

# VCF on VxRail Multirack Deployment using BGP EVPN

Adding a Virtual Infrastructure workload domain with NSX-T

## Abstract

This document provides step-by-step deployment instructions for Dell EMC OS10 Enterprise Edition (EE) L2 VXLAN tunnels using BGP EVPN. This guide contains the foundation for multirack VxRail host discovery and deployment. Also, the VMware Cloud Foundation on Dell EMC VxRail with NSX-T is deployed, providing the initial building block for a workload domain in the Software Defined Data Center (SDDC).

August 2019

# Revisions

Date	Description
August 2019	Initial release

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.

© 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.

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

<b>1</b>	<b>Introduction</b>	<b>6</b>
1.1	VMware Cloud Foundation on VxRail	6
1.2	VMware Validated Design for SDDC on VxRail	8
1.3	VMware NSX Data Center	9
1.4	Prerequisites	10
1.5	Supported switches and operating systems	10
1.6	Typographical conventions	10
1.7	Attachments	10
<b>2</b>	<b>Hardware overview</b>	<b>11</b>
2.1	Dell EMC VxRail E560	11
2.2	Dell EMC PowerSwitch S5248F-ON	11
2.3	Dell EMC PowerSwitch Z9264F-ON	11
2.4	Dell EMC PowerSwitch S3048-ON	12
<b>3</b>	<b>Network transport</b>	<b>13</b>
3.1	Layer 3 leaf and spine topology	13
3.2	BGP EVPN VXLAN overview	13
3.2.1	The VXLAN protocol	15
3.2.2	BGP EVPN VXLAN operation	15
<b>4</b>	<b>Topology</b>	<b>16</b>
4.1	Leaf-spine underlay	16
4.1.1	BGP ASNs and router IDs	17
4.1.2	Point-to-point IP networks	17
4.2	Underlay network connections	19
4.3	BGP EVPN VXLAN overlay	20
4.4	VxRail node connections	21
<b>5</b>	<b>Planning and preparation</b>	<b>22</b>
5.1	VLAN IDs and IP subnets	22
5.2	External services	22
5.3	DNS	23
5.3.1	NTP	23
5.3.2	DHCP	23
5.4	Switch preparation	24
5.5	Check switch OS version	24
5.6	Verify license installation	24
5.7	Factory default configuration	25

5.8	Switch settings.....	26
6	Configure and verify the underlay network.....	27
6.1	Configure leaf switch underlay networking.....	27
6.2	Configure leaf switch NSX-T overlay networking .....	34
6.3	Configure spine switches.....	36
6.4	Verify establishment of BGP between leaf and spine switches .....	40
6.5	Verify BGP EVPN and VXLAN between leaf switches.....	42
7	Create a VxRail Virtual Infrastructure workload domain .....	44
7.1	Create a local user in the workload domain vCenter Server .....	45
7.2	VxRail initialization.....	45
7.3	VxRail deployment values .....	46
7.4	Add the primary VxRail cluster to the workload domain.....	46
8	Configure NSX-T north-south connectivity.....	48
8.1	Create transport zones .....	48
8.2	Create uplink profiles and the network I/O control profile.....	49
8.3	Create the NSX-T segments for system, uplink, and overlay traffic.....	49
8.4	Create an NSX-T edge cluster profile.....	49
8.5	Deploy the NSX-T edge appliances .....	50
8.6	Join the NSX-T edge nodes to the management plane .....	50
8.7	Create anti-affinity rules for NSX-T edge nodes.....	50
8.8	Add the NSX-T edge nodes to the transport zones.....	52
8.9	Create and configure the Tier-0 gateway .....	53
8.10	Create and configure the Tier-1 gateway .....	53
8.11	Verify BGP peering and route redistribution .....	54
9	Validate connectivity between virtual machines.....	55
9.1	Ping from Web01 to Web02 .....	56
9.2	Ping from Web01 to App01 .....	56
9.3	Ping from Web01 to 10.0.1.2.....	57
9.4	Ping from App01 to 10.0.1.2.....	57
9.5	Traceflow App01 to 10.0.1.2.....	58
A	Validated components.....	59
A.1	Dell EMC PowerSwitch models .....	59
A.2	VxRail E560 nodes .....	59
A.3	Appliance software .....	60
B	Technical resources .....	61
B.1	VxRail, VCF, and VVD Guides .....	61

B.2	Dell EMC Networking Guides .....	61
C	Fabric Design Center .....	62
D	Support and feedback .....	63

# 1 Introduction

Our vision at Dell EMC is to be the essential infrastructure company from the edge to the core, and the cloud. Dell EMC Networking ensures modernization for today's applications and the emerging cloud-native world. Dell EMC is committed to disrupting the fundamental economics of the market with a clear strategy that gives you the freedom of choice for networking operating systems and top-tier merchant silicon. The Dell EMC strategy enables business transformations that maximize the benefits of collaborative software and standards-based hardware, including lowered costs, flexibility, freedom, and security. Dell EMC provides further customer enablement through validated deployment guides which demonstrate these benefits while maintaining a high standard of quality, consistency, and support.

At the physical layer of a Software Defined Data Center (SDDC), the Layer 2 or Layer 3 transport services provide the switching fabric. A leaf-spine architecture using Layer 3 IP supports a scalable data network. In a Layer 3 network fabric, the physical network configuration terminates Layer 2 networks at the leaf switch pair at the top of each rack. However, VxRail management and NSX Controller instances and other virtual machines rely on VLAN-backed Layer 2 networks.

Discovery or virtual machine migration cannot be completed because the IP subnet is available only in the rack where the virtual machine resides. To resolve this challenge, a Border Gateway Protocol (BGP) Ethernet VPN (EVPN) is implemented. The implementation creates control plane backed tunnels between the separate IP subnets creating Layer 2 networks that span multiple racks.

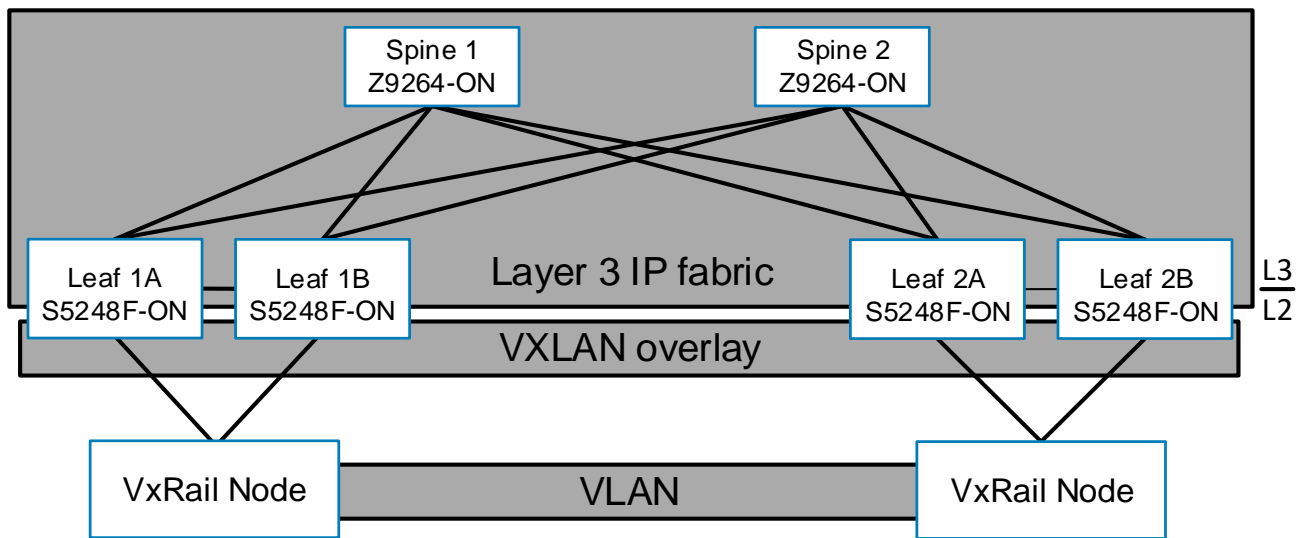


Figure 1 Illustration of stretched layer 2 segments between VxRail nodes in separate racks

## 1.1 VMware Cloud Foundation on VxRail

VMware Cloud Foundation on Dell EMC VxRail, part of Dell Technologies Cloud Platform, provides the simplest path to the hybrid cloud through a fully integrated hybrid cloud platform that leverages native VxRail hardware and software capabilities and other VxRail unique integrations (such as vCenter plugins and Dell EMC networking integration) to deliver a turnkey hybrid cloud user experience with full stack integration. Full stack integration means that customers get both the HCI infrastructure layer and cloud software stack in one, complete, automated lifecycle, turnkey experience. The platform delivers a set of software defined services for compute (with vSphere and vCenter), storage (with vSAN), networking (with NSX), security, and cloud management (with vRealize Suite) in both private or public environments making it the operational hub for their hybrid cloud as shown in Figure 2.

VMware Cloud Foundation on VxRail makes operating the data center fundamentally simpler by bringing the ease and automation of the public cloud in-house by deploying a standardized and validated network flexible architecture with built-in lifecycle automation for the entire cloud infrastructure stack including hardware. SDDC Manager orchestrates the deployment, configuration, and lifecycle management (LCM) of vCenter, NSX, and vRealize Suite above the ESXi and vSAN layers of VxRail. It unifies multiple VxRail clusters as workload domains or as multi-cluster workload domains. Integrated with the SDDC Manager management experience, VxRail Manager is used to deploy, configure vSphere clusters powered by vSAN. It is also used to execute the lifecycle management of ESXi, vSAN, and HW firmware using a fully integrated and seamless SDDC Manager orchestrated process. It monitors the health of hardware components and provides remote service support as well. This level of integration is what gives customers a truly unique turnkey hybrid cloud experience not available on any other infrastructure. All of this with available single vendor support through Dell EMC.

VMware Cloud Foundation on Dell EMC VxRail provides a consistent hybrid cloud unifying customer public and private cloud platforms under a common operating environment and management framework. Customers can operate both their public and private platforms using one set of tools and processes, with a single management view and provisioning process across both platforms. This consistency allows for easy portability of applications.

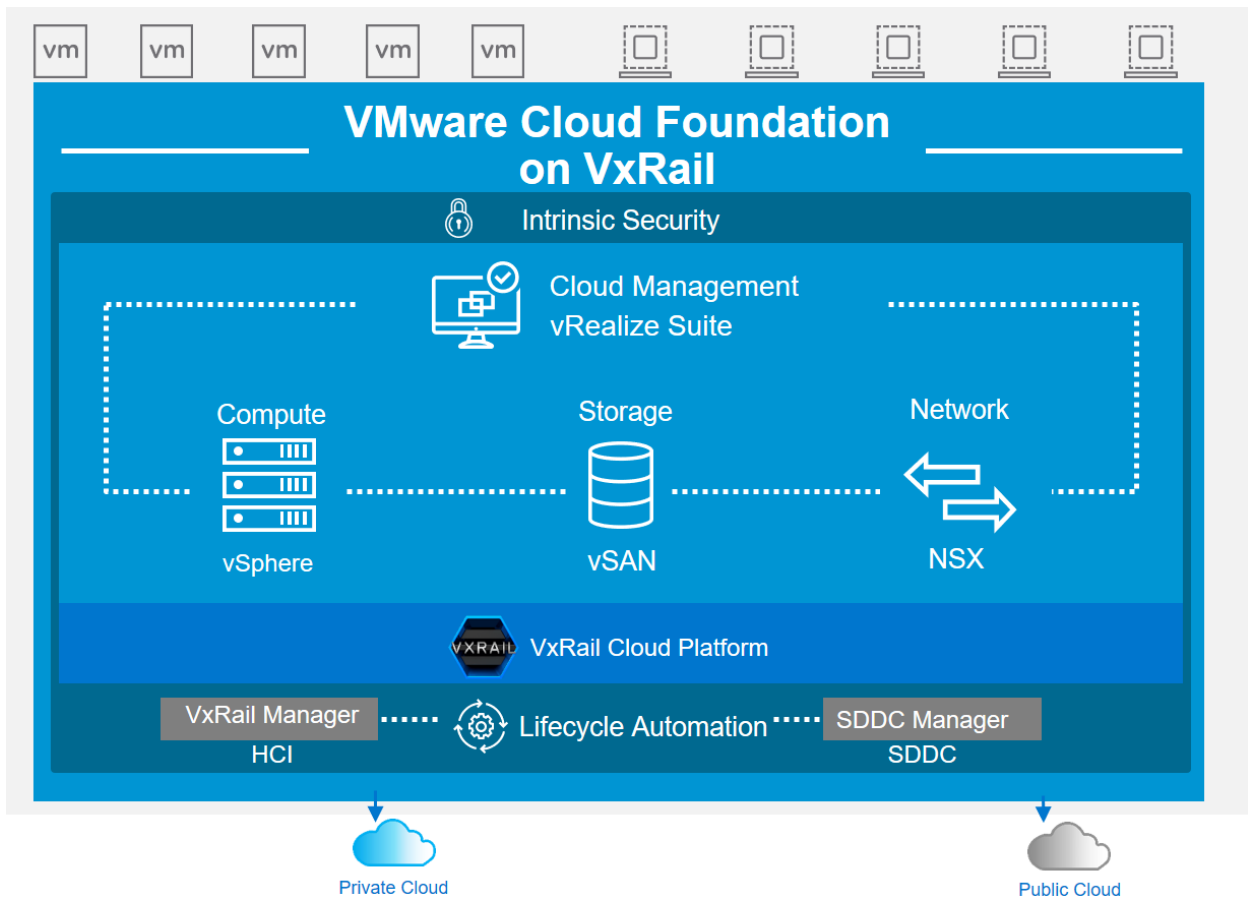


Figure 2 VMware Cloud Foundation on VxRail (VCF on VxRail) high-level architecture

To learn more about VMware Cloud Foundation on VxRail, see:

[VMware Cloud Foundation on VxRail Architecture Guide](#)

## 1.2 VMware Validated Design for SDDC on VxRail

[VMware Validated Designs \(VVD\)](#) simplify the process of deploying and operating an SDDC. They are comprehensive, solution-oriented designs that provide a consistent and repeatable production-ready approach to the SDDC. They are prescriptive blueprints that include comprehensive deployment and operational practices for the SDDC. They are an option available for customers, who are not ready for or do not value the complete approach to SDDC automation available in VCF on VxRail.

A VMware Validated Design is composed of a standardized, scalable architecture that is backed by the technical expertise of VMware and a software bill of materials (BOM) comprehensively tested for integration and interoperability that spans compute, storage, networking, and management. Detailed guidance that synthesizes best practices on how to deploy, integrate, and operate the SDDC is provided to aid users to achieve performance, availability, security, and operational efficiency.

With the VVD for SDDC on VxRail, customers can quickly architect, implement, and operate the complete SDDC faster and with less risk. Customers also get the benefits of best of breed HCI infrastructure platform. The latest available version at the time of writing this document is 5.0.1.

Customers can realize the following benefits by using VVD on VxRail:

- Accelerated time-to-market - streamline and simplify the complex design process of the SDDC, shortening deployment and provisioning cycles
- Increased efficiency – provide detailed, step-by-step guidance to reduce the time and the effort that is spent on operational tasks
- Lessen the uncertainty of deployments and operations - reduce uncertainty and potential risks that are associated with implementing and operating the SDDC
- IT agility – designed for expandability and to support a broad set of use cases and diverse types of applications that help IT respond faster to the business needs

To learn more about VVD on VxRail, see [Dell EMC VxRail - Accelerating the Journey to VMware Software-Defined Data Center \(SDDC\)](#).



## 1.3 VMware NSX Data Center

VMware NSX Data Center delivers virtualized networking and security entirely in software, completing a vital pillar of the Software Defined Data Center (SDDC), and enabling the virtual cloud network to connect and protect across data centers, clouds, and applications.

With NSX Data Center, networking and security are brought closer to the application wherever it is running, from virtual machines (VMs) to containers to bare metal. Like the operational model of VMs, networks can be provisioned and managed independent of the underlying hardware. NSX Data Center reproduces the entire network model in software, enabling any network topology—from simple to complex multitier networks—to be created and provisioned in seconds. Users can create multiple virtual networks with diverse requirements, using a combination of the services that are offered through NSX or from a broad ecosystem of third-party integrations ranging for next-generation firewalls to performance management solutions to build inherently more agile and secure environments. These servers can be extended to various endpoints within and across clouds.

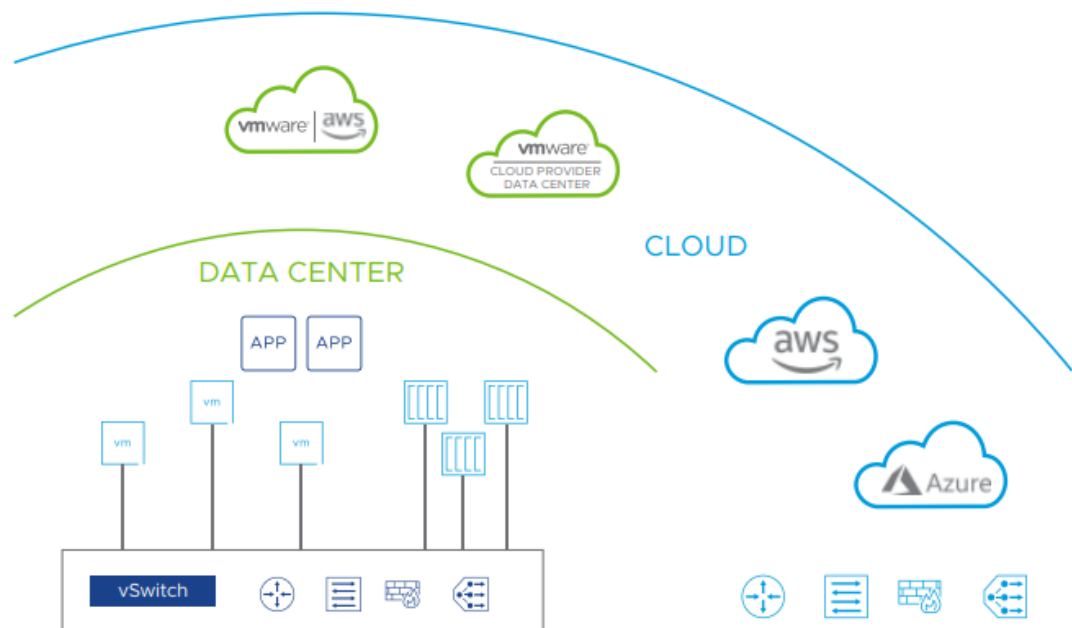


Figure 3 NSX Data Center: Network Virtualization and Security Platform

VMware NSX Data Center delivers an entirely new operational model for networking that is defined in software, forming the foundation of the SDDC and extending to a virtual cloud network. Data center operators can now achieve levels of agility, security, and economics that were previously unreachable when the data center network was tied solely to physical hardware components. NSX Data Center provides a complete set of logical networking and security capabilities and services, including logical switching, routing, firewalling, load balancing, a virtual private network (VPN), quality of service (QoS), and monitoring. These services are provisioned in virtual networks through any cloud management platform using NSX Data Center APIs. Virtual networks are deployed non-disruptively over any existing networking hardware and can extend across data centers, public and private clouds, container platforms, and bare-metal servers.

## 1.4 Prerequisites

This deployment guide is a continuation of the deployment guide, [VCF on VxRail multitrack deploying using BGP EVPN](#). That guide provides step-by-step instructions on creating a VCF on VxRail multitrack management domain.

## 1.5 Supported switches and operating systems

The examples provided in this deployment guide use VxRail 4.7.211 nodes that are connected to Dell EMC PowerSwitch S5248F-ON switches running Dell EMC OS10 EE 10.4.3.5. The following list of Dell PowerSwitch devices supports L2 VXLAN (Static VXLAN with VLT or BGP EVPN). It is a reasonable assumption that any switch listed below can be substituted with minor changes and create a generally similar outcome outlined in this deployment guide.

### Dell EMC PowerSwitch S Series 25-100GbE Switches

- S5224F-ON, S5232F-ON, S5248F-ON ([S5200-ON Spec Sheet](#))

### Dell EMC PowerSwitch S Series 10-40GbE Switches



- [S4048T-ON](#)
- S4112F-ON, S4112T-ON, S4128F-ON, S4128T-ON, S4148F-ON, S4148FE-ON, S4148T-ON, S4148U-ON ([S4100-ON Spec Sheet](#))
- S4248U, S4248FB-ON, S4248FBL-ON ([S4200-ON Spec Sheet](#))
- [S6010-ON](#)

## 1.6 Typographical conventions

The CLI and GUI examples in this document use the following conventions:

Monospace Text	CLI examples
<u>Underlined Monospace Text</u>	CLI examples that wrap the page
<i>Italic Monospace Text</i>	Variables in CLI examples
<b>Bold Monospace Text</b>	Commands entered at the CLI prompt, or to highlight information in CLI output
<b>Bold text</b>	UI elements and information that is entered in the GUI

## 1.7 Attachments

This document in .pdf format includes one or more file attachments. To access attachments in Adobe Acrobat Reader, click the  icon in the left pane halfway down the page, then click the  icon.

## 2 Hardware overview

This section briefly describes the hardware that is used to validate the deployment examples in this document. Appendix A contains a complete listing of hardware and software that is validated for this guide.

---

**Note:** While the steps in this document were validated using the specified Dell EMC PowerSwitch models and operating systems, they may be used for other Dell EMC PowerSwitch models using the same networking OS version or later assuming the switch has the available port numbers, speeds, and types.

---

### 2.1 Dell EMC VxRail E560

The Dell EMC VxRail E series consists of nodes that are best suited for remote office or entry workloads. The E series nodes support up to 40 CPU cores, 1536GB memory, and 16TB hybrid or 30TB all-flash storage in a 1-Rack Unit (RU) form factor. Each node has 4x 25 GbE ports for upstream connectivity. Two ports are attached to the Network Daughter Card (NDC), and the other two ports are provided through a PCI-E expansion card. The example within this document uses four VxRail E560 nodes.



Figure 4 Dell EMC VxRail 1-RU node

### 2.2 Dell EMC PowerSwitch S5248F-ON

The Dell EMC PowerSwitch S5248F-ON is a 1-RU fixed switch with 48x 25 GbE, 4x multirate 100 GbE, and 2x 200 GbE ports. The S5248F-ON supports L2 static VXLAN with VLT. The example within this document uses four S5248F-ON switches in VLT pairs as leaf switches.

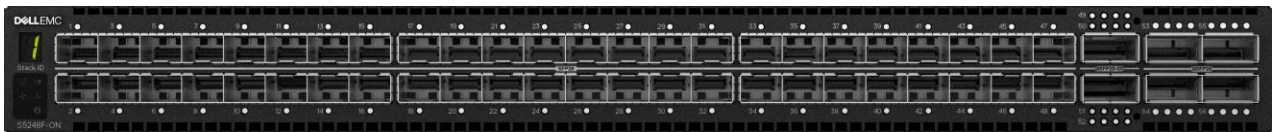


Figure 5 Dell EMC PowerSwitch S5248F-ON

### 2.3 Dell EMC PowerSwitch Z9264F-ON

The Dell EMC PowerSwitch Z9264F-ON is a 2-RU 100 GbE aggregation/spine switch. The Z9264F-ON has up to 64 ports of multirate 100 GbE, or up to 128 ports of 10/25/40/50 GbE ports using supported breakout cables. The example within this document uses two Z9264F-ON switches as spine switches.

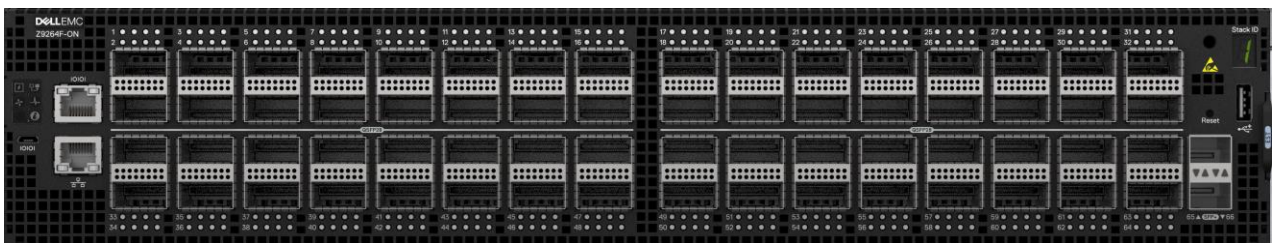


Figure 6 Dell EMC PowerSwitch Z9264F-ON

## 2.4 Dell EMC PowerSwitch S3048-ON

The Dell EMC PowerSwitch S3048-ON is a 1-Rack Unit (RU) switch with forty-eight 1GbE BASE-T ports and four 10GbE SFP+ ports. In this document, one S3048-ON supports out-of-band (OOB) management traffic for all examples.



Figure 7 Dell EMC PowerSwitch S3048-ON

## 3 Network transport

VMware Validated Design supports both Layer 2 and Layer 3 network transport. In this section, the details of the Layer 3 leaf-spine topology are provided.

---

**Note:** Most of the steps in this section may already be done if all of the configuration steps from the VCF on VxRail multitrack deploying using BGP EVPN deployment guide were followed. To ensure completion, the necessary steps are included in this section.

---

### 3.1 Layer 3 leaf and spine topology

In this document, a Clos leaf-spine topology is used for each availability zone. Individual switch configuration shows how to set up end-to-end Virtual Extensible Local Area Networks (VXLANs). External Border Gateway Protocol (eBGP) is used for exchanging IP routes in the IP underlay network, and EVPN routes in the VXLAN overlay network. Virtual Link Trunking (VLT) is deployed between leaf pairs and internal BGP (iBGP) to provide Layer 3 path redundancy in the event a leaf switch loses connectivity to the spine switches.

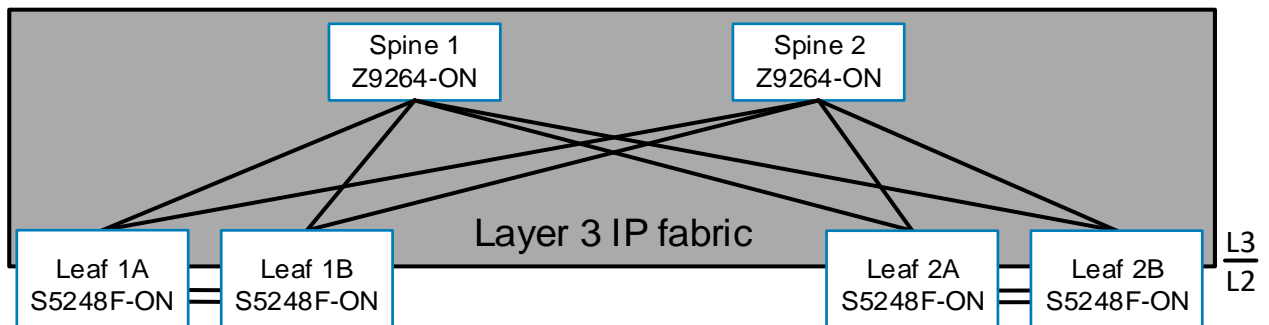


Figure 8 Layer 3 IP network transport

### 3.2 BGP EVPN VXLAN overview

EVPN is a control plane for VXLAN that is used to reduce flooding in the network and resolve scalability concerns. EVPN uses multiprotocol BGP (MP-BGP) to exchange information between VXLAN tunnel endpoints (VTEPs). EVPN was introduced in [RFC 7432](#), and [RFC 8365](#) describes VXLAN-based EVPN.

VXLAN-based EVPN is a next-generation VPN. It is intended to replace previous generation VPNs like Virtual Private LAN Service (VPLS). Some of its key features are:

- Support for multitenancy
- Layer 2 and 3 integrated routing and bridging (IRB)
- Multihoming
- Minimization of ARP propagation
- MAC mobility (simplified VM migration)

The primary use cases for EVPN are:

- Expanding the potential number of Layer 2 domains
- Service provider multitenant hosting
- Data center interconnect (DCI)

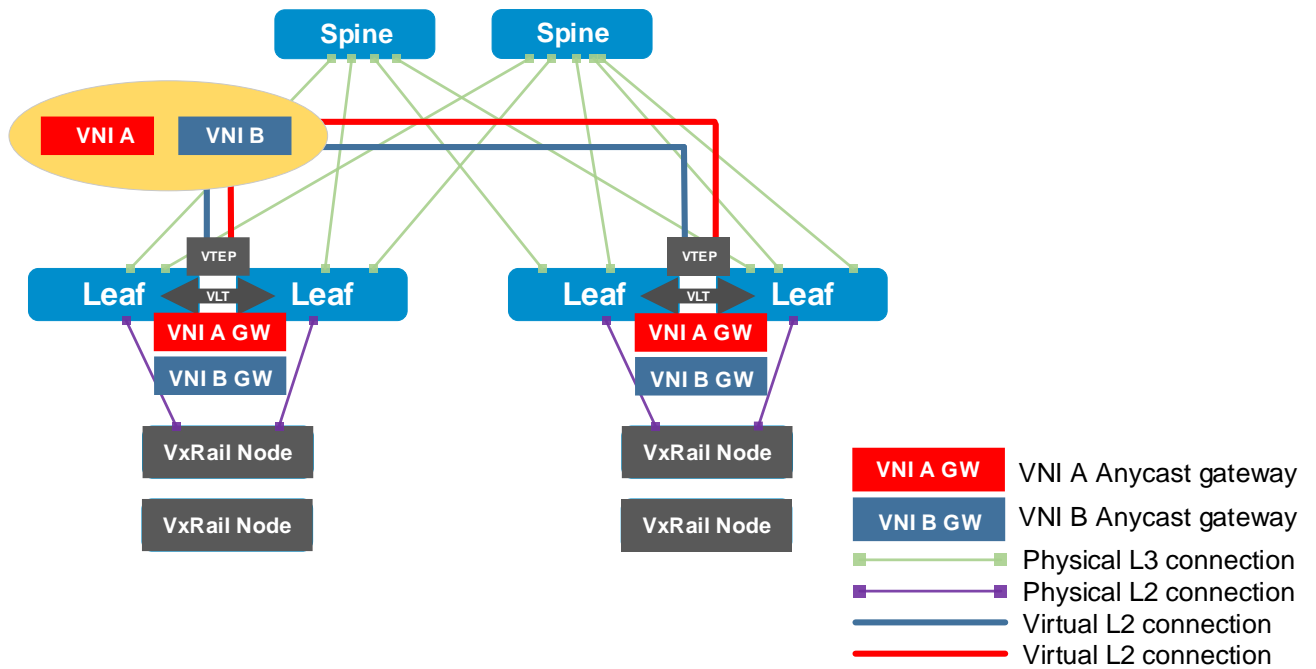


Figure 9 BGP EVPN topology

This deployment guide uses EVPN/VXLAN to achieve the following:

- Tunneling of Layer 2 overlay virtual networks through a physical Layer 3 leaf-spine underlay network using VXLAN-based EVPN to allow VxRail to communicate across four networks:
  - ESXi Management
  - vSAN
  - vMotion
  - VxRail management

### 3.2.1 The VXLAN protocol

VXLAN allows a Layer 2 network to scale across the data center by overlaying an existing Layer 3 network and is described in Internet Engineering Task Force document [RFC 7348](#). Each overlay is seen as a VXLAN segment.

Each segment is identified through a 24-bit segment ID seen as a VNI. This allows up to 16 Million VNIs, far more than the traditional 4,094 VLAN IDs that are allowed on a physical switch.

VXLAN is a tunneling scheme that encapsulates Layer 2 frames in User Datagram Protocol (UDP) segments, as shown in Figure 10.

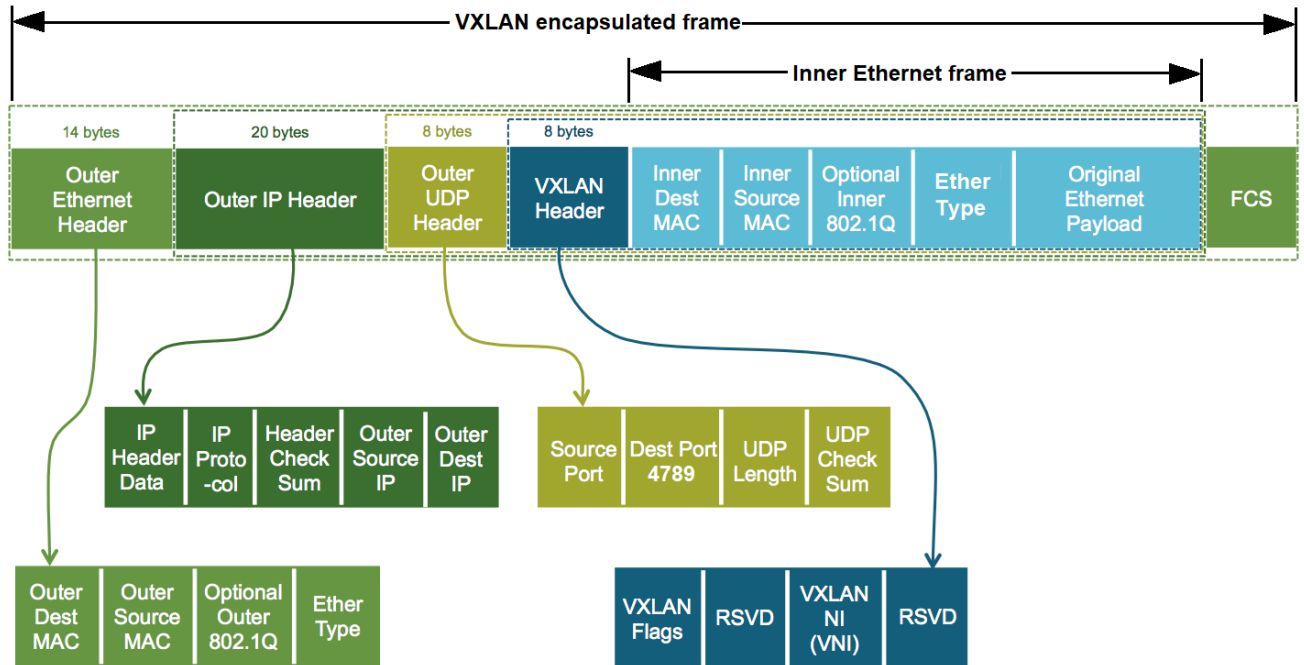


Figure 10 VXLAN encapsulated frame

VXLAN encapsulation adds approximately 50 bytes of overhead to each Ethernet frame. As a result, all switches in the underlay (physical) network must be configured to support an MTU of at least 1600 bytes on all participating interfaces.

---

**Note:** In this deployment example, switch interfaces are set to their maximum supported MTU size of 9216 bytes.

---

VTEPs handle VXLAN encapsulation and de-encapsulation. In this implementation, the leaf switches are the VTEPs.

### 3.2.2 BGP EVPN VXLAN operation

EVPN uses BGP to exchange endpoint MAC and IP address information between VTEPs. When a host sends a packet to an endpoint, the switch looks up the routing table for a match. If it finds a match that exists behind another VTEP, the packet is encapsulated with VXLAN and UDP headers and encapsulated again with outer IP and Ethernet headers for transport over the leaf-spine network. When the packet arrives at the destination VTEP, the outer Ethernet, IP, UDP, and VXLAN headers are removed, and the switch sends the original packet to the endpoint.

## 4 Topology

### 4.1 Leaf-spine underlay

In a Layer 3 leaf-spine network, the traffic between leaf switches and spine switches are routed. Equal cost multipath routing (ECMP) is used to load balance traffic across the Layer 3 connections. BGP is used to exchange routes. The Layer 3/Layer 2 (L3/L2) boundary is at the leaf switches.

Two leaf switches are configured as Virtual Link Trunking (VLT) peers at the top of each rack. VLT allows all connections to be active while also providing fault tolerance. As administrators add racks to the data center, two leaf switches configured for VLT are added to each new rack. Connections within racks from hosts to leaf switches are Layer 2, and each host is connected using a VLT port-channel.

In this example, two Z9264F-ON switches are used as spines, and four S5248F-ON switches are used as leaf switches in Rack 1 and Rack 2.

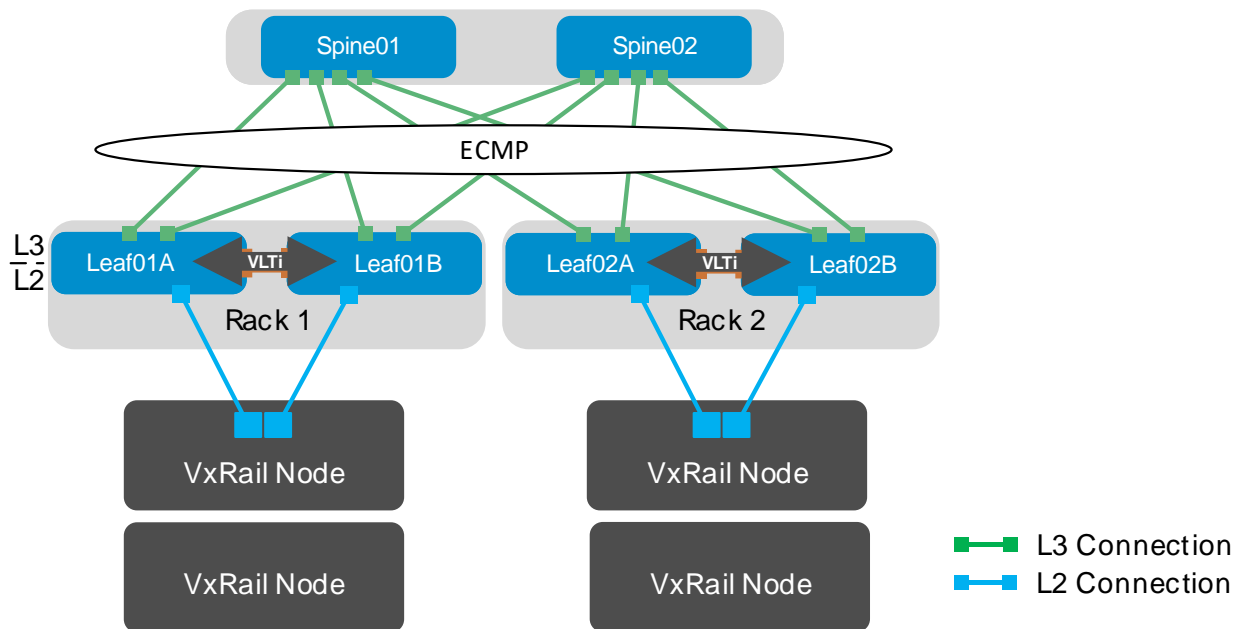


Figure 11 Leaf-spine underlay network

**Note:** Using a leaf-spine network in the data center is considered a best practice. With Z9264F-ON switches as spines and two leaf switches per rack, this topology scales to 32 racks. For more leaf-spine network information, see Dell EMC PowerSwitch Layer 3 Leaf-Spine Deployment and Best Practices with OS10. There are some BGP configuration differences in this guide to enable the BGP EVPN VXLAN feature.



### 4.1.1 BGP ASNs and router IDs

Figure 12 shows the autonomous system numbers (ASNs) and router IDs used for the leaf and spine switches in this guide. Spine switches share a common ASN, and each pair of leaf switches shares a common ASN.

ASNs should follow a logical pattern for ease of administration and allow for growth as switches are added. Using private ASNs in the data center is the best practice. Private, 2-byte ASNs range from 64512 through 65534.

In this example, 65100 is used on both switches at the spine layer. Leaf switches use the next available ASNs, 65101, for example, and the last digit is used to identify the leaf pair. Extra spine switches would be assigned the existing ASN for the spine layer, 65100. Extra leaf switches would be added in pairs with the next pair assigned an ASN of 65103.

The IP addresses shown are loopback addresses that are used as BGP router IDs. Loopback addresses should follow a logical pattern to make it easier to manage and allow for growth. In this example, the 10.0.0.0/16 IP address space is used. The third octet in the address represents the layer, “1” for the spine layer and “2” for the leaf layer, and the fourth octet is the counter for the appropriate layer

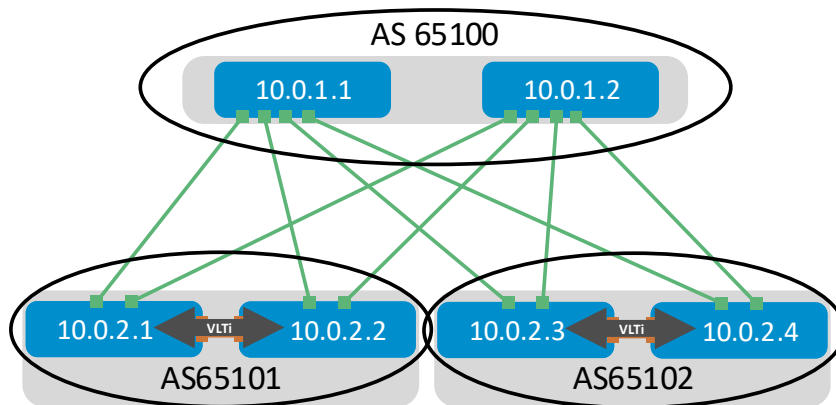


Figure 12 BGP ASNs and router IDs

### 4.1.2 Point-to-point IP networks

Establishing a logical, scalable IP address scheme is important before deploying a leaf-spine topology. The point-to-point links used in this deployment are labeled A-H in Figure 13.

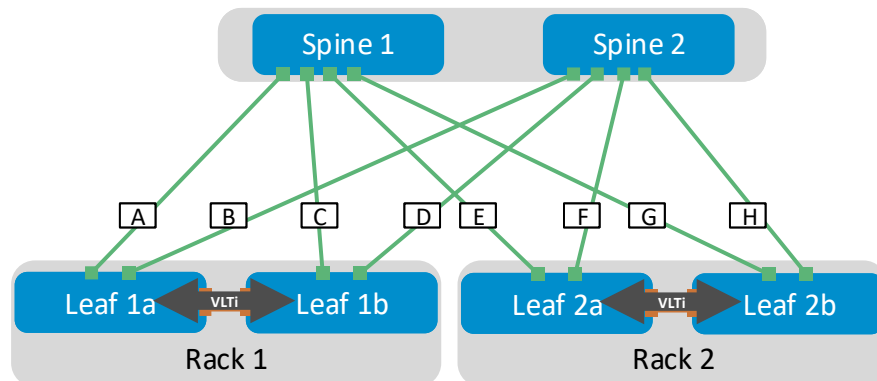


Figure 13 Point-to-point networks

Each link is a separate, point-to-point IP network. Table 1 details the links labeled in Figure 13. The IP addresses in the table are used in the switch configuration examples.

Table 1 Point-to-point network IP addresses

Link label	Source switch	Source IP address	Destination switch	Destination IP address	Network
A	Spine 1	192.168.1.0	Leaf 1a	192.168.1.1	192.168.1.0/31
B	Spine 2	192.168.2.0	Leaf 1a	192.168.2.1	192.168.2.0/31
C	Spine 1	192.168.1.2	Leaf 1b	192.168.1.3	192.168.1.2/31
D	Spine 2	192.168.2.2	Leaf 1b	192.168.2.3	192.168.2.2/31
E	Spine 1	192.168.1.4	Leaf 2a	192.168.1.5	192.168.1.4/31
F	Spine 2	192.168.2.4	Leaf 2a	192.168.2.5	192.168.2.4/31
G	Spine 1	192.168.1.6	Leaf 2b	192.168.1.7	192.168.1.6/31
H	Spine 2	192.168.2.6	Leaf 2b	192.168.2.7	192.168.2.6/31

---

**Note:** As with all examples in this guide, any valid IP address scheme can be used. The earlier example point-to-point addresses use a 31-bit mask to save address space. This is optional and covered in RFC 3021.

---

## 4.2 Underlay network connections

Figure 14 shows the wiring configuration for the six switches that include the leaf-spine network. The solid colored lines are 100 GbE links, and the light blue dashed lines are two QSFP28-DD 200 GbE cable pairs that are used for the VLT interconnect (VLTi). The use of QSFP28-DD offers a 400 GbE VLTi to handle any potential traffic increases resulting from failed interconnects to the spine layers. As a rule, it is suggested to maintain at minimum a 1:1 ratio between available bandwidth to the spine and bandwidth for the VLTi.

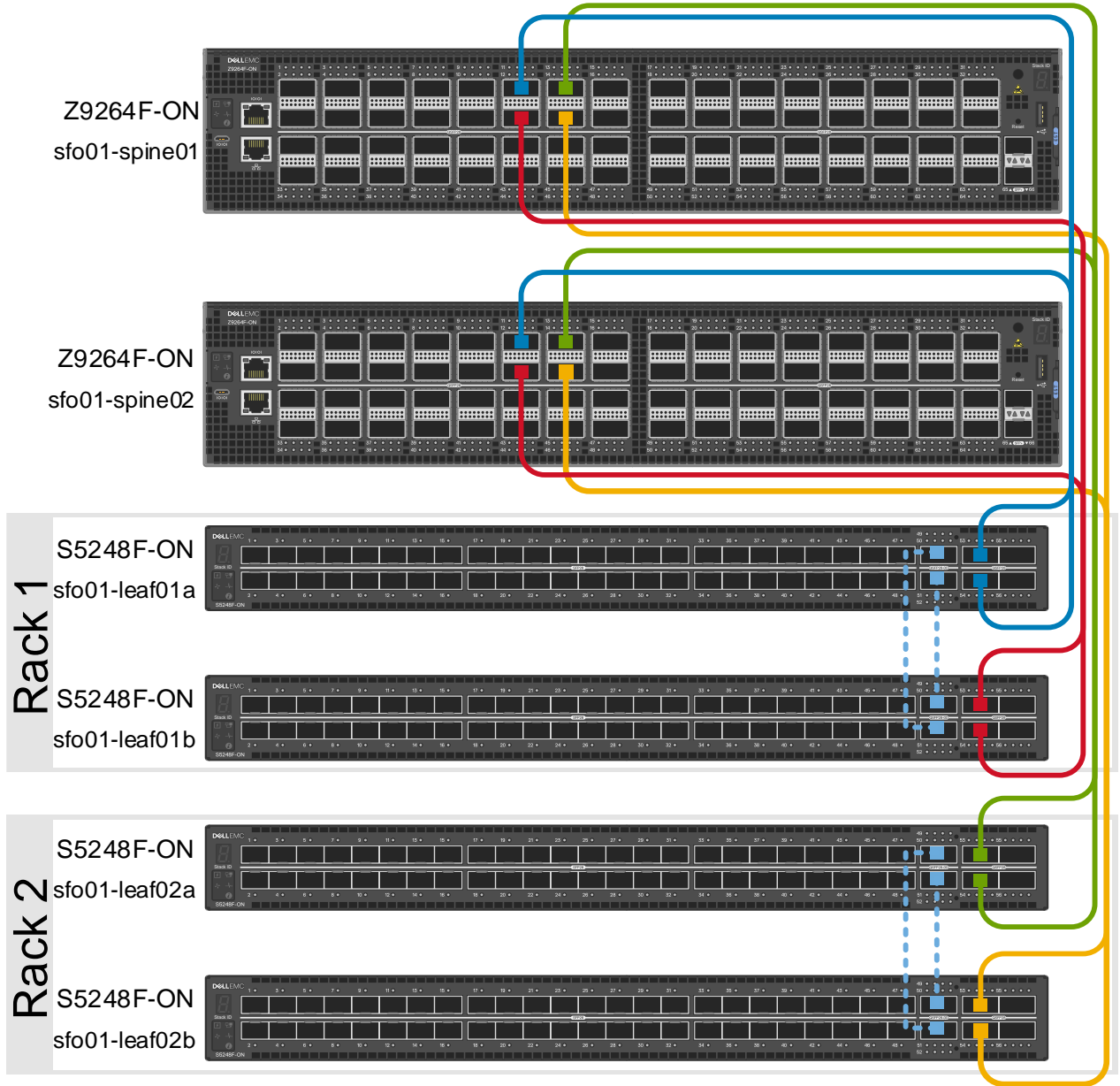


Figure 14 Physical switch topology

**Note:** All switch configuration commands are provided in the file attachments. See Section 1.7 for instructions on accessing the attachments.

## 4.3 BGP EVPN VXLAN overlay

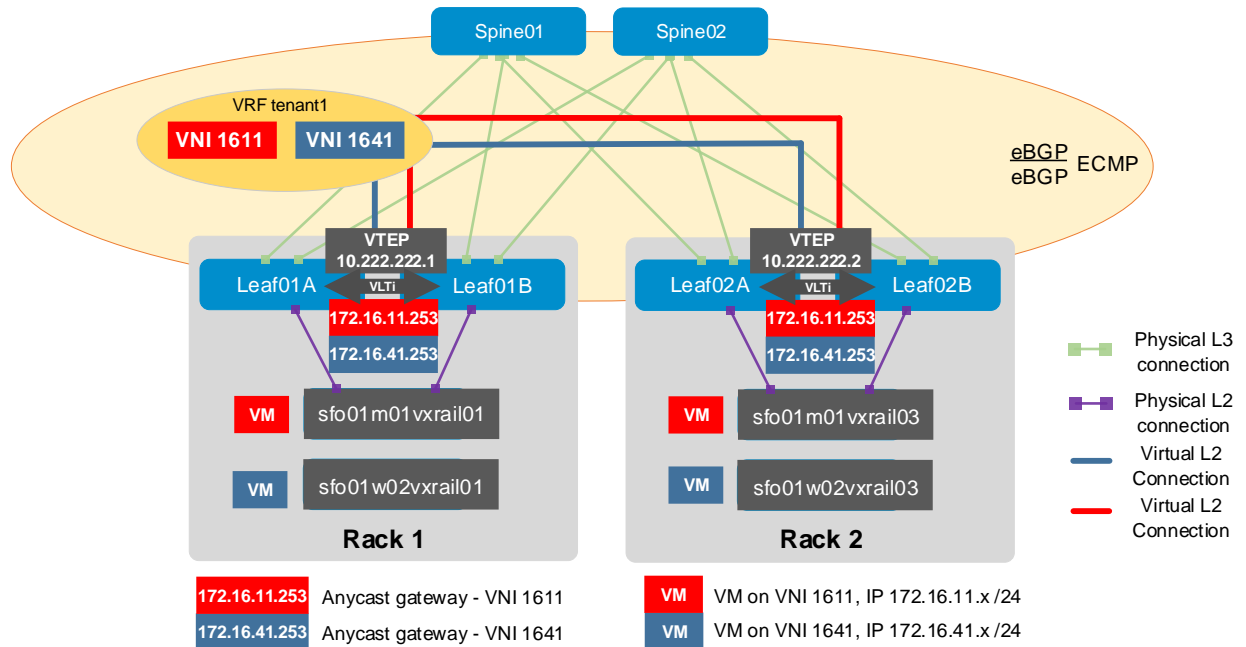


Figure 15 BGP EVPN topology with anycast gateways and an indirect gateway

In this deployment example, four VNIs are used: 1641, 1642, and 1643, and 3939. All VNIs are configured all on the four leaf switches. However, only VNIs 1641, 1642, and 1643 are configured with anycast gateways. Because these VNIs have anycast gateways, VMs on those VNIs which are routing to other networks can use the same gateway information while behind different leaf pairs. When those VMs route, their local leaf switches will always be doing the routing. This replaces VRRP and enables VMs to migrate from one leaf pair to another without the need to change the network configuration. It also eliminates hair pinning and improves link utilization since routing is performed closer to the source.

---

**Note:** VNI 1611 is used in the management domain and hosts all management VMs, including vCenter, PSCs, and the NSX-T management cluster.

---

## 4.4 VxRail node connections

Workload domains include combinations of ESXi hosts and network equipment which can be set up with varying levels of hardware redundancy. Workload domains are connected to a network core that distributes data between them.

Figure 16 shows a physical view of Rack 1. On each VxRail node, the NDC links carry traditional VxRail network traffic such as management, vMotion, vSAN, and VxRail management traffic. The 2x 25 GbE PCIe, shown here in slot 2, is dedicated to NSX-T overlay and NSX-T uplink traffic. Resiliency is achieved by providing redundant leaf switches at the ToR.

Each VxRail node has an iDRAC connected to an S3048-ON OOB management switch. This connection is used for the initial node configuration. The S5248F-ON leaf switches are connected using two QSFP28-DD 200 GbE direct-access cables (DAC) forming a VLT interconnect (VLTi) for a total throughput of 400 GbE. Upstream connections to the spine switches are not shown but are configured using two QSFP28 100 GbE uplinks.

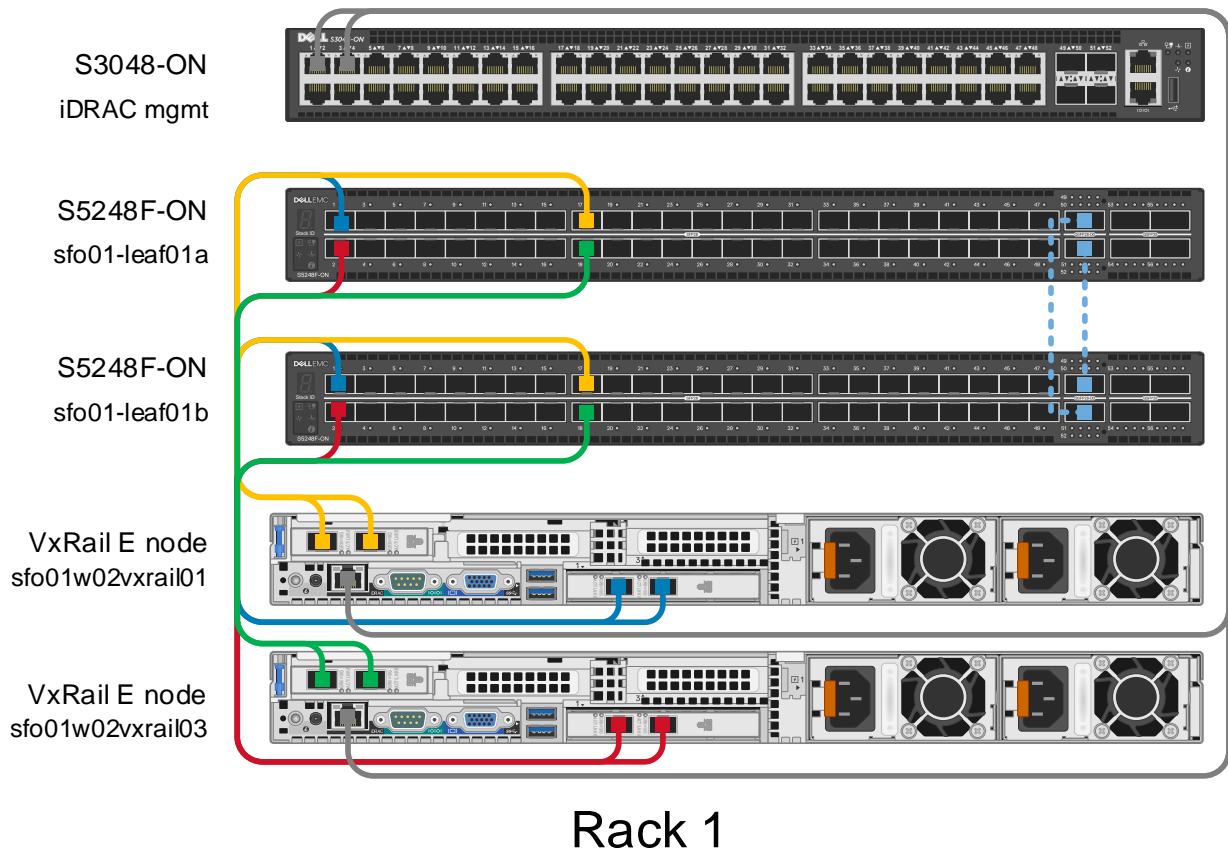


Figure 16 Dell EMC VxRail multirack Rack 1 physical connectivity

## 5 Planning and preparation

Before creating the IP underlay that drives the SDDC, it is essential to plan out the networks, IP subnets, and external services required. Also, planning of the prerequisites on all required switching hardware is recommended.

### 5.1 VLAN IDs and IP subnets

VCF on VxRail requires that specific VLAN IDs and IP subnets for the traffic types in the SDDC are defined ahead of time. Table 2 shows the values that are used in this document. The NSX-V TEP VLAN is configured to use the VLAN ID in each rack while the IP subnet is changed based on the rack. For example, 172.25.101.0/24 represents Rack 1 with a gateway of 172.25.101.253.

Table 2 VLAN and IP subnet configuration

Cluster	VLAN function	VLAN ID	VNI	Subnet	Gateway
Workload cluster	ESXi Management	1641	1641	172.16.41.0/24	172.16.41.253
	vSphere vMotion	1642	1642	172.16.42.0/24	172.16.42.253
	vSAN	1643	1643	172.16.43.0/24	172.16.43.253
	VxRail management	3939	3939	-	-
	NSX-T TEPs	2500	-	172.25.n.0/24	172.25.n.253
	NSX-T Uplink01	1647	-	172.16.47.0/24	-
	NSX-T Uplink02	1648	-	172.16.48.0/24	-

---

**Note:** Use these VLAN IDs and IP subnets as samples. Configure the VLAN IDs and IP subnets according to the environment.

---

### 5.2 External services

In this section, the following are discussed, along with the guidelines on placement:

- DNS
- NTP
- DHCP

## 5.3 DNS

In this document, the Active Directory (AD) servers provide DNS services. Other DNS records that are used in this document follow the VVD examples. The examples can be found in the VVD documentation section, [Prerequisites for the NSX-T Deployment](#).

Table 3 Hostnames and IP addresses for the external services

Component group	Hostname	DNS zone	IP address	Description
AD/DNS	dc01rpl	rainpole.local	172.16.11.4	Windows 2016 host containing AD and DNS server for rainpole.local
	dc01sfo	sfo01.rainpole.local	172.16.11.5	AD and DNS server in a child domain

### 5.3.1 NTP

Synchronized systems over NTP are essential for the validity of vCenter Single Sign-On and other certificates. Consistent system clocks are essential for the proper operation of the components in the SDDC because, in some instances, they rely on vCenter Single Sign-on. Using NTP also makes it easier to correlate log files from multiple sources during troubleshooting, auditing, or inspection of log files to detect attacks.

Table 4 shows the DNS Canonical Name (CNAME) record that maps the two time sources to one DNS name.

Table 4 NTP server FQDN and IP configuration

NTP server FQDN	Mapped IP address
ntp.sfo01.rainpole.local	<ul style="list-style-type: none"><li>172.16.11.5</li><li>172.16.11.4</li></ul>
0.ntp.sfo01.rainpole.local	172.16.11.5
1.ntp.sfo01.rainpole.local	172.16.11.4

### 5.3.2 DHCP

DHCP is required for each VMkernel port of the ESXi hosts with an IPv4 address. A Microsoft Windows Server 2016 virtual machine that is associated with external services on subnet 10.10.14.0/24 is used in this deployment. DHCP relay (`ip helper-address`) is used on the leaf switches to route DHCP requests on behalf of the NSX VTEPs to the DHCP server. Table 5 outlines the DHCP values that are used in this document.

The VVD outlines the example usage of VLAN 1644 and the IP subnet of 172.16.44.0/24. In this paper, this has been modified to accommodate multiple subnets. VLAN ID 2500 is used, and the corresponding IP subnets are reserved in the underlay network for these subnets. The third octet increases by 1 to represent the rack ID. For example, rack 1 is 172.25.101.0/24.

---

**Note:** This scheme can be expanded to include multiple available zones, a topic that is not covered in this workload domain deployment.

---

Table 5 shows the IP address ranges used in this document. The DHCP servers in either availability zone are assumed to be configured correctly and are outside of the scope of this document.

Table 5 DHCP scope values

ID	DHCP server IP address	Start IP address	End IP address	Gateway	Subnet mask
Rack 1	10.10.14.5	172.25.101.1	172.25.101.199	172.25.101.253	/24
Rack 2	10.10.14.5	172.25.102.1	172.25.102.199	172.25.102.253	/24

## 5.4 Switch preparation

### 5.5 Check switch OS version

Dell EMC PowerSwitches must be running OS10EE version 10.4.3.5 or later for this deployment.

Run the `show version` command to check the OS version. Dell EMC recommends upgrading to the latest release available on [Dell Digital Locker](#) (account required).

```
OS10# show version
Dell EMC Networking OS10-Enterprise
Copyright (c) 1999-2019 by Dell Inc. All Rights Reserved.
OS Version: 10.4.3.5
Build Version: 10.4.3.5
```

---

**Note:** For information about installing and upgrading OS10EE, see the Dell EMC Networking OS10 Enterprise Edition Quick Start Guide.

---

### 5.6 Verify license installation

Run the `show license status` command to verify the license installation. Verify that the **License Type:** field shows **PERPETUAL**, as shown in the following example.

---

**Note:** If an evaluation license is installed, licenses purchased from Dell EMC are available for download on Dell Digital Locker. Installation instructions are provided in the OS10 Enterprise Edition User Guide Release 10.4.3.0.

---

```
OS10# show license status
System Information
-----
Vendor Name       : Dell EMC
Product Name     : S5248F-ON
Hardware Version  : A01
Platform Name    : x86_64-dellemc_s5248f_c3538-r0
PPID             : CN046MRJCES0089K0004
Service Tag      : 68X00Q2
Product Base     :
Product Serial Number:
Product Part Number :
License Details
-----
Software         : OS10-Enterprise
```



```
Version      :      10.4.3.5
License Type :      PERPETUAL
License Duration:      Unlimited
License Status :      Active
License location:      /mnt/license/68X00Q2.lic
```

---

**Note:** A perpetual license is already on the switch if OS10EE was factory installed.

---

## 5.7 Factory default configuration

The switch configuration commands in the chapters that follow begin with the leaf switches at their factory default settings. Dell EMC PowerSwitches running OS10EE can be reset to their default configuration as follows:

```
OS10# delete startup-configuration
Proceed to delete startup-configuration [confirm yes/no(default)]:y

OS10# reload
System configuration has been modified. Save? [yes/no]:n
Proceed to reboot the system? [confirm yes/no]:y
```

The switch reboots to its factory default configuration.

---

**Note:** OS10EE at its default settings has Telnet disabled, SSH enabled, and the OOB management interface that is configured to get its IP address using DHCP. The default username and password are both admin. Dell EMC recommends changing the admin password to a complex password during the first login.

---

## 5.8 Switch settings

Table 6 shows the unique values for the four S5248F-ON switches. The table provides a summary of the configuration differences between each switch and each VLT switch pair.

Table 6 Unique switch settings for leaf switches

Setting	S5248F-Leaf1A	S5248F-Leaf1B	S5248F-Leaf2A	S5248F-Leaf2B
Hostname	sfo01-Leaf01A	sfo01-Leaf01B	sfo01-Leaf02A	sfo01-Leaf02B
OOB IP address	100.67.198.32/24	100.67.198.31/24	100.67.198.30/24	100.67.198.29/24
Autonomous System Number (ASN)	65101	65101	65102	65102
Point-to-point interface IP addresses	192.168.1.1/31 192.168.2.1/31	192.168.1.3/31 192.168.2.3/31	192.168.1.5/31 192.168.2.5/31	192.168.1.7/31 192.168.2.7/31
Loopback0 address (router ID)	10.0.2.1/32	10.0.2.2/32	10.0.2.3/32	10.0.2.4/32
Loopback1 address (EVPN)	10.2.2.1/32	10.2.2.2/32	10.2.2.3/32	10.2.2.4/32
Loopback2 address (NVE)	10.222.222.1/32	10.222.222.1/32	10.222.222.2/32	10.222.222.2/32
VLAN 4000 IP address	192.168.3.0/31	192.168.3.1/31	192.168.3.2/31	192.168.3.3/31
VLAN 2500 IP addresses (interface and VIP)	172.25.101.251/24 172.25.101.253/24	172.25.101.252/24 172.25.101.253/24	172.25.102.251/24 172.25.102.253/24	172.25.102.252/24 172.25.102.253/24
VLAN 1647 IP addresses (ESG)	172.16.47.1/24	-	-	-
VLAN 1648 IP addresses (ESG)	-	172.16.48.1/24	-	-
virtual-network 1641 IP addresses (interface and anycast)	172.16.41.252/24 172.16.41.253/24	172.16.41.251/24 172.16.41.253/24	172.16.41.250/24 172.16.41.253/24	172.16.41.249/24 172.16.41.253/24
virtual-network 1642 IP addresses (interface and anycast)	172.16.42.252/24 172.16.42.253/24	172.16.42.251/24 172.16.42.253/24	172.16.42.250/24 172.16.42.253/24	172.16.42.249/24 172.16.42.253/24
virtual-network 1643 IP addresses (interface and anycast)	172.16.43.252/24 172.16.43.253/24	172.16.43.251/24 172.16.43.253/24	172.16.43.250/24 172.16.43.253/24	172.16.43.249/24 172.16.43.253/24

**Note:** Use these VLAN IDs and IP subnets as samples. Configure the VLAN IDs and IP subnets according to your environment.



1. Configure general switch settings, including management and NTP source.

```
OS10# configure terminal
OS10(config)# interface mgmt 1/1/1
OS10(conf-if-ma-1/1/1)# no ip address dhcp
OS10(conf-if-ma-1/1/1)# ip address 100.67.198.32/24
OS10(conf-if-ma-1/1/1)# exit
OS10(config)# management route 100.67.0.0/16 managementethernet
OS10(config)# hostname sfo01-Leaf01A
sfo01-Leaf01A(config)# ntp server 100.67.10.20
sfo01-Leaf01A(config)# bfd enable
sfo01-Leaf01A(config)# ipv6 mld snooping enable
```

2. Configure a loopback interface for the Router ID using the following command:

```
sfo01-Leaf01A(config)# interface loopback 0
sfo01-Leaf01A(conf-if-lo-0)# description Router-ID
sfo01-Leaf01A(conf-if-lo-0)# no shutdown
sfo01-Leaf01A(conf-if-lo-0)# ip address 10.0.2.1/32
sfo01-Leaf01A(conf-if-lo-0)# exit
```

3. Configure a loopback interface for NVE.

```
sfo01-Leaf01A(config)# interface loopback 2
sfo01-Leaf01A(conf-if-lo-2)# description nve_loopback
sfo01-Leaf01A(conf-if-lo-2)# no shutdown
sfo01-Leaf01A(conf-if-lo-2)# ip address 10.222.222.1/32
sfo01-Leaf01A(conf-if-lo-2)# exit
```

4. Configure the loopback interface for the VXLAN source tunnel interface.

```
sfo01-Leaf01A(config)# nve
sfo01-Leaf01A(config-nve)# source-interface loopback2
sfo01-Leaf01A(config-nve)# exit
```

5. Using the following commands, configure the VXLAN virtual networks:

```
sfo01-Leaf01A(config)# virtual-network 1641
sfo01-Leaf01A(config-vn)# vxlan-vni 1641
sfo01-Leaf01A(config-vn)# exit
sfo01-Leaf01A(config)# virtual-network 1642
sfo01-Leaf01A(config-vn)# vxlan-vni 1642
sfo01-Leaf01A(config-vn)# exit
sfo01-Leaf01A(config)# virtual-network 1643
sfo01-Leaf01A(config-vn)# vxlan-vni 1643
sfo01-Leaf01A(config-vn)# exit
sfo01-Leaf01A(config)# virtual-network 3939
sfo01-Leaf01A(config-vn)# vxlan-vni 3939
sfo01-Leaf01A(config-vn)# exit
```

6. Assign the VLAN member interfaces to virtual networks.

```
sfo01-Leaf01A(config)# interface vlan1641
sfo01-Leaf01A(config-if-vl-1641)# description sfo-w02-mgmt
sfo01-Leaf01A(config-if-vl-1641)# virtual-network 1641
sfo01-Leaf01A(config-if-vl-1641)# no shutdown
sfo01-Leaf01A(config-if-vl-1641)# mtu 9216
sfo01-Leaf01A(config-if-vl-1641)# exit
sfo01-Leaf01A(config)# interface vlan1642
sfo01-Leaf01A(config-if-vl-1642)# virtual-network 1642
sfo01-Leaf01A(config-if-vl-1642)# description sfo-w02-vmotion
sfo01-Leaf01A(config-if-vl-1642)# no shutdown
sfo01-Leaf01A(config-if-vl-1642)# mtu 9216
sfo01-Leaf01A(config-if-vl-1642)# exit
sfo01-Leaf01A(config)# interface vlan1643
sfo01-Leaf01A(config-if-vl-1643)# virtual-network 1643
sfo01-Leaf01A(config-if-vl-1643)# description sfo-w02-san
sfo01-Leaf01A(config-if-vl-1643)# no shutdown
sfo01-Leaf01A(config-if-vl-1643)# mtu 9216
sfo01-Leaf01A(config-if-vl-1643)# exit
sfo01-Leaf01A(config)# interface vlan3939
sfo01-Leaf01A(config-if-vl-3939)# description vxrail-mgmt
sfo01-Leaf01A(config-if-vl-3939)# virtual-network 3939
sfo01-Leaf01A(config-if-vl-3939)# ipv6 mld snooping querier
sfo01-Leaf01A(config-if-vl-3939)# no shutdown
sfo01-Leaf01A(config-if-vl-3939)# mtu 9216
sfo01-Leaf01A(config-if-vl-3939)# exit
```

---

**Note:** Enable the ipv6 mld snooping querier on any logical network that requires multicast support.

---

7. Configure access ports as VLAN members for a switch-scoped VLAN-to-VNI mapping. VLAN 1641 is untagged to accommodate ESXi management, which is configured by default as untagged on the management VMkernel. The remaining VLANs are tagged to support the tagging from the default VxRail created vDS.

```
sfo01-Leaf01A(config)# interface ethernet1/1/3
sfo01-Leaf01A(conf-if-eth1/1/1)# description sfo01w02vxrail01
sfo01-Leaf01A(conf-if-eth1/1/1)# no shutdown
sfo01-Leaf01A(conf-if-eth1/1/1)# switchport mode trunk
sfo01-Leaf01A(conf-if-eth1/1/1)# switchport access vlan 1641
sfo01-Leaf01A(conf-if-eth1/1/1)# switchport trunk allowed vlan 1642-
1643,3939
sfo01-Leaf01A(conf-if-eth1/1/1)# mtu 9216
sfo01-Leaf01A(conf-if-eth1/1/1)# spanning-tree port type edge
sfo01-Leaf01A(conf-if-eth1/1/1)# flowcontrol receive on
sfo01-Leaf01A(conf-if-eth1/1/1)# flowcontrol transmit off
sfo01-Leaf01A(conf-if-eth1/1/1)# exit
sfo01-Leaf01A(config)# interface ethernet1/1/4
sfo01-Leaf01A(conf-if-eth1/1/2)# description sfo01w02vxrail02
sfo01-Leaf01A(conf-if-eth1/1/2)# no shutdown
sfo01-Leaf01A(conf-if-eth1/1/2)# switchport mode trunk
```

```

sfo01-Leaf01A(conf-if-eth1/1/2) # switchport access vlan 1641
sfo01-Leaf01A(conf-if-eth1/1/2) # switchport trunk allowed vlan 1642-1643,3939
sfo01-Leaf01A(conf-if-eth1/1/2) # mtu 9216
sfo01-Leaf01A(conf-if-eth1/1/2) # spanning-tree port type edge
sfo01-Leaf01A(conf-if-eth1/1/2) # flowcontrol receive on
sfo01-Leaf01A(conf-if-eth1/1/2) # flowcontrol transmit off
sfo01-Leaf01A(conf-if-eth1/1/2) # exit

```

8. Enter the following commands to configure upstream network-facing ports:

```

sfo01-Leaf01A(config) # interface ethernet1/1/53
sfo01-Leaf01A(conf-if-eth1/1/53) # description sfo01-spine01
sfo01-Leaf01A(conf-if-eth1/1/53) # no shutdown
sfo01-Leaf01A(conf-if-eth1/1/53) # no switchport
sfo01-Leaf01A(conf-if-eth1/1/53) # mtu 9216
sfo01-Leaf01A(conf-if-eth1/1/53) # ip address 192.168.1.1/31
sfo01-Leaf01A(conf-if-eth1/1/53) # exit
sfo01-Leaf01A(config) # interface ethernet1/1/54
sfo01-Leaf01A(conf-if-eth1/1/54) # description sfo01-spine02
sfo01-Leaf01A(conf-if-eth1/1/54) # no shutdown
sfo01-Leaf01A(conf-if-eth1/1/54) # no switchport
sfo01-Leaf01A(conf-if-eth1/1/54) # mtu 9216
sfo01-Leaf01A(conf-if-eth1/1/54) # ip address 192.168.2.1/31
sfo01-Leaf01A(conf-if-eth1/1/54) # exit

```

9. Add a route map. This example route map is used to illustrate how to allow IP traffic to be passed on the switch.

```

sfo01-Leaf01A(config) # ip prefix-list spine-leaf seq 10 permit 10.0.2.0/24 ge 32
sfo01-Leaf01A(config) # ip prefix-list spine-leaf seq 20 permit 10.2.2.0/24 ge 32
sfo01-Leaf01A(config) # ip prefix-list spine-leaf seq 30 permit 10.222.222.0/24 ge 32
sfo01-Leaf01A(config) # route-map spine-leaf permit 10
sfo01-Leaf01A(config-route-map) # match ip address prefix-list spine-leaf
sfo01-Leaf01A(config-route-map) # exit

```

10. Configure eBGP

```

sfo01-Leaf01A(config) # router bgp 65101
sfo01-Leaf01A(config-router-bgp-65101) # router-id 10.0.2.1
sfo01-Leaf01A(config-router-bgp-65101) # bfd all-neighbors interval 200 min_rx 200 multiplier 3 role active
sfo01-Leaf01A(config-router-bgp-65101) # address-family ipv4 unicast
sfo01-Leaf01A(config-router-bgpv4-af) # redistribute connected route-map spine-leaf
sfo01-Leaf01A(config-router-bgpv4-af) # exit
sfo01-Leaf01A(config-router-bgp-65101) # bestpath as-path multipath-relax
sfo01-Leaf01A(config-router-bgp-65101) # maximum-paths ebgp 2

```

---

**Note:** If more than two ESGs are being used, update the maximum-paths ebgp value accordingly.

---

11. Configure eBGP for the IPv4 point-to-point peering using the following commands:

```
sfo01-Leaf01A(config-router-bgp-65101)# neighbor 192.168.1.0
sfo01-Leaf01A(config-router-neighbor)# advertisement-interval 5
sfo01-Leaf01A(config-router-neighbor)# bfd
sfo01-Leaf01A(config-router-neighbor)# fall-over
sfo01-Leaf01A(config-router-neighbor)# remote-as 65100
sfo01-Leaf01A(config-router-neighbor)# no shutdown
sfo01-Leaf01A(config-router-neighbor)# address-family ipv4 unicast
sfo01-Leaf01A(config-router-neighbor-af)# exit
sfo01-Leaf01A(config-router-neighbor)# exit
sfo01-Leaf01A(config-router-bgp-65101)# neighbor 192.168.2.0
sfo01-Leaf01A(config-router-neighbor)# advertisement-interval 5
sfo01-Leaf01A(config-router-neighbor)# bfd
sfo01-Leaf01A(config-router-neighbor)# fall-over
sfo01-Leaf01A(config-router-neighbor)# remote-as 65100
sfo01-Leaf01A(config-router-neighbor)# no shutdown
sfo01-Leaf01A(config-router-neighbor)# address-family ipv4 unicast
sfo01-Leaf01A(config-router-neighbor-af)# exit
sfo01-Leaf01A(config-router-neighbor)# exit
sfo01-Leaf01A(config-router-bgp-65101)# exit
```

12. Configure a loopback interface for BGP EVPN peering.

```
sfo01-Leaf01A(config)# interface loopback 1
sfo01-Leaf01A(conf-if-lo-1)# description evpn_loopback
sfo01-Leaf01A(conf-if-lo-1)# no shutdown
sfo01-Leaf01A(conf-if-lo-1)# ip address 10.2.2.1/32
sfo01-Leaf01A(conf-if-lo-1)# exit
```

13. Enter the following commands to configure BGP EVPN peering:

```
sfo01-Leaf01A(config)# router bgp 65101
sfo01-Leaf01A(config-router-bgp-65101)# neighbor 10.2.1.1
sfo01-Leaf01A(config-router-neighbor)# remote-as 65100
sfo01-Leaf01A(config-router-neighbor)# ebgp-multihop 2
sfo01-Leaf01A(config-router-neighbor)# send-community extended
sfo01-Leaf01A(config-router-neighbor)# update-source loopback1
sfo01-Leaf01A(config-router-neighbor)# no shutdown
sfo01-Leaf01A(config-router-neighbor)# address-family ipv4 unicast
sfo01-Leaf01A(config-router-neighbor-af)# no activate
sfo01-Leaf01A(config-router-neighbor-af)# exit
sfo01-Leaf01A(config-router-neighbor)# address-family l2vpn evpn
sfo01-Leaf01A(config-router-neighbor-af)# activate
sfo01-Leaf01A(config-router-neighbor-af)# exit
sfo01-Leaf01A(config-router-neighbor)# exit
sfo01-Leaf01A(config-router-bgp-65101)# neighbor 10.2.1.2
sfo01-Leaf01A(config-router-neighbor)# remote-as 65100
sfo01-Leaf01A(config-router-neighbor)# ebgp-multihop 2
sfo01-Leaf01A(config-router-neighbor)# send-community extended
```

```

sfo01-Leaf01A(config-router-neighbor)# update-source loopback1
sfo01-Leaf01A(config-router-neighbor)# no shutdown
sfo01-Leaf01A(config-router-neighbor)# address-family ipv4 unicast
sfo01-Leaf01A(config-router-neighbor-af)# no activate
sfo01-Leaf01A(config-router-neighbor-af)# exit
sfo01-Leaf01A(config-router-neighbor)# address-family l2vpn evpn
sfo01-Leaf01A(config-router-neighbor-af)# activate
sfo01-Leaf01A(config-router-neighbor-af)# exit
sfo01-Leaf01A(config-router-neighbor)# exit
sfo01-Leaf01A(config-router-bgp-65101)# exit

```

#### 14. Configure EVPN

```

sfo01-Leaf01A(config)# evpn
sfo01-Leaf01A(config-evpn)# evi 1641
sfo01-Leaf01A(config-evpn-evi-1641)# vni 1641
sfo01-Leaf01A(config-evpn-evi-1641)# rd 10.222.222.1:1641
sfo01-Leaf01A(config-evpn-evi-1641)# route-target 1641:1641 both
sfo01-Leaf01A(config-evpn-evi-1641)# exit
sfo01-Leaf01A(config-evpn)# evi 1642
sfo01-Leaf01A(config-evpn-evi-1642)# vni 1642
sfo01-Leaf01A(config-evpn-evi-1642)# rd 10.222.222.1:1642
sfo01-Leaf01A(config-evpn-evi-1642)# route-target 1642:1642 both
sfo01-Leaf01A(config-evpn-evi-1642)# exit
sfo01-Leaf01A(config-evpn)# evi 1643
sfo01-Leaf01A(config-evpn-evi-1643)# vni 1643
sfo01-Leaf01A(config-evpn-evi-1643)# rd 10.222.222.1:1643
sfo01-Leaf01A(config-evpn-evi-1643)# route-target 1643:1643 both
sfo01-Leaf01A(config-evpn-evi-1643)# exit

```

#### 15. Configure the dedicated L3 underlay path to reach a VLT peer in the event of network failure:

```

sfo01-Leaf01A(config)# interface vlan4000
sfo01-Leaf01A(config-if-vl-4000)# no shutdown
sfo01-Leaf01A(config-if-vl-4000)# mtu 9216
sfo01-Leaf01A(config-if-vl-4000)# ip address 192.168.3.0/31
sfo01-Leaf01A(config-if-vl-4000)# exit

```

#### 16. Configure the VLTi member links.

```

sfo01-Leaf01A(config)# interface range ethernet1/1/49-1/1/52
sfo01-Leaf01A(conf-range-eth1/1/49-1/1/52)# description VLTi
sfo01-Leaf01A(conf-range-eth1/1/49-1/1/52)# no shutdown
sfo01-Leaf01A(conf-range-eth1/1/49-1/1/52)# no switchport
sfo01-Leaf01A(conf-range-eth1/1/49-1/1/52)# exit

```

#### 17. Configure the VLT domain.

```

sfo01-Leaf01A(config)# vlt-domain 1
sfo01-Leaf01A(conf-vlt-1)# backup destination 100.67.198.31
sfo01-Leaf01A(conf-vlt-1)# discovery-interface ethernet1/1/49-1/1/52
sfo01-Leaf01A(conf-vlt-1)# peer-routing
sfo01-Leaf01A(conf-vlt-1)# vlt-mac 00:00:01:02:03:01

```



```
sfo01-Leaf01A(conf-vlt-1) # exit
```

18. Configure the iBGP IPv4 peering between the VLT peers.

```
sfo01-Leaf01A(config) # router bgp 65101
sfo01-Leaf01A(config-router-bgp-65101) # neighbor 192.168.3.1
sfo01-Leaf01A(config-router-neighbor) # remote-as 65101
sfo01-Leaf01A(config-router-neighbor) # no shutdown
sfo01-Leaf01A(config-router-neighbor) # exit
```

19. Create a tenant VRF.

---

**Note:** An OS10 best practice is to isolate any virtual network traffic in a non-default VRF.

---

```
sfo01-Leaf01A(config) # ip vrf tenant1
sfo01-Leaf01A(conf-vrf) # exit
```

20. Configure the anycast gateway MAC address.

```
sfo01-Leaf01A(config) # ip virtual-router mac-address 00:01:01:01:01:01
```

21. Configure routing on virtual networks.

```
sfo01-Leaf01A(config) # interface virtual-network1641
sfo01-Leaf01A(conf-if-vn-1641) # no shutdown
sfo01-Leaf01A(conf-if-vn-1641) # mtu 9216
sfo01-Leaf01A(conf-if-vn-1641) # ip vrf forwarding tenant1
sfo01-Leaf01A(conf-if-vn-1641) # ip address 172.16.41.252/24
sfo01-Leaf01A(conf-if-vn-1641) # ip virtual-router address 172.16.41.253
sfo01-Leaf01A(conf-if-vn-1641) # exit
sfo01-Leaf01A(config) # interface virtual-network1642
sfo01-Leaf01A(conf-if-vn-1642) # no shutdown
sfo01-Leaf01A(conf-if-vn-1642) # mtu 9216
sfo01-Leaf01A(conf-if-vn-1642) # ip vrf forwarding tenant1
sfo01-Leaf01A(conf-if-vn-1642) # ip address 172.16.42.252/24
sfo01-Leaf01A(conf-if-vn-1642) # ip virtual-router address 172.16.42.253
sfo01-Leaf01A(conf-if-vn-1642) # exit
sfo01-Leaf01A(config) # interface virtual-network1643
sfo01-Leaf01A(conf-if-vn-1643) # no shutdown
sfo01-Leaf01A(conf-if-vn-1643) # mtu 9216
sfo01-Leaf01A(conf-if-vn-1643) # ip vrf forwarding tenant1
sfo01-Leaf01A(conf-if-vn-1643) # ip address 172.16.43.252/24
sfo01-Leaf01A(conf-if-vn-1643) # ip virtual-router address 172.16.43.253
sfo01-Leaf01A(conf-if-vn-1643) # exit
sfo01-Leaf01A(config) # interface virtual-network3939
sfo01-Leaf01A(conf-if-vn-3939) # no shutdown
sfo01-Leaf01A(conf-if-vn-3939) # ip vrf forwarding tenant1
sfo01-Leaf01A(conf-if-vn-3939) # exit
```

## 6.2 Configure leaf switch NSX-T overlay networking

In this section, the specific networking required to support the NSX-T overlay networks are configured on sfo01-Leaf1A. Figure 18 shows three networks, VLANs 2500, 1647, and 1648. VLAN 2500 is used to support NSX-T TEPs and VLANs 1647 and 1648 are used for north-south traffic into the NSX-T overlay.

**Note:** The physical connections from the VxRail nodes to the leaf switches use the PCIe card in slot2.

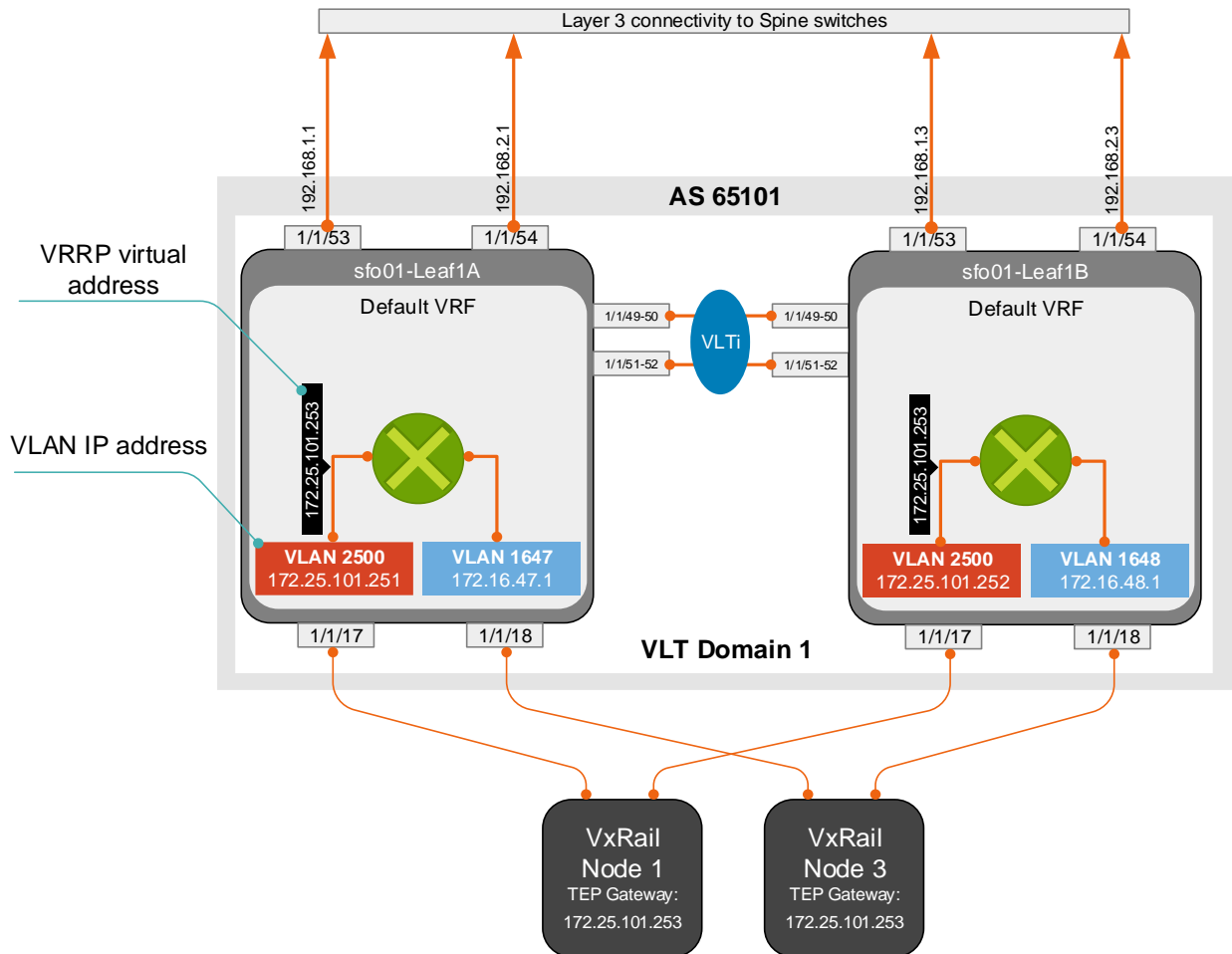


Figure 18 NSX-T networking

1. Configure the interface VLAN 2500 to carry east-west overlay traffic. This VLAN uses the `ip helper-address` command to forward DHCP requests to the DHCP server.

```
sfo01-Leaf01A(config)# interface vlan 2500
sfo01-Leaf01A(conf-if-vl-2500)# description sfo01-w-host-overlay
sfo01-Leaf01A(conf-if-vl-2500)# no shutdown
sfo01-Leaf01A(conf-if-vl-2500)# mtu 9216
sfo01-Leaf01A(conf-if-vl-2500)# ip address 172.25.101.251/24
sfo01-Leaf01A(conf-if-vl-2500)# ip helper-address 10.10.14.5
sfo01-Leaf01A(conf-if-vl-2500)# vrrp-group 250
sfo01-Leaf01A(conf-vlan2500-vrid-250)# virtual-address 172.25.101.253
sfo01-Leaf01A(conf-vlan2500-vrid-250)# exit
sfo01-Leaf01A(conf-if-vl-2500)# exit
```

2. Create VLAN 1647 and assign an IP address. This VLAN is used to carry north-south traffic from the Edge service cluster configured in Section 8.

```
sfo01-Leaf01A(config)# interface vlan1647
sfo01-Leaf01A(config-if-vl-1647)# description sfo01-w-uplink01
sfo01-Leaf01A(config-if-vl-1647)# no shutdown
sfo01-Leaf01A(config-if-vl-1647)# mtu 9216
sfo01-Leaf01A(config-if-vl-1647)# ip address 172.27.11.1/24
sfo01-Leaf01A(config-if-vl-1647)# exit
```

3. Create VLAN 1649. This VLAN provides north-south traffic flow between the transport nodes and the Tier-1 Gateway you created in Section 8.

```
sfo01-Leaf01A(config)# interface vlan1649
sfo01-Leaf01A(config-if-vl-1649)# description sfo01-w-edge-overlay
sfo01-Leaf01A(config-if-vl-1649)# no shutdown
sfo01-Leaf01A(config-if-vl-1649)# mtu 9216
sfo01-Leaf01A(config-if-vl-1649)# exit
```

4. Configure VxRail node ports as VLAN members for VLANs 1647, 1649, and 2500.

```
sfo01-Leaf01A(config)# interface range ethernet1/1/17-1/1/18
sfo01-Leaf01A(conf-range-eth1/1/17-1/1/18)# switchport trunk allowed vlan
1647,1649,2500
sfo01-Leaf01A(conf-range-eth1/1/17-1/1/18)# end
```

5. Configure eBGP for the peering with the ESGs.

```
sfo01-Leaf01A(config)# router bgp 65101
sfo01-Leaf01A(config-router-bgp-65101)# neighbor 172.16.47.2
sfo01-Leaf01A(config-router-neighbor)# advertisement-interval 5
sfo01-Leaf01A(config-router-neighbor)# bfd
sfo01-Leaf01A(config-router-neighbor)# fall-over
sfo01-Leaf01A(config-router-neighbor)# password <bgp-password>
sfo01-Leaf01A(config-router-neighbor)# remote-as 65000
sfo01-Leaf01A(config-router-neighbor)# timers 4 12
sfo01-Leaf01A(config-router-neighbor)# no shutdown
sfo01-Leaf01A(config-router-neighbor)# exit
sfo01-Leaf01A(config-router-bgp-65101)# neighbor 172.16.47.3
sfo01-Leaf01A(config-router-neighbor)# advertisement-interval 5
sfo01-Leaf01A(config-router-neighbor)# bfd
sfo01-Leaf01A(config-router-neighbor)# fall-over
sfo01-Leaf01A(config-router-neighbor)# password <bgp-password>
sfo01-Leaf01A(config-router-neighbor)# remote-as 65000
sfo01-Leaf01A(config-router-neighbor)# timers 4 12
sfo01-Leaf01A(config-router-neighbor)# no shutdown
sfo01-Leaf01A(config-router-neighbor)# end
```

---

**Note:** The BGP password is set during the configuration of the Tier-0 gateway in Section 8.

---

6. Update the existing route map to allow the new networks to pass through the leaf-spine fabric.

```

sfo01-Leaf01A(config)# ip prefix-list spine-leaf seq 70 permit
172.25.101.0/24
sfo01-Leaf01A(config)# ip prefix-list spine-leaf seq 70 permit
172.25.102.0/24
sfo01-Leaf01A(config)# ip prefix-list spine-leaf seq 90 permit
172.16.49.0/24

```

- Repeat the steps using the appropriate values from 5.8, for the remaining leaf switches.

---

**Note:** If any of the workload subnets need to have access to the underlay network, additional IP prefix lists will need to be added.

---

## 6.3 Configure spine switches

This section covers the configuration of the Z9264-ON switch with the hostname `sfo01-spine01` that is shown in Figure 19.

---

**Note:** All switch configuration commands are provided in the file attachments. See Section 1.7 for instructions on accessing the attachments.

---

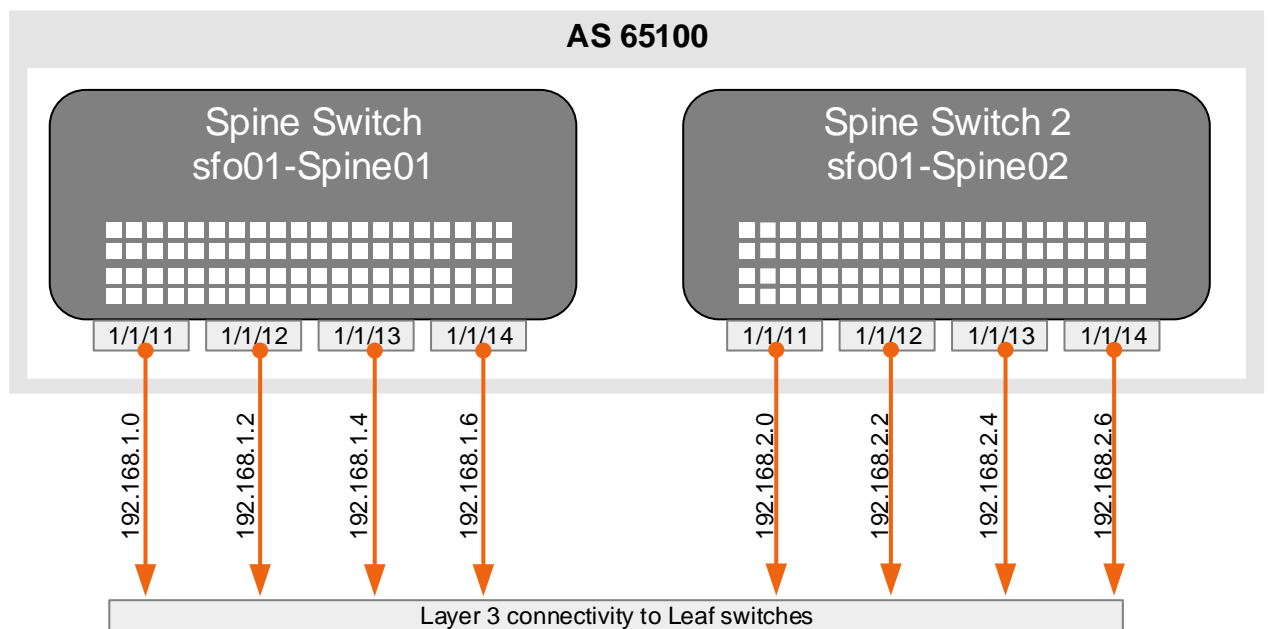


Figure 19 Spine layer diagram

---

**Note:** Some of the steps in this section may already be done if all configuration steps from the VCF on VxRail multitrack deploying using BGP EVPN deployment guide were followed. All steps are included here to ensure completion.

---

- Configure general switch settings, including OOB management and NTP source.

```

OS10# configure terminal
OS10(config)# interface mgmt 1/1/1
OS10(conf-if-ma-1/1/1)# no ip address dhcp
OS10(conf-if-ma-1/1/1)# ip address 100.67.198.36/24

```

```
OS10(conf-if-ma-1/1/1)# exit
OS10(config)# management route 100.67.0.0/16 managementethernet
OS10(config)# hostname sfo01-Spine01
sfo01-Spine01(config)# ntp server 100.67.10.20
sfo01-Spine01(config)# hardware forwarding-table mode scaled-l3-routes
sfo01-Spine01(config)# bfd enable
```

2. Configure a loopback interface for the Router ID.

```
sfo01-Spine01(config)# interface loopback 0
sfo01-Spine01(conf-if-lo-0)# description Router-ID
sfo01-Spine01(conf-if-lo-0)# no shutdown
sfo01-Spine01(conf-if-lo-0)# ip address 10.0.1.1/32
sfo01-Spine01(conf-if-lo-0)# exit
```

3. Using the following commands, configure downstream ports on underlay links to leaf switches.

```
sfo01-Spine01(config)# interface ethernet1/1/11
sfo01-Spine01(conf-if-eth1/1/11)# description sfo01-Leaf01A
sfo01-Spine01(conf-if-eth1/1/11)# no shutdown
sfo01-Spine01(conf-if-eth1/1/11)# no switchport
sfo01-Spine01(conf-if-eth1/1/11)# mtu 9216
sfo01-Spine01(conf-if-eth1/1/11)# ip address 192.168.1.0/31
sfo01-Spine01(conf-if-eth1/1/11)# exit
sfo01-Spine01(config)# interface ethernet1/1/12
sfo01-Spine01(conf-if-eth1/1/12)# description sfo01-Leaf01B
sfo01-Spine01(conf-if-eth1/1/12)# no shutdown
sfo01-Spine01(conf-if-eth1/1/12)# no switchport
sfo01-Spine01(conf-if-eth1/1/12)# mtu 9216
sfo01-Spine01(conf-if-eth1/1/12)# ip address 192.168.1.2/31
sfo01-Spine01(conf-if-eth1/1/12)# exit
sfo01-Spine01(config)# interface ethernet1/1/13
sfo01-Spine01(conf-if-eth1/1/13)# description sfo01-Leaf02A
sfo01-Spine01(conf-if-eth1/1/13)# no shutdown
sfo01-Spine01(conf-if-eth1/1/13)# no switchport
sfo01-Spine01(conf-if-eth1/1/13)# mtu 9216
sfo01-Spine01(conf-if-eth1/1/13)# ip address 192.168.1.4/31
sfo01-Spine01(conf-if-eth1/1/13)# exit
sfo01-Spine01(config)# interface ethernet1/1/14
sfo01-Spine01(conf-if-eth1/1/14)# description sfo01-Leaf02B
sfo01-Spine01(conf-if-eth1/1/14)# no shutdown
sfo01-Spine01(conf-if-eth1/1/14)# no switchport
sfo01-Spine01(conf-if-eth1/1/14)# mtu 9216
sfo01-Spine01(conf-if-eth1/1/14)# ip address 192.168.1.6/31
sfo01-Spine01(conf-if-eth1/1/14)# exit
```

4. Add a route map using the following commands:

```
sfo01-Spine01(config)# ip prefix-list spine-leaf seq 10 permit 10.0.1.0/24
ge 32
sfo01-Spine01(config)# ip prefix-list spine-leaf seq 20 permit 10.2.1.0/24
ge 32
sfo01-Spine01(config)# route-map spine-leaf permit 10
```

```
sfo01-Spine01 (config-route-map) # match ip address prefix-list spine-leaf  
sfo01-Spine01 (config-route-map) # exit
```

5. Enter the following commands to configure eBGP.

```
sfo01-Spine01 (config) # router bgp 65100  
sfo01-Spine01 (config-router-bgp-65100) # bfd all-neighbors interval 200  
min_rx 200 multiplier 3 role active  
sfo01-Spine01 (config-router-bgp-65100) # router-id 10.0.1.1  
sfo01-Spine01 (config-router-bgp-65100) # address-family ipv4 unicast  
sfo01-Spine01 (config-router-bgpv4-af) # redistribute connected route-map  
spine-leaf  
sfo01-Spine01 (config-router-bgp-65100) # bestpath as-path multipath-relax  
sfo01-Spine01 (config-router-bgp-65100) # maximum-paths ebgp 2  
sfo01-Spine01 (config-router-bgpv4-af) # exit
```

6. Configure eBGP for IPv4 point-to-point peering.

```
sfo01-Spine01 (config-router-bgp-65100) # neighbor 192.168.1.1  
sfo01-Spine01 (config-router-neighbor) # remote-as 65101  
sfo01-Spine01 (config-router-neighbor) # no shutdown  
sfo01-Spine01 (config-router-neighbor) # advertisement-interval 5  
sfo01-Spine01 (config-router-neighbor) # bfd  
sfo01-Spine01 (config-router-neighbor) # fall-over  
sfo01-Spine01 (config-router-neighbor) # exit  
sfo01-Spine01 (config-router-bgp-65100) # neighbor 192.168.1.3  
sfo01-Spine01 (config-router-neighbor) # remote-as 65101  
sfo01-Spine01 (config-router-neighbor) # no shutdown  
sfo01-Spine01 (config-router-neighbor) # advertisement-interval 5  
sfo01-Spine01 (config-router-neighbor) # bfd  
sfo01-Spine01 (config-router-neighbor) # fall-over  
sfo01-Spine01 (config-router-neighbor) # exit  
sfo01-Spine01 (config-router-bgp-65100) # neighbor 192.168.1.5  
sfo01-Spine01 (config-router-neighbor) # remote-as 65102  
sfo01-Spine01 (config-router-neighbor) # no shutdown  
sfo01-Spine01 (config-router-neighbor) # advertisement-interval 5  
sfo01-Spine01 (config-router-neighbor) # bfd  
sfo01-Spine01 (config-router-neighbor) # fall-over  
sfo01-Spine01 (config-router-neighbor) # exit  
sfo01-Spine01 (config-router-bgp-65100) # neighbor 192.168.1.7  
sfo01-Spine01 (config-router-neighbor) # remote-as 65102  
sfo01-Spine01 (config-router-neighbor) # no shutdown  
sfo01-Spine01 (config-router-neighbor) # advertisement-interval 5  
sfo01-Spine01 (config-router-neighbor) # bfd  
sfo01-Spine01 (config-router-neighbor) # fall-over  
sfo01-Spine01 (config-router-neighbor) # exit  
sfo01-Spine01 (config-router-bgp-65100) # exit
```

7. Configure a loopback interface for BGP EVPN peering.

```
sfo01-Spine01 (config) # interface loopback 1  
sfo01-Spine01 (conf-if-lo-1) # description evpn_loopback  
sfo01-Spine01 (conf-if-lo-1) # no shutdown
```

```
sfo01-Spine01 (conf-if-lo-1) # ip address 10.2.1.1/32
sfo01-Spine01 (conf-if-lo-1) # exit
```

## 8. Configure BGP EVPN peering.

```
sfo01-Spine01 (config) # router bgp 65100
sfo01-Spine01 (config-router-bgp-65100) # neighbor 10.2.2.1
sfo01-Spine01 (config-router-neighbor) # remote-as 65101
sfo01-Spine01 (config-router-neighbor) # send-community extended
sfo01-Spine01 (config-router-neighbor) # update-source loopback1
sfo01-Spine01 (config-router-neighbor) # ebgp-multihop 2
sfo01-Spine01 (config-router-neighbor) # no shutdown
sfo01-Spine01 (config-router-neighbor) # address-family ipv4 unicast
sfo01-Spine01 (config-router-neighbor-af) # no activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # address-family l2vpn evpn
sfo01-Spine01 (config-router-neighbor-af) # activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # exit
sfo01-Spine01 (config-router-bgp-65100) # neighbor 10.2.2.2
sfo01-Spine01 (config-router-neighbor) # remote-as 65101
sfo01-Spine01 (config-router-neighbor) # send-community extended
sfo01-Spine01 (config-router-neighbor) # update-source loopback1
sfo01-Spine01 (config-router-neighbor) # ebgp-multihop 2
sfo01-Spine01 (config-router-neighbor) # no shutdown
sfo01-Spine01 (config-router-neighbor) # address-family ipv4 unicast
sfo01-Spine01 (config-router-neighbor-af) # no activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # address-family l2vpn evpn
sfo01-Spine01 (config-router-neighbor-af) # activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # exit
sfo01-Spine01 (config-router-bgp-65100) # neighbor 10.2.2.3
sfo01-Spine01 (config-router-neighbor) # remote-as 65102
sfo01-Spine01 (config-router-neighbor) # send-community extended
sfo01-Spine01 (config-router-neighbor) # update-source loopback1
sfo01-Spine01 (config-router-neighbor) # ebgp-multihop 2
sfo01-Spine01 (config-router-neighbor) # no shutdown
sfo01-Spine01 (config-router-neighbor) # address-family ipv4 unicast
sfo01-Spine01 (config-router-neighbor-af) # no activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # address-family l2vpn evpn
sfo01-Spine01 (config-router-neighbor-af) # activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # exit
sfo01-Spine01 (config-router-bgp-65100) # neighbor 10.2.2.4
sfo01-Spine01 (config-router-neighbor) # remote-as 65102
sfo01-Spine01 (config-router-neighbor) # send-community extended
sfo01-Spine01 (config-router-neighbor) # update-source loopback1
sfo01-Spine01 (config-router-neighbor) # ebgp-multihop 2
sfo01-Spine01 (config-router-neighbor) # no shutdown
```

```

sfo01-Spine01 (config-router-neighbor) # address-family ipv4 unicast
sfo01-Spine01 (config-router-neighbor-af) # no activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # address-family l2vpn evpn
sfo01-Spine01 (config-router-neighbor-af) # activate
sfo01-Spine01 (config-router-neighbor-af) # exit
sfo01-Spine01 (config-router-neighbor) # exit
sfo01-Spine01 (config-router-bgp-65100) # exit

```

9. Repeat the steps using the appropriate values from Section 5.8 for the remaining spine switch.

## 6.4 Verify establishment of BGP between leaf and spine switches

The leaf switches must establish a connection to the spine switches before BGP updates can be exchanged. Verify that peering is successful and BGP routing has been established.

1. Run the `show ip bgp summary` command to display information about the BGP and TCP connections to neighbors. In Figure 20, all three BGP sessions for each leaf switch are shown. The last session, 192.168.3.1, is the iBGP session between the leaf pairs if there is a leaf to spine layer failure.

```

100.67.198.32 - PuTTY
sfo01-Leaf01A# show ip bgp summary
BGP router identifier 10.0.2.1 local AS number 65101
Global BFD is enabled
Neighbor          AS           MsgRcvd    MsgSent    Up/Down        State/Pfx
192.168.1.0       65100       36488     36451     3w:1d:00:07:10 41
192.168.2.0       65100       36493     36448     3w:1d:00:07:21 38
192.168.3.1       65101       36560     36586     3w:1d:00:07:11 51
sfo01-Leaf01A#

```

Figure 20 The output of show ip bgp summary



- Run the `show ip route bgp` command to verify that all routes using BGP are being received. The command also confirms that the multiple gateway entries show the multiple routes to the BGP learned networks. Figure 21 shows two different routes to the remote loopback addresses for 10.0.2.3/32 and 10.2.2.3/32.

```

100.67.198.32 - PuTTYNG
sfo01-Leaf01A# show ip route bgp
Codes: C - connected
       S - static
       B - BGP, IN - internal BGP, EX - external BGP
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, * - candidate default,
       + - summary route, > - non-active route
Gateway of last resort is not set

  Destination                                Gateway                                     Dist/Metric   Last Change
-----
B EX 10.0.1.1/32                            via 192.168.1.0                             20/0          20:08:50
B EX 10.0.1.2/32                            via 192.168.2.0                             20/0          20:08:50
B IN 10.0.2.2/32                            via 192.168.3.1                             200/0         20:08:50
B EX 10.0.2.3/32                            via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.0.2.4/32                            via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.2.1.1/32                            via 192.168.1.0                             20/0          20:08:50
B EX 10.2.1.2/32                            via 192.168.2.0                             20/0          20:08:50
B IN 10.2.2.2/32                            via 192.168.3.1                             200/0         20:08:50
B EX 10.2.2.3/32                            via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.2.2.4/32                            via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.2.2.21/32                           via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.2.2.22/32                           via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.10.11.0/24                          via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.10.12.0/24                          via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.10.13.0/24                          via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.10.14.0/24                          via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.222.222.2/32                       via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 10.222.222.21/32                      via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 192.168.1.20/31                       via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 192.168.1.22/31                       via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 192.168.2.20/31                       via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 192.168.2.22/31                       via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0
B EX 192.168.3.20/31                       via 192.168.1.0                             20/0          20:08:50
                                           via 192.168.2.0

sfo01-Leaf01A#

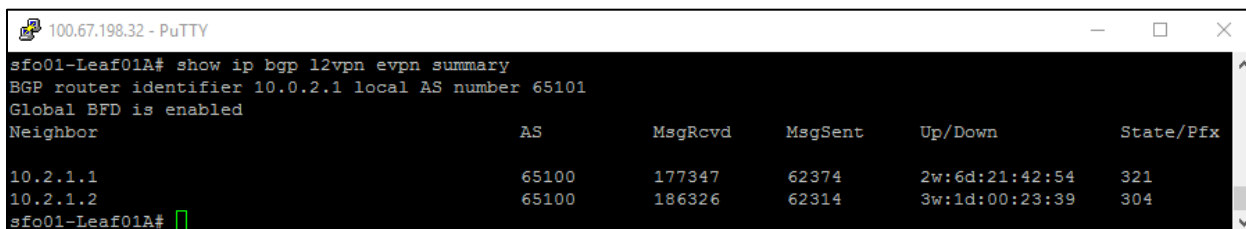
```

Figure 21 The output of `show ip route`

## 6.5 Verify BGP EVPN and VXLAN between leaf switches

For the L2 VXLAN virtual networks to communicate, each leaf must be able to establish a connection to the other leaf switches before host MAC information can be exchanged. Verify that peering is successful and BGP EVPN routing has been established.

1. Run the `show ip bgp l2vpn evpn summary` command to display information about the BGP EVPN and TCP connections to neighbors. Figure 22 shows the BGP states between leaf switch sfo01-Leaf01A and sfo01-spine01 (10.2.1.1) and sfo01-spine02 (10.2.1.2).



```
100.67.198.32 - PuTTY
sfo01-Leaf01A# show ip bgp l2vpn evpn summary
BGP router identifier 10.0.2.1 local AS number 65101
Global BFD is enabled
Neighbor                AS           MsgRcvd    MsgSent    Up/Down      State/Pfx
10.2.1.1                65100       177347    62374     2w:6d:21:42:54 321
10.2.1.2                65100       186326    62314     3w:1d:00:23:39 304
sfo01-Leaf01A#
```

Figure 22 Output of `show ip bgp l2vpn evpn neighbors`

2. Run the `show evpn evi` command to verify the current state of all configured virtual networks. Figure 23 shows the state of each virtual network as `Up` and that the Integrated Routing and Bridging (IRB) VRF is set to `tenant1`.

---

**Note:** EVIs 1611-1613 were previously configured. See Section 1.4.

---

```
100.67.198.32 - PuTTYNG
sfo01-Leaf01A# show evpn evi
EVI : 1611, State : up
  Bridge-Domain      : Virtual-Network 1611, VNI 1611
  Route-Distinguisher : 1:10.222.222.1:1611
  Route-Targets      : 0:1611:1611 both
  Inclusive Multicast : 10.222.222.2, 10.222.222.21
  IRB                : Enabled(tenant1)

EVI : 1612, State : up
  Bridge-Domain      : Virtual-Network 1612, VNI 1612
  Route-Distinguisher : 1:10.222.222.1:1612
  Route-Targets      : 0:1612:1612 both
  Inclusive Multicast : 10.222.222.2
  IRB                : Enabled(tenant1)

EVI : 1613, State : up
  Bridge-Domain      : Virtual-Network 1613, VNI 1613
  Route-Distinguisher : 1:10.222.222.1:1613
  Route-Targets      : 0:1613:1613 both
  Inclusive Multicast : 10.222.222.2
  IRB                : Enabled(tenant1)

EVI : 1641, State : up
  Bridge-Domain      : Virtual-Network 1641, VNI 1641
  Route-Distinguisher : 1:10.222.222.1:1641
  Route-Targets      : 0:1641:1641 both
  Inclusive Multicast : 10.222.222.2, 10.222.222.21
  IRB                : Enabled(tenant1)

EVI : 1642, State : up
  Bridge-Domain      : Virtual-Network 1642, VNI 1642
  Route-Distinguisher : 1:10.222.222.1:1642
  Route-Targets      : 0:1642:1642 both
  Inclusive Multicast : 10.222.222.2
  IRB                : Enabled(tenant1)

EVI : 1643, State : up
  Bridge-Domain      : Virtual-Network 1643, VNI 1643
  Route-Distinguisher : 1:10.222.222.1:1643
  Route-Targets      : 0:1643:1643 both
  Inclusive Multicast : 10.222.222.2
  IRB                : Enabled(tenant1)

EVI : 3939, State : up
  Bridge-Domain      : Virtual-Network 3939, VNI 3939
  Route-Distinguisher : 1:10.222.222.1:3939
  Route-Targets      : 0:3939:3939 both
  Inclusive Multicast : 10.222.222.2
  IRB                : Enabled(tenant1)

sfo01-Leaf01A#
```

Figure 23 The output of show evpn evi

---

**Note:** For more validation and troubleshooting commands, see the OS10 Enterprise Edition User Guide.

---

## 7 Create a VxRail Virtual Infrastructure workload domain

This chapter provides guidance on creating a VxRail Virtual Infrastructure (VI) workload domain before adding a cluster. Deploy the vCenter server and make the domain ready for the cluster addition.

---

**Note:** You can only perform one workload domain operation at a time. For example, when creating a workload domain, you cannot add a cluster to any other workload domain.

---

1. On the SDDC Manager Dashboard, click **+ Workload Domain** and then select VI-VxRail Virtual Infrastructure Setup.
2. Type a name for the VI workload domain, such as **W02**. The name must contain between 3 and 20 characters.
3. Type a name for the organization that will use the virtual infrastructure, such as **Dell**. The name must contain between 3 and 20 characters.
4. Click **Next**.
5. On the **Computer** page, enter the vCenter IP address, **172.16.11.67**, and DNS name, **sfo01w02vc01.sfo01.rainpole.local**

---

**Note:** Before updating the IP address in the wizard, ensure that you have reserved the IP addresses in DNS.

---

6. **Type 255.255.255.0** and **172.16.11.253** the vCenter subnet mask and default gateway.
7. Type and retype the vCenter root password.
8. Click **Next**.
9. At the Review step of the wizard, shown in Figure 24, scroll down the page to review the information.
10. Click **Finish** to start the creation process.

Review	
Name	
VI Workload Domain Name	W02
Organization Name	Dell
Compute	
vCenter Root Password	*****
vCenter IP Address	172.16.11.67
vCenter DNS Name	sfo01w02vc01.sfo01.rainpole.local
vCenter Subnet Mask	255.255.255.0
vCenter Default Gateway	172.16.11.253

Figure 24 VxRail VI Configuration review

The **Workload Domains** page displays with a notification that the VI workload domain is being added.

11. Click **View Task Status** to view the domain creation tasks and sub-tasks. The status is active until the primary cluster is added to the domain.

## 7.1 Create a local user in the workload domain vCenter Server

Before adding the VxRail cluster, image the workload domain nodes. Once complete, perform the VxRail first run of the workload domain nodes using the external vCenter server.

Create a local user in the vCenter Server as this is an external server that the VMware Cloud Foundation deploys. This is required for the first run of VxRail.

1. Log in to the workload domain vCenter Server Appliance through VMware vSphere Web Client.
2. Select **Menu > Administration > Single Sign-On**
3. Click **Users and Groups**.
4. Click **Users**.
5. Select domain **vSphere.local**.
6. Click **Add User**.
7. In the **Add User** pop-up window, enter the values for the mandatory fields.
8. Enter *vxadmin* as the **Username** and **Password**. Confirm the **Password**.
9. Click **Add**.
10. Wait for the task to complete.

## 7.2 VxRail initialization

This section outlines the general steps that are needed to initialize a VxRail cluster.

1. Install the VxRail nodes into the two racks in the data center.
2. Attach the appropriate cabling between the ports of the VxRail nodes and the switch ports.
3. Power on the four primary E-series nodes in both racks to form the initial VxRail cluster.
4. To access the VxRail ESXi management on VLAN 1641, connect a workstation or laptop that is configured for VxRail.
5. Using a web browser, go to the default VxRail IP address, 192.168.10.200, to begin the VxRail initialization process.
6. Complete the steps provided within the initialization wizard.

Using the values provided, VxRail performs the verification process. Once the validation is complete, the initialization process builds a new VxRail cluster. The building progress of the cluster displays in the status window provided. When the **Hooray!** message displays, the VxRail initialization is complete, and the new VxRail cluster is built.

## 7.3 VxRail deployment values

Table 7 lists the values that are used during the VxRail Manager initialization and expansion operation.

**Note:** The values are listed in order as they are entered in the GUI.

Table 7 VxRail network configuration values

	Parameter	Value
Appliance Settings	NTP server	172.16.11.5
	Domain	sfo01.rainpole.local
ESXi hostname and IP addresses	ESXi hostname prefix	sfo01w02vxrail
	Separator	none
	Iterator	Num 0x
	Offset	1
	Suffix	none
	ESXi beginning address	172.16.41.101
	ESXi ending address	172.16.41.104
External vCenter Server	vCenter Server FQDN	sfo01w02vc01.sfo01.rainpole.local
	admin username/password	administrator@vsphere.local
	vCenter Server SSO domain	vsphere.local
	PSC FQDN	sfo01w01psc01.sfo01.rainpole.local
	Data center name	VxRail-DataCenter
	Cluster name	VxRail-Cluster
VxRail Manager	VxRail Manager hostname	sfo01w02vxrail-mgr
	VxRail IP address	172.16.41.72
	Subnet mask	255.255.255.0
	Gateway	172.16.41.253
vMotion	Starting address for IP pool	172.16.42.101
	Ending address for IP pool	172.16.42.104
	Subnet mask	255.255.255.0
	VLAN ID	1642
vSAN	Starting address for IP pool	172.16.43.101
	Ending address for IP pool	172.16.43.104
	Subnet mask	255.255.255.0
	VLAN ID	1643
Solutions Logging	vRealize Log Insight hostname	-

Table 7 lists the values that are used during the VxRail Manager initialization and expansion operation

## 7.4 Add the primary VxRail cluster to the workload domain

1. On the **SDDC Manager** Dashboard, click **Inventory > Workload Domains**. The **Workload Domains** page displays information for all workload domains.
2. In the workload domains table, use your cursor to hover over the workload domain in the activating state. A set of three dots displays on the left of the workload domain name.

3. Click the three dots, then click **Add VxRail Cluster**. The **Add VxRail cluster to Workload Domain** page displays.
4. On the **VxRail Manager** page, a single VxRail cluster is discovered. Select the **VxRail-Cluster** object and click **Next**.
5. The **Discovered Host** page displays a list of the discovered hosts for that cluster. Update the SSH password for the discovered hosts for that cluster and then click **Next**. The Networking page displays the networking details for the cluster.
  - a. Choose **NSX-T** for the NSX Platform
  - b. Enter the VLAN ID **2500** for the overlay network.
  - c. Enter the following NSX manager details, see Figure 25:
    - i. Enter **172.16.11.81** for the **Cluster Virtual IP address**. Enter **sfo01wnsx01.sfo01.rainpole.local** for the **FQDN**.
    - ii. Enter **172.16.11.82**, **172.16.11.83**, and **172.16.11.84** for the three IP addresses and the corresponding FQDNs **sfo01w01a**, **sfo01w01b**, and **sfo01w01c** (prefix **sfo01.rainpole.local**).
    - iii. For the subnet mask, enter **255.255.255.0**
    - iv. Enter the IP address, **172.16.11.253**, for the default gateway.
    - v. Enter the admin password for NSX-T.
  - d. Click **Next**.

Networking

NSX Platform  NSX-V  NSX-T

Ensure NSX-T is downloaded. Refer to product documentation for more information.

**Overlay Networking**

VLAN ID 2500

**NSX-T Manager**

NSX-T Manager Cluster IP 172.16.11.81

NSX-T Manager Cluster FQDN sfo01wnsx01.sfo01.rainpole.local

NSX-T Manager 1 IP 172.16.11.82

NSX-T Manager 1 FQDN sfo01wnsx01a.sfo01.rainpole.local

NSX-T Manager 2 IP 172.16.11.83

NSX-T Manager 2 FQDN sfo01wnsx01b.sfo01.rainpole.local

NSX-T Manager 3 IP 172.16.11.84

NSX-T Manager 3 FQDN sfo01wnsx01.sfo01.rainpole.local

CANCEL BACK NEXT

Figure 25 VI-WLD NSX-T configuration details

**Note:** The configuration of IP network 172.16.11.0/24 is covered in the creation of the management workload domain. See, VCF on VxRail multitrack deploying using BGP EVPN.

## 8 Configure NSX-T north-south connectivity

The necessary components to facilitate NSX-T overlay networking are automatically created by VCF on VxRail including VLAN and overlay transport zones as well as Uplink, NIOC, and Transport Node profiles.

This chapter provides the steps that are required to establish north-south connectivity from the NSX-T overlay network to the leaf switches. This includes the deployment of NSX-T Edge appliances and the related NSX-T objects including Edge Cluster profiles, assigning Edge Transport Node roles, and creating an Edge Cluster.

Using the [VMware Validated Design 5.0.1 Deployment of VMware NSX-T for Workload Domains](#), the following sections were completed:

- [Create Transport Zones](#)
- [Create Uplink Profiles and the Network I/O Control Profile](#)
- [Create the NSX-T Segments for System, Uplink, and Overlay Traffic](#)
- [Create an NSX-T Edge Cluster Profile](#)
- [Deploy the NSX-T Edge Appliances](#)
- [Join the NSX-T Edge Nodes to the Management Plane](#)
- [Create an Anti-Affinity Rule for the NSX-T Edge Nodes in the VI-WLD cluster](#)
- [Add the NSX-T Edge Anodes to the Transport Zones](#)
- [Create an NSX-T Edge Cluster](#)
- [Create and Configure the Tier-0 Gateway](#)
- [Create and Configure the Tier-1 Gateway](#)

The rest of this chapter contains information to use to augment the previous steps, which include the tables with the specific values used in this guide. If a value is not specified in the following example, then the exact steps and values from the VVD are used.

### 8.1 Create transport zones

Table 8 shows the values that are used when creating two extra transport zones for uplinks. VCF on VxRail automatically created a transport zone that is associated with the transport zone profile called `vlan-tz-<UID>` as well as the `overlay-tz -<UID>`, where UID is a system-generated identifier. Two uplinks transport zones need to be defined. Each associated a separate NSX-T Virtual Distributed Switch (N-VDS). Each N-VDS is configured to carry VLAN traffic. See [Create transport zones](#) for step by step instructions.

Table 8 Edge transport zones

Name	N-VDS name	N-VDS mode	Traffic type
sfo01-w-uplink01	sfo01-w-uplink01	Standard	VLAN
sfo01-w-uplink02	sfo01-w-uplink02	Standard	VLAN



## 8.2 Create uplink profiles and the network I/O control profile

Table 9 shows the values that are used for the corresponding uplink profiles that the uplink transport zones use.

Table 9 Uplink profiles

Name	Teaming – teaming policy	Teaming – active uplink	Transport VLAN	MTU
esxi-w02-uplink-profile	Load Balance Source	uplink-1, uplink-2	2500	9000
sfo01-w-uplink01-profile	Failover Order	uplink-1	1647	9000
sfo01-w-uplink02-profile	Failover Order	uplink-2	1648	9000

## 8.3 Create the NSX-T segments for system, uplink, and overlay traffic

Table 10 shows the values that are used that are for the required uplink segments. The system and overlay segments were automatically created by VCF on VxRail.

Table 10 Uplink profiles

Segment name	Uplink and type	Transport zone	VLAN
sfo01-w-nvds01-uplink01	Isolated – No logical connections	vlan-tz-<GUID>	0-4094
sfo01-w-nvds01-uplink02	Isolated – No logical connections	vlan-tz-<GUID>	0-4094
sfo01-w-uplink01	Isolated – No logical connections	sfo01-w-uplink01	1647
sfo01-w-uplink02	Isolated – No logical connections	sfo01-w-uplink02	1648

## 8.4 Create an NSX-T edge cluster profile

Table 11 shows the values that are used for the required edge cluster profile.

Table 11 Edge cluster profile settings

Setting	Value
Name	sfo01-w-edge-cluster01-profile
BFD Probe	1000
BFD Allowed Hops	255
BFD Declare Dead Multiple	3

## 8.5 Deploy the NSX-T edge appliances

To provide tenant workloads with routing services and connectivity to networks that are external to the workload domain, deploy two NSX-T Edge nodes. Table 12 shows the values that are used for the edge nodes.

Table 12 Uplink profiles

Setting	Value for sfo01wesg01	Value for sfo01esg02
Hostname	sfo01wesg01.sfo01.rainpole.local	sfo01wesg02.sfo01.rainpole.local
Port Groups	sfo01-w-nvds01-management	sfo01-w-nvds01-management
Primary IP Address	172.16.41.21	172.16.41.22

Table 13 shows the networks that the four uplinks on each ESG are attached to both ESG use the same values.

Table 13 Edge cluster profile settings

Source network	Destination network
Network 3	sfo01-w-nvds01-uplink02
Network 2	sfo01-w-nvds01-uplink01
Network 1	sfo01-w-overlay
Network 0	sfo01-w-nvds01-management

## 8.6 Join the NSX-T edge nodes to the management plane

Table 14 shows the values that are used to connect to the edge nodes to the management plane.

Table 14 Uplink profiles

Setting	Value for sfo01wesg01	Value for sfo01esg02
Name	sfo01wesg01	sfo01wesg02
Port Groups	sfo01-w-nvds01-management	sfo01-w-nvds01-management
Primary IP Address	172.16.41.21	172.16.41.22

## 8.7 Create anti-affinity rules for NSX-T edge nodes

In this environment, the underlying VxRail hosts are spread out among numerous racks in the data center. In a simple example, all north-south peering can be established in a single rack, an edge rack. A VM-Host affinity rule is created to ensure that the ESG nodes are always running on VxRail nodes in that designated rack, for example, Rack 1.

Table 15, along with the following steps, is used to create two rules. The first rule designates the hosts that the edge nodes can use. The second rule designates the ESG nodes themselves.

1. Browse to the cluster in the vSphere Client.
2. Click the **Configure** tab, click **VM/Host Groups**.

3. Click **Add**.
4. In the **Create VM/Host Rules** dialog box, type a name for the rule.
5. From the **Type** drop-down menu, select the appropriate type.
6. Click **Add** and in the **Add Group Member** window select either VxRail nodes or Edge nodes to which the rule applies and click **OK**.
7. Click **OK**.
8. Repeat the steps in this section for the remaining rule.

Table 15 VM/Host groups

VM/Host group name	Type	Members
NSX-T Edge Hosts	Host Group	sfo01w02vxrail01 and sfo01w02vxrail03
NSX-T Edge Nodes	VM Group	sfo01wesg01 and sfo01wesg02

Once the group rules are in place, create a VM/Host rule to bind the VM group Edge nodes to the Host group.

1. Browse to the cluster in the vSphere Client.
2. Click the **Configure** tab, click **VM/Host Rules**.
3. Click **Add**.
4. In the **Create VM/Host Rules** dialog box, type `host-group-rule-edgeCluster`
5. From the **Type** drop-down menu, select **Virtual Machines to Hosts**.
6. From the VM Group drop-down, select **Edge Service Gateways** and choose **Must run on hosts in group**.
7. From the **Host Group** drop-down, select **Edge Hosts**.
8. Click **OK**.

## 8.8 Add the NSX-T edge nodes to the transport zones

After you deploy the NSX-T edge nodes and join them to the management plane, to connect the nodes to the workload domain. Next, add the nodes to the transport zones for uplink and overlay traffic and configure the N-VDS on each edge node. Table 16 shows the values that are used for both edge nodes.

Table 16 Edge node transport zone settings

Setting	Value for sfo01wesg01	Value for sfo01esg02
Transport Zones	sfo01-w-uplink01(VLAN)	sfo01-w-uplink01(VLAN)
	sfo01-w-uplink02(VLAN)	sfo01-w-uplink02(VLAN)
	sfo01-w-overlay(Overlay)	sfo01-w-overlay(Overlay)

Table 17 shows the Edge node N-VDS settings.

Table 17 Edge node N-VDS settings

Setting	Value for sfo01wesg01	Value for sfo01esg02
Edge Switch Name	sfo01-w-nvds01	sfo01-w-nvds01
Uplink Profile	sfo01-w-overlay-profile	sfo01-w-overlay-profile
IP Assignment	Use Static IP List	Use Static IP List
Static IP List	172.16.49.21	172.16.49.22
Gateway	172.16.49.253	172.16.49.253
Subnet Mask	255.255.255.0	255.255.255.0
Virtual NICs	fp-eth0 uplink-1	fp-eth0 uplink-1

## 8.9 Create and configure the Tier-0 gateway

The Tier-0 gateway in the NSX-T Edge cluster provides a gateway service between the logical and physical network. The NSX-T Edge cluster can back multiple Tier-0 gateways. In this example, each edge node hosts a Tier-0 gateway, and ECMP is used to create multiple paths to the two leaf switches in the rack. See [Create and configure the Tier-0 gateway](#) for step-by-step instructions. In this example, no changes were made to the values found in the VVD.

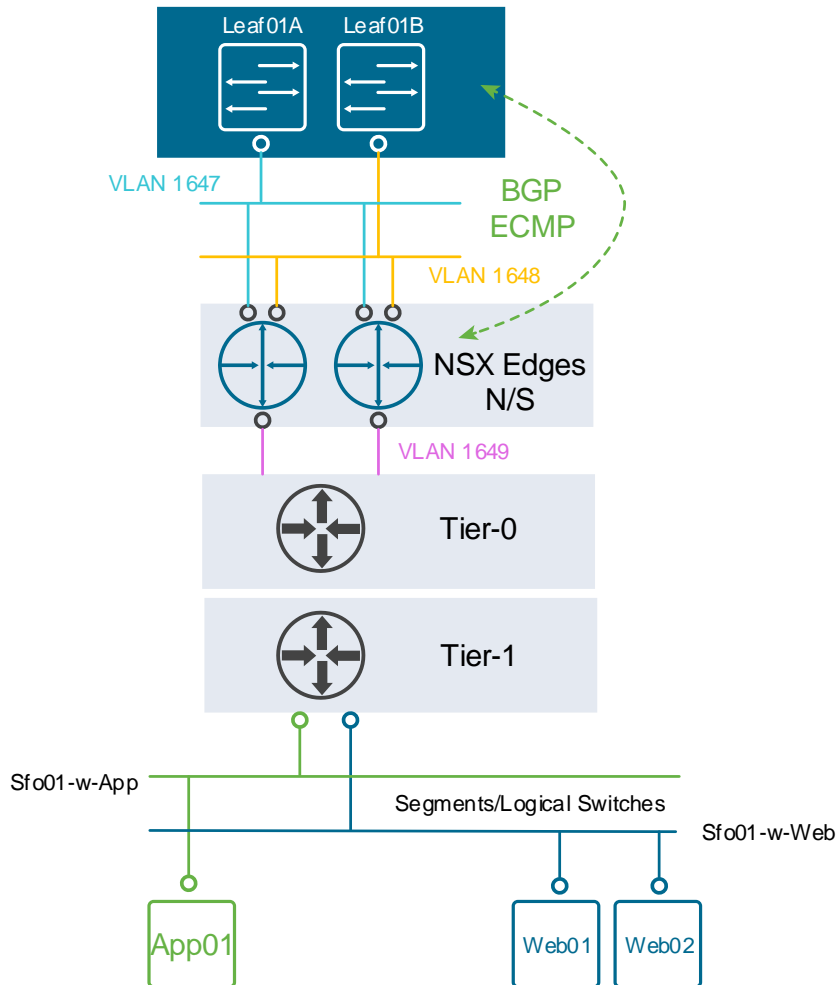


Figure 26 NSX-T Edge Cluster

---

**Note:** For step-by-step CLI configuration of the leaf switches, see Section 6.2.

---

## 8.10 Create and configure the Tier-1 gateway

Create and configure the Tier-1 gateway to redistribute routes to the Tier-0 gateway and provide routing between tenant workloads. Tier-1 gateways have downlink ports to connect to NSX-T segments and uplink ports to connect to NSX-T Tier-0 gateways.

---

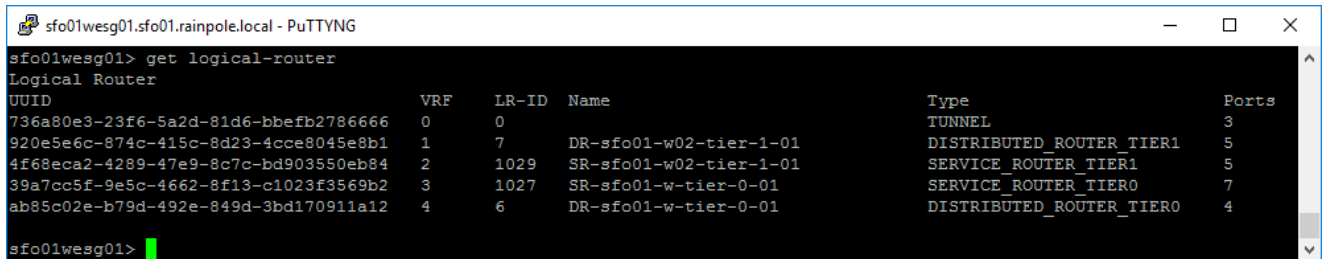
**Note:** Refer to Create and Configure the Tier-1 Gateway.

---

## 8.11 Verify BGP peering and route redistribution

The Tier-0 gateway must establish a connection to each of the upstream Layer 3 devices before BGP updates can be exchanged. Verify that the NSX-T Edge nodes are successfully peering and that BGP routing is established.

1. Open an SSH connection to sfo01wesg01.
2. Log in using the previously defined credentials.
3. Use the `get logical-router` command to get information about the Tier-0 and Tier-1 service routes and distributed routers. Figure 27 shows the logical routers and the corresponding VRF values for each.



```
sfo01wesg01.sfo01.rainpole.local - PuTTYNG
sfo01wesg01> get logical-router
Logical Router
UUID                VRF  LR-ID  Name                                     Type                Ports
736a80e3-23f6-5a2d-81d6-bbefb2786666  0    0      DR-sfo01-w02-tier-1-01                 TUNNEL              3
920e5e6c-874c-415c-8d23-4cce8045e8b1  1    7      DR-sfo01-w02-tier-1-01                 DISTRIBUTED_ROUTER_TIER1  5
4f68eca2-4289-47e9-8c7c-bd903550eb84  2    1029   SR-sfo01-w02-tier-1-01                 SERVICE_ROUTER_TIER1    5
39a7cc5f-9e5c-4662-8f13-cl023f3569b2  3    1027   SR-sfo01-w-tier-0-01                   SERVICE_ROUTER_TIER0    7
ab85c02e-b79d-492e-849d-3bd170911a12  4    6      DR-sfo01-w-tier-0-01                   DISTRIBUTED_ROUTER_TIER0  4
sfo01wesg01>
```

Figure 27 The output of `get logical-router`

4. Using the `vrf <VRF value>` for **SERVICE\_ROUTER\_TIER0**, connect to the service router for Tier 0. In this example, command `vrf 3` is issued.

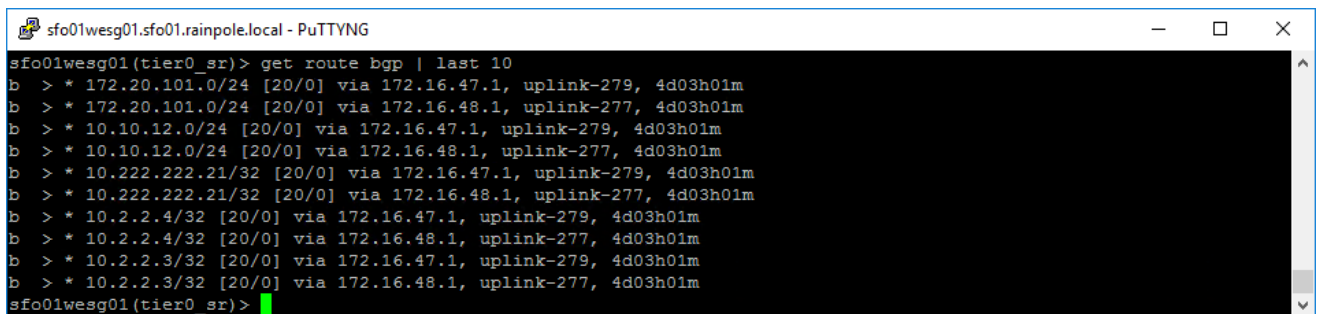
---

**Note:** The prompt changes to `hostname(tier0_sr)>`. All commands are associated with this object.

---

5. Use the `get bgp neighbor` command to verify the BGP connections to the neighbors of the service router for Tier 0.

Use the `get route bgp` command to verify that you are receiving routes by using BGP and that multiple routes to BGP-learned networks exist. Figure 28 shows the truncated output of `get route bgp`. The networks that are shown are BGP routes (represented by the lowercase `b`) and are available using 172.16.47.1 (sfo01-leaf1a) and 172.16.48.1 (sfo01-leaf1b). This verifies successful BGP: Peering and ECMP enablement for all external networks.



```
sfo01wesg01.sfo01.rainpole.local - PuTTYNG
sfo01wesg01(tier0_sr)> get route bgp | last 10
b > * 172.20.101.0/24 [20/0] via 172.16.47.1, uplink-279, 4d03h01m
b > * 172.20.101.0/24 [20/0] via 172.16.48.1, uplink-277, 4d03h01m
b > * 10.10.12.0/24 [20/0] via 172.16.47.1, uplink-279, 4d03h01m
b > * 10.10.12.0/24 [20/0] via 172.16.48.1, uplink-277, 4d03h01m
b > * 10.222.222.21/32 [20/0] via 172.16.47.1, uplink-279, 4d03h01m
b > * 10.222.222.21/32 [20/0] via 172.16.48.1, uplink-277, 4d03h01m
b > * 10.2.2.4/32 [20/0] via 172.16.47.1, uplink-279, 4d03h01m
b > * 10.2.2.4/32 [20/0] via 172.16.48.1, uplink-277, 4d03h01m
b > * 10.2.2.3/32 [20/0] via 172.16.47.1, uplink-279, 4d03h01m
b > * 10.2.2.3/32 [20/0] via 172.16.48.1, uplink-277, 4d03h01m
sfo01wesg01(tier0_sr)>
```

Figure 28 The output of `get route bgp`

6. Repeat the procedure on sfo01wesg02.

## 9 Validate connectivity between virtual machines

This chapter covers a quick validation of the entire solution. A combination of ping and traceflow are used between a combination of three virtual machines and a loopback interface. Two of the VMs are associated with one segment, Web, and the other VM is associated with the segment, App. The loopback interface represents all external networks.

Figure 29 shows the three virtual machines, Web01, Web02, and App01, are running on two separate hosts. Web01 and Web02 are associated with the Web segment, and App01 is configured in the App segment. Web01 is in rack 1, and the other two virtual machines are in rack 2. In this test, only two of the four VxRail nodes are used. The router ID for Spine02, 10.0.1.2/32, is used to test connectivity from the NSX-T overlay to the switch underlay.

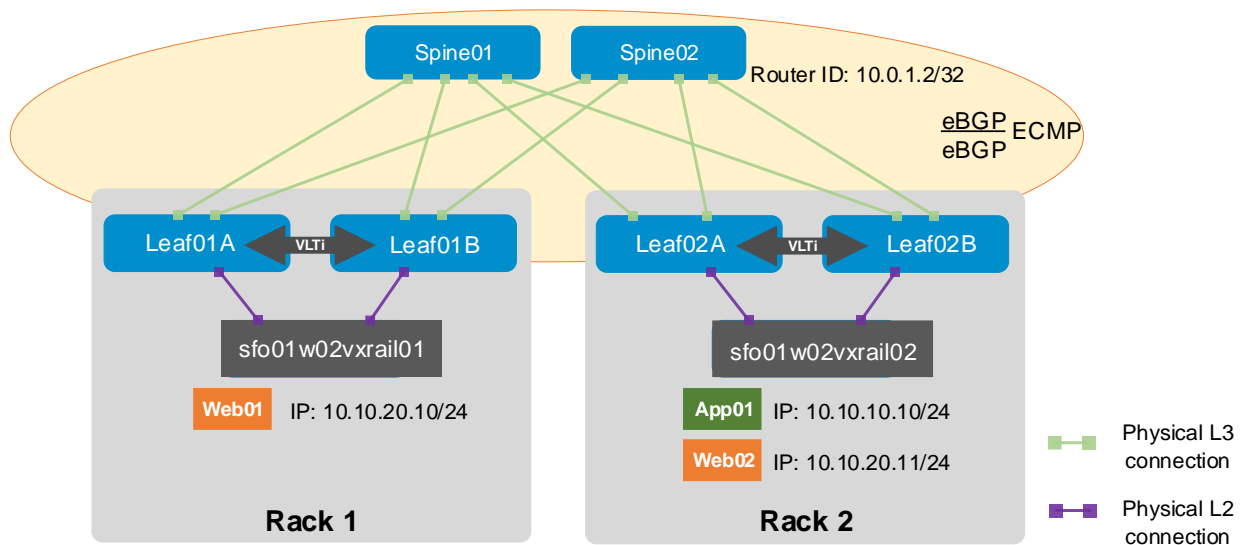


Figure 29 Validation topology

The tests that are performed are:

- Ping from Web01 to Web02
- Ping from Web01 to App01
- Ping from Web01 to Router ID on Spine02
- Ping from App01 to Router ID on Spine02
- Traceflow from App01 to Router ID on Spine02

---

**Note:** The steps required to create the two segments are beyond the scope of this document. See the NSX-T Administration Guide: Segments.

---

## 9.1 Ping from Web01 to Web02

Figure 30 shows the results of the ping that is issued from Web01 (10.10.20.10/24) to Web02 (10.10.20.11/24)

```
administrator@Web-01:~$ ping 10.10.20.11
PING 10.10.20.11 (10.10.20.11) 56(84) bytes of data.
64 bytes from 10.10.20.11: icmp_seq=1 ttl=64 time=0.303 ms
64 bytes from 10.10.20.11: icmp_seq=2 ttl=64 time=0.260 ms
64 bytes from 10.10.20.11: icmp_seq=3 ttl=64 time=0.273 ms
64 bytes from 10.10.20.11: icmp_seq=4 ttl=64 time=0.261 ms
^C
--- 10.10.20.11 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 78ms
rtt min/avg/max/mdev = 0.260/0.274/0.303/0.020 ms
administrator@Web-01:~$
```

Figure 30 Ping results from Web01 to Web02

## 9.2 Ping from Web01 to App01

Figure 31 shows the results of the ping that is issued from Web01 (10.10.20.10/24) to App01 (10.10.10.10/24).

```
administrator@Web-01:~$ ping 10.10.10.10
PING 10.10.10.10 (10.10.10.10) 56(84) bytes of data.
64 bytes from 10.10.10.10: icmp_seq=1 ttl=63 time=3.73 ms
64 bytes from 10.10.10.10: icmp_seq=2 ttl=63 time=0.356 ms
64 bytes from 10.10.10.10: icmp_seq=3 ttl=63 time=0.331 ms
64 bytes from 10.10.10.10: icmp_seq=4 ttl=63 time=0.301 ms
^C
--- 10.10.10.10 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 32ms
rtt min/avg/max/mdev = 0.301/1.180/3.732/1.473 ms
administrator@Web-01:~$
```

Figure 31 Ping results from Web01 to App01



## 9.3 Ping from Web01 to 10.0.1.2

Figure 32 shows the results of the ping that is issued from Web01 (10.10.20.10/24) to the destination address 10.0.1.2, the loopback address on sfo01-spine02.

```
administrator@Web-01:~$ ping 10.0.1.2
PING 10.0.1.2 (10.0.1.2) 56(84) bytes of data.
64 bytes from 10.0.1.2: icmp_seq=1 ttl=61 time=4.22 ms
64 bytes from 10.0.1.2: icmp_seq=2 ttl=61 time=0.774 ms
64 bytes from 10.0.1.2: icmp_seq=3 ttl=61 time=0.741 ms
64 bytes from 10.0.1.2: icmp_seq=4 ttl=61 time=0.700 ms
^C
--- 10.0.1.2 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 29ms
rtt min/avg/max/mdev = 0.700/1.608/4.219/1.507 ms
administrator@Web-01:~$
```

Figure 32 Ping results from Web01 to 10.0.1.2

---

**Note:** If connectivity is needed to the underlay network from the workload tenants, additional IP prefix list will need to be added to the leaf switches. Refer to Section 6.2.

---

## 9.4 Ping from App01 to 10.0.1.2

Figure 33 shows the results of the ping that is issued from App01 (10.10.10.10/24) to the destination address 10.0.1.2, the loopback address on sfo01-spine02.

```
administrator@App-01:~$ ping 10.0.1.2
PING 10.0.1.2 (10.0.1.2) 56(84) bytes of data.
64 bytes from 10.0.1.2: icmp_seq=1 ttl=61 time=647 ms
64 bytes from 10.0.1.2: icmp_seq=2 ttl=61 time=0.644 ms
64 bytes from 10.0.1.2: icmp_seq=3 ttl=61 time=0.724 ms
64 bytes from 10.0.1.2: icmp_seq=4 ttl=61 time=0.675 ms
^C
--- 10.0.1.2 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 37ms
rtt min/avg/max/mdev = 0.644/162.308/647.189/279.946 ms
administrator@App-01:~$
```

Figure 33 Ping results from App01 to 10.0.1.2

## 9.5 Traceflow App01 to 10.0.1.2

Figure 34 shows the results of the traceflow tool from VM App01 (10.10.10.10) to the destination address 10.0.1.2, the loopback address on sfo01-spine02.

Trace Results			
Show: ALL 1 DELIVERED 0 DROPPED			
Physical Hop Count	Observation Type	Transport Node	Component
0	Injected	sfo01w02vxrail02.sf...	Network adapter
0	Received	sfo01w02vxrail02.sf...	Distributed Firewall
0	Forwarded	sfo01w02vxrail02.sf...	Distributed Firewall (Rule ID: 2)
0	Forwarded	sfo01w02vxrail02.sf...	App
0	Received	sfo01w02vxrail02.sf...	w-tier-1-01
0	Forwarded	sfo01w02vxrail02.sf...	w-tier-1-01
0	Received	sfo01w02vxrail02.sf...	transit-bp-71a1c986-1964-4ac7-94...
0	Forwarded	sfo01w02vxrail02.sf...	Remote IP : 172.16.49.21
1	Received	sfo01wesg01.sfo01.r...	Remote IP : 172.25.102.104
1	Received	sfo01wesg01.sfo01.r...	Edge Tunnel
1	Forwarded	sfo01wesg01.sfo01.r...	transit-bp-71a1c986-1964-4ac7-94...
1	Received	sfo01wesg01.sfo01.r...	w-tier-1-01
1	Received	sfo01wesg01.sfo01.r...	Distributed Firewall
1	Forwarded	sfo01wesg01.sfo01.r...	Distributed Firewall (Rule ID: 1025)
1	Forwarded	sfo01wesg01.sfo01.r...	w-tier-1-01
1	Received	sfo01wesg01.sfo01.r...	transit-rl-71a1c986-1964-4ac7-94b...
1	Forwarded	sfo01wesg01.sfo01.r...	transit-rl-71a1c986-1964-4ac7-94b...
1	Received	sfo01wesg01.sfo01.r...	w-tier-0-01
1	Received	sfo01wesg01.sfo01.r...	Distributed Firewall
1	Forwarded	sfo01wesg01.sfo01.r...	Distributed Firewall (Rule ID: 1024)
1	Delivered	sfo01wesg01.sfo01.r...	sfo01wesg01-uplink01

Figure 34 Traceflow results from App01 to 10.0.1.2/32

**Note:** See NSX-T Administration Guide: Traceflow for more information.

## A Validated components

### A.1 Dell EMC PowerSwitch models

Table 18 Switches and operating system versions

Qty	Item	Version
4	Dell EMC PowerSwitch S5248F-ON leaf switches	10.4.3.5
2	Dell EMC PowerSwitch Z9264F-ON spine switches	10.4.3.5
2	Dell EMC PowerSwitch S3048-ON OOB mgmt switches	10.4.3.5

### A.2 VxRail E560 nodes

A cluster of four VxRail E560 nodes was used to validate the VI-WLD in this guide. The nodes were each configured using the information that is provided in Table 19. The values in Table 19 can also be found in the [VxRail Support Matrix](#).

Table 19 VxRail node components

Qty per node	Item	Firmware version
2	Intel Xeon Gold 6136 CPU @ 3.00GHz, 12 cores	-
12	16GB DDR4 DIMMs (192GB total)	-
2	800GB SAS SSD	-
8	1.2TB SAS HDD	EF03
1	Dell HBA330 Storage Controller	15.17.09.06
1	Boot Optimized Storage Solution (BOSS) Controller w/ 2x240GB SATA SSDs	2.5.13.3016
1	Broadcom 57414 NDC – 2x25GbE SFP28 ports	21.40.22.21
1	Broadcom 57414 NIC – 2x25GbE SFP29 ports	21.40.22.21
-	BIOS	2.1.8
-	iDRAC with Lifecycle Controller	3.32.32.32

---

**Note:** For information on the management world load components see the deployment guide, VCF on VxRail multirack deploying using BGP EVPN.

---

## A.3 Appliance software

This Deployment Guide uses the VxRail appliance software 4.7.211 for development. The software consists of the component versions that are provided in Table 20.

Table 20 VxRail appliance software component versions

Item	Version
VxRail Manager	4.7.211.13893929
ESRS	3.28.0006
Log Insight	4.6.0.8080673
VMware vCenter	6.7 U2a 13643870
VMware ESXi	6.7 EP09 13644319
Platform Service	4.7.211
NSX-T	2.4.1.0.0-13716575

## B Technical resources

### B.1 VxRail, VCF, and VVD Guides

[VMware Cloud Foundation on VxRail Planning and Preparation Guide](#)

[VMware Cloud Foundation on VxRail Architecture Guide](#)

[VMware Cloud Foundation on VxRail Administrator Guide](#)

[VMware Cloud Foundation on VxRail Technical FAQ](#)

[Dell EMC VxRail Network Guide](#)

[VMware Validated Design 5.0.1 NSX-T Workload Domains](#)

[NSX-T Data Center Administration Guide](#)

[VxRail Support Matrix](#)

### B.2 Dell EMC Networking Guides

[Dell EMC PowerSwitch Guides](#)

[OS10 Enterprise Edition Users Guide 10.4.3.0 Edition](#)

[Dell EMC Networking Layer 3 Leaf-Spine Deployment and Best Practices with OS10EE](#)

[Dell EMC Networking Virtualization Overlay with BGP EVPN](#)

[Dell EMC OS10EE BGP EVPN Configuration Cheat Sheet](#)

[Dell EMC VxRail Multirack Deployment Guide](#)

## C Fabric Design Center

The Dell EMC Fabric Design Center (FDC) is a cloud-based application that automates the planning, design, and deployment of network fabrics that power Dell EMC compute, storage, and HCI solutions. The FDC is ideal for turnkey solutions and automation based on validated deployment guides.

FDC allows for design customization and flexibility to go beyond validated deployment guides. For additional information, go to the [Dell EMC Fabric Design Center](#).

## D 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**

Dell EMC encourages readers to provide feedback on the quality and usefulness of this publication by sending an email to [Dell\\_Networking\\_Solutions@Dell.com](mailto:Dell_Networking_Solutions@Dell.com).