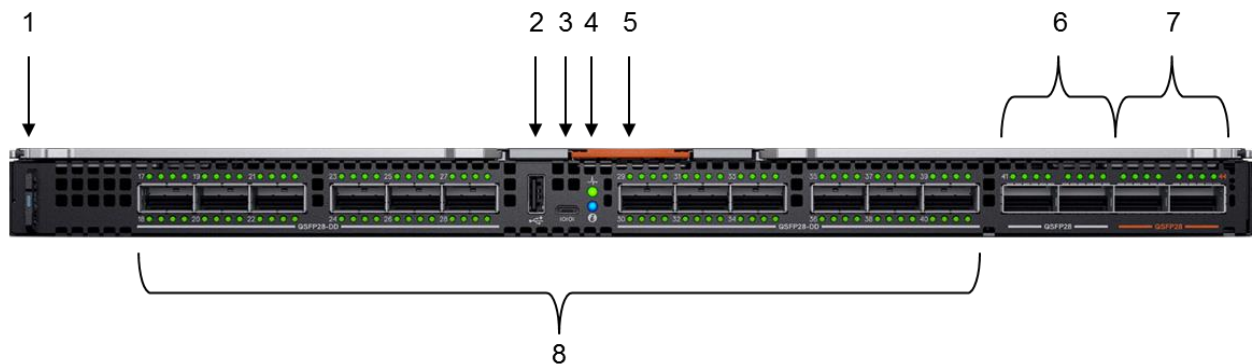
The following MX7116n FEM components are labeled in Figure 9:

1. Express service tag
2. Supported optic LED
3. Power and indicator LEDs
4. Module insertion/removal latch
5. Two QSFP28-DD fabric expander ports

Note: The MX7116n FEM cannot act as a stand-alone switch and must be connected to the MX9116n FSE to function.

1.2.4 Dell EMC Networking MX5108n Ethernet switch

The Dell EMC Networking MX5108n Ethernet switch is targeted at small PowerEdge MX7000 deployments of one or two chassis. While not a scalable switch, it still provides high-performance and low latency with a non-blocking switching architecture. The MX5108n provides line-rate 25 Gbps Layer 2 and Layer 3 forwarding capacity to all connected servers with no oversubscription.

Besides eight internal 25GbE ports, the MX5108n provides:

- One 40GbE QSFP+ port
- Two 100GbE QSFP28 ports
- Four 10GbE RJ45 BASE-T ports

The ports can be used to provide a combination of network uplinks, VLT interconnects (VLTi), or for FCoE connectivity. The MX5108n supports FCoE Initilization Protocol (FIP) Snooping Bridge (FSB) mode but does not support NPG or direct attach FC capabilities. The MX7000 chassis supports up to four MX5106n Ethernet switches in Fabric A and/or B. See Appendix C.4 for more information.
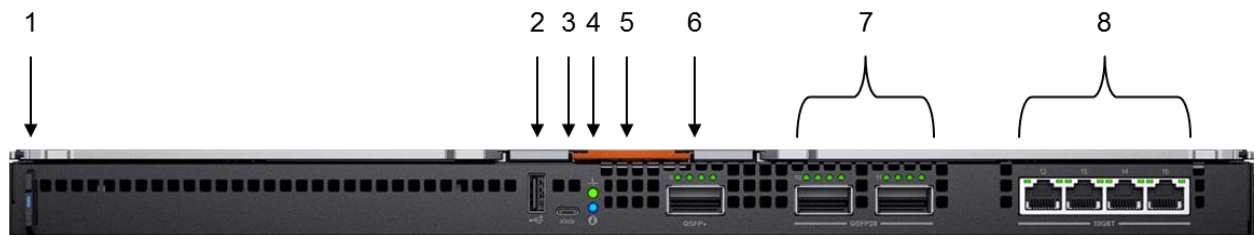
Figure 10   Dell EMC Networking MX5108n Ethernet switch

The following MX5108n components are labeled in Figure 10:

1. Express service tag
2. Storage USB port
3. Micro-B USB console port
4. Power and indicator LEDs
5. Module insertion/removal latch
6. One QSFP+ port
7. Two QSFP28 ports
8. Four 10GBASE-T ports

DELLEMC

1.2.5 Dell EMC PowerEdge MX pass-through modules

There are two Ethernet Pass-Through Modules (PTM) providing non-switched Ethernet connections to Top of Rack switches. Each PTM provides 16 internal 10GbE and/or 25GbE ports. Figure 11 shows the 25GbE Ethernet PTM. The 25GbE PTM provides 16 external SFP+ ports that can operate at 10/25GbE.

The 10GBASE-T Ethernet PTM, shown in Figure 12, provides 16 external RJ45 BASE-T ports that can operate at 10GbE. An MX7000 chassis supports a mix of four PTMs in Fabric A and/or B. See Appendix C.4 for more information.
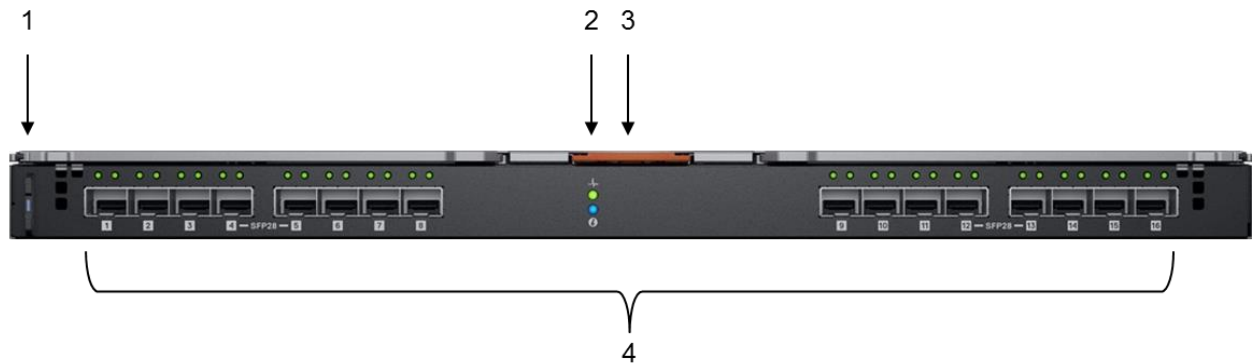
Figure 11   Dell EMC Networking 25GbE PTM

The following 25GbE PTM components are labeled in Figure 11:

1. Express service tag
2. Power and indicator LEDs
3. Module insertion/removal latch
4. 16 SFP28 ports
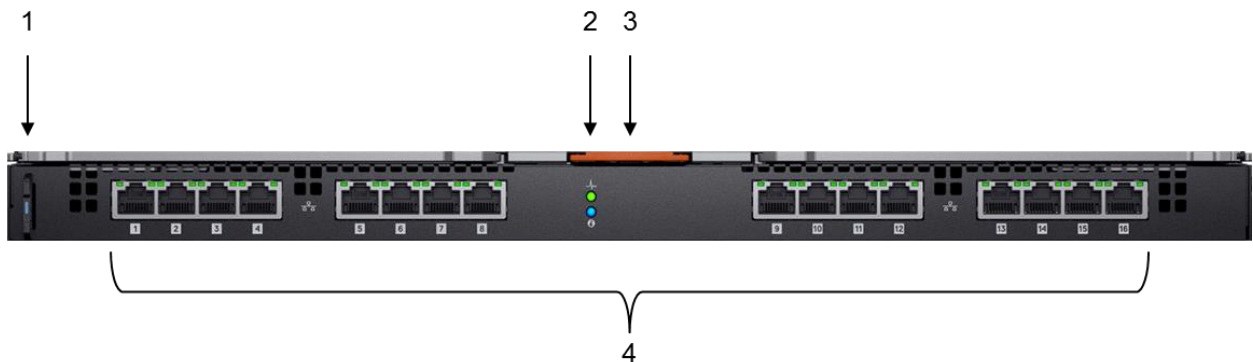
Figure 12   Dell EMC Networking 10GBASE-T PTM

The following 10GBASE-T PTM components are labeled in Figure 12:

1. Express service tag
2. Power and indicator LEDs
3. Module insertion/removal latch
4. 16 10GBASE-T ports

Note: For information about PTM port to compute sled-mapping, see Appendix C.5.

1.2.6 Dell EMC Networking MXG610s Fibre Channel switch

The Dell EMC Networking MXG610s is a high-performance, 32 Gb Fibre Channel switch based on Brocade Gen 6 technology. The MXG610s features 8/16/32 GbE capable interfaces. It is ideal for connectivity to all-flash SAN storage solutions and is designed for maximum flexibility and value with "pay-as-you-grow" scalability using a Port on Demand, or PoD, license model. The MXG610s is compatible with Brocade and Cisco fabric/director class switches.

Besides sixteen internal 32Gb FC ports, the MXG610s provides:

- Eight 32Gb FC SFP ports
- Two 4x 32Gb FC QSFP ports

The MX7000 chassis supports redundant MXG610s modules in Fabric C and must be purchased in pairs.
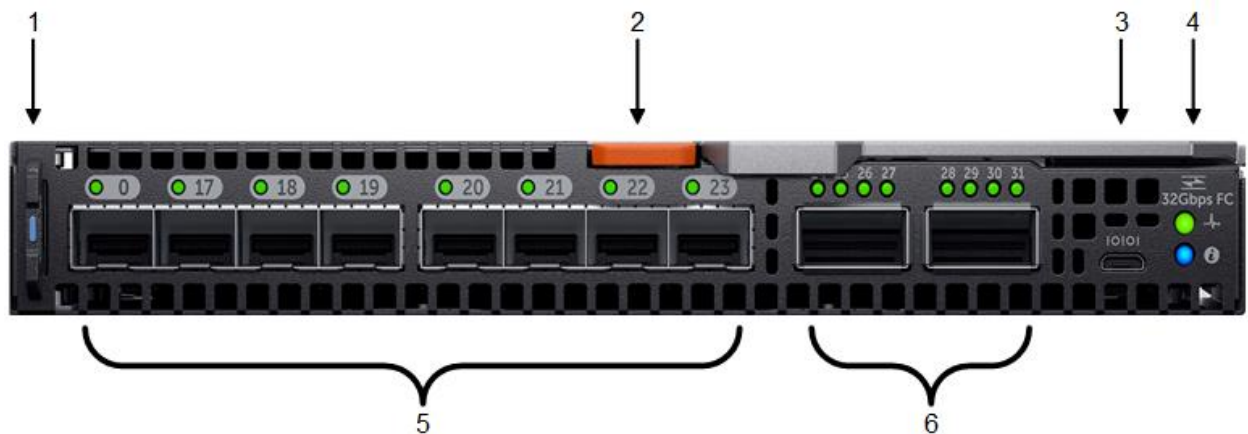
Figure 13   Dell EMC Networking MXG610s Fibre Channel switch module

The following MXG610s components are labeled in Figure 13:

1. Express service tag
2. Module insertion/removal latch
3. Micro-B USB console port
4. Power and indicator LEDs
5. Eight 32Gb FC SFP ports
6. Two 4x 32Gb FC QSFP ports

1.2.7 Dell EMC PowerEdge MX5000s SAS module

The Dell EMC PowerEdge MX5000s SAS module supports x4 SAS internal connections to all eight front-facing slots in the PowerEdge MX7000 chassis. The MX5000s uses T10 SAS zoning to provide multiple SAS zones/domains for the compute sleds. Storage management is conducted through the OEM-M console.

The MX5000s provides Fabric C SAS connectivity to each compute and one or more MX5016s storage sleds. Compute sleds connect to the MX5000s using either SAS Host Bus Adapters (HBA) or a PowerEdge Raid Controller (PERC) in the mini-mezzanine PCI-e slot.

The MX5000s switches are deployed as redundant pairs to offer multiple SAS paths to the individual SAS disk drives. An MX7000 chassis supports redundant MX5000s in Fabric C.

Note: A PowerEdge MX5000s SAS module and a MXG610s is not supported in the same MX7000 chassis.
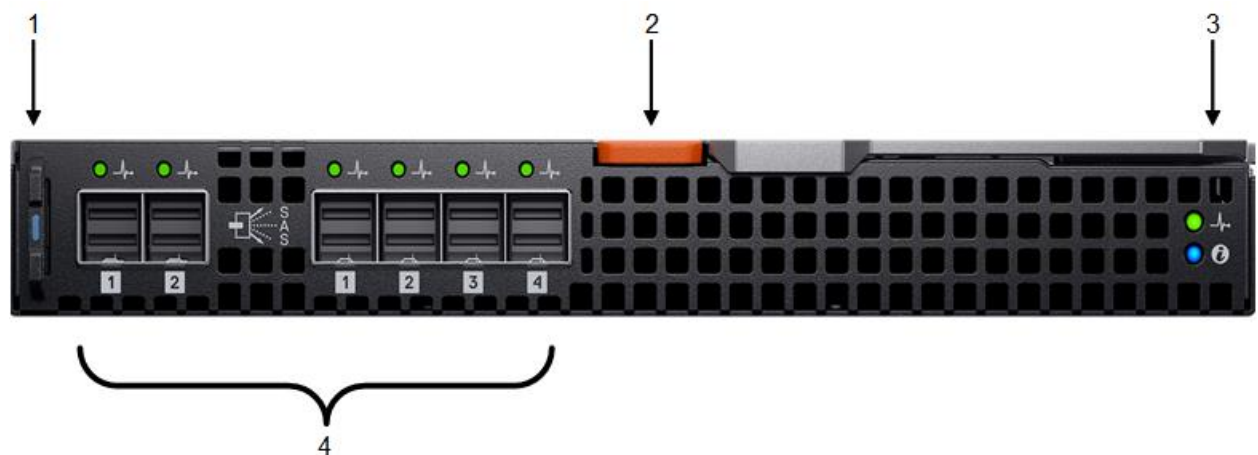
Figure 14   Dell EMC PowerEdge MX5000s SAS module

The following MX5000s components are labeled in Figure 14:

1. Express service tag
2. Module insertion/removal latch
3. Power and indicator LEDs
4. Six SAS ports

1.3 PowerEdge MX7000 chassis fabric slots

The PowerEdge MX7000 chassis includes two general-purpose I/O fabrics, Fabric A and B. The vertically aligned compute sleds, in slots one through eight, connect to the horizontally aligned I/O modules (IOMs), in fabric slots A1, A2, B1, and B2. This orthogonal connection method results in a mid-plane free design and enables the adoption of new I/O technologies without the burden of having to upgrade the mid-plane.

The MX740c supports two mezzanine cards, which are installed in slots A1 and B1 and the MX840c supports four mezzanine cards, which are installed in slots A1, A2, B1, and B2. Each mezzanine card provides redundant connections to each fabric, A or B, shown in Figure 15. A mezzanine card connects orthogonally to the pair of IOMs installed in the corresponding fabric slot. For example, port one of mezzanine card A1 connects to fabric slot A1, a MX9116n FSE (not shown). The second port of mezzanine card A1 connects to fabric slot A2, a MX7116n FEM (not shown).
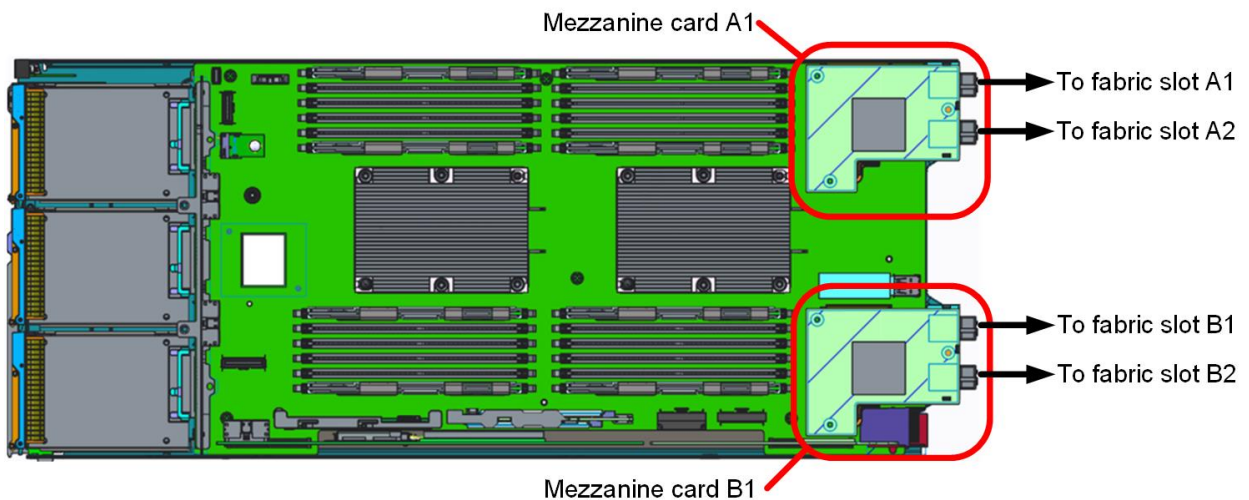
Figure 15   Dell EMC PowerEdge MX740c mezzanine cards

The MX7000 chassis also provides Fabric C, shown in Figure 16, supporting redundant MXG610s FC switches, or MX5000s SAS modules. This fabric uses a mid-plane connecting the C1 and C2 modules to each compute or storage sled. The MX740c supports one mini-mezzanine card, which is installed in slot C1, and the MX840c supports two mini-mezzanine cards, which are installed in slots C1 and C2.
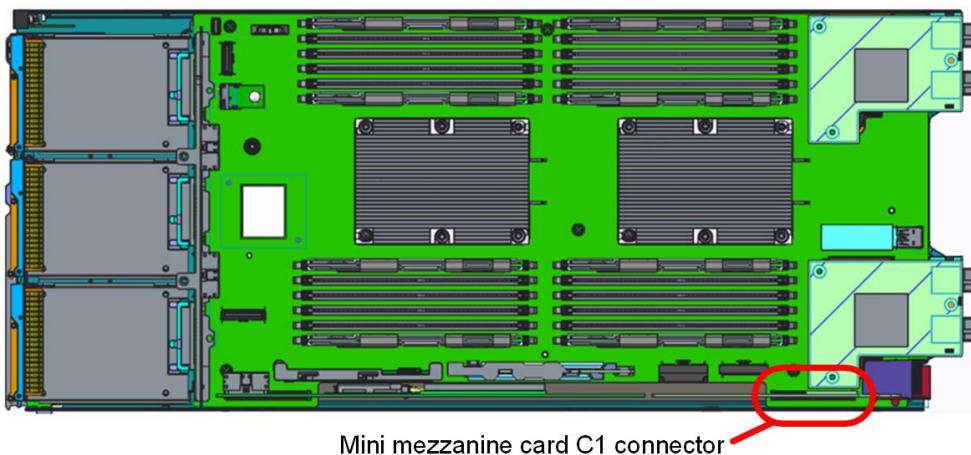
Figure 16   Dell EMC PowerEdge MX740c mini-mezzanine card

DELLEMC

2 OpenManage Enterprise - Modular console

The PowerEdge MX9002m module hosts the OpenManage Enterprise (OME-M) console. OME-M is the latest addition to the Dell OpenManage Enterprise suite of tools and provides a centralized management interface for the PowerEdge MX platform. The OME-M console features include:

- End-to-end life cycle management for servers, storage, and networking
- Touch LCD for initial setup, and fielding error conditions
- Use of iDRAC9 intelligent automation and security features
- Manage one or multiple chassis from a single web or REST API using multi-chassis management groups
- OpenManage Mobile for configuration and troubleshooting including wireless server vKVM

2.1 PowerEdge MX7000 initial deployment

Initial PowerMX7000 chassis configuration consists of initial deployment, assigning network settings for OME-M, and completing the **Chassis Deployment Wizard**.

There are three methods for initial configuration:

- Using the LCD touchscreen on the front-left of the MX7000 chassis
- Setting the initial OME-M console IP address through the KVM ports on the front-right side of the MX7000 chassis
- Setting the initial OME-M console IP address through the serial port on the MX9002m module

On first logging into the OME-M console, a getting started wizard is displayed. The **Chassis Deployment Wizard** enables configuration of the following:

- System time
- Alert settings
- iDRAC9 Quick deployment settings
- Network IOM access settings
- Firmware updates
- Proxy settings
- MCM group definition

2.2 PowerEdge MX7000 Multi-Chassis Management groups

Multiple PowerEdge MX9002m modules are grouped to form domains called Multi-Chassis Management (MCM) groups. A single MCM group can have up to 10 chassis, where a single chassis serves as the lead and the remaining chassis are members. The OME-M console supports a daisy chain topology using the redundant 1GbE ports on the MX9002m module.

The following tasks can be done with an MCM group:

- View the rollup health status of the OME-M chassis group
- Automatically propagate lead chassis settings to member chassis
- Perform any chassis operating on the entire group
- Create a Scalable Fabric Architecture with the MX9116n FSE and MX7116n FEM

Note: See **Chapter 3** for more information about Scalable Fabric Architectures.

Creating an MCM group consists of physically cabling all MX9002m modules in the supported daisy chain topology, and then using the OME-M console to create the group. Figure 17 shows the supported MX9002m module cabling method. Chassis one, two, and three are shown while chassis *n* represents the last chassis in the group.
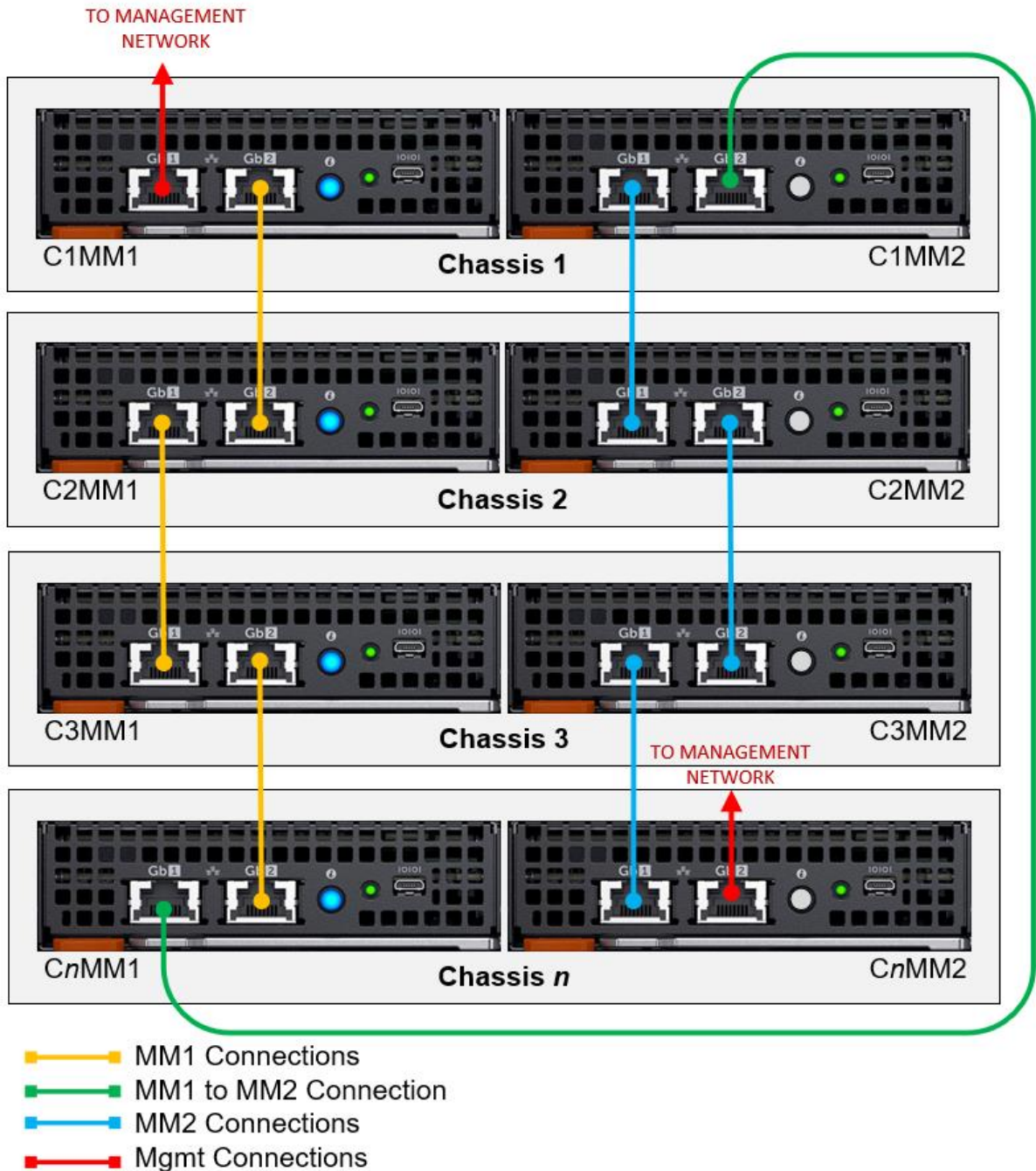
Figure 17   Dell EMC PowerEdge MX9002m module daisy chain cabling

Table 1 lists each MX9002m module and the associated connections that are shown in Figure 17.

D∕LLEMC

Table 1   MX9002m module cable requirements

C*x*MM*x*	Port Gb1	Port Gb2
C1MM1	To management network	C2MM1 Port Gb1
C2MM1	C1MM1 Port Gb2	C3MM2 Port Gb1
C3MM1	C2MM1 Port Gb2	C*n*MM1 Port Gb1
C*n*MM1	C3MM1 Port Gb2	C1MM2 Port Gb2
C1MM2	C2MM2 Port Gb2	C*n*MM1 Port Gb2
C2MM2	C3MM2 Port Gb2	C1MM2 Port Gb1
C3MM2	C*n*MM2 Port Gb2	C2MM2 Gb1
C*n*MM2	To management network	C3MM2 Port Gb1

Note: For additional information about MCM connections, see Appendix A.4.

Once cabling is complete, the OME-M console is used to create the MCM group. This is done by either the **Chassis Deployment Wizard** or by going to to the designated lead chassis and clicking **Configure → Create Chassis Group** and completing the MCM group wizard. Figure 18 shows a two-member MX7000 MCM group, **Group1,** with two participating chassis members.
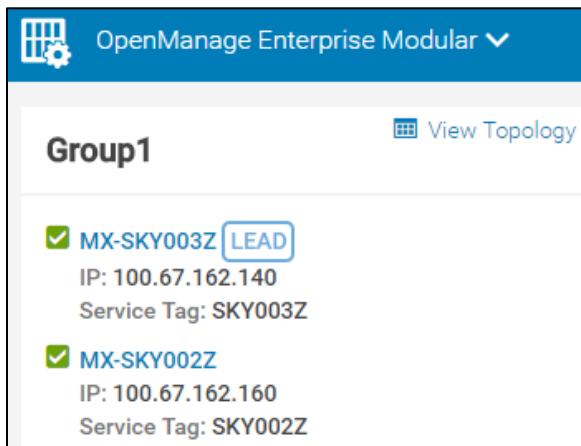
Figure 18   OME-M console – A healthy MCM group
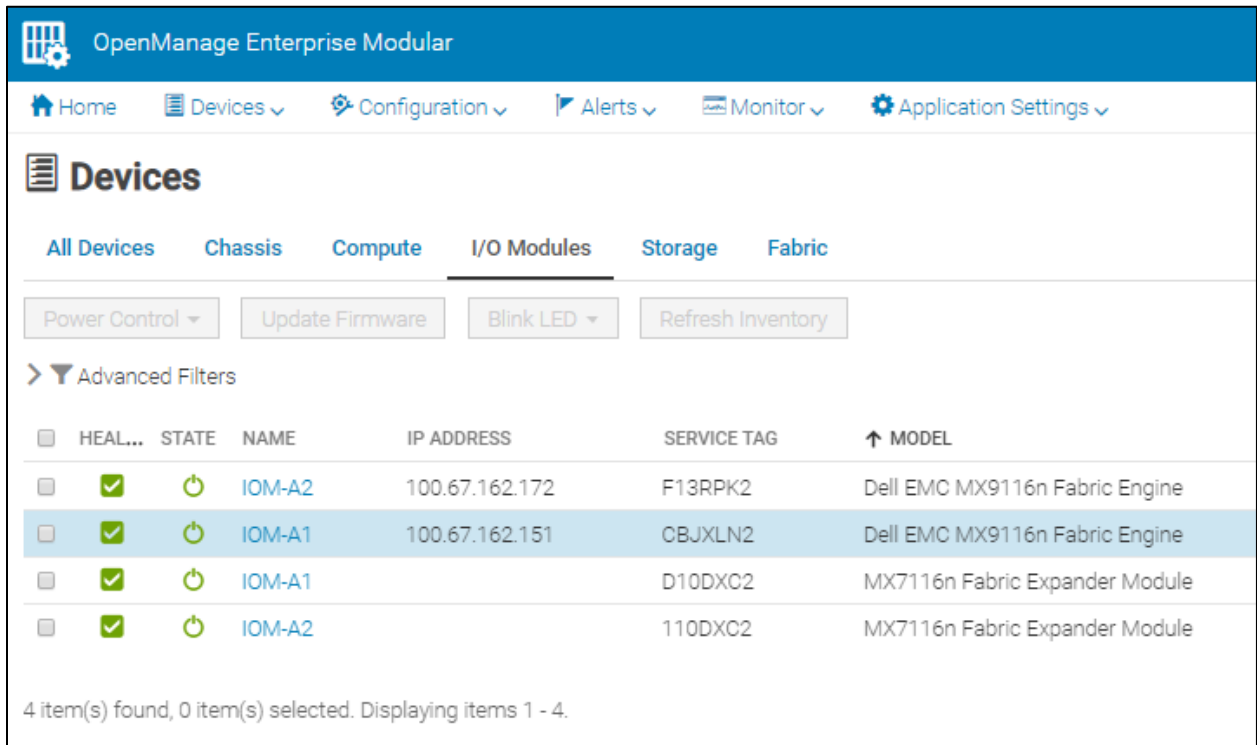
Note: In the event the lead chassis fails, an administrator can access a member chassis through its assigned IP address to determine which chassis has taken over the lead role. In the event the chassis is not in the lead role, a caution banner is displayed.

2.3 PowerEdge MX7000 OME-M component management

All switches running OS10EE Enterprise Edition (OS10EE) automatically form a redundant management cluster that provides a single REST API endpoint to OME-M to manage all switches in a chassis or across all chassis in an MCM group. This SmartFabric management endpoint exists even if the switches are not running in SmartFabric mode.

Figure 19 shows all the PowerEdge MX networking IOMs in the MCM group. The REST API endpoint is used for switch life cycle management from the OME-M console in Full Switch mode, and most configuration operations for SmartFabric Services mode. Some administrative tasks include:

- View IOM-specific alerts
- Power cycle the IOM
- Configure or change device management information
- Configure supported breakout port modes
- Configure port administrative states
- Configure MTU
- Toggle Auto Negotiation
- Update IOM firmware

Figure 19   OME-M console – I/O Modules inventory

DELLEMC

Figure 20 shows using the OME-M console to change the **Breakout Type** for **port-group1/1/1** on the **IOM-A1**, a MX9116n FSE. This behavior is similar for the MX5108n.
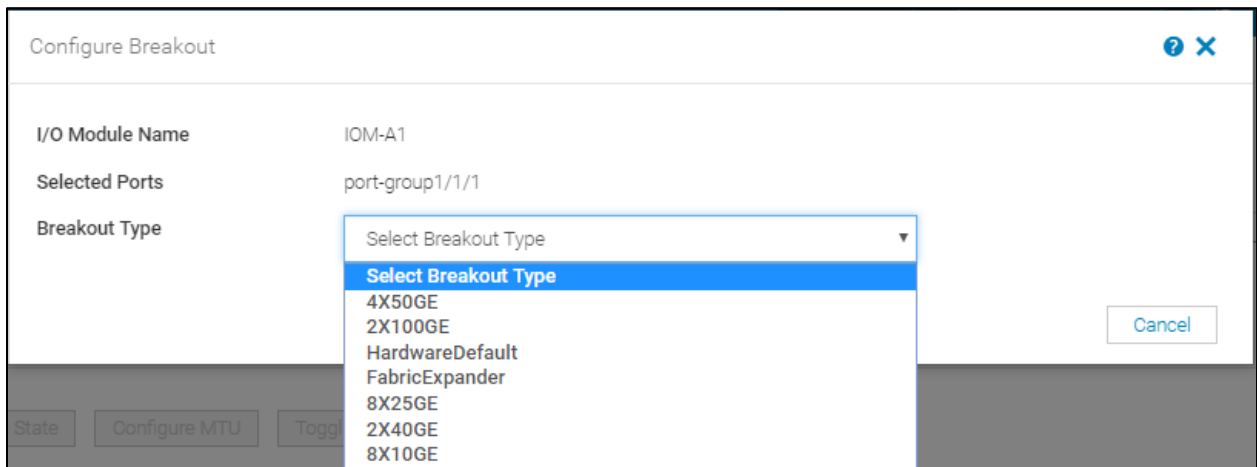
Figure 20   OME-M console – MX9116n FSE breakout mode options

---

Note: To change the breakout on a port, the **HardwareDefault** option must be selected and applied before choosing the appropriate breakout type.

DELLEMC

3 PowerEdge MX Scalable Fabric architecture

A Scalable Fabric spans multiple chassis and enables them to behave like a single chassis from a networking perspective. In contrast, an MCM group enables multiple chassis to be managed like a single chassis.

A Scalable Fabric consists of two main components - one pair of MX9116n FSEs in the first two chassis, and extra pairs of MX7116n FEMs in the remaining chassis. Each MX7116n FEM connects to the MX9116n FSE corresponding to its fabric and slot. This hardware-enabled architecture applies regardless of whether the switch is running in Full Switch or SmartFabric mode.

Note: See Chapter 4 for more information about Full Switch and SmartFabric modes.

Figure 21 shows up to 10 MX7000 chassis in a single Scalable Fabric architecture. The first two chassis hold a combination of the MX9116n FSE and MX7116n FEM, while chassis 3 through 10 houses only the MX7116n FEM. All connections in Figure 21 use QSFP28-DD connections.
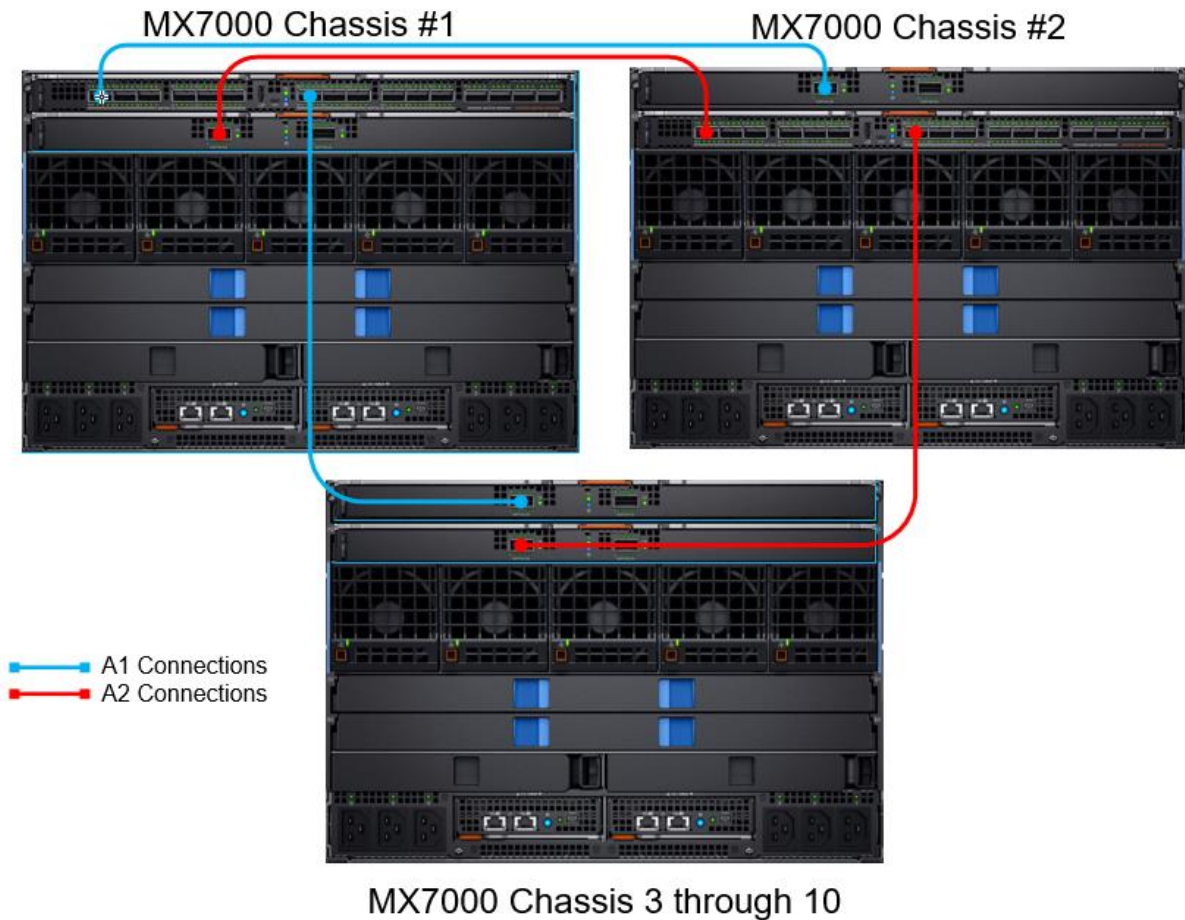
Figure 21   Scalable Fabric example using Fabric A

Table 2 shows the recommended IOM slot placement when creating a Scalable Fabric Architecture.

Table 2   Scalable Fabric Architecture maximum recommended design

MX7000 chassis	Fabric Slot	IOM Module
Chassis 1	A1	MX9116n FSE
	A2	MX7116n FEM
Chassis 2	A1	MX7116n FEM
	A2	MX9116n FSE
Chassis 3–10	A1	MX7116n FEM
	A2	MX7116n FEM

To provide further redundancy and throughput to each compute sled, Fabric B can be used to create an extra Scalable Fabric Architecture. Utilizing Fabric A and B can provide up to four 25 GbE connections to each MX740c or eight 25 GbE connections to each MX840c. Chassis 3 through 10 will be connected similarly to the example shown in Figure 21.

▬▬▬▬▬ A Fabric Connections
▬▬▬▬▬ B Fabric Connections
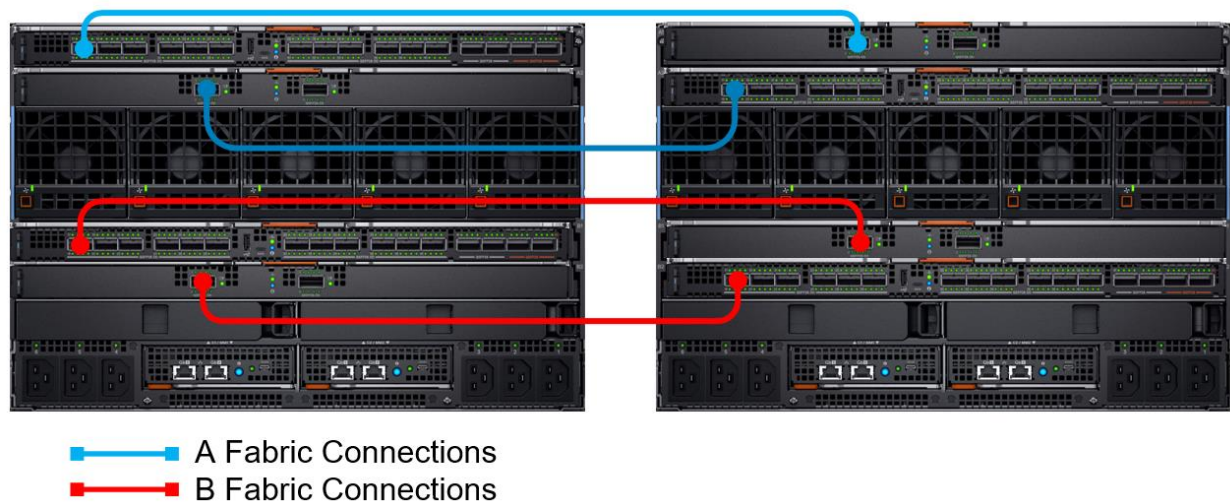
Figure 22   Two Scalable Fabrics spanning two PowerEdge MX7000 chassis

The following restrictions and guidelines are in place when building a Scalable Fabric:

- Mixing IOM types in the same Scalable Fabric (for example, MX9116n FSE in fabric slot A1 and MX5108n in fabric slot A2) is not supported. See Appendix C.6 for information on **unsupported** configurations.
- All participating MX9116n FSEs and MX7116n FEMs must be in MX7000 chassis that are part of the same MCM group. See Chapter 2.2 for more information.
- When using both Fabric A and B for a Scalable Fabric Architecture, the following restrictions apply:

    – IOM placement for each fabric must be the same in each chassis. For instance, with an MX9116n FSE in chassis 1 fabric slot A1, then the second MX9116n FSE should be in chassis 1 fabric slot B1. This only applies if the **2xMX9116n in different chassis** topology is selected.

- For chassis 3 through 10, that only contain MX7116n FEMs, they must connect to the MX9116n FSE that is in the same slot and in the same MCM chassis group.

Note: For information about the recommended MX9116n FSE port connectivity order, see Appendix A1.

3.1 QSFP28 double density connectors

A key technology that enables the Scalable Fabric Architecture is the QSFP28 double-density (DD) connector. The QSFP28-DD form factor expands on the QSFP28 pluggable form factor by doubling the number of available lanes from four to eight, with each lane operating at 25 GbE, the result is 200 GbE for each connection.

Figure 23 shows that the QSFP28-DD connector is slightly longer than the QSFP28 connector. This is to enable for the second row of pads that carry the additional four 25 GbE lanes. This allows the form factor of a single QSFP28-DD port to present as two QSFP28 logical ports.

Note: A 100GbE QSFP28 optic can be inserted into a QSFP28-DD port resulting in 100GbE of available bandwidth. The other 100GbE will not be available.
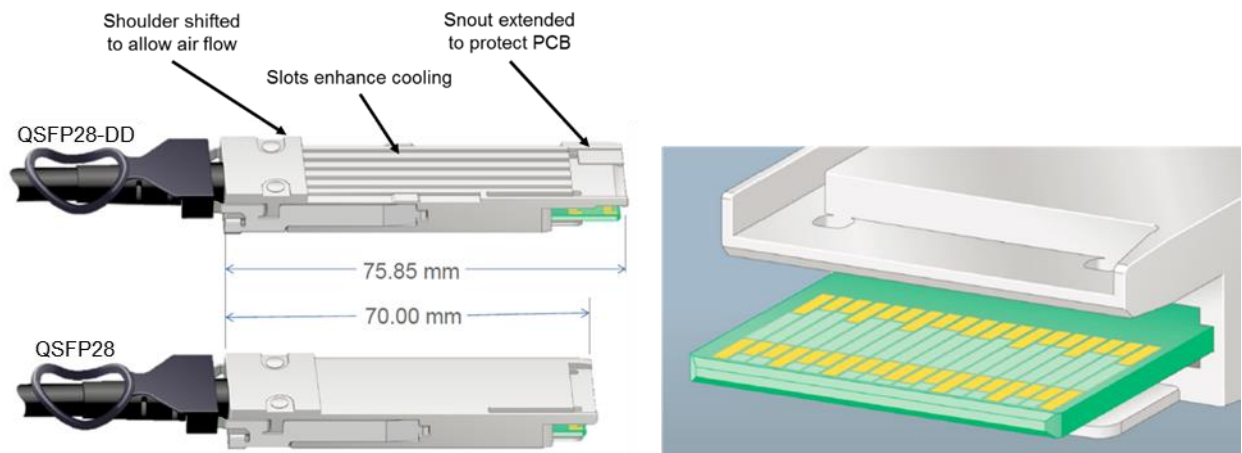
Figure 23   QSFP28-DD and QSFP28 physical interfaces

QSFP28-DD cables and optics build on the current QSFP28 naming convention. For example, the current 100GbE short range transceiver has the following description:

```
Q28-100G-SR4: Dell Networking Transceiver, 100GbE QSFP28, SR4, MPO12, MMF
```

The equivalent QSFP28-DD description is easily identifiable:

```
Q28DD-200G-2SR4: Dell Networking Transceiver, 2x100GbE QSFP28-DD, 2SR4, MPO12-
DD, MMF
```

Note: See Appendix B for more information about the various media types and their use in the PowerEdge MX7000.

3.2 Interfaces and port groups

On the MX9116n FSE and MX5108n, server-facing interfaces are internal and are enabled by default. To view the backplane port connections to servers, use the `show inventory media` command.

In the output, a server-facing interface displays `INTERNAL` as its media. A `FIXED` port does not use external transceivers and always displays as `Dell EMC Qualified true`.

```
OS10# show inventory media
 -------------------------------------------------------------------------------
                             System Inventory Media
 -------------------------------------------------------------------------------
Node/Slot/Port  Category    Media                        Serial-Number  Dell EMC-
                                                                        Qualified
 -------------------------------------------------------------------------------
1/1/1              FIXED      INTERNAL                                      true
1/1/2              FIXED      INTERNAL                                      true
1/1/3              FIXED      INTERNAL                                      true
1/1/4              FIXED      INTERNAL                                      true
1/1/5              FIXED      INTERNAL                                      true
1/1/6              FIXED      INTERNAL                                      true
1/1/7              FIXED      INTERNAL                                      true
1/1/8              FIXED      INTERNAL                                      true
1/1/9              FIXED      INTERNAL                                      true
1/1/10             FIXED      INTERNAL                                      true
1/1/11             FIXED      INTERNAL                                      true
1/1/12             FIXED      INTERNAL                                      true
1/1/13             FIXED      INTERNAL                                      true
1/1/14             FIXED      INTERNAL                                      true
1/1/15             FIXED      INTERNAL                                      true
1/1/16             FIXED      INTERNAL                                      true
1/1/17          QSFP28-DD  QSFP28-DD 200GBASE 2SR4 AOC  TW04829489D0007    true
1/1/18          QSFP28-DD  QSFP28-DD 200GBASE 2SR4 AOC  TW04829489D0007    true
1/1/19          Not Present
1/1/20          Not Present
1/1/21          Not Present

-------------------- Output Truncated -------------------------------------

1/1/37          QSFP28-DD  QSFP28-DD 200GBASE 2SR4 AOC  TW04829489J0021    true
1/1/38          QSFP28-DD  QSFP28-DD 200GBASE 2SR4 AOC  TW04829489J0021    true
1/1/39          QSFP28-DD  QSFP28-DD 200GBASE 2SR4 AOC  TW04829489J0024    true
1/1/40          QSFP28-DD  QSFP28-DD 200GBASE 2SR4 AOC  TW04829489J0024    true
1/1/41          QSFP28     QSFP28     100GBASE CR4 2M   CN0APX0084G1F05     true
1/1/42          QSFP28     QSFP28     100GBASE CR4 2M   CN0APX0084G1F49     true

-------------------- Output Truncated -------------------------------------
```

To view the server-facing interface port status, use the `show interface status` command. Server-facing ports are numbered `ethernet1/1/1` up to `ethernet1/1/16`.

Note: Even-numbered Ethernet ports between 1/1/1-1/1/16 are reserved for quad port NICs.

For the MX9116n FSE, servers that have a dual-port NIC connect only to odd-numbered internal Ethernet interfaces; for example, a MX740c in slot one would be `ethernet1/1/1`, and a MX840c in slots five and six occupies `ethernet1/1/9 and ethernet1/1/11`.

A port group is a logical port that consists of one or more physical ports, and provides a single interface. On MX IOMs, only the MX9116n FSE supports port groups:

- QSFP28-DD – Port groups 1 through 12
- QSFP28 – Port groups 13 and 14
- Unified port groups – Port groups 15 and 16

Figure 24 shows these port groups along the top and the bottom shows the physical ports in each port group. For instance, QSFP28-DD port group 1 has members Ethernet 1/1/17 and Ethernet 1/1/18, and unified port 15 has a single member, Ethernet 1/1/43.
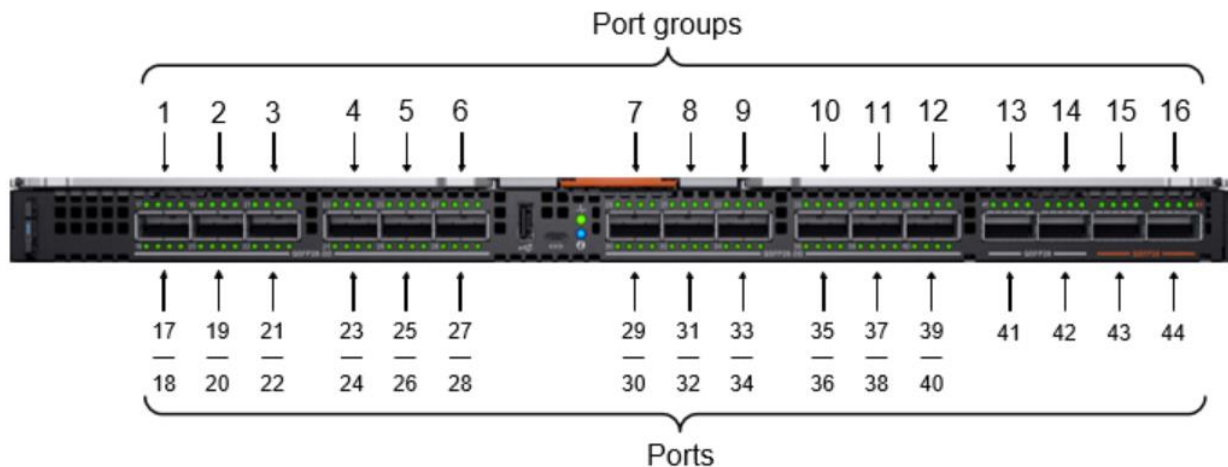
Figure 24   MX9116n FSE port groups

3.2.1   QSFP28-DD port groups

On the MX9116n FSE, QSFP28-DD port groups are 1 through 12 and contain ports `ethernet 1/1/17` through `1/1/40` and are used to:

- Connect to a MX7116n FEM to extend the Scalable Fabric
- Connect to an Ethernet rack server or storage device
- Connect to another networking device

By default, QSFP28-DD port groups 1 through 9 are in `fabric-expander-mode`, which is an 8x25GbE interface, used only to connect to MX7116n FEMs in extra chassis. The interfaces from the MX7116n FEM appear as standard Ethernet interfaces from the perspective of the MX9116n FSE.

Figure 25 illustrates how the QSFP28-DD cable provides 8x25GbE lanes between the MX9116n FSE and a MX7116n FEM.
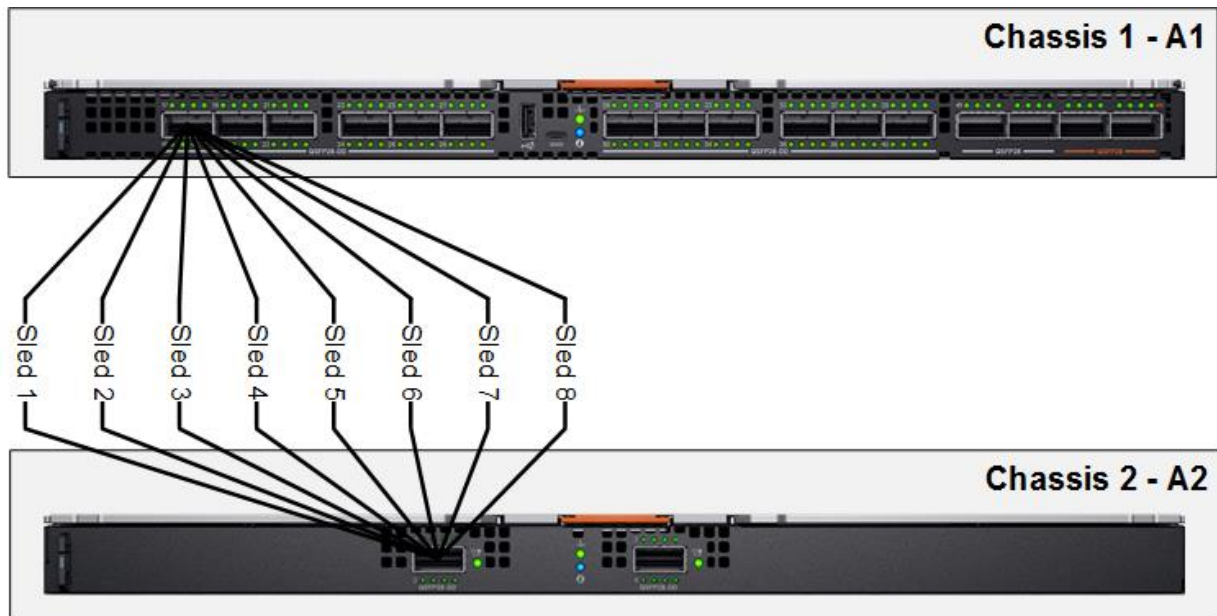
Figure 25   QSFP28-DD connection between MX9116n FSE and MX7116n FEM

Note: Compute sleds with dual-port NICs require only MX7116n FEM port 1 to be connected.

Besides `fabric-expander-mode,` QSFP28-DD port groups support the following Ethernet breakout configurations:

- 8x 10GbE – Breakout a QSFP28-DD port into eight 10GbE interfaces
- 8x 25GbE – Breakout a QSFP28-DD port into eight 25GbE interfaces
- 2x 40GbE – Breakout a QSFP28-DD port into two 40GbE interfaces
- 2x 100GbE – Breakout a QSFP28-DD port into two 100GbE interfaces

By default, QSFP28-DD port groups 10 through 12 are in 2x 100 GbE breakout mode.

Note: Before changing the port breakout configuration from one setting to another, the port must first be set back to the default setting.

3.2.2   Single-density QSFP28 port groups

On the MX9116n FSE, single-density QSFP28 port groups are 13 and 14 and contain ports `Ethernet 1/1/41` and `1/1/42` and are used to connect to upstream networking devices. By default, both port groups are set to 1x100GbE. Port groups 13 and 14 support the following Ethernet breakout modes:

The MX9116n FSE single-density QSFP28 port groups support the following Ethernet breakout configurations:

- 4x 10GbE – Breakout a QSFP28 port into four 10GbE interfaces
- 1x 40GbE – Set a QSFP28 port to 40GbE mode
- 4x 25GbE – Breakout a QSFP28 port into four 25GbE interfaces
- 2x 50GbE – Breakout a QSFP28 port into two 50GbE interfaces
- 1x 100GbE – Reset the unified port back to the default, 100GbE mode

DELLEMC

When connecting a pair of PowerEdge MX IOMs in SmartFabric mode to an upstream switch pair, ensure the following:

- Both upstream switches must be connected using technologies such as Dell EMC VLT or Cisco Nexus vPC.
- The upstream switch ports must be in a port channel using LACP.
- A compatible Spanning Tree Protocol (STP) is configured. For more information see Appendix C.2.

3.2.3   Unified port groups

In OS10EE, unified port groups operate as either Ethernet or FC. By default, both unified port groups, 15 and 16, are set to 1x 100GbE. To activate the two port groups as FC interfaces, use the command `mode fc`.

The MX9116n FSE unified port groups support the following Ethernet breakout configurations:

- 4x 10GbE – Breakout a QSFP29 port into four 10GbE interfaces
- 1x 40GbE – Set a QSFP28 port to 40GbE mode
- 4x 25GbE – Breakout a QSFP port into four 25GbE interfaces
- 2x 50GbE – Breakout a QSFP28 port into two 50GbE interfaces
- 1x 100GbE – Reset the unified port back to the default, 100GbE mode

The MX9116n FSE unified port groups support the following FC breakout configurations:

- 4x 8Gb – Breakout a unified port group into four 8Gb FC interfaces
- 2x 16Gb – Breakout a unified port group into two 16Gb FC interfaces
- 4x 16Gb – Breakout a unified port group into four 16Gb FC interfaces
- 1x 32Gb – Breakout a unified port group into one 32Gb FC interface
- 2x 32Gb – Breakout a unified port group into two 32Gb FC interfaces
- 4x 32Gb – Breakout a unified port group into four 32Gb FC interfaces, rate limited

Note: See Appendix C.1 for more information about rate limited 32Gb FC.

DELLEMC

3.2.4   Virtual ports

A virtual port is a logical OS10EE port that connects to a downstream server and has no physical hardware location on the switch. Virtual ports are created when an MX9116n FSE onboards a MX7116n FEM. The onboarding process consists of discovery and configuration. An attached MX9116n FSE in Full Switch mode automatically discovers the MX7116n FEM when the following conditions are met:

- The MX7116n FEM is connected to the MX9116n FSE by attaching a Dell qualified cable between the QSFP28-DD ports on both devices.
- The interface for the QSFP28-DD port group connected to the MX9116n FSE is in 8x25GbE FEM mode.
- At least one blade server is inserted into the MX7000 chassis containing the MX7116n FEM.

To verify that a MX7116n FEM is communicating with the MX9116n FSE, enter the `show discovered-expanders` command.

```
MX9116n-FSE # show discovered-expanders
Service    Model    Type  Chassis         Chassis-slot Port-group Virtual
tag                        service-tag                             Slot-Id
--------------------------------------------------------------------
D10DXC2    MX7116n  1     SKY002Z         A1           1/1/1
           FEM
```

Once the MX7116n is discovered, the logical interfaces must be provisioned. SmartFabrics does this automatically. However, in full switch mode the `unit-provision node/unit-id provision_name` command must be used.

Note: For more information about virtual ports, see the *OS10EE Enterprise Edition Users Guide*.

DELLEMC

3.3 Embedded Top of Rack switching

Most environments with blade servers also have rack servers. Figure 26 shows a typical design having rack servers connecting to their respective Top of Rack (ToR) switches and blade chassis connecting to a different set of ToR switches. If the storage array was Ethernet-based, it can be connected to the core/spine.

4x Dell EMC PowerEdge 740 (compute)

Dell EMC PowerEdge M1000e (compute)

Dell EMC Unity 500F (storage)

VLTi
100 GbE link
25 GbE link
Storage link
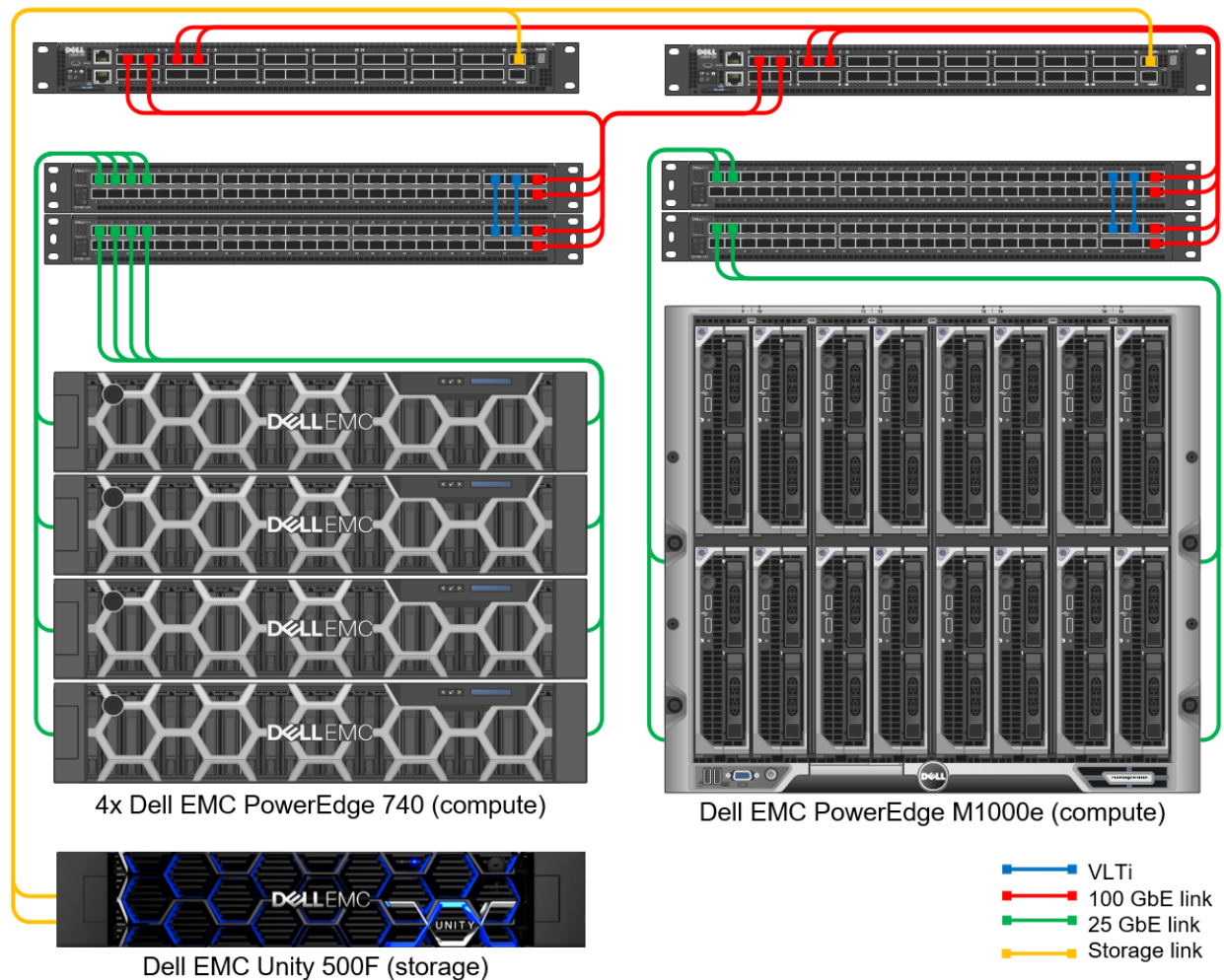
Figure 26   Traditional mixed blade/rack networking

Note: Depending on where the L2/L3 boundary is at, an additional pair of leaf switches may be added for the storage connections.

This design is inefficient and expensive. Communication between rack and blade servers must traverse the core, increasing latency, and the storage array consumes expensive core switch ports. All of this results in increased operations cost from the increased number of managed switches.

Embedded ToR functionality is built into the MX9116n FSE. Configure any QSFP28-DD port to break out into 8x10GbE or 8x25GbE and connect the appropriate cables and optics. This enables all servers and storage connect directly to the MX9116n FSE and communication between all devices that are kept within the switch. This provides a single point of management and network security while reducing cost and improving performance and latency.

In Figure 26, above, there are eight switches in total. In the following figure, using embedded ToR, switch count is reduced to the two MX9116n FSE in the two chassis:
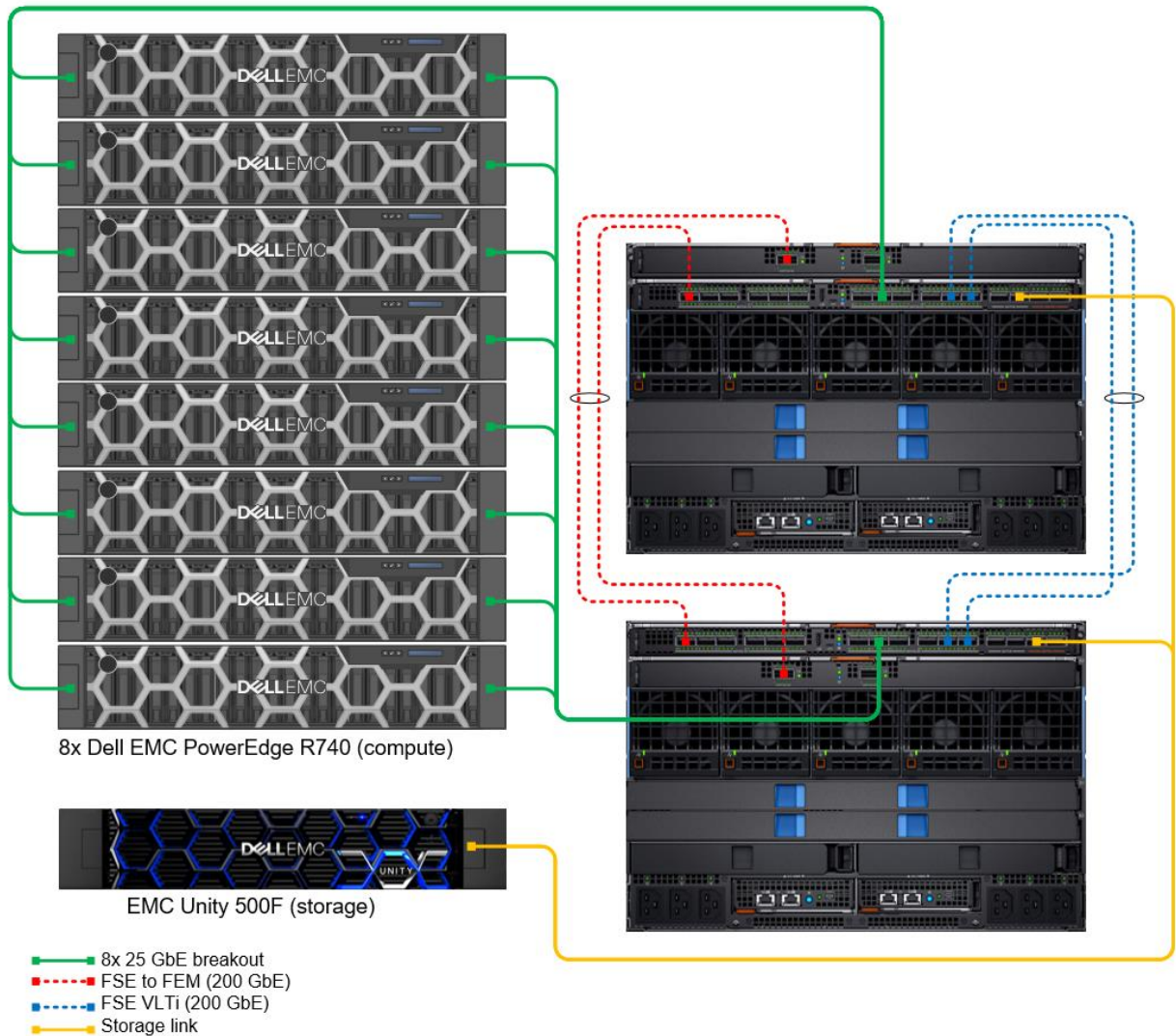
Figure 27   MX9116n FSE embedded ToR

Note: Embedded ToR is only supported in Full Switch mode. Support for Embedded ToR in SmartFabric mode will be added in a future release.

3.4   Scalable Fabric OS10EE CLI Commands

Additional OS10EE CLI commands are available on MX9116n to assist with MX7116n FEM management:

- `show discovered-expanders` – displays the MX7116n FEMs attached to the MX9116n FSEs
- `show unit-provision` – displays the unit ID and service tag of the MX7116n FEM attached to an MX9116n FSE
- `unit-provision` – Provisions the unit ID of a Fabric Expander Module that is attached to a Fabric Switching Engine in Full Switch mode

4 PowerEdge MX Ethernet switch operating modes

PowerEdge MX Ethernet switches support Dell EMC Networking OS10EE Enterprise Edition (OS10EE). OS10EE is a Network Operating System supporting multiple architectures and environments.

When running OS10EE, the MX9116n FSE and the MX5108n operate in one of two modes:

- Full Switch Mode – Enabled by default, all switch-specific OS10EE capabilities are available
- SmartFabric Mode – IOMs operate as Layer 2 I/O aggregation devices and are managed through the OME-M console

Table 3 outlines the differences between the two operating modes. The differences between operating modes apply to both the MX9116n FSE and the MX5108n.

Table 3   IOM operating mode differences

Full Switch mode	SmartFabric mode

Configuration changes are persistent during power cycle events.	Only the configuration changes made using the following OS10EE commands, are persistent across power cycle events. All other CLI configuration commands are disabled.
	
	`clock`
	`hostname`
	`interface`
	`ip nameserver`
	`logging`
	`management route`
	`ntp`
	`snmp-server`
	`username`
	`spanning-tree`
	`vlan`

All switch interfaces are assigned to VLAN 1 by default and are in the same Layer 2 bridge domain.	Layer 2 bridging is disabled, and interfaces must join a bridge domain (VLAN) before being able to forward frames.
All configurations changes are saved in the running configuration by default. To display the current configuration, use the `show running-configuration` command.	Verify configuration changes using feature-specific show commands, such as `show interface` and `show vlan`, instead of `show running-configuration`.

4.1   Full Switch mode

PowerEdge MX switches operate in Full Switch mode by default. In Full Switch mode, all OS10EE features and functions that are supported by the hardware are available. This mode is useful when there is a dedicated network team that maintains control over all switching devices.

From the OME-M console, specific software features can be configured:

- Configure hostname and IP management protocols, such as SNMP and NTP
- Set port status up/down and configure breakout interfaces
- Monitor system logs, alerts, and events
- Update and manage firmware
- View physical topology
- Power management

Full Switch mode has two primary use cases:

1. A network design that is not available with SmartFabric Services.
2. A network team is already in place using traditional CLI or API methods to manage all switching devices.

4.2   SmartFabric mode

A SmartFabric is a logical entity that consists of a collection of physical resources, such as servers and switches, and logical resources, such as networks, templates, and uplinks, that are managed as a single unit. When PowerEdge MX networking components are placed in SmartFabric mode, the switches operate as layer 2 I/O aggregation devices.

The OME-M console supports most switch configuration settings. SmartFabric mode supports all OS10EE `show` commands and the following subset of CLI configuration commands:

- `clock` – Configure clock parameters
- `end` – Exit to the EXEC mode
- `exit` – Exit from the current mode
- `help` – Display available commands
- `hostname` – Set the system hostname
- `interface` – Configure or select an interface
- `ip nameserver` – Configure nameserver
- `logging` – Configure system logging
- `management route` – Configure the IPV4/IPv6 management route
- `no` – Delete or disable commands in Configuration mode
- `ntp` – Configure the network time protocol
- `snmp-server` – Configure the SNMP server
- `username` – Create or modify user credentials
- `spanning-tree`

  - `disable` – Disable spanning tree globally
  - `mac-flush-timer` – Set the time used to flush MAC address entries
  - `mode` – Enable a spanning-tree mode, such as RSTP or MST
  - `mst` – Configure multiple spanning-tree (MST) modes
  - `rstp` – Configure rapid spanning-tree protocol (RSTP) mode
  - `vlan` – Configure spanning-tree on a VLAN range

Note: For additional information about `spanning-tree` and available options, see Appendix C.2.

4.3 PowerEdge MX7000 SmartFabric topologies

This section contains the supported SmartFabric topologies. The topology is chosen while deploying a SmartFabric and the deployment wizard validates that the appropriate switches and installed in the appropriate Fabric A or Fabric B slots.

4.3.1 Two MX9116n FSEs in different chassis

Figure 28 shows the supported topology for the Scalable Fabric Architecture. The blue links represent two QSFP28-DD ports in a VLTi configuration. Placing the MX9116n FSEs in separate chassis provides redundancy in the event of a hardware failure. This configuration supports placement in chassis1, fabric slot A1, and chassis 2, slot A2. Alternatively, chassis1, slot B1, and chassis 2, slot B2 is supported. The red links represent the FEM to FSE connections.

Note: A SmartFabric cannot include an IOM in Fabric A and an IOM in Fabric B.
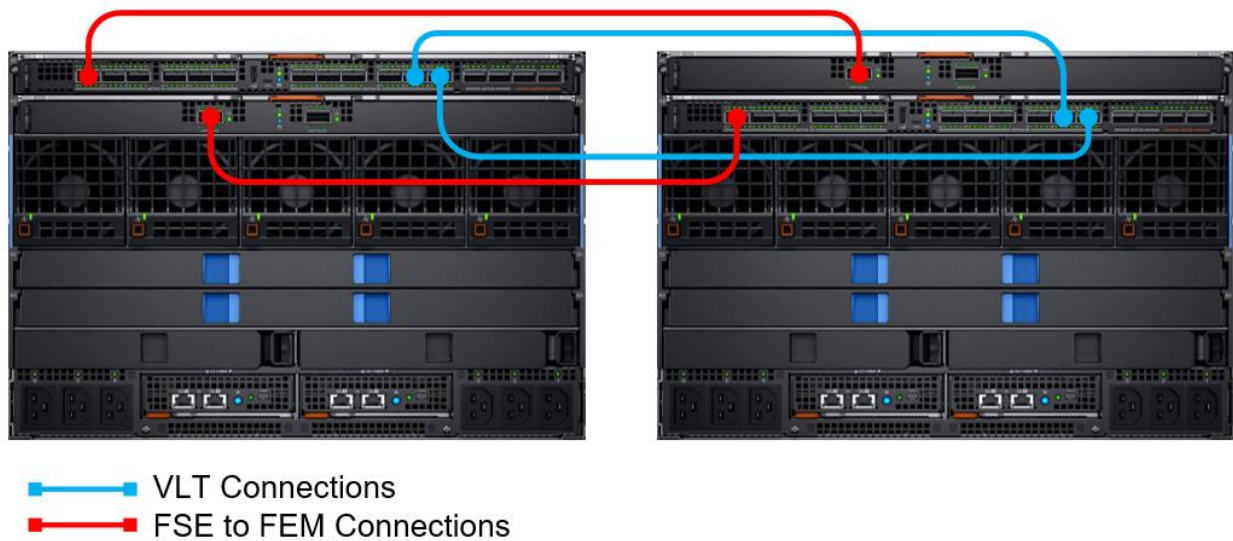

Figure 28   Two MX9116n FSEs, different chassis topology

4.3.2 Two MX9116n FSEs in the same chassis

Figure 29 shows two MX9116n FSEs in either fabric slots A1/A2 or B1/B2. The blue links represent two QSFP28-DD ports in a VLTi configuration.
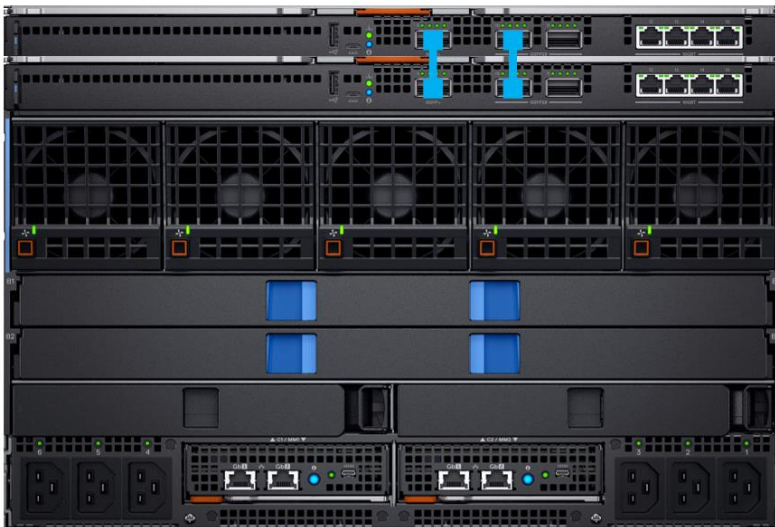

VLT Connections

Figure 29   Two MX9116n FSEs in a single chassis topology

4.3.3 Two M5108n Ethernet switches in the same chassis

Figure 30 shows two MX5108n switches in single chassis configuration. The blue links represent two ports in a VLTi configuration. The first port is a QSFP+ port and the second port is a QSFP28 port. In SmartFabric mode the QSFP28 port is configured to operate at 40GbE allowing for a valid port channel.

Note: See Appendix A.3 for more information.

VLT Connections

Figure 30   Two MX5108n in a single chassis topology

Note: A SmartFabric using the MX5108n cannot span multiple MX7000 chassis

DELLEMC

4.4 SmartFabric uplink topologies

This section outlines the supported SmartFabric topologies.

4.4.1 MX9116n FSE–Ethernet only

Figure 31 shows the wiring topology in an Ethernet only environment. The MX9116n FSE pair and the leaf pair are each configured for VLT. Uplinks are shown in green, with both links from each MX9116n FSE in a VLT port channel. In addition to the uplinks shown, if extra uplink bandwidth is needed, the QSFP28-DD port groups can be used as uplinks by changing the breakout type and cabling as described in Section 3.2.1.

Note: See Appendix A.2 for information about connecting to the Dell EMC Networking Z9100-ON ToR switch.
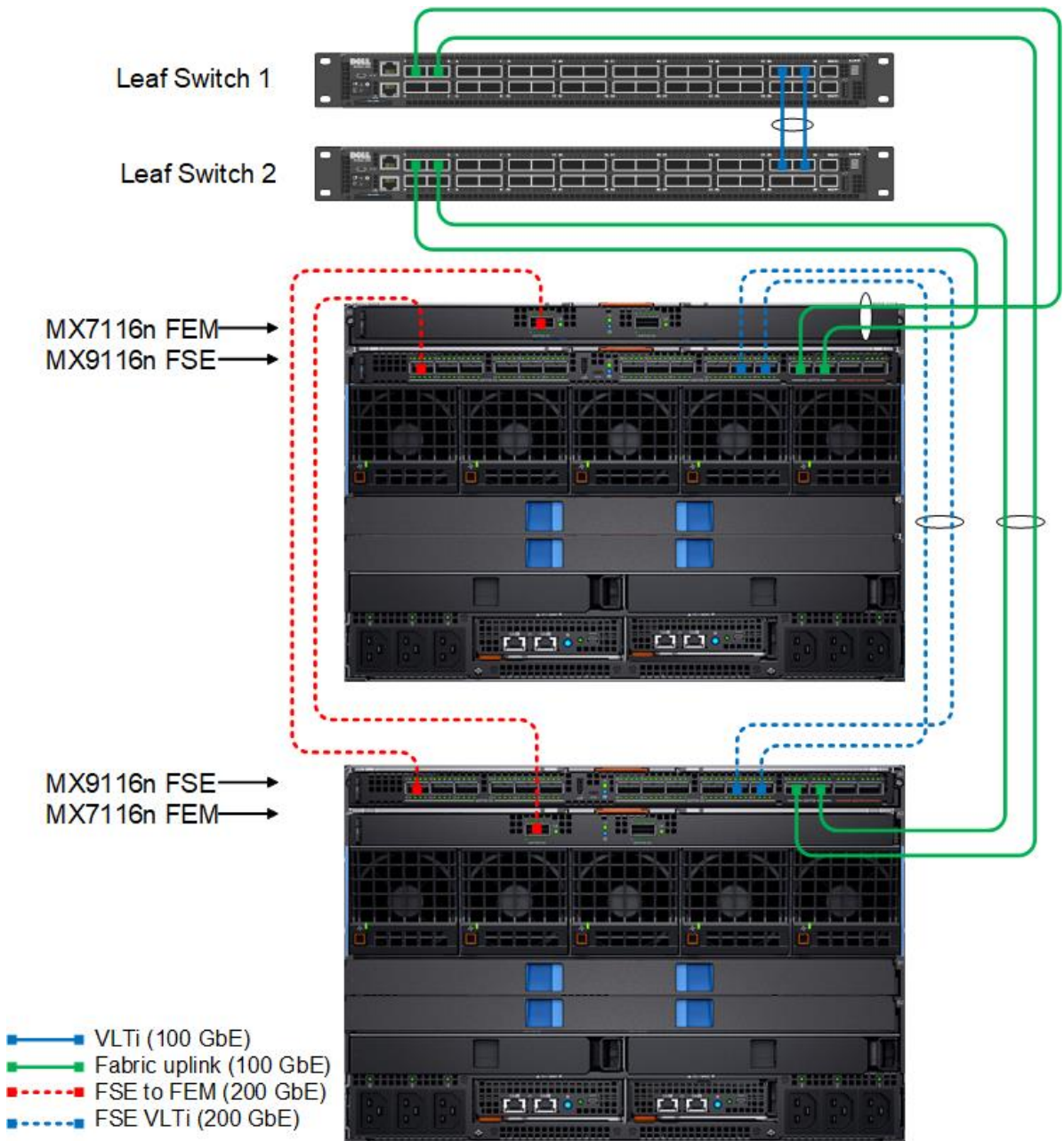

Figure 31   MX9116n FSE – Ethernet only

4.4.2 MX9116n FSE–Fibre Channel connectivity

Figure 32 builds on the Ethernet only topology and shows the two universal ports on each MX9116n FSE being used for FC connectivity. The unified ports groups, 15 and 16, support both NPIV Proxy Gateway (NPG) and direct attach F_PORT mode, not shown. In this example, two Dell Networking S4148U-ON switches are shown but any FC-capable switch can be used.

Note: FCoE uplinks must be physically separate from Ethernet uplinks.
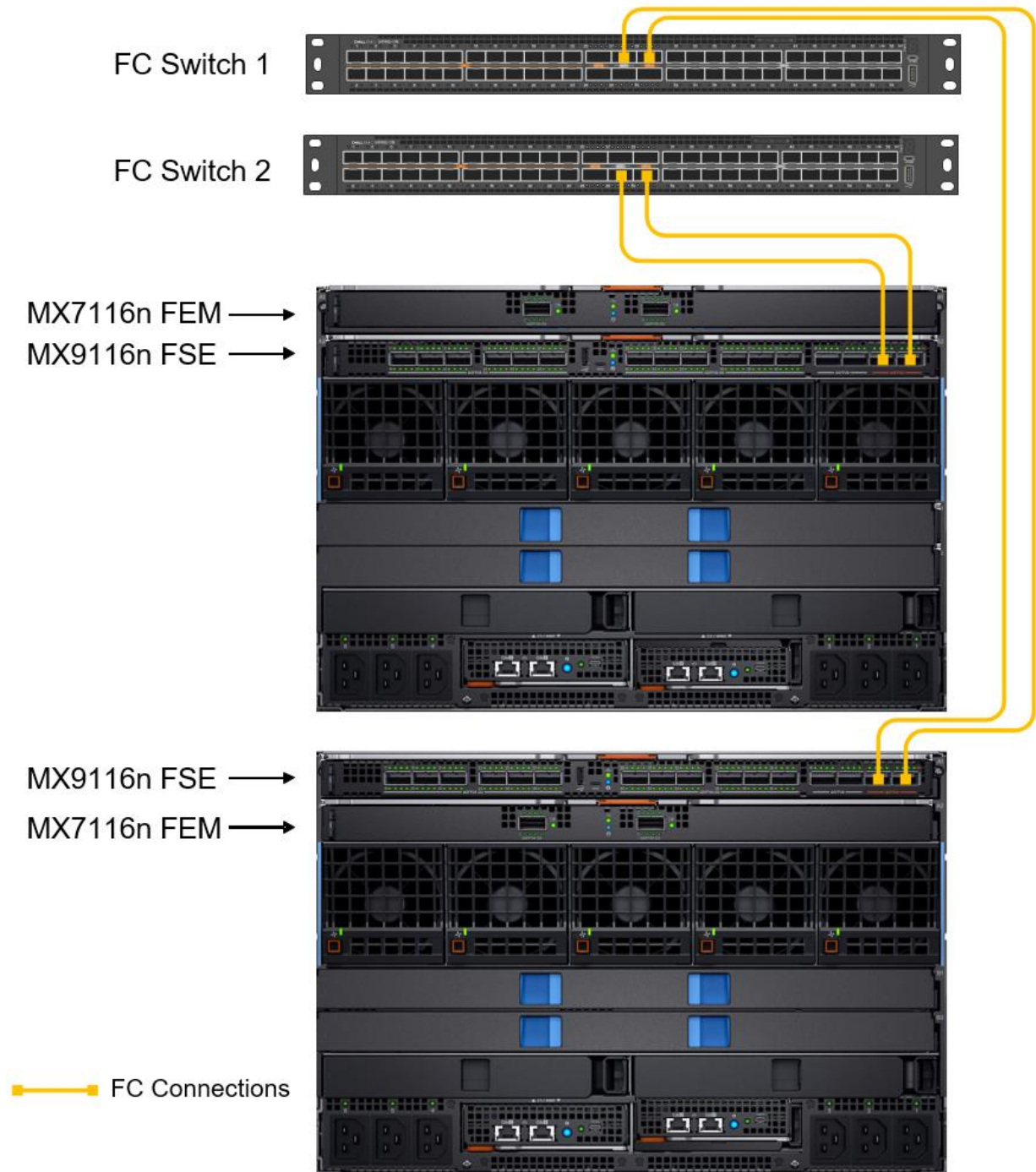
Figure 32   MX9116n FSE – Fibre Channel connectivity

4.4.3 MX5108n–Ethernet only

Figure 33 and Figure 34 show two methods for using the MX5108n in a LAN. The two diagrams highlight using either 10GBASE-T ports or QSFP28 ports with breakout for 10GbE or 25GbE uplinks. Apart from the ports and cables that are used, the configuration is the same.
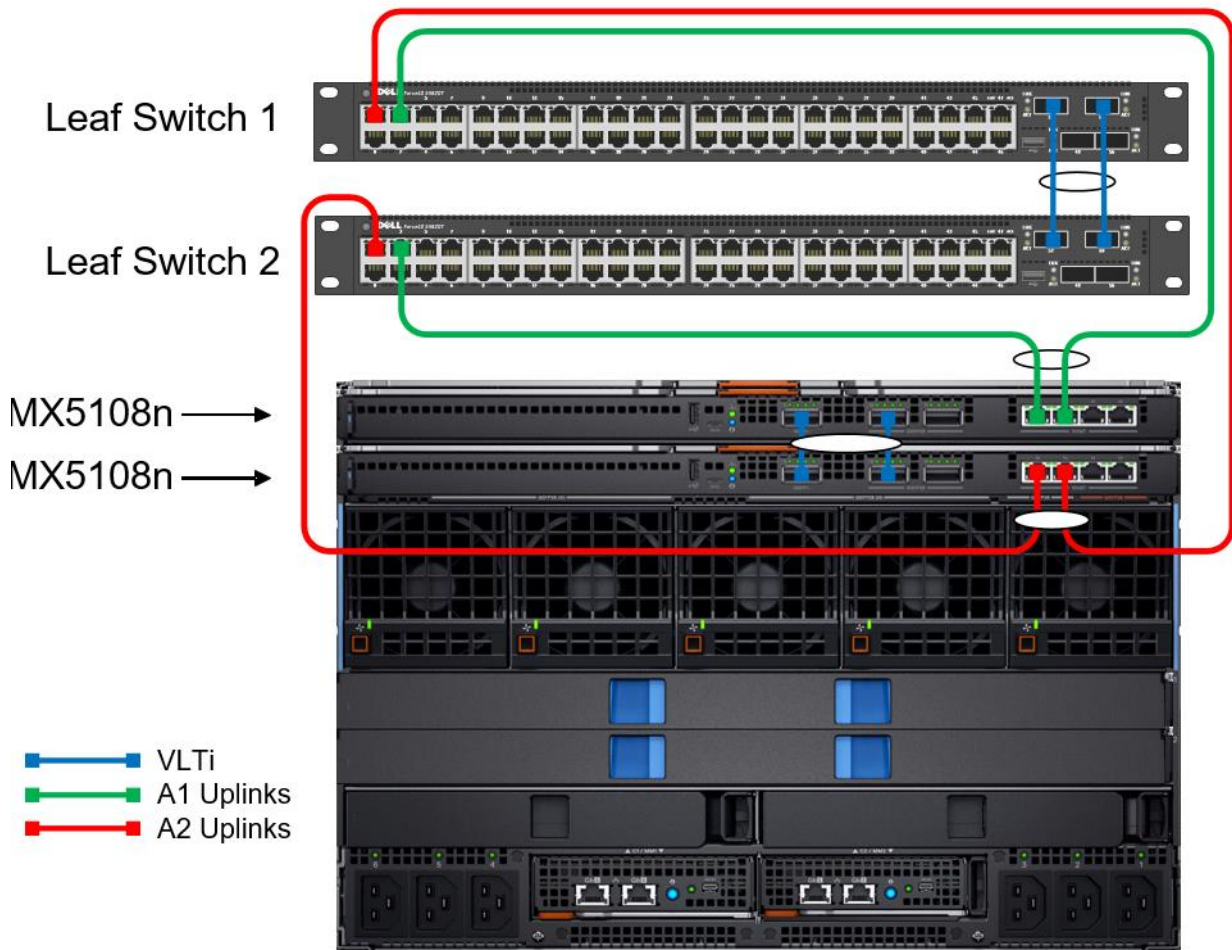
Figure 33   MX5108n – Ethernet using 10GBASE-T ports
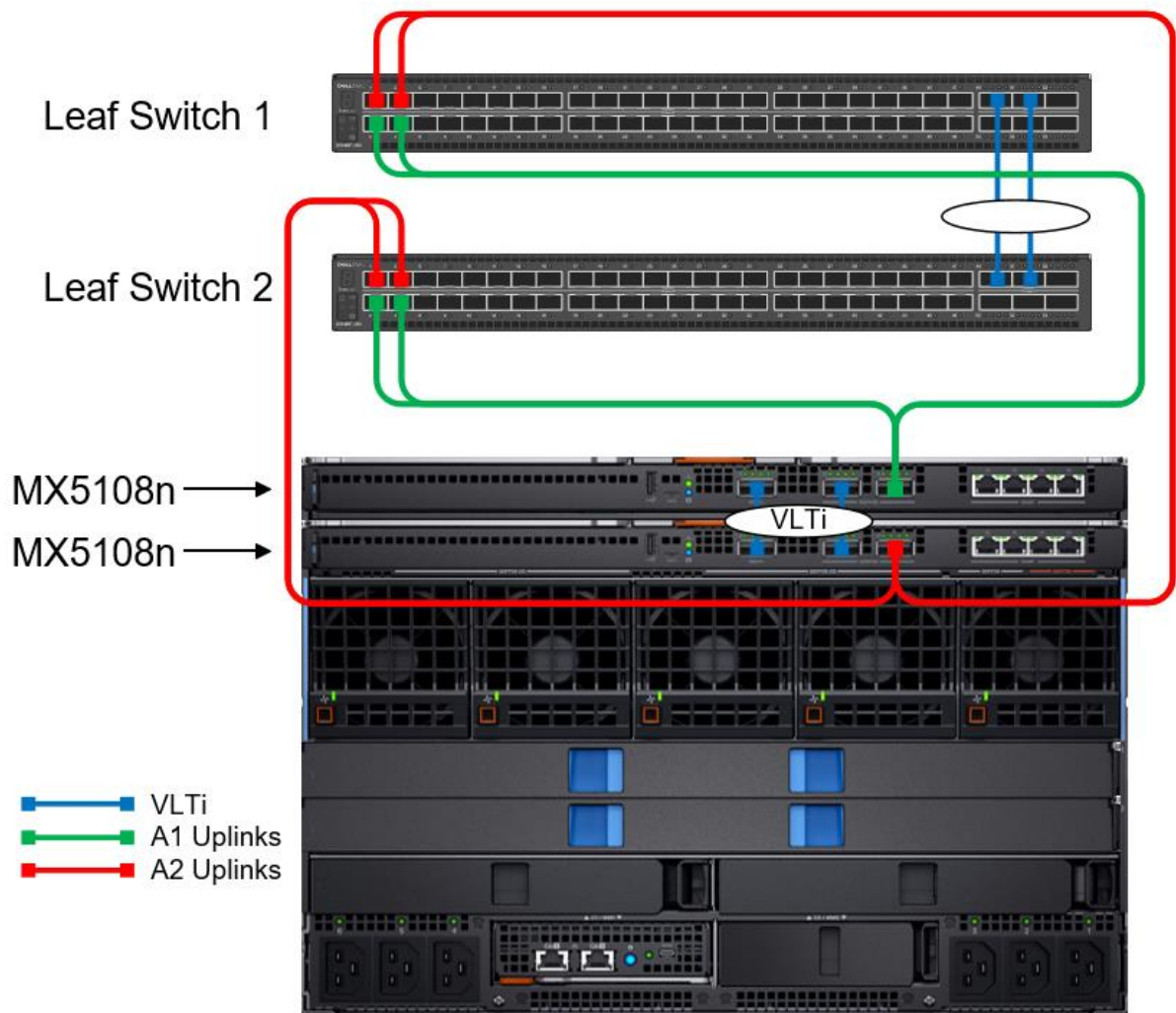
Figure 34   MX5108n – LAN using QSFP28 ports in breakout mode

4.4.4   MX5108n–FCoE FIP Snooping Bridge

Figure 35 shows how the MX5108n can add FCoE functionality to an existing LAN-only topology. The MX5108n is configured to operate in FSB mode. The MX5108n snoops FIP packets on FCoE-enabled VLANs and discovers the following information:

- End nodes (ENodes)
- Fibre channel forwarder (FCF)
- Connections between ENodes and FCFs
- Sessions between ENodes and FCFs

Figure 35   MX5108n – FSB mode

DELLEMC

4.5 Deploy a SmartFabric

SmartFabric deployment consists of four steps that are completed using the OME-M console:

1. Create the VLANs to be used in the fabric.
2. Select switches and create the fabric based on the physical topology desired.
3. Create uplinks from the fabric to the existing network and assign VLANs to those uplinks.
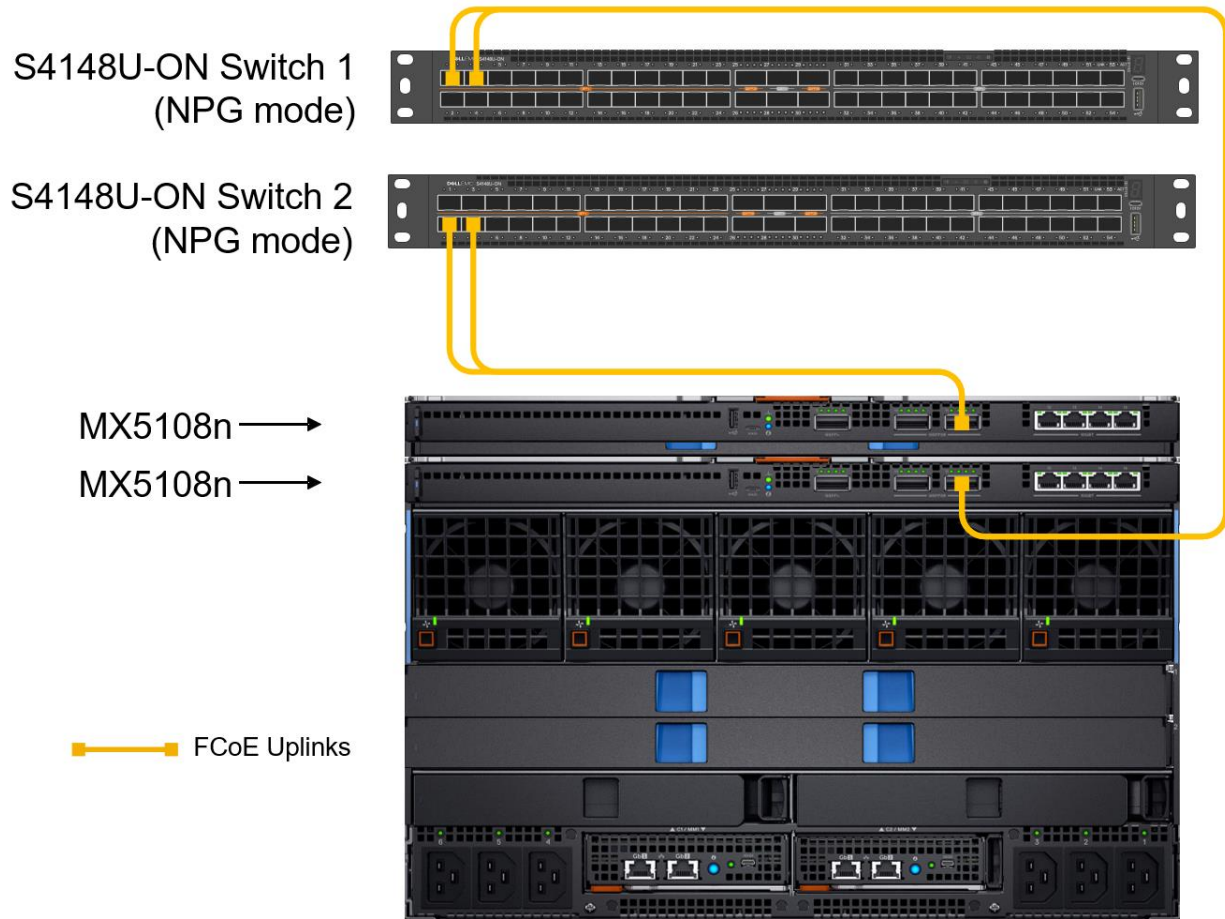4. Deploy the appropriate server templates to the compute sleds.

Note: For more information about deploying a SmartFabric, see the *Dell EMC PowerEdge MX SmartFabric Deployment Guide*.

4.5.1 VLAN management and automated QoS

Besides being able to create and assign VLANs, SmartFabric automates QoS settings based on chosen configuration options. Figure 36 shows that when defining a VLAN, one of several QoS options must be chosen. Each of these options represents a traffic class ID that is mapped to a data queue.   .

Figure 36   Defining a VLAN and choosing Network Type

Table 4 lists the traffic types and related settings.

Table 4   Traffic types and QoS settings

Traffic Type	Description	QoS Group
General Purpose (Bronze)	Used for low-priority data traffic	2
General Purpose (Silver)	Used for standard/default priority data traffic	3
General Purpose (Gold)	Used for high priority data traffic	4
General Purpose (Platinum)	Used for extremely high priority data traffic	5
Cluster Interconnect	Used for cluster heartbeat VLANs	5
Hypervisor Management	Used for hypervisor management connections such as the ESXi management VLAN	5
Storage - iSCSI	Used for iSCSI VLANs	5
Storage - FCoE	Used for FCoE VLANs	5
Storage - Data Replication	Used for VLANs supporting storage data replication such as for VMware VSAN	5
VM Migration	Used for VLANs supporting vMotion and similar technologies	5
VMware FT Logging	Used for VLANs supporting VMware Fault Tolerance	5

When the VLAN is deployed through a server template, the SmartFabric Services update all participating switches with a QoS class map. This class map uses VLAN ID classification. A QoS policy map and then assigns all traffic matching the class map to the appropriate data queue.

Note: For more information about QoS, see the *OS10EE Enterprise Edition Users Guide*.

4.5.2   VLAN scaling guidelines

Because SmartFabric mode provides network automation capabilities that Full Switch mode does not, the number of recommended VLANs differs between the modes. Table 5 provides the recommended maximum number of VLANs per fabric, uplink, and server port.

Table 5   SmartFabric VLAN scaling guidelines

OS10EE Version	Parameter	Value
10.4.0E(R3S)	Recommended max VLANs per fabric	128
	Recommended max VLANs per uplink	128
	Recommended max VLANs per server port	32

Note: These maximums are considered a best practice, not an enforced limit.

DELLEMC

A Guidelines and Best Practices

The following restrictions and guidelines are in place when building a Scalable Fabric:

- Mixing IOM types in the same Scalable Fabric (for example, MX9116n FSE in fabric slot A1 and MX5108n in fabric slot A2) is not a supported configuration.
- All MX9116n FSEs and MX7116n FEMs that participate in a Smart Fabric must be in MX7000 chassis that are part of the same MCM group.
- When using both Fabric A and B for a Scalable Fabric Architecture, the following restrictions apply:

    - IOM placement for each fabric must be the same in each chassis. For instance, with an MX9116n FSE in fabric slot A1, the second MX9116n FSE should be in fabric slot B1.
    - Chassis 3 through 10, that only contain MX7116n FEMs, connections must be made to the MX9116n FSE that is in the same slot and in the same MCM chassis group.

- Interconnecting switches in A1/A2 with switches in B1/B2, regardless of chassis, is not supported.
- When operating with multiple chassis, switches in A1/A2 or B1/B2 in one chassis must be interconnected only with other A1/A2 or B1/B2 switches respectively.
- Create symmetrical uplinks. For example, if one MX9116n FSE has two uplinks, the other MX9116n FSE must have two uplinks of the same speed.
- Enable LACP on upstream switches. LACP is required on the connected port channel.
- Ensure that switches in SmartFabric mode do not link to another pair of switches in SmartFabric mode. However, a SmartFabric can uplink to a pair of switches in Full Switch mode.

A.1 Recommended port order for MX7116n FEM connectivity

The table below outlines the recommended port order for connecting chassis with MX7116n FEM modules to the MX9116n FSE to optimize buffer utilization. Table 6 references IOMs in Fabric A, but the same guidelines apply to IOMs in Fabric B.

Table 6   Recommended PowerEdge MX7000 chassis connection order

Chassis	MX9116n FSE port group	Physical Port Numbers
1/2	Port group 1	17 and 18
3	Port group 7	29 and 30
4	Port group 2	19 and 20
5	Port group 8	31 and 32
6	Port group 3	21 and 22
7	Port group 9	33 and 34
8	Port group 4	23 and 24
9	Port group 10	35 and 36
10	Port group 5	25 and 26

Figure 37   Recommended MX7000 chassis connection order

A.2   Connecting to the Dell EMC Networking Z9100-ON spine

The Dell EMC PowerEdge MX7000 SmartFabric requires a minimum of two uplink connections to a Top of Rack (ToR) switch. This example shows how to optimize connections from the MX9116n FSEs on the PowerEdge MX7000 to the Dell EMC Networking Z9100-ON ToR switch.

The Z9100-ON groups interfaces into "tiles" with shared memory buffers as shown in Figure 38. This design distributes uplinks from each Scalable Fabric to optimal ports. As shown in Figure 39, it is recommended that each chassis fabric uplink from the MX9116n FSEs be connected to a different tile on the Z9100-ON.

Note: Figure 39 only displays the connections for Fabric A. Fabric B will be connected to the red "tiles". Any subsequent chassis will be connected in a similar fashion using the other "tiles" on the switch. See Table 7 for port connection information.
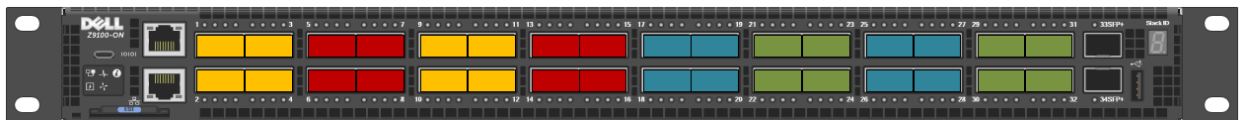
Figure 38   Z9100-ON Tile Distribution

Figure 39 displays the uplink connections from two MX9116n FSEs in two chassis. The two chassis provide the uplinks for a group of up to 10 chassis. Chassis 3-10 each have four MX7116n FEMs, which connect only to the MX9116n FSEs in the first two chassis of the respective group. A second set of 10 chassis would start with chassis 11 and 12, if up to 20 chassis are required. Chassis 1-2 and 11-12 each have two MX9116n FSEs and two MX7116n FEMs.

While any QSFP28 port can be used for Ethernet connections, the last two QSFP28 ports, interfaces 43 and 44, can be used for Ethernet or FC connectivity.

Figure 39 shows the use of all four QSFP28 ports on each of the MX9116n FSEs in fabric A. If ports 43 and 44 are needed for other connections, only ports 41 and 42 need to be connected to the Z9100-ON spine switches.

Figure 39   Uplink connections for fabric A

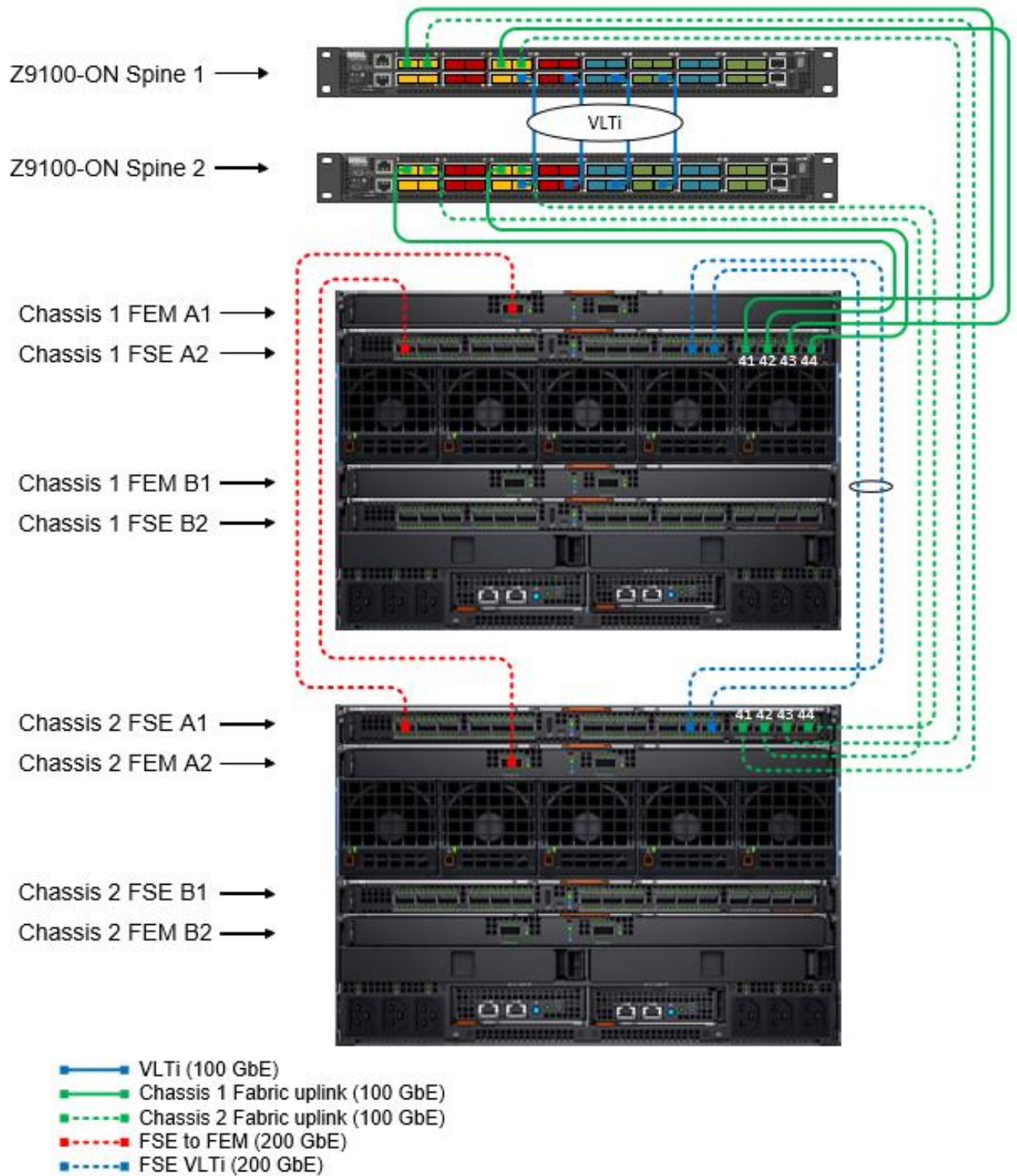Table 7 outlines the recommended port order for connecting chassis with MX9116n FSEs with a pair of Dell EMC Z9100-ON spine switches to optimize buffer utilization. Chassis one and two are in the first group of 10 chassis, the second set of 10 chassis starts with chassis 11 and 12.

Table 7   Uplink connections

Chassis number / FSE slot number	FSE port number	Z9100-ON Spine 1 port number	Z9100-ON Spine 2 port number
Chassis 1 FSE A2	41	1/1/1	
	42		1/1/1
	43	1/1/3	
	44		1/1/3
Chassis 1 FSE B2	41	1/1/5	
	42		1/1/5
	43	1/1/7	
	44		1/1/7
Chassis 2 FSE A1	41	1/1/9	
	42		1/1/9
	43	1/1/11	
	44		1/1/11
Chassis 2 FSE B1	41	1/1/13	
	42		1/1/13
	43	1/1/15	
	44		1/1/15
Chassis 11 FSE A2	41	1/1/17	
	42		1/1/17
	43	1/1/19	
	44		1/1/19
Chassis 11 FSE B2	41	1/1/21	
	42		1/1/21
	43	1/1/23	
	44		1/1/23
Chassis 12 FSE A1	41	1/1/25	
	42		1/1/25
	43	1/1/27	
	44		1/1/27
Chassis 12 FSE B1	41	1/1/29	
	42		1/1/29
	43	1/1/31	
	44		1/1/31

A.3   MX5108n VLT interconnects

When selecting the **2xMX5108n Ethernet Switches in same chassis** option in the SmartFabric wizard section of the GUI, it is important to note that the Virtual Link Trunk interconnect (VLTi) uses port 9, a 40GbE port, and port 10, a 100GbE port. The SmartFabric wizard then automatically configures a 1x40GbE breakout on port 10 so that the VLTi has two connections that are running at the same speed. The VLTi connections using ports 9 and 10 are shown in Figure 40.
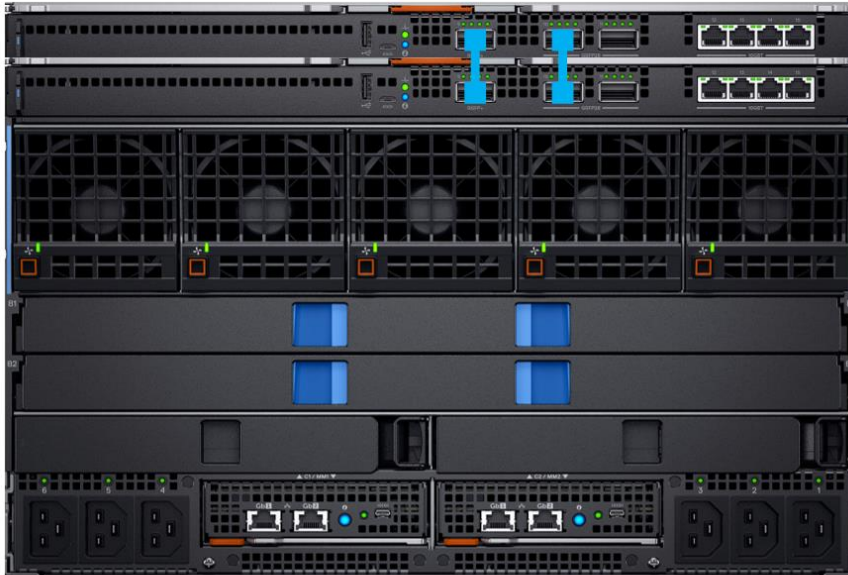
▬▬▬▬■   VLT Connections

Figure 40   Two MX5108n in a single chassis

When using the SmartFabric wizard for **2xMX5108n Ethernet Switches in same chassis,** there is a printout button on the **Summary** tab of the wizard. The **Print** option is shown in Figure 41.
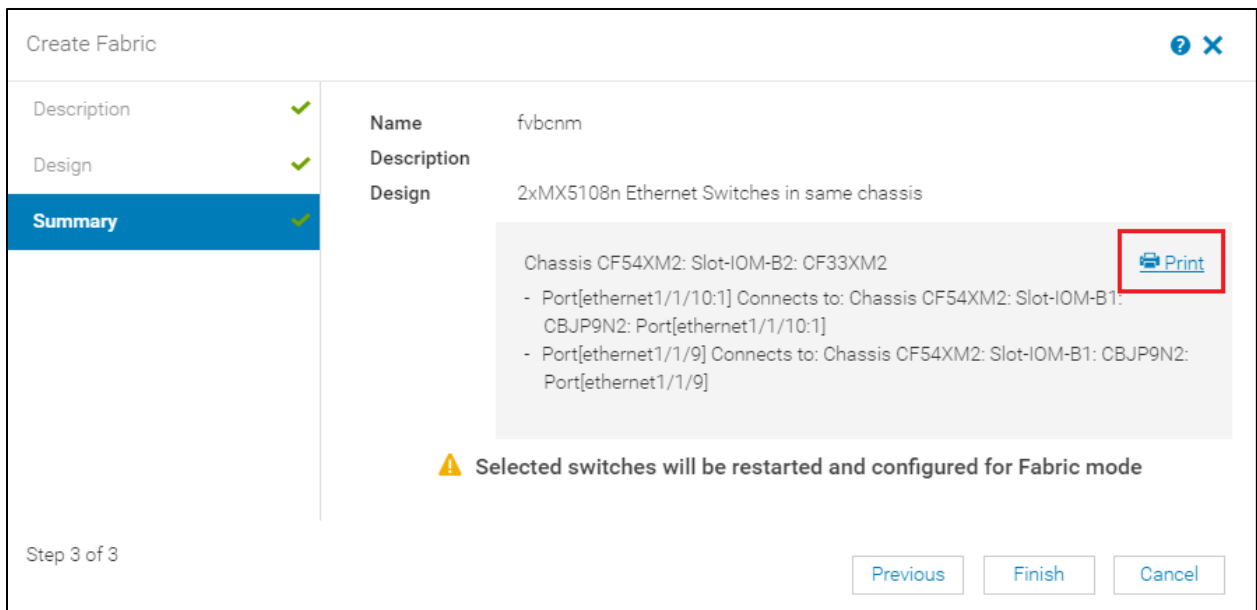
Figure 41   SmartFabric wizard 2xMX5108n Ethernet Switches in same chassis

Once this is selected, it presents the option to print out a topology view of the specified SmartFabric topology. The expected VLTi connections for the **2xMX5108n Ethernet Switches in same chassis** are shown in Figure 42. This option helps to ensure that the VLTi connections expected in the SmartFabrics are correctly cabled.
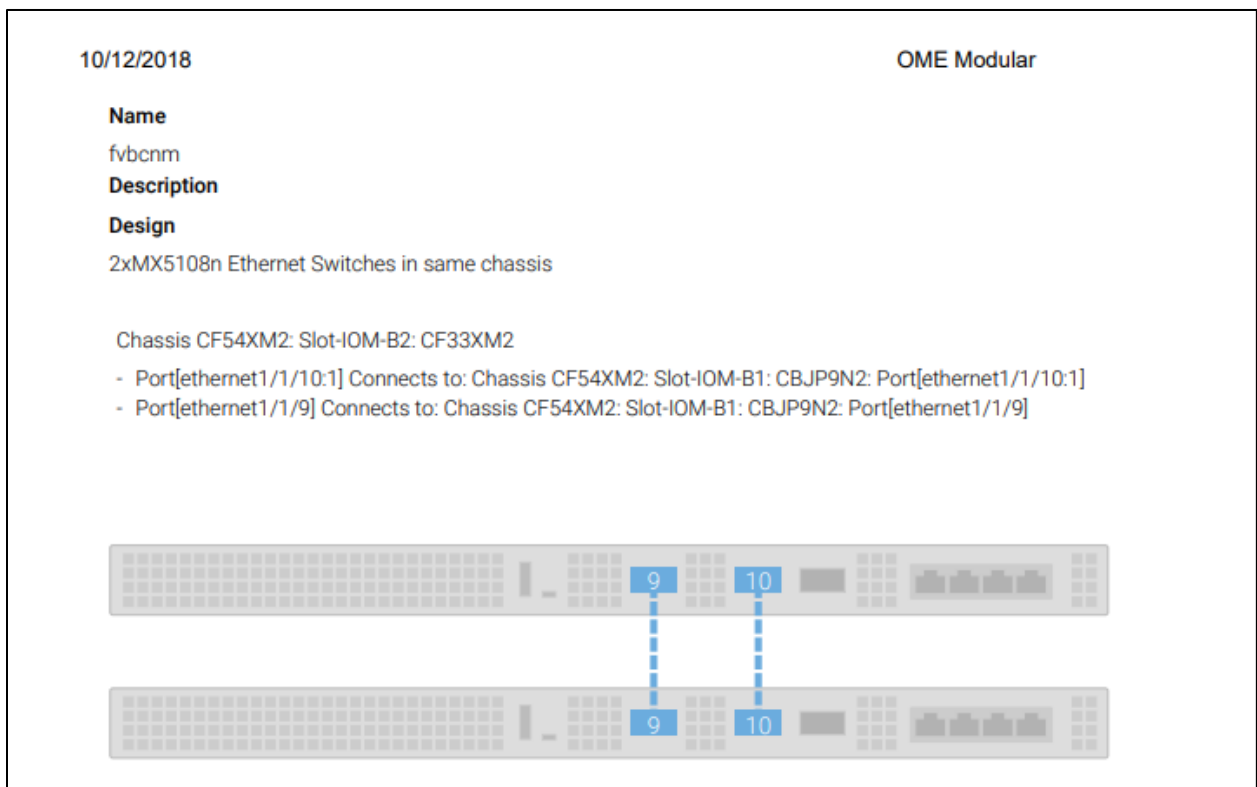

Figure 42   MX5108n VLTi topology

A.4 Multi-Chassis Management(MCM)

Following are the prerequisites to create a wired or daisy-chained chassis group:

- Each chassis must be physically connected as part of the Multi-Chassis Management (MCM) daisy-chain as shown in Figure 43. A password is not needed as the machine to machine authentication trust is used.
- Ensure that each member chassis is added to the group using the automatic or manual method.
- Ensure that the chassis settings are selected for applying to the other chassis. Settings include power, user authentication, alert destination, time, proxy, security, network services, and local access.

The daisy-chain configuration helps in surviving:

- A single network cable failure
- A single management module failure
- Power loss of any chassis in the stack

The wired chassis displays under **Available Chassis** in the **Group Deployment Wizard**.

Note: More information about the configuration of MCM groups can be found in the *Dell EMC OpenManage Enterprise-Modular Edition Version 1.00.01 for PowerEdge MX7000 Chassis* as well as *PowerEdge MX7000 Chassis Management Networking Cabling*.
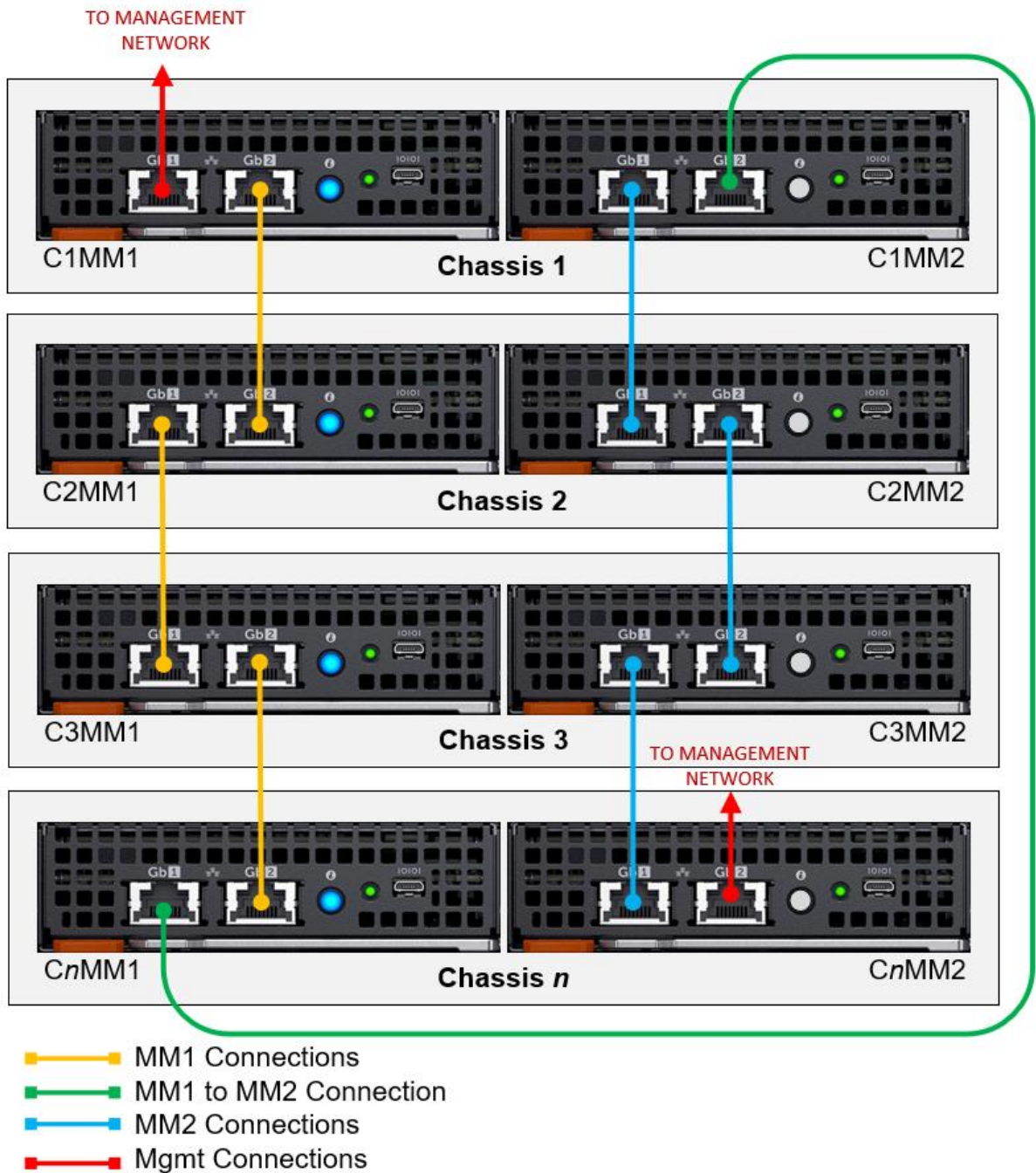
DELLEMC

Figure 43   Dell EMC PowerEdge MX9002m module daisy chain cabling

D&LLEMC

B Supported optics and cables

The PowerEdge MX7000 supports various optics and cables. The sections in this appendix provide a brief summary of the specified industry standards and the use case regarding the chassis. Table 8 shows the various cable types that are supported.

Note: Additional information on supported cables and optics can be found in the *PowerEdge MX IO Guide*.

Table 8   Cable types

Cable Type	Description
DAC (Copper)	• Direct attach copper • Copper wires and shielding • 2-wires/channel
AOC (Optical)	Active Optical Cable
MMF (Optical)	• Multi-Mode Fiber • Large core fiber (~50um) • 100m (/300m to match SMF) reach • Transceivers are low cost • Fiber 3x cost of SMF
SMF (Optical)	• Single-Mode Fiber • Tiny core fiber (~9um) • 2/10Km reach • Transceivers are expensive

DELLEMC

Table 9 shows the different optical connectors as well as a brief description of the standard.

Table 9   Optical connectors

Connector	Description
Small Form-factor Pluggable (SFP) • SFP = 1Gb • SFP+= 10Gb • SFP28 = 25Gb	SFP • 1 Channel • 2 Fibers or wires • 1-1.5W • Duplex LC optical connector • MMF or SMF
Quad Small Form-factor Pluggable (QSFP) • QSFP+ = 40Gb • QSFP28 = 100Gb	QSFP • 4 Channels • 8 Fibers or wires • 3.5W-5W • MPO12 8 fiber parallel optical connector
Quad Small Form-factor Pluggable Double – Density (QSFP-DD) • QSFP28-DD = 2x100Gb • QSFP56-DD = 2x200Gb	QSFP-DD • 8 channels • 16 fibers or wires • 10W • MPO12DD 16 fiber parallel optical connector

Table 10 shows the location within the chassis, where each type of media is relevant.

Table 10   Media associations

Media Type	Mx9116n	MX7116n	MX5108n	25GbE PTM	25GbE Switch	100GbE Switch
SFP+				x	x	
SFP28				x	x	
QSFP+	X		x			x
QSFP28	X		x			x
QSFP28-DD	X	x				

Each type of media has a specific use case regarding the MX7000, with each type of media there are various applications. The following sections outline where in the chassis each type of media is relevant.

Note: See the _Dell EMC Networking Transceivers and Cables_ document for more information about supported optics and cables.

B.1 SFP+/SFP28

As seen in Table 9, SFP+ is a 10GbE transceiver and SFP28 is a 25GbE transceiver both of which can use either fiber or copper media to achieve 10GbE or 25GbE communication in each direction. While the MX5108n has four 10GBASE-T copper interfaces, the focus is on optical connectors.

The SFP+ media type is typically seen in the PowerEdge MX7000 with the use of the 25GbE Pass Through Module(PTM) and using breakout cables form the QSFP+ and QSFP28 ports. The following are supported on the PowerEdge MX7000:

- Direct Attach Copper (DAC)
- LC fiber optic cable with SFP+ transceivers

The use of SFP+/SFP28 as it relates to QSFP+ and QSFP28, is discussed in those sections.

Note: The endpoints of the connection need to be set to 10GbE if SFP+ media is being used.
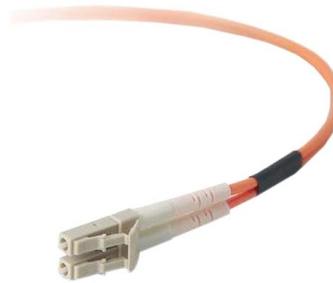
Direct Attach Copper (DAC)	LC Fiber Optic Cable	SFP+/SFP28 Transceivers

Figure 44   SFP+/SFP28 Media

Figure 44 shows an example of SFP+ cables and transceivers. Also, the SFP+ form factor can be seen referenced in the QSFP+ and QSFP28 sections with the use of breakout cables.

B.2 QSFP+

QSFP+ is a 40Gb standard that uses either fiber or copper media to achieve communication in each direction. This standard has four individual 10Gb lanes which can be used together to achieve 40GbE throughput or separately as four individual 10GbE connections (using breakout connections). One variant of the Dell QSFP+ transceiver is shown in Figure 45.

Figure 45   QSFP+ Transceiver

The QSFP+ media type has several uses in the PowerEdge MX7000. While the MX9116n does not have interfaces that are dedicated to QSFP+, ports 41 through 44 can be broken out to 1x40GbE that enables QSFP+ media to be used in those ports. The MX5108n has one dedicated QSFP+ port and two QSFP28 ports that can be configured for 1x40GbE.

Figure 46 shows an example of QSFP+ Coppers. The Direct Attach Copper (DAC) is a copper cable with a QSFP+ transceiver on either end. The Multi-fiber Push On (MPO) cable is a fiber cable that has MPO connectors on either end; these connectors attach to QSFP+ transceivers. The third variant is an Active Optical Cable (AOC) that is similar to the DAC with a fixed fiber optic cable in between the attached QSFP+ transceivers.

Direct Attach Copper (DAC)   **Multi-fiber Push On (MPO) Cable**   **Active Optical Cable (AOC)**

Figure 46   QSFP+ cables

The PowerEdge MX7000 also supports the use of QSFP+ to SFP+ breakout cables. This offers the ability to use a QSFP+ port and connect to four SFP+ ports on the terminating end.

DELLEMC

Figure 47 shows the DAC and MPO cables, which are two variations of breakout cables. The MPO cable in this example attaches to one QSFP+ transceiver and four SFP+ transceivers.

Direct Attach Copper (DAC) Breakout   **Multi-fiber Push On (MPO) Breakout Cable**
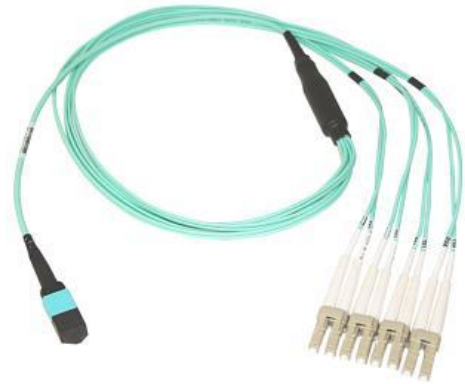
Figure 47   QSFP+ to SFP+ Breakout cables

Note: The MPO breakout cables uses a QSFP+ transceiver on one end and four SFP+ transceivers on the terminating end.

B.3   QSFP28

The QSFP28 standard is 100Gb that uses either fiber or copper media to achieve communication in each direction. The QSFP28 transceiver has four individual 25Gb lanes which can be used together to achieve 100GbE throughput or separately as four individual 25GbE connections (using 4 SFP28 modules). One variant of the Dell QSFP28 transceiver is shown in Figure 48.

Figure 48   QSFP28 transceiver

There are three variations of cables for QSFP28 connections. The variations are shown in Figure 49.

| **Direct Attach Copper (DAC)** | **Multi-fiber Push On (MPO) Cable** | **Active Optical Cable (AOC)** |

Figure 49   QSFP28 cables

Note: The QSFP28 form factor can use the same MPO cable as QSFP+. The DAC and AOC cables are different in that the attached transceiver is a QSFP28 transceiver rather than QSFP+.

QSFP28 supports the following breakout configurations:

- **1x40Gb** with QSFP+ connections, using either a DAC, AOC, or MPO cable and transceiver.
- **2x50Gb** with a fully populated QSFP28 end and two de-populated QSFP28 ends, each with 2x25GbE lanes. This product is only available as DAC cables.
- **4x25Gb** with a QSFP28 connection and using 4 SFP28 connections, using either a DAC, AOC, or MPO breakout cable with associated transceivers.
- **4x10Gb** with a QSFP28 connection and using 4 SFP+ connections, using either a DAC, AOC, or MPO breakout cable with associated transceivers.

DELLEMC

C   Additional Information

This section contains more details to clarify the concepts covered in the main body of the document.

C.1   Rate limited 32 Gb Fibre Channel

When using 32 Gb FC, the actual data rate is 28 Gbps due to 64b/66b encoding. Figure 50 shows one of the unified interfaces, port groups 15 and 16. The port group is set to 4x 32 Gb FC mode. However, each of the four lanes is 25 Gbps, not 28 Gbps. When these lanes are mapped from the Network Processing Unit (NPU) to the FC ASIC, for conversion to FC signaling, the 32GFC interfaces are mapped to four 25 Gbps lanes. With each lane operating at 25 Gbps, not 28 Gbps, the result is rate limited 32 Gb FC.
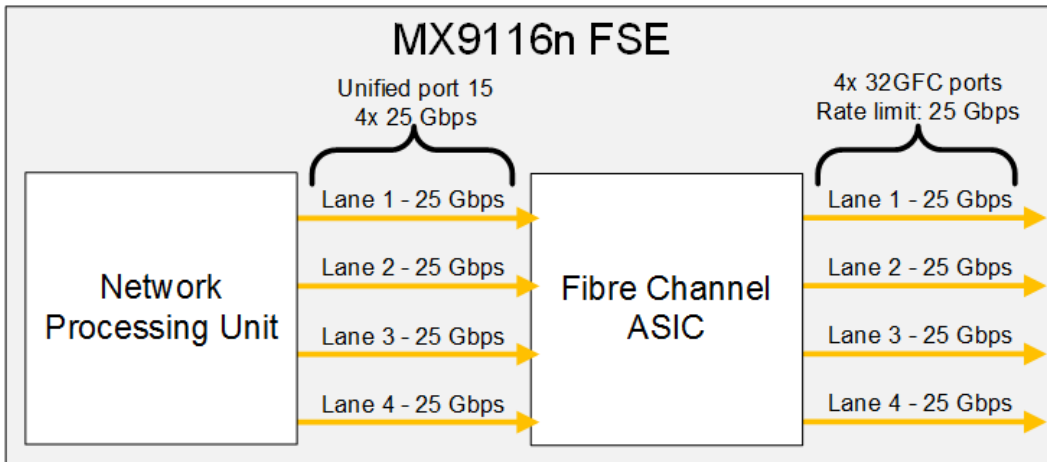

Figure 50   4x 32 Gb FC breakout mode, rate limit of 25 Gbps

While each 32 Gb FC connection is providing 25 Gbps, the overall FC bandwidth available is 100 Gbps per unified port group, or 200 Gbps for both ports. However, if an application requires the maximum 28 Gbps throughput per port, use the 2x 32 Gb breakout mode. This mode configures the connections between the NPU and the FC ASIC as shown in Figure 51.
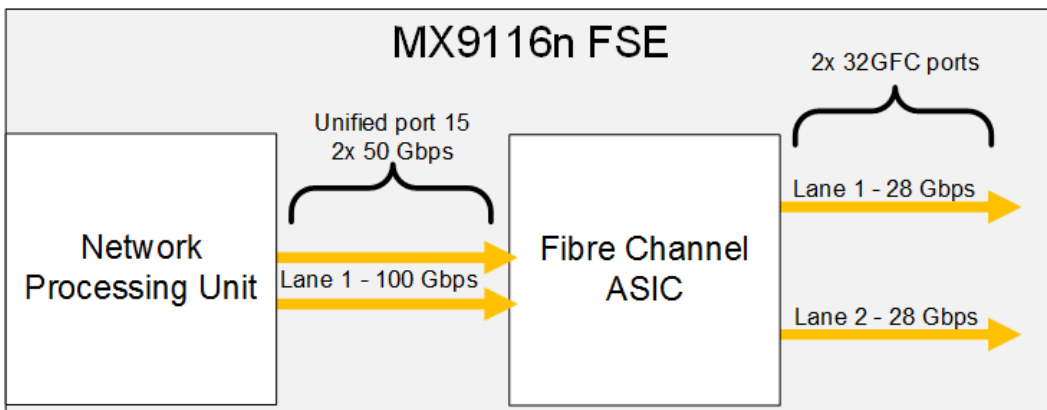

Figure 51   2x 32 Gb FC breakout mode

In 2x 32 Gb FC breakout mode, the MX9116n FSE binds two 50 Gbps links together to provide a total of 100 Gbps bandwidth per lane to the FC ASIC. This results in the two FC ports operating at 28 Gbps. The overall FC bandwidth available is 56 Gbps per unified port, or 112 Gbps for both. Compared to the 200 Gbps using 4x 32 Gb FC.

DELLEMC

Note: Rate limited ports are not oversubscribed ports. There is no FC frame drop on these ports and buffer to buffer credit exchanges ensure flow consistency.

C.2 Spanning Tree Protocol recommendations

By default, OS10EE uses Rapid per-VLAN Spanning Tree Plus (RPVST+) across all switching platforms including PowerEdge MX networking IOMs. OS10EE also supports Rapid Spanning Tree protocol (RSTP) and Multiple Spanning Tree (MST).

Exercise caution when connecting a RPVST+ to an existing RSTP environment. RPVST+ creates a single topology per VLAN with the default VLAN, by default VLAN 1, being used for the Common Spanning Tree (CST) with RSTP.

For non-native VLANs, all Bridge Protocol Data Unit (BPDU) traffic is tagged and forwarded by the upstream, RSTP-enable switch, with the associated VLAN. These BPDUs are sent to a protocol-specific multicast address. Any other RPVST+ tree that is attached to the RSTP tree processes these packets accordingly leading to the potential of unexpected trees.

Note: Dell EMC Networking recommends, that when connecting to an existing environment that is not using RPVST+, to change to the existing STP protocol before connecting a OS10EE switch.

In the following example, RSTP is enabled globally on an MX9116n FSE, with a similar MST configuration:

```
MX9116n-FSE(config)# spanning-tree mode rstp
MX9116n-FSE(config)# end
MX9116n-FSE#show spanning-tree brief

Spanning tree enabled protocol rstp with force-version rstp
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 0, Address 4c76.25e8.f2c0
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 32768, Address 2004.0f00.cd1e
Configured hello time 2, max age 20, forward delay 15
Flush Interval 200 centi-sec, Flush Invocations 95
Flush Indication threshold 0 (MAC flush optimization is disabled)

<output truncated>
```

C.3 NIC teaming restrictions

While NIC teaming is not required, it is recommended for redundancy unless an implementation recommends against it.

There are two main types of NIC teaming:

- Switch dependent: Also seen as 802.3ad or Dynamic Link Aggregation. This teaming method uses the LACP protocol.
- Switch independent: This method uses the operating system and NIC device drivers on the server to team the NICs. Each NIC vendor may provide slightly different implementations.

NIC Partitioning (NPAR) can impact how NIC teaming operates. Based on restrictions that are implemented by the NIC vendors that are related to NIC partitioning, certain configurations preclude certain types of teaming. The following restrictions are in place for both Full Switch and SmartFabric modes.

- If NPAR is not in use, both switch dependent (LACP) and switch independent teaming methods are supported.
- If NPAR is in use, only switch independent teaming methods are supported. Switch dependent teaming is NOT supported.

If switch-dependent (LACP) teaming is going to be used the IDRAC shared LAN on Motherboard (LOM) feature can only be used if the **Failover** option on IDRAC is enabled. Also, if Microsoft Windows is the host operating system, the LACP timer MUST be set to **Slow**.

To change the timing, see the following:

- Windows 2012R2 – Update adds support for the slow timer in LACP in Windows Server 2012 R2 - Instructions 2012R2
- Windows 2016 – LACP Timer is configured incorrectly when you create a NIC Team for LBFO - Instructions 2016

C.4 PowerEdge MX IOM slot support matrix

Table 11 shows the recommended PowerEdge MX IOM slot configurations.

Table 11   Recommended IOM slot configurations

Slot A1	Slot A2	Slot B1	Slot B2
MX9116n	MX9116n		
MX5108n	MX5108n		
MX7116n	MX7116n		
25G PTM	25G PTM		
10GBASE-T PTM	10GBASE-T PTM		
MX9116n	MX9116n	MX9116n	MX9116n
MX5108n	MX5108n	MX5108n	MX5108n
MX7116n	MX7116n	MX7116n	MX7116n
MX9116n	MX7116n		
MX7116n	MX9116n		
MX9116n	MX7116n	MX9116n	MX7116n
MX7116n	MX9116n	MX7116n	MX9116n
25G PTM	25G PTM	25G PTM	25G PTM
10GBASE-T PTM	10GBASE-T PTM	10GBASE-T PTM	10GBASE-T PTM

Table 12 lists all other supported IOM slot configurations. These configurations are either single IOMs configurations or each chassis contains dual MX9116n FSEs. In either configuration, redundancy is not a requirement.

Table 12   Other supported IOM configurations

Slot A1	Slot A2	Slot B1	Slot B2
MX9116n			
MX5108n			
MX7116n			
25G PTM			
10GBASE-T PTM			
MX9116n		MX9116n	
MX5108n		MX5108n	
MX7116n		MX7116n	
25G PTM		25G PTM	
10GBASE-T PTM		10GBASE-T PTM	
MX9116n	MX9116n	MX9116n	
MX5108n	MX5108n	MX5108n	
MX7116n	MX7116n	MX7116n	
25G PTM	25G PTM	25G PTM	
10GBASE-T PTM	10GBASE-T PTM	10GBASE-T PTM	
MX9116n	MX9116n	MX5108n	MX5108n
MX9116n	MX9116n	25G PTM	25G PTM
MX9116n	MX9116n	10GBASE-T PTM	10GBASE-T PTM
MX9116n	MX7116n	MX5108n	MX5108n
MX7116n	MX9116n	MX5108n	MX5108n
MX9116n	MX7116n	25G PTM	25G PTM
MX7116n	MX9116n	25G PTM	25G PTM
MX9116n	MX7116n	10GBASE-T PTM	10GBASE-T PTM
MX7116n	MX9116n	10GBASE-T PTM	10GBASE-T PTM
MX7116n	MX7116n	MX5108n	MX5108n

Slot A1	Slot A2	Slot B1	Slot B2
MX7116n	MX7116n	25G PTM	25G PTM
MX7116n	MX7116n	10GBASE-T PTM	10GBASE-T PTM
MX5108n	MX5108n	MX9116n	MX9116n
MX5108n	MX5108n	MX7116n	MX7116n
MX5108n	MX5108n	MX9116n	MX7116n
MX5108n	MX5108n	MX7116n	MX9116n
MX5108n	MX5108n	25G PTM	25G PTM
MX5108n	MX5108n	10GBASE-T PTM	10GBASE-T PTM
25G PTM	25G PTM	MX9116n	MX9116n
25G PTM	25G PTM	MX7116n	MX7116n

25G PTM	25G PTM	MX9116n	MX7116n
25G PTM	25G PTM	MX7116n	MX9116n
25G PTM	25G PTM	10GBASE-T PTM	10GBASE-T PTM
10GBASE-T PTM	10GBASE-T PTM	MX9116n	MX9116n
10GBASE-T PTM	10GBASE-T PTM	MX7116n	MX7116n
10GBASE-T PTM	10GBASE-T PTM	MX9116n	MX7116n
10GBASE-T PTM	10GBASE-T PTM	MX7116n	MX9116n
10GBASE-T PTM	10GBASE-T PTM	25G PTM	25G PTM

C.5 PTM port mapping

The following figures show the port mapping between compute sleds and PTM interfaces. This mapping applies to both 25 GbE and 10 GbE PTMs.

Figure 52   Ethernet PTM Dual Port NIC Mapping

Figure 53   Ethernet PTM Quad Port NIC Mapping

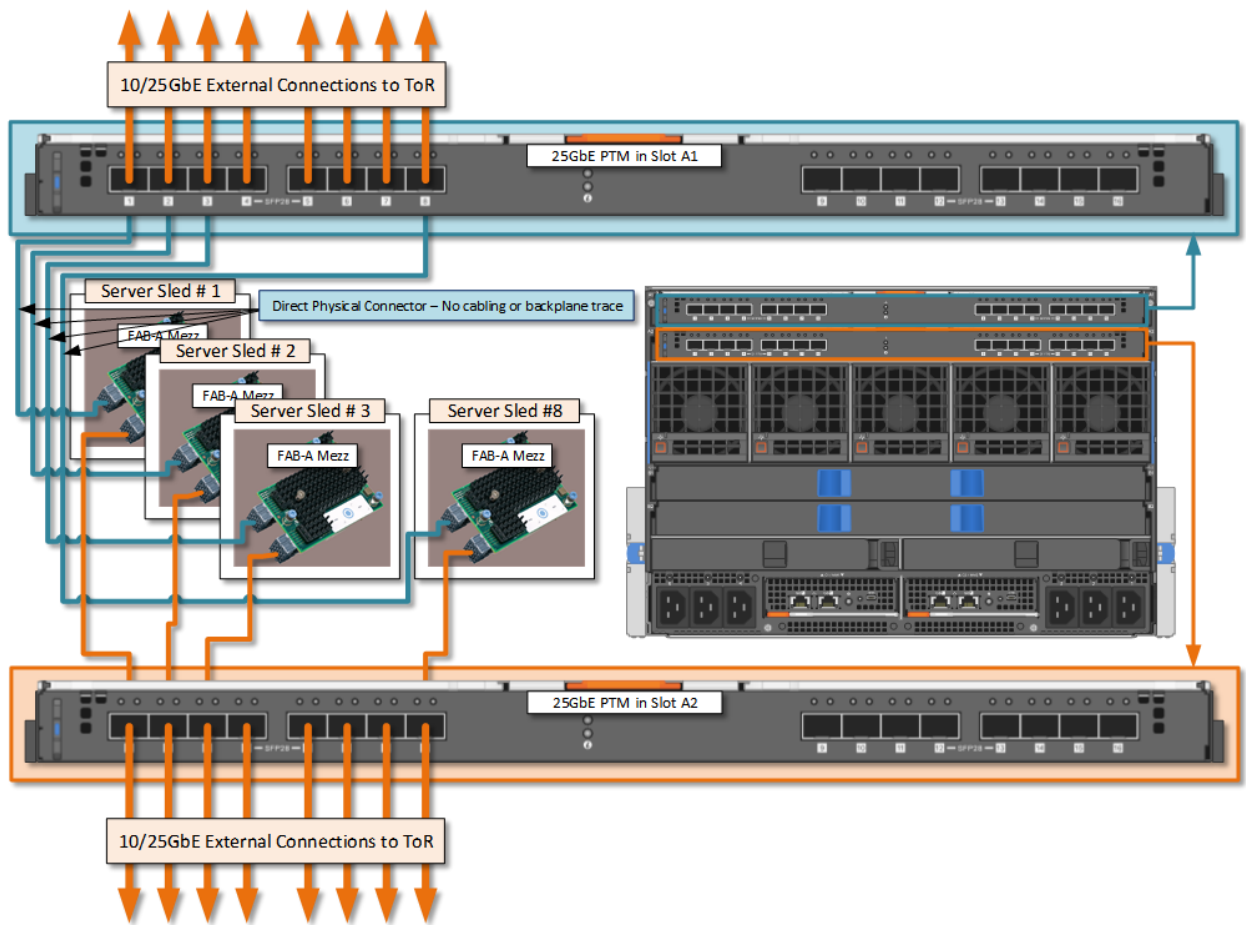**Note:** Ports 9 through 16 are reserved for future expansion.

Figure 54   Ethernet PTM mezzanine mapping

C.6 Examples of unsupported Scalable Fabric topologies

The Scalable Fabric Architecture currently requires that a MX7116n FEM in slot A1 be connected to an MX9116n FSE in slot A1, an MX7116n in slot A2 connected to an MX9116n FSE in slot A2, the same rule applies to devices in B1 and B2.

It is not supported to have an MX7116n FEM in slot A2 connected to an MX9116n FSE in slot A1, for example.

The following examples are not supported configurations.
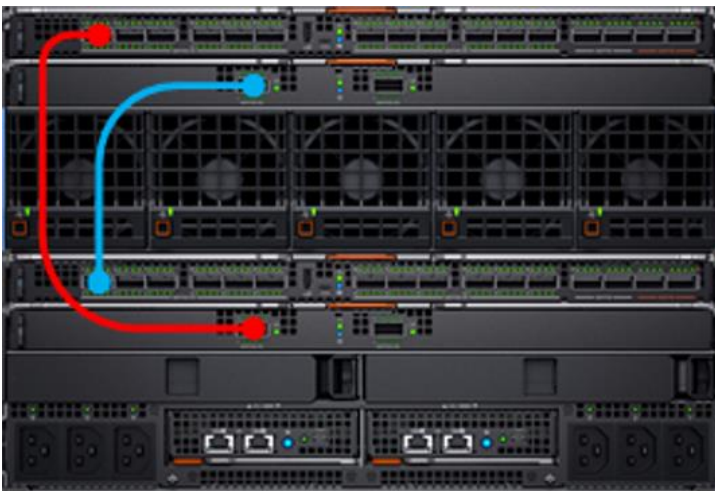

Figure 55   Unsupported A1-A2, B1-B2
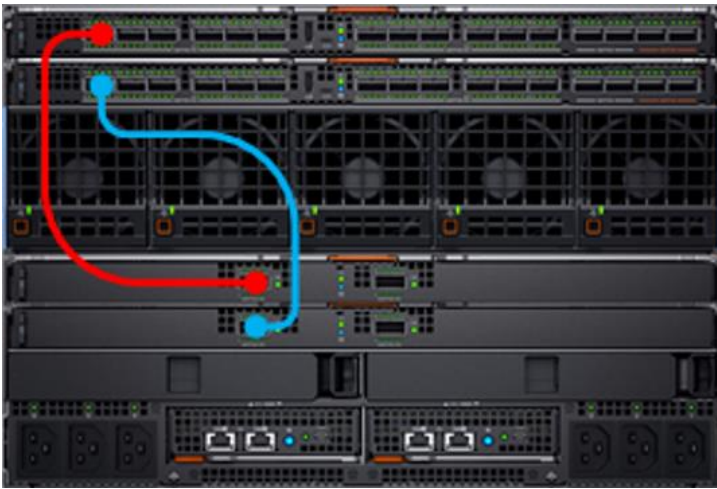

Figure 56   Unsupported A1-B2, B1-A2

DELLEMC

Figure 57   Unsupported A1-B1, A2-B2

The following configurations are supported in full switch mode; however, they are **not supported** configurations for SmartFabric mode.
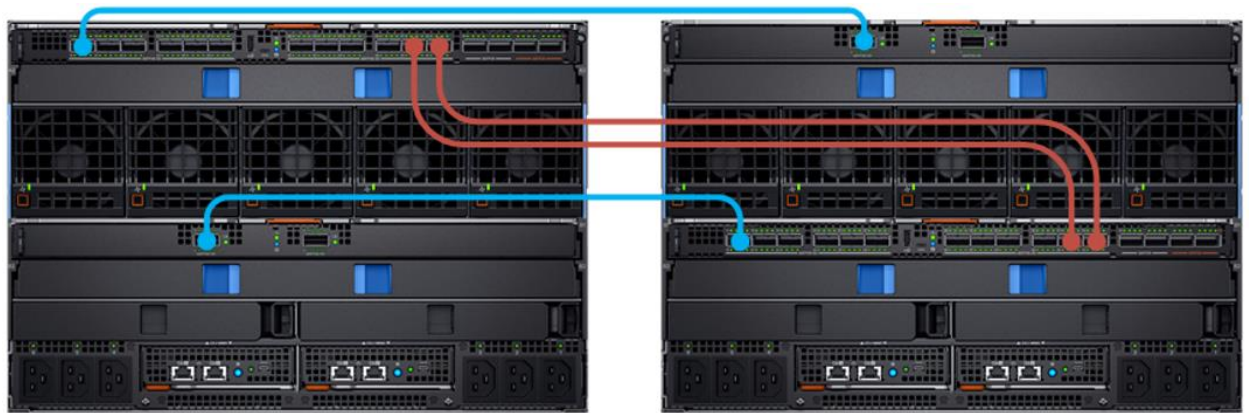

MX9116n A1/B1   MX5108n A1/B1

Figure 58   Unsupported VLT configuration

MX9116n Chassis 1:A1/Chassis2:B1

Figure 59   Unsupported inter-chassis configuration

D Technical resources

[Dell EMC TechCenter Networking Guides](#)

[Dell EMC PowerEdge MX SmartFabric Deployment Guide](#)

[Dell EMC PowerEdge MX SmartFabric Deployment Video](#)

[Dell EMC OS10EE Enterprise Edition Users Guide](#)

[PowerEdge MX IO Guide](#)

[Dell EMC Networking Transceivers and Cables](#)

[Manuals and documentation for Dell EMC PowerEdge MX7000](#)

DELLEMC

E Support and feedback

Contacting Technical Support

Support Contact Information   Web: http://www.dell.com/support

                                              Telephone: USA: 1-800-945-3355

Feedback for this document

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

DELLEMC