Leaf-Spine Deployment and Best Practices Guide

Deploying leaf-spine networks in the data center with Dell EMC Networking and Cisco Nexus multilayer switches

Dell EMC Networking Solutions Engineering
July 2017

Revisions

Date	Rev.	Description	Authors
July 2017	1.0	Initial release	Jim Slaughter, Shree Rathinasamy

Table of contents

Revisions 2

1 Introduction 6

1.1 Typographical Conventions 7

2 Hardware overview 8

2.1 Dell EMC Networking S3048-ON 8

2.2 Dell EMC Networking S4048-ON 8

2.3 Dell EMC Networking S6010-ON 8

2.4 Dell EMC Networking Z9100-ON 9

2.5 Cisco Nexus 5600 and 7000 series switches 9

3 Leaf-spine overview 10

3.1 Layer 3 leaf-spine topology 10

3.2 Layer 2 leaf-spine topology 11

3.3 Design considerations 11

4 Management network 12

5 Protocols used in the leaf-spine examples 13

5.1 VLT 13

5.2 UFD 14

5.3 RSTP 14

5.4 Routing protocols 14

5.4.1 BGP 14

5.4.2 OSPF 15

5.4.3 IS-IS 15

5.5 BFD 15

5.6 ECMP 16

6 Layer 3 configuration planning 17

6.1 BGP ASN configuration 17

6.2 IP addressing 17

6.2.1 Loopback addresses 17

6.2.2 Point-to-point addresses 18

7 Example 1: Layer 3 with Dell EMC leaf and spine switches 20

7.1 S4048-ON leaf switch configuration 21

7.1.1 S4048-ON BGP configuration 25

7.1.2 S4048-ON OSPF configuration 26
7.2 Z9100-ON spine switch configuration....27
7.2.1 Z9100-ON BGP configuration....30
7.2.2 Z9100-ON OSPF configuration 31
7.3 Example 1 validation 31
7.3.1 show ip bgp summary....32
7.3.2 show ip ospf neighbor....32
7.3.3 show ip route bgp 33
7.3.4 show ip route ospf....34
7.3.5 show bfd neighbors 35
7.3.6 show vlt brief....36
7.3.7 show vlt detail 36
7.3.8 show vlt mismatch 36
7.3.9 show uplink-state-group 37
7.3.10 show spanning-tree rstp brief....37
8 Example 2: Layer 3 with Dell EMC leaf and Cisco Nexus spine switches 38
8.1 Nexus 5600 series spine switch configuration 39
8.1.1 Nexus 5600 series BGP configuration 40
8.1.2 Nexus 5600 series OSPF configuration 41
8.2 Example 2 validation 42
8.2.1 show ip bgp summary....42
8.2.2 show ip ospf neighbor....43
8.2.3 show ip route bgp 43
8.2.4 show ip route ospf....44
8.2.5 show bfd neighbors 46
8.2.6 Dell EMC Networking leaf validation commands previously covered 46
9 Example 3: Layer 2 with Dell EMC leaf and spine switches 47
9.1 S4048-ON leaf switch configuration 47
9.2 S6010-ON spine configuration 50
9.3 Example 3 validation 53
9.3.1 show vlt brief....53
9.3.2 show vlt detail 54
9.3.3 show vlt mismatch 54

9.3.4 show uplink-state-group 55
9.3.5 show spanning-tree rstp brief 55
10 Example 4: Layer 2 with Dell EMC leaf and Cisco Nexus spine switches 57
10.1 Z9100-ON leaf switch configuration 58
10.2 Nexus 7000 series spine switch configuration 61
10.3 Example 4 validation 64
10.3.1 show vpc 64
10.3.2 show vpc consistency-parameters 65
10.3.3 show spanning-tree 66
10.3.4 Dell EMC Networking leaf validation commands previously covered 67
A Dell EMC Networking switch factory default settings 68
B Validated hardware and operating systems 69
C Technical support and resources 70
D Support and Feedback 71

1 Introduction

Data center networks have traditionally been built in a three-layer hierarchical tree consisting of access, aggregation and core layers as shown in Figure 1.

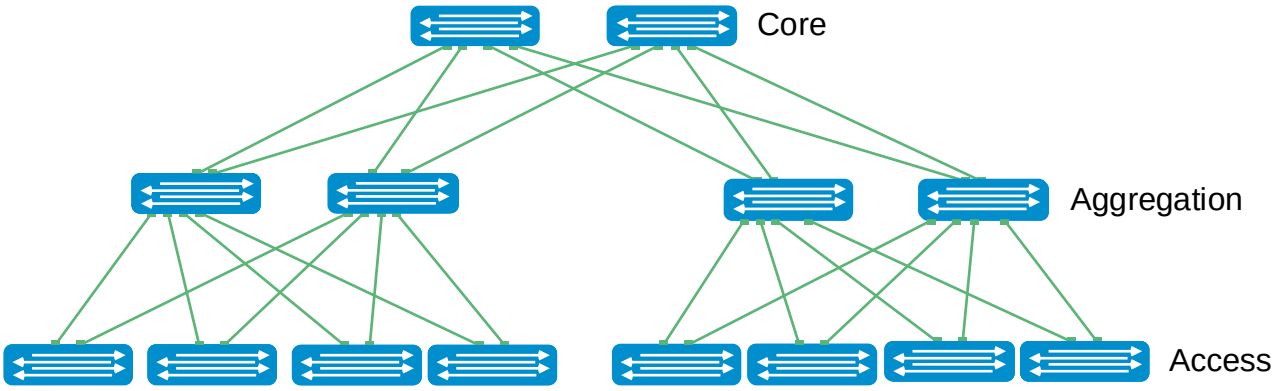

Figure 1 Traditional network architecture

Due to increasing east-west traffic within the data center (server-server, server-storage, etc.), an alternative to the traditional access-aggregation-core network model is becoming more widely used. This architecture, shown in Figure 2, is known as a Clos or leaf-spine network and is designed to minimize the number of hops between hosts.

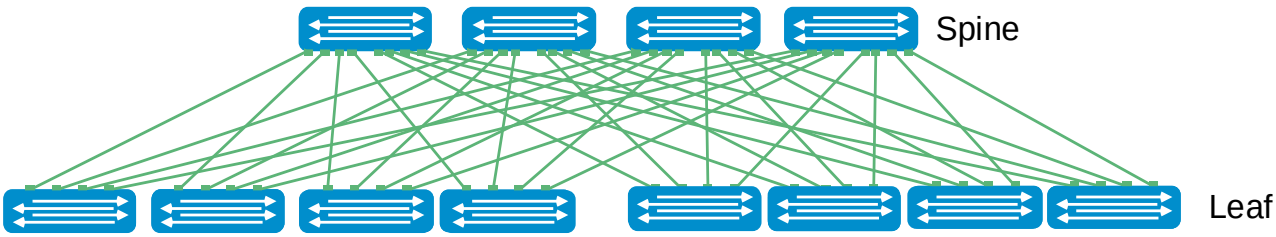

Figure 2 Leaf-spine architecture

In a leaf-spine architecture, the access layer is referred to as the leaf layer. Servers and storage devices connect to leaf switches at this layer. At the next level, the aggregation and core layers are condensed into a single spine layer. Every leaf switch connects to every spine switch to ensure that all leaf switches are no more than one hop away from one another. This minimizes latency and the likelihood of bottlenecks in the network.

A leaf-spine architecture is highly scalable. As administrators add racks to the data center, a pair of leaf switches are added to each new rack. Spine switches may be added as bandwidth requirements increase.

The connections between leaf and spine switches can be layer 2 (switched) or layer 3 (routed). This deployment guide provides step-by-step configuration examples of both topologies. It includes examples using Dell EMC Networking switches at both the leaf and spine layers and examples with Cisco Nexus switches at the spine layer. The objective is to enable a network administrator or engineer with traditional networking experience to deploy a layer 2 or layer 3 leaf-spine architecture using the examples provided.

1.1 Typographical Conventions

The command line examples in this document use the following conventions:

`Monospace Text`	CLI examples
`Underlined Monospace Text`	CLI examples that wrap the page. This text is entered as a single command.
`Italic Monospace Text`	Variables in CLI examples
`Bold Monospace Text`	Used to distinguish CLI examples from surrounding text.

2 Hardware overview

This section briefly describes the hardware used to validate the examples in this guide. A complete listing of hardware and components used is provided in Appendix B.

2.1 Dell EMC Networking S3048-ON

The Dell EMC Networking S3048-ON is a 1-Rack Unit (RU) switch with forty-eight 1GbE Base-T ports and four 10GbE SFP+ ports. In this guide, one S3048-ON switch supports out-of-band (OOB) management traffic in each rack.

Figure 3 Dell EMC Networking S3048-ON

2.2 Dell EMC Networking S4048-ON

The Dell EMC Networking S4048-ON is a 1-RU, multilayer switch with forty-eight 10GbE SFP+ ports and six 40GbE QSFP+ ports. S4048-ON switches are used as leaf switches in the examples in this guide.

Figure 4 Dell EMC Networking S4048-ON

2.3 Dell EMC Networking S6010-ON

The Dell EMC Networking S6010-ON is a 1-RU, multilayer switch with thirty-two 40GbE QSFP+ ports. S6010-ON switches are used as spine switches in an example in this guide.

Figure 5 Dell EMC Networking S6010-ON

2.4 Dell EMC Networking Z9100-ON

The Dell EMC Networking Z9100-ON is a 1-RU, multilayer switch with thirty-two ports supporting 10/25/40/50/100GbE plus two 10GbE ports. The Z9100-ON is used both as a leaf and as a spine in examples in this guide. As a leaf, the Z9100-ON is able to support servers with a wide range of network adapter speeds such as 25GbE. As a spine, the Z9100-ON provides top performance with 100GbE support.

Figure 6 Dell EMC Networking Z9100-ON

2.5 Cisco Nexus 5600 and 7000 series switches

Some deployment examples in this guide use Cisco Nexus 5600 or 7000 series switches as spines. The Nexus switches are configured for interoperability with Dell EMC Networking leaf switches in layer 2 and layer 3 topology examples.

3 Leaf-spine overview

The connections between leaf and spine switches can be layer 2 (switched) or layer 3 (routed). The terms “layer 3 topology” and “layer 2 topology” in this guide refer to these connections. In both topologies, downstream connections to servers, storage and other endpoint devices within the racks are layer 2 and connections to external networks are layer 3.

The following concepts apply to layer 2 and layer 3 leaf-spine topologies:

- Each leaf switch connects to every spine switch in the topology.
- Servers, storage arrays, edge routers and similar devices always connect to leaf switches, never to spines.

The layer 2 and layer 3 topologies each use two leaf switches at the top of each rack configured as a Virtual Link Trunking (VLT) pair. 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.

The total number of leaf-spine connections is equal to the number of leaf switches multiplied by the number of spine switches. Bandwidth of the fabric may be increased by adding connections between leafs and spines as long as the spine layer has capacity for the additional connections.

3.1 Layer 3 leaf-spine topology

In a layer 3 leaf-spine network, traffic between leafs and spines is routed. The layer 3 / layer 2 boundary is at the leaf switches. Spine switches are never connected to each other in a layer 3 topology. Equal cost multi-path routing (ECMP) is used to load balance traffic across the layer 3 network. Connections within racks from hosts to leaf switches are layer 2. Connections to external networks are made from a pair of edge or border leafs as shown in Figure 7.

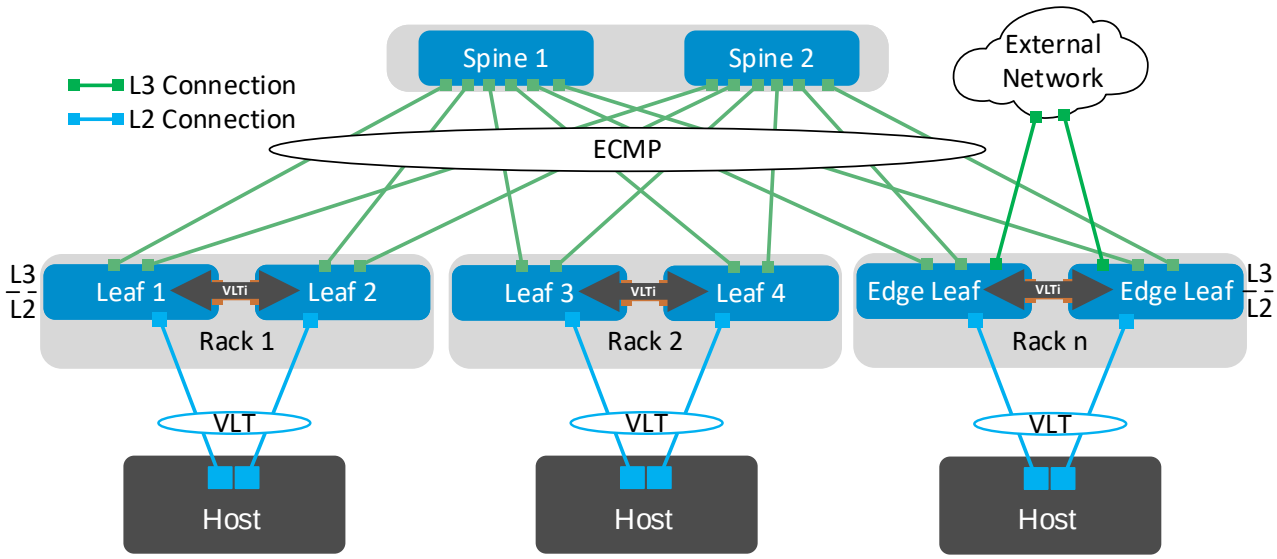

Figure 7 Layer 3 leaf-spine network

3.2 Layer 2 leaf-spine topology

In a layer 2 leaf-spine network, traffic between leafs and spines is switched (except for a pair of edge leafs) as shown in Figure 8. VLT is used for multipathing and load balancing traffic across the layer 2 leaf-spine fabric. Connections from hosts to leaf switches are also layer 2.

For connections to external networks, layer 3 links are added between the spines and a pair of edge leafs.

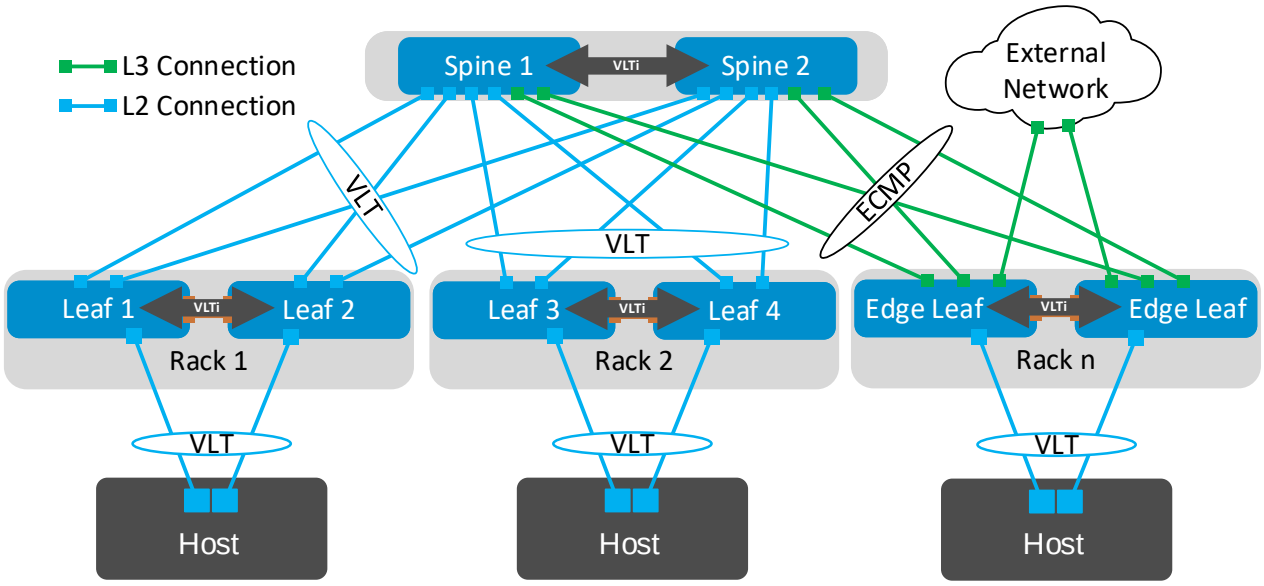

Figure 8 Layer 2 leaf-spine network

3.3 Design considerations

When compared to a layer 3 topology, a layer 2 topology is generally less complex but has some limitations that must be considered. These include:

- For each VLAN, the layer 2 topology creates one large broadcast domain across the fabric. The layer 3 topology has the benefit of containing broadcast domains to each rack.
- The layer 2 topology is limited to 4094 VLANs across the fabric. The layer 3 topology allows up to 4094 VLANs per rack.
- The layer 2 topology is limited to two physical switches at the spine layer (configured as VLT peers). In a layer 3 topology, additional spines may be added as needed to provide additional paths and bandwidth. Therefore, a layer 3 topology is more scalable and is better suited for very large networks.
- Overlay networks utilizing VXLAN (such as VMware NSX) require a layer 3 underlay network.

If none of the layer 2 limitations are a concern, it may ultimately come down to a matter of preference. This guide provides examples of both topologies.

4 Management network

The OOB management network is isolated from the leaf-spine production network. It is the same for the layer 2 and layer 3 leaf-spine topologies.

An S3048-ON installed in each rack provides 1GbE connectivity to the management network. The RJ-45 OOB management ports on each spine and leaf switch are connected to the S3048-ON switches as shown in Figure 9. PowerEdge server iDRACs and Chassis Management Controllers (CMCs) are also connected for server administration.

For the S3048-ON management switches, all ports used are in layer 2 mode and are in the default VLAN. Rapid Spanning Tree Protocol (RSTP) is enabled as a precaution against loops. No additional configuration is required.

Note: The management network used in the examples in this guide is 100.67.187.0/24.

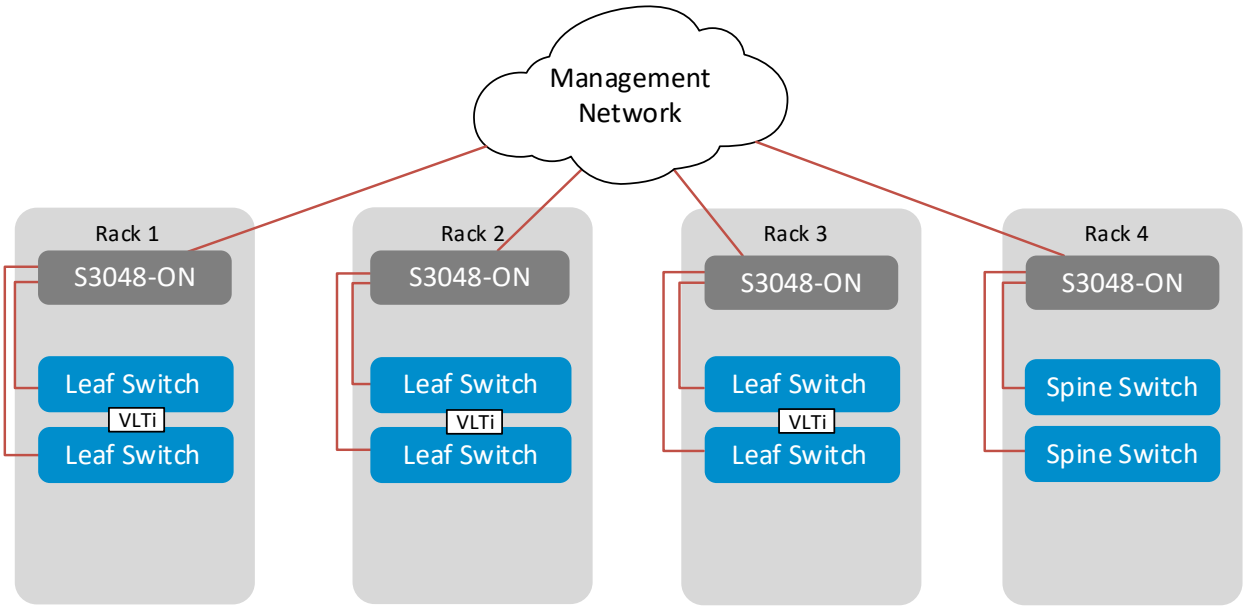

Figure 9 Management Network

Note: A management network is not a requirement to configure or use a leaf-spine network, but is recommended to efficiently manage servers, switches and storage devices.

5 Protocols used in the leaf-spine examples

This section provides an overview of the protocols used in constructing the leaf-spine network examples in this guide.

The first three protocols are used in all layer 2 and layer 3 topology examples:

- VLT, Section 5.1
- Uplink Failure Detection (UFD), Section 5.2
- RSTP, Section 5.3

The remaining protocols are only used in the layer 3 topology examples:

- Routing protocols, Section 5.4:
  - Border Gateway Protocol (BGP)
  - Open Shortest Path First (OSPF)
- Bidirectional Forwarding Detection (BFD), Section 5.5
- ECMP, Section 5.6

5.1 VLT

VLT allows link aggregation group (LAG) terminations on two separate switches and supports a loop-free topology. The two switches are referred to as VLT peers and are kept synchronized via an inter-switch link called the VLT interconnect (VLTi). A separate backup link maintains heartbeat messages across the OOB management network.

VLT provides layer 2 multipathing and load-balances traffic. VLT offers the following additional benefits:

- Eliminates spanning tree-blocked ports
- Uses all available uplink bandwidth
- Provides fast convergence if either a link or device fails
- Assures high availability

In layer 2 leaf-spine topologies, VLT is used at both the leaf and spine layers.

In layer 3 topologies, VLT is only used at the leaf layer. An additional feature called VLT peer routing is enabled on the leaf switches for connections to layer 3 networks. VLT peer routing:

- Enables one VLT node to act as the default gateway for its VLT peer
- Eliminates the need to use Virtual Router Redundancy Protocol (VRRP)
- Enables active-active load sharing

With peer routing enabled, traffic is routed through either VLT peer and is passed directly to the next hop without needing to traverse the VLTi.

Note: Downstream connections from leaf switches configured for VLT do not necessarily have to be configured as LAGs if other fault tolerant methods are preferred (e.g. multipath IO). In this guide, examples 1 and 2 use LAGs to downstream servers while examples 3 and 4 do not.

5.2 UFD

If a leaf switch loses all connectivity to the spine layer, by default the attached hosts continue to send traffic to that leaf without a direct path to the destination. The VLTi link to the peer leaf switch handles traffic during such a network outage, but this is not considered a best practice.

Dell EMC recommends enabling UFD, which detects the loss of upstream connectivity. An uplink-state group is configured on each leaf switch, which creates an association between the uplinks to the spines and the downlink interfaces.

In the event all uplinks fail on a switch, UFD automatically shuts down the downstream interfaces. This propagates to the hosts attached to the leaf switch. The host then uses its link to the remaining switch to continue sending traffic across the leaf-spine network.

5.3 RSTP

As a precautionary measure, Dell EMC recommends enabling RSTP on all switches that have layer 2 interfaces. Because VLT environments are loop-free, simultaneously running spanning tree is optional though considered a best practice in case of switch misconfiguration or improperly connected cables. In properly configured and connected leaf-spine networks, there are no ports blocked by spanning tree.

5.4 Routing protocols

Any of the following routing protocols may be used on layer 3 connections when designing a leaf-spine network:

- BGP
- OSPF
- Intermediate System to Intermediate System (IS-IS)

5.4.1 BGP

BGP may be selected for scalability and is well suited for very large networks. BGP can be configured as External BGP (EBGP) to route between autonomous systems or Internal BGP (IBGP) to route within a single autonomous system.

Layer 3 leaf-spine networks use ECMP routing. EBGP and IBGP handle ECMP differently. By default, EBGP supports ECMP without any adjustments. IBGP requires a BGP route reflector and the use of the AddPath feature to fully support ECMP. To keep configuration complexity to a minimum, Dell EMC recommends EBGP in leaf-spine fabric deployments.

BGP tracks IP reachability to the peer remote address and the peer local address. Whenever either address becomes unreachable, BGP brings down the session with the peer. To ensure fast convergence with BGP,

Dell EMC recommends enabling fast fall-over with BGP. Fast fall-over terminates external BGP sessions of any directly adjacent peer if the link to reach the peer goes down without waiting for the hold-down timer to expire.

Note: BGP fast fall-over is enabled manually on Dell EMC switches and is enabled by default on Cisco Nexus 5600 series switches.

Examples using EBGP (BGPv4) are provided in the layer 3 topology examples in this guide.

5.4.2 OSPF

OSPF is an interior gateway protocol that provides routing inside an autonomous network. OSPF routers send link-state advertisements to all other routers within the same autonomous system areas. While generally more memory and CPU intensive than BGP, OSPF may offer faster convergence. OSPF is often used in smaller networks. Examples using OSPF (OSPFv2 for IPv4) are provided in the layer 3 topology examples in this guide.

5.4.3 IS-IS

IS-IS, like OSPF, is a link-state routing protocol that computes the best path through the network. Though supported, the protocol is not as widely used and is not covered in this document. To use IS-IS for routing, consult the applicable switch's user guide.

5.5 BFD

BFD is a protocol used to rapidly detect communication failures between two adjacent systems over a layer 3 link. It is a simple and lightweight replacement for existing routing protocol link state detection mechanisms. Though optional, use of BFD is considered a best practice for optimizing a leaf-spine network.

BFD provides forwarding path failure detection times on the order of milliseconds rather than seconds as with conventional routing protocols. It is independent of routing protocols and provides a consistent method of failure detection when used across a network. Networks converge faster because BFD triggers link state changes in the routing protocol sooner and more consistently.

Dell EMC Networking has implemented BFD at layer 3 with user datagram protocol (UDP) encapsulation. BFD is supported with routing protocols including BGP, OSPF and IS-IS.

5.6 ECMP

The nature of a leaf-spine topology is that leaf switches are no more than one hop away from each other. As shown in Figure 10, Leaf 1 has two equal cost paths to Leaf 4, one through each spine. The same is true for all leafs.

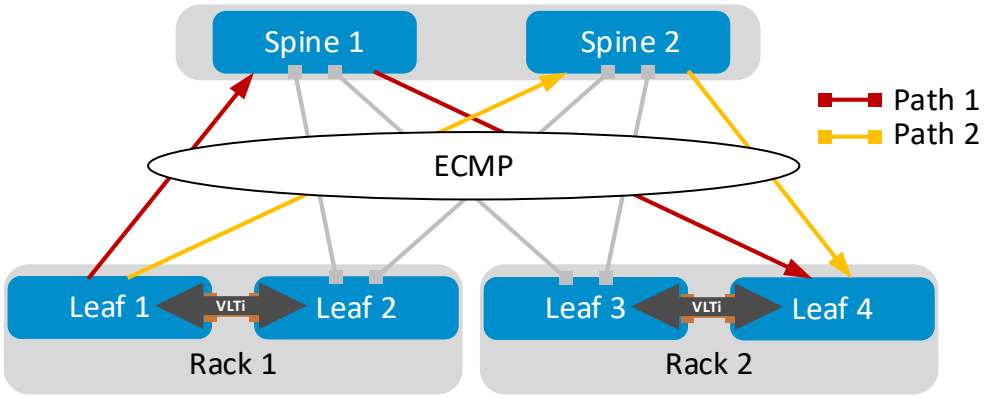

Figure 10 Use of ECMP in a layer 3 topology

ECMP is a routing technique used in a layer 3 leaf-spine topology for load balancing packets along these multiple equal cost paths. ECMP is enabled on all leaf and spine switches, allowing traffic between leafs to be load balanced across the spines.

6 Layer 3 configuration planning

6.1 BGP ASN configuration

When EBGP is used, an autonomous system number (ASN) is assigned to each switch. Valid private, 2-byte ASNs range from 64512 through 65534. Figure 11 shows the ASN assignments used for leaf and spine switches in the BGP examples in this guide.

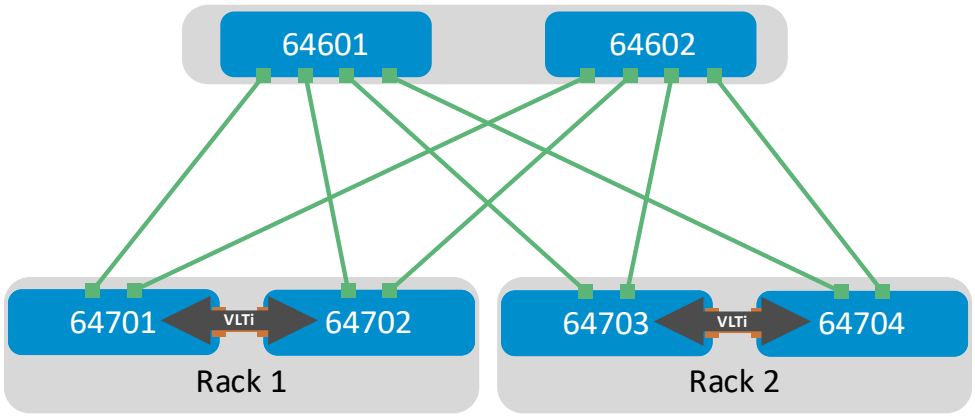

Figure 11 BGP ASN assignments

ASNs should follow a logical pattern for ease of administration and allow for growth as additional leaf and spine switches are added. In this example, an ASN with a "6" in the hundreds place represents a spine switch (e.g. 64**6**01) and an ASN with a "7" in the hundreds place represents a leaf switch (e.g. 64**7**01).

6.2 IP addressing

Establishing a logical, scalable IP address scheme is important before deploying a leaf-spine topology. This section covers the IP addressing used in the layer 3 examples in this guide.

6.2.1 Loopback addresses

Loopback addresses may be used as router IDs when configuring routing protocols. As with ASNs, loopback addresses should follow a logical pattern that will make it easier for administrators to manage the network and allow for growth. Figure 12 shows the loopback addresses used as router IDs in the BGP and OSPF examples in this guide.

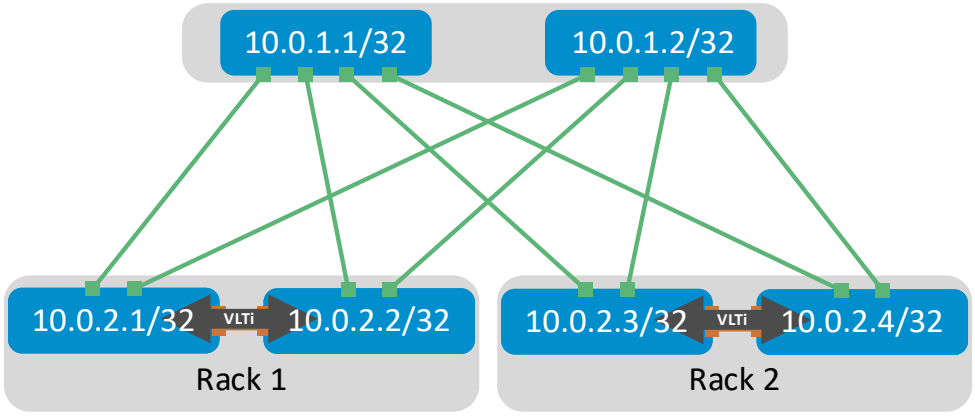

Figure 12 Loopback addressing

All loopback addresses used are part of the 10.0.0.0/8 address space with each address using a 32-bit mask. In this example, the third octet represents the layer, “1” for spine and “2” for leaf. The fourth octet is the counter for the appropriate layer. For example, 10.0.1.1/32 is the first spine switch in the topology while 10.0.2.4/32 is the fourth leaf switch.

6.2.2 Point-to-point addresses

Table 1 lists layer 3 connection details for each leaf and spine switch.

All addresses come from the same base IP prefix, 192.168.0.0/16 with the third octet representing the spine number. For example, 192.168.1.0/31 is a two host subnet connected to Spine 1 while 192.168.2.0/31 is connected to Spine 2. This IP scheme is easily extended as leaf and spine switches are added to the network.

Link labels are provided in the table for quick reference with Figure 13.

Table 1 Interface and IP configuration

Link Label	Source switch	Source interface	Source IP	Network	Destination switch	Destination interface	Destination IP
A	Leaf 1	fo1/49	.1	192.168.1.0/31	Spine 1	fo1/1/1	.0
B	Leaf 1	fo1/50	.1	192.168.2.0/31	Spine 2	fo1/1/1	.0
C	Leaf 2	fo1/49	.3	192.168.1.2/31	Spine 1	fo1/2/1	.2
D	Leaf 2	fo1/50	.3	192.168.2.2/31	Spine 2	fo1/2/1	.2
E	Leaf 3	fo1/49	.5	192.168.1.4/31	Spine 1	fo1/3/1	.4
F	Leaf 3	fo1/50	.5	192.168.2.4/31	Spine 2	fo1/3/1	.4
G	Leaf 4	fo1/49	.7	192.168.1.6/31	Spine 1	fo1/4/1	.6
H	Leaf 4	fo1/50	.7	192.168.2.6/31	Spine 2	fo1/4/1	.6

The point-to-point IP addresses used in this guide are shown in Figure 13:

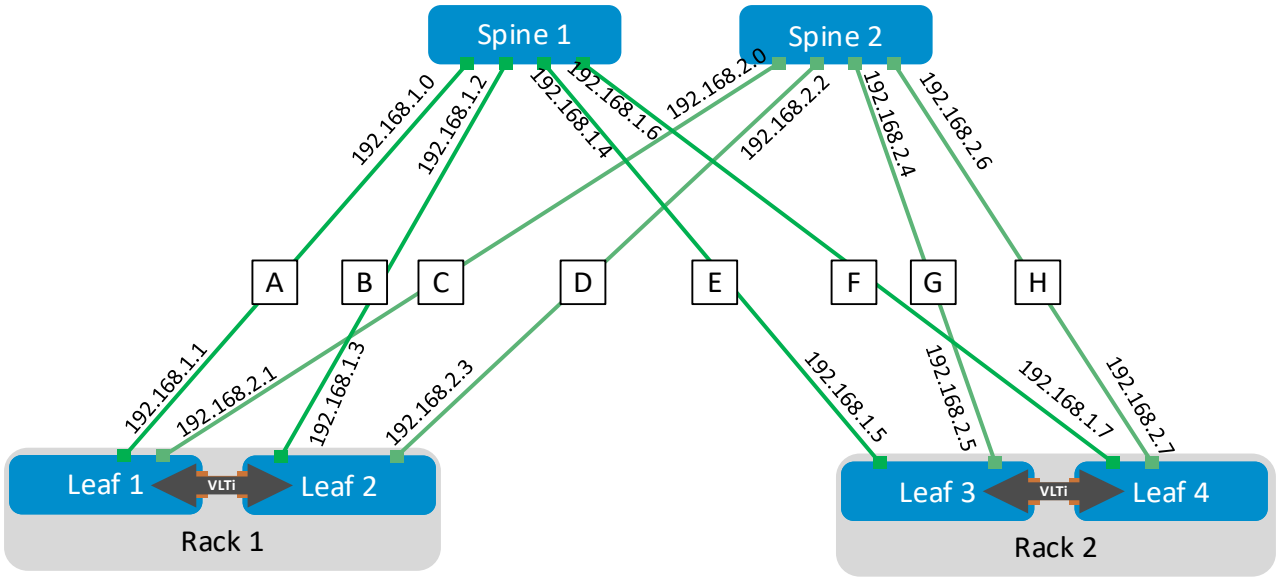

Figure 13 Point-to-point IP addresses

Note: The example point-to-point addresses use a 31-bit mask to save address space. This is optional and covered in RFC 3021. Below is an example when setting an IP address with a 31-bit mask on a Dell EMC S4048-ON. The warning message can be safely ignored on point-to-point interfaces.

```
S4048-Leaf-1(conf-if-fo-1/49)#ip address 192.168.1.1/31
% Warning: Use /31 mask on non point-to-point interface cautiously.
```

7 Example 1: Layer 3 with Dell EMC leaf and spine switches

This section provides BGP and OSPF configuration examples to build the layer 3 leaf-spine topology shown in Figure 14. Dell EMC Networking S4048-ON switches are used at the leaf layer and Dell EMC Networking Z9100-ON switches are used at the spine layer.

Note: The BGP ASNs and IP addresses defined in Section 6 are used here.

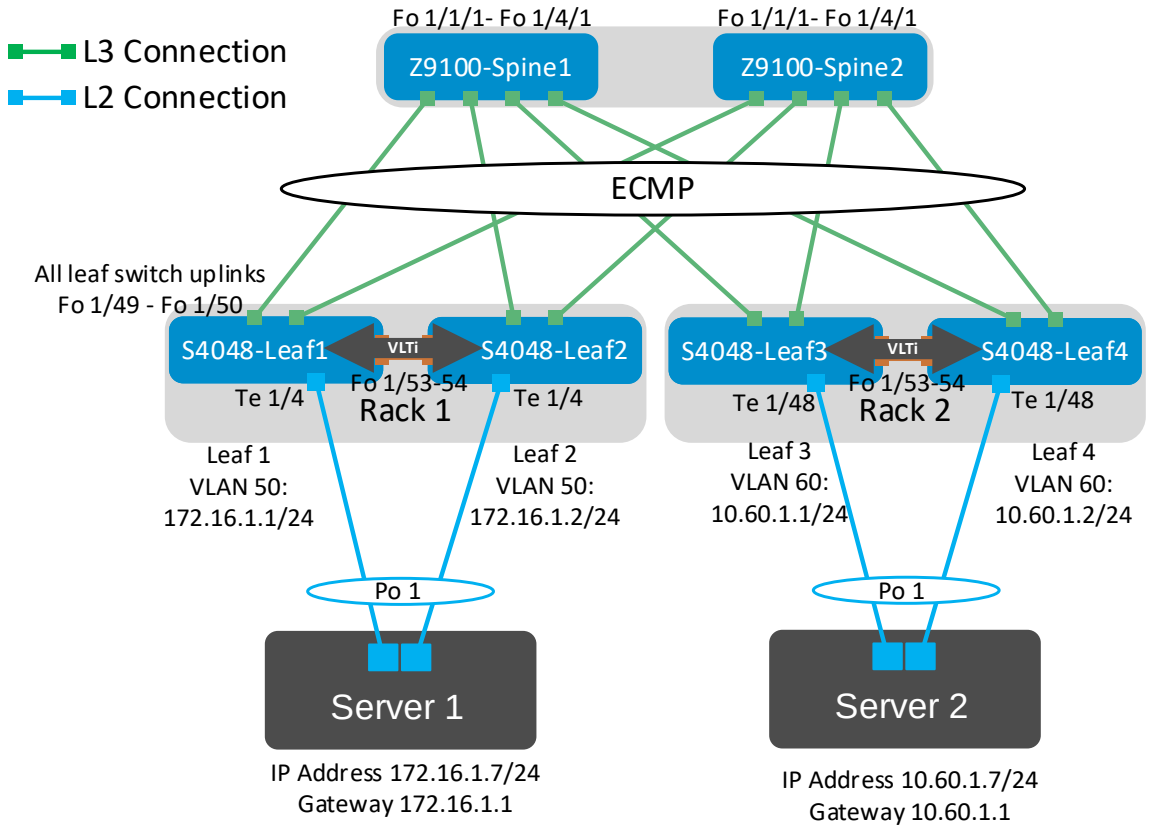

Figure 14 Example 1: Layer 3 leaf-spine topology with Dell EMC leaf and spine switches

In this topology, there is one broadcast domain in each rack.

In Rack 1, VLAN 50 is used and devices in VLAN 50 are assigned IP addresses on the 172.16.1.0/24 network. With VLT peer routing enabled on S4048-Leaf1 and S4048-Leaf2, Server 1 may specify the IP address assigned to VLAN 50 on either leaf, 172.16.1.1 or 172.16.1.2, as its default gateway. Traffic is load balanced across both leafs.

Rack 2 is configured in an identical manner, except VLAN 60 is used and devices in VLAN 60 are assigned IP addresses on the 10.60.1.0/24 network. Server 2 may specify the VLAN 60 IP address of either leaf, 10.60.1.1 or 10.60.1.2, as its default gateway.

Note: All switch configuration files for the topology in Figure 14 are contained in the attachment named **Example1_config_files.pdf.** The files may be edited as needed in a plain text editor and commands pasted directly into switch consoles. Dell EMC Networking switches start at their factory default settings per Appendix A.

7.1 S4048-ON leaf switch configuration

The following configuration details are for S4048-Leaf1 and S4048-Leaf2 in Figure 14. The configuration commands for S4048-Leaf3 and S4048-Leaf4 are similar and are provided in the attachments.

Note: On S4048-ON switches, Telnet is enabled and SSH is disabled by default. Both services require the creation of a non-root user account to login. If needed, it is a best practice to use SSH instead of Telnet for security. SSH can optionally be enabled with the command: `(conf)#ip ssh server enable.`
A user account can be created to access the switch via SSH with the command
`(conf)#username ssh_user sha256-password ssh_password`

First, configure the serial console enable password and disable Telnet.

S4048-Leaf1	S4048-Leaf2
enable configure enable sha256-password *enable_password* no ip telnet server enable	enable configure enable sha256-password *enable_password* no ip telnet server enable

Set the hostname, configure the OOB management interface and default gateway. Enable LLDP and BFD. Enable RSTP as a precaution. S4048-Leaf1 is configured as the primary RSTP root bridge using the `bridge-priority 0` command. S4048-Leaf2 is configured as the secondary RSTP root bridge using the `bridge-priority 4096` command.

S4048-Leaf1	S4048-Leaf2
hostname S4048-Leaf1 interface ManagementEthernet 1/1 ip address 100.67.187.35/24 no shutdown management route 0.0.0.0/0 100.67.187.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc bfd enable protocol spanning-tree rstp no disable bridge-priority 0	hostname S4048-Leaf2 interface ManagementEthernet 1/1 ip address 100.67.187.34/24 no shutdown management route 0.0.0.0/0 100.67.187.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc bfd enable protocol spanning-tree rstp no disable bridge-priority 4096

Configure the VLT interconnect between S4048-Leaf1 and S4048-Leaf2. In this configuration, add interfaces fortyGigE 1/53-54 to static port channel 127 for the VLT interconnect. The backup destination is the management IP address of the VLT peer switch. Enable peer routing.

Note: Dell EMC recommends that the VLTi is configured as a static LAG, without LACP, per the commands shown below.

S4048-Leaf1	S4048-Leaf2
interface port-channel 127 description VLTi channel-member fortyGigE 1/53 - 1/54 no shutdown interface range fortyGigE 1/53 - 1/54 description VLTi no shutdown vlt domain 127 peer-link port-channel 127 back-up destination 100.67.187.34 unit-id 0 peer-routing exit	interface port-channel 127 description VLTi channel-member fortyGigE 1/53 - 1/54 no shutdown interface range fortyGigE 1/53 - 1/54 description VLTi no shutdown vlt domain 127 peer-link port-channel 127 back-up destination 100.67.187.35 unit-id 1 peer-routing exit

Configure each downstream server-facing interface with an LACP port channel. Configure each port channel for VLT. Port channel 1 connects downstream to Server 1 and is configured as an RSTP edge port.

S4048-Leaf1	S4048-Leaf2
interface tengigabitethernet 1/4 description Server 1 port-channel-protocol LACP port-channel 1 mode active no shutdown interface port-channel 1 description Server 1 portmode hybrid switchport spanning-tree rstp edge-port vlt-peer-lag port-channel 1 no shutdown	interface tengigabitethernet 1/4 description Server 1 port-channel-protocol LACP port-channel 1 mode active no shutdown interface port-channel 1 description Server 1 portmode hybrid switchport spanning-tree rstp edge-port vlt-peer-lag port-channel 1 no shutdown

Create a VLAN interface containing the server-facing port channel(s). Use the same VLAN ID on both leafs. Create a switched virtual interface (SVI) by assigning an IP address to the VLAN interface. The address must be unique but on the same network on both leaf switches.

Note: In this example, Server 1's NIC is configured as an LACP NIC team. It is assigned the IP address 172.16.1.7/24. The SVI on either leaf, 172.16.1.1 or 172.16.1.2, is specified as Server 1's default gateway. The VLT peer routing feature will load balance traffic and provide fault tolerance if either leaf goes offline, regardless of which leaf's address is specified as the server's gateway.

S4048-Leaf1	S4048-Leaf2
interface Vlan 50 ip address 172.16.1.1/24 untagged Port-channel 1 no shutdown	interface Vlan 50 ip address 172.16.1.2/24 untagged Port-channel 1 no shutdown

The two upstream layer 3 interfaces connected to the spine switches are configured. Assign IP addresses per Table 1. Configure a loopback interface to be used as the router ID. This is used with BGP or OSPF.

Note: If multiple loopback interfaces exist on a system, the interface with the highest numbered IP address is used as the router ID. This configuration only uses one loopback interface.

S4048-Leaf1	S4048-Leaf2
interface fortyGigE 1/49 description Spine-1 ip address 192.168.1.1/31 no shutdown interface fortyGigE 1/50 description Spine-2 ip address 192.168.2.1/31 no shutdown interface loopback 0 description Router ID ip address 10.0.2.1/32 no shutdown	interface fortyGigE 1/49 description Spine-1 ip address 192.168.1.3/31 no shutdown interface fortyGigE 1/50 description Spine-2 ip address 192.168.2.3/31 no shutdown interface loopback 0 description Router ID ip address 10.0.2.2/32 no shutdown

Configure a route map and IP prefix-list to redistribute all loopback addresses and leaf networks via BGP or OSPF.

The command `seq 10 permit 10.0.0.0/8 ge 24` includes all addresses in the 10.0.0.0/8 address range with a mask greater than or equal to 24. This includes all loopback addresses used as router IDs as well as the 10.60.1.0/24 network used on Leafs 3 and 4 as shown in Figure 14.

The command `seq 20 permit 172.16.0.0/16 ge 24` includes the 172.16.1.0/24 network used on Leafs 1 and 2 as shown in Figure 14.

S4048-Leaf1	S4048-Leaf2
route-map spine-leaf permit 10 match ip address spine-leaf ip prefix-list spine-leaf description Redistribute loopback and leaf networks seq 10 permit 10.0.0.0/8 ge 24 seq 20 permit 172.16.0.0/16 ge 24	route-map spine-leaf permit 10 match ip address spine-leaf ip prefix-list spine-leaf description Redistribute loopback and leaf networks seq 10 permit 10.0.0.0/8 ge 24 seq 20 permit 172.16.0.0/16 ge 24

Include the point-to-point interfaces to each leaf pair in an ECMP group. Enable link bundle monitoring to report when traffic is unevenly distributed across multiple links.

Note: ECMP is not enabled until BGP or OSPF is configured.

S4048-Leaf1	S4048-Leaf2
ecmp-group 1 interface fortyGigE 1/49 interface fortyGigE 1/50 link-bundle-monitor enable	ecmp-group 1 interface fortyGigE 1/49 interface fortyGigE 1/50 link-bundle-monitor enable

Configure UFD. This shuts the downstream interfaces if all uplinks fail. The hosts attached to the switch use the remaining LACP port member to continue sending traffic across the fabric.

Finally, exit configuration mode and save the configuration.

S4048-Leaf1	S4048-Leaf2
uplink-state-group 1 description Disable downstream ports in event all uplinks fail downstream TenGigabitEthernet 1/1-1/48 upstream fortyGigE 1/49,1/50 end write	uplink-state-group 1 description Disable downstream ports in event all uplinks fail downstream TenGigabitEthernet 1/1-1/48 upstream fortyGigE 1/49,1/50 end write

7.1.1 S4048-ON BGP configuration

Use these commands to configure BGP. If OSPF is used, skip to section 7.1.2.

First, enable BGP with the `router bgp` *ASN* command. The ASN is from Figure 11.

The `bgp bestpath as-path multipath-relax` enables ECMP. The `maximum-paths ebgp 2` command specifies the maximum number of parallel paths to a destination to add to the routing table. This number should be equal to or greater than the number of spines, up to 64.

BGP neighbors are configured and fast fall-over is enabled.

BFD settings are configured to 100 millisecond send/receive intervals. The multiplier is the number of packets that must be missed to declare a session down.

Finally, exit configuration mode and save the configuration with the `end` and `write` commands.

S4048-Leaf1	S4048-Leaf2
enable configure router bgp 64701 bgp bestpath as-path multipath-relax maximum-paths ebgp 2 redistribute connected route-map spine-leaf bgp graceful-restart neighbor spine-leaf peer-group neighbor spine-leaf fall-over neighbor spine-leaf advertisement-interval 1 neighbor spine-leaf no shutdown neighbor spine-leaf bfd neighbor 192.168.1.0 remote-as 64601 neighbor 192.168.1.0 peer-group spine-leaf neighbor 192.168.1.0 no shutdown neighbor 192.168.2.0 remote-as 64602 neighbor 192.168.2.0 peer-group spine-leaf neighbor 192.168.2.0 no shutdown bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write	enable configure router bgp 64702 bgp bestpath as-path multipath-relax maximum-paths ebgp 2 redistribute connected route-map spine-leaf bgp graceful-restart neighbor spine-leaf peer-group neighbor spine-leaf fall-over neighbor spine-leaf advertisement-interval 1 neighbor spine-leaf no shutdown neighbor spine-leaf bfd neighbor 192.168.1.2 remote-as 64601 neighbor 192.168.1.2 peer-group spine-leaf neighbor 192.168.1.2 no shutdown neighbor 192.168.2.2 remote-as 64602 neighbor 192.168.2.2 peer-group spine-leaf neighbor 192.168.2.2 no shutdown bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write

7.1.2 S4048-ON OSPF configuration

Use these commands to configure OSPF. Skip this section if BGP is used.

First, enable OSPF is enabled with the `router ospf` *`process-id`* command (valid range is 1-65535).

Add the connected networks to OSPF area 0.

The `maximum-paths 2` command enables ECMP and specifies the maximum number of parallel paths to a destination to add to the routing table. This number should be equal to or greater than the number of spines, up to 64.

BFD settings are configured to 100 millisecond send/receive intervals. The multiplier is the number of packets that must be missed to declare a session down.

Finally, exit configuration mode and save the configuration with the `end` and `write` commands.

S4048-Leaf1	S4048-Leaf2
enable configure router ospf 1 log-adjacency-changes network 192.168.1.0/31 area 0 network 192.168.2.0/31 area 0 maximum-paths 2 redistribute connected route-map spine-leaf bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write	enable configure router ospf 1 log-adjacency-changes network 192.168.1.2/31 area 0 network 192.168.2.2/31 area 0 maximum-paths 2 redistribute connected route-map spine-leaf bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write

7.2 Z9100-ON spine switch configuration

The following configuration details are for Z9100-Spine1 and Z9100-Spine2 in Figure 14.

Note: On Z9100-ON switches, Telnet is enabled and SSH is disabled by default. Both services require the creation of a non-root user account to login. If needed, it is a best practice to use SSH instead of Telnet for security. SSH can optionally be enabled with the command: `(conf)#ip ssh server enable.`
A user account can be created to access the switch via SSH with the command
`(conf)#username ssh_user sha256-password ssh_password`

First, configure the serial console enable password and disable Telnet.

Z9100-Spine1	Z9100-Spine2
enable configure enable sha256-password *enable_password* no ip telnet server enable	enable configure enable sha256-password *enable_password* no ip telnet server enable

Set the hostname, configure the OOB management interface and default gateway. Set the hostname, configure the OOB management interface and default gateway. Enable LLDP and BFD.

Set the port speed of the four ports connected to the leaf switches to 40GbE.

Z9100-Spine1	Z9100-Spine2
hostname Z9100-Spine1 interface ManagementEthernet 1/1 ip address 100.67.187.39/24 no shutdown management route 0.0.0.0/0 100.67.187.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc bfd enable stack-unit 1 port 1 portmode single speed 40G no-confirm stack-unit 1 port 2 portmode single speed 40G no-confirm stack-unit 1 port 3 portmode single speed 40G no-confirm stack-unit 1 port 4 portmode single speed 40G no-confirm	hostname Z9100-Spine2 interface ManagementEthernet 1/1 ip address 100.67.187.38/24 no shutdown management route 0.0.0.0/0 100.67.187.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc bfd enable stack-unit 1 port 1 portmode single speed 40G no-confirm stack-unit 1 port 2 portmode single speed 40G no-confirm stack-unit 1 port 3 portmode single speed 40G no-confirm stack-unit 1 port 4 portmode single speed 40G no-confirm

Configure the four point-to-point interfaces connected to leaf switches. Assign IP addresses per Table 1. Configure a loopback interface to be used as the router ID. This is used with BGP or OSPF.

Note: If multiple loopback interfaces exist on a system, the interface with the highest numbered IP address is used as the router ID. This configuration only uses one loopback interface.

Z9100-Spine1	Z9100-Spine2
interface fortyGigE 1/1/1 description Leaf 1 fo1/49 ip address 192.168.1.0/31 no shutdown interface fortyGigE 1/2/1 description Leaf 2 fo1/49 ip address 192.168.1.2/31 no shutdown interface fortyGigE 1/3/1 description Leaf 3 fo1/49 ip address 192.168.1.4/31 no shutdown interface fortyGigE 1/4/1 description Leaf 4 fo1/49 ip address 192.168.1.6/31 no shutdown interface loopback 0 description Router ID ip address 10.0.1.1/32 no shutdown	interface fortyGigE 1/1/1 description Leaf 1 fo1/50 ip address 192.168.2.0/31 no shutdown interface fortyGigE 1/2/1 description Leaf 2 fo1/50 ip address 192.168.2.2/31 no shutdown interface fortyGigE 1/3/1 description Leaf 3 fo1/50 ip address 192.168.2.4/31 no shutdown interface fortyGigE 1/4/1 description Leaf 4 fo1/50 ip address 192.168.2.6/31 no shutdown interface loopback 0 description Router ID ip address 10.0.1.2/32 no shutdown

Configure a route map and IP prefix-list to redistribute all loopback addresses and leaf networks via BGP or OSPF.

The command `seq 10 permit 10.0.0.0/8 ge 24` includes all addresses in the 10.0.0.0/8 address range with a mask greater than or equal to 24. This includes all loopback addresses used as router IDs as well as the 10.60.1.0/24 network used on Leafs 3 and 4 as shown in Figure 14.

The command `seq 20 permit 172.16.0.0/16 ge 24` includes the 172.16.1.0/24 network used on Leafs 1 and 2 as shown in Figure 14.

Z9100-Spine1	Z9100-Spine2
route-map spine-leaf permit 10 match ip address spine-leaf ip prefix-list spine-leaf description Redistribute loopback and leaf networks seq 10 permit 10.0.0.0/8 ge 24 seq 20 permit 172.16.0.0/16 ge 24	route-map spine-leaf permit 10 match ip address spine-leaf ip prefix-list spine-leaf description Redistribute loopback and leaf networks seq 10 permit 10.0.0.0/8 ge 24 seq 20 permit 172.16.0.0/16 ge 24

Include the point-to-point interfaces to each leaf pair in an ECMP group. Enable link bundle monitoring to report when traffic is unevenly distributed across multiple links.

Note: ECMP is not actually enabled until BGP or OSPF is configured.

Finally, exit configuration mode and save the configuration with the `end` and `write` commands.

Z9100-Spine1	Z9100-Spine2
ecmp-group 1 interface fortyGigE 1/1/1 interface fortyGigE 1/2/1 link-bundle-monitor enable ecmp-group 2 interface fortyGigE 1/3/1 interface fortyGigE 1/4/1 link-bundle-monitor enable end write	ecmp-group 1 interface fortyGigE 1/1/1 interface fortyGigE 1/2/1 link-bundle-monitor enable ecmp-group 2 interface fortyGigE 1/3/1 interface fortyGigE 1/4/1 link-bundle-monitor enable end write

7.2.1 Z9100-ON BGP configuration

Use these commands to configure BGP. If OSPF is used, skip to section 7.2.2.

First, enabled BGP with the `router bgp` *ASN* command. The ASN is from Figure 11.

The `bgp bestpath as-path multipath-relax` enables ECMP. The `maximum-paths ebgp 2` command specifies the maximum number of parallel paths to a destination to add to the routing table. In this topology, there are two equal cost best paths from a spine to a host, one to each leaf that the host is connected.

BGP neighbors are configured and fast fall-over is enabled.

BFD settings are configured to 100 millisecond send/receive intervals. The multiplier is the number of packets that must be missed to declare a session down. Finally, exit configuration mode and save the configuration.

Z9100-Spine1	Z9100-Spine2
enable configure router bgp 64601 bgp bestpath as-path multipath-relax maximum-paths ebgp 2 redistribute connected route-map spine-leaf bgp graceful-restart neighbor spine-leaf peer-group neighbor spine-leaf fall-over neighbor spine-leaf advertisement-interval 1 neighbor spine-leaf no shutdown neighbor spine-leaf bfd neighbor 192.168.1.1 remote-as 64701 neighbor 192.168.1.1 peer-group spine-leaf neighbor 192.168.1.1 no shutdown neighbor 192.168.1.3 remote-as 64702 neighbor 192.168.1.3 peer-group spine-leaf neighbor 192.168.1.3 no shutdown neighbor 192.168.1.5 remote-as 64703 neighbor 192.168.1.5 peer-group spine-leaf neighbor 192.168.1.5 no shutdown neighbor 192.168.1.7 remote-as 64704 neighbor 192.168.1.7 peer-group spine-leaf neighbor 192.168.1.7 no shutdown bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write	enable configure router bgp 64602 bgp bestpath as-path multipath-relax maximum-paths ebgp 2 redistribute connected route-map spine-leaf bgp graceful-restart neighbor spine-leaf peer-group neighbor spine-leaf fall-over neighbor spine-leaf advertisement-interval 1 neighbor spine-leaf no shutdown neighbor spine-leaf bfd neighbor 192.168.2.1 remote-as 64701 neighbor 192.168.2.1 peer-group spine-leaf neighbor 192.168.2.1 no shutdown neighbor 192.168.2.3 remote-as 64702 neighbor 192.168.2.3 peer-group spine-leaf neighbor 192.168.2.3 no shutdown neighbor 192.168.2.5 remote-as 64703 neighbor 192.168.2.5 peer-group spine-leaf neighbor 192.168.2.5 no shutdown neighbor 192.168.2.7 remote-as 64704 neighbor 192.168.2.7 peer-group spine-leaf neighbor 192.168.2.7 no shutdown bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write

7.2.2 Z9100-ON OSPF configuration

Use these commands to configure OSPF. Skip this section if BGP is used.

First, enable OSPF with the `router ospf` *`process-id`* command (valid range is 1-65535).

Add the connected networks to OSPF area 0.

The `maximum-paths 2` command enables ECMP and specifies the maximum number of parallel paths to a destination to add to the routing table. In this topology, there are two equal cost best paths from a spine to a host, one to each leaf that the host is connected.

BFD settings are configured to 100 millisecond send/receive intervals. The multiplier is the number of packets that must be missed to declare a session down.

Finally, exit configuration mode and save the configuration with the `end` and `write` commands.

Z9100-Spine1	Z9100-Spine2
enable configure router ospf 1 log-adjacency-changes network 192.168.1.0/31 area 0 network 192.168.1.2/31 area 0 network 192.168.1.4/31 area 0 network 192.168.1.6/31 area 0 maximum-paths 2 redistribute connected route-map spine-leaf bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write	enable configure router ospf 1 log-adjacency-changes network 192.168.2.0/31 area 0 network 192.168.2.2/31 area 0 network 192.168.2.4/31 area 0 network 192.168.2.6/31 area 0 maximum-paths 2 redistribute connected route-map spine-leaf bfd all-neighbors interval 100 min_rx 100 multiplier 3 role active end write

7.3 Example 1 validation

In addition to sending traffic between hosts, the configuration shown in Figure 14 can be validated with the commands shown in this section. For more information on commands and output, see the Command Line Reference Guide for the applicable switch (links to documentation are provided in Appendix C).

Command and output examples are provided for one spine and one leaf. Command output on other switches is similar.

7.3.1 show ip bgp summary

When BGP is configured, this command shows the status of all BGP connections. Each spine has four neighbors (the four leafs) and each leaf has two neighbors (the two spines). This command also confirms BFD is enabled on the 6th line of output.

```
Z9100-Spine-1#show ip bgp summary
BGP router identifier 10.0.1.1, local AS number 64601
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
8 network entrie(s) using 608 bytes of memory
13 paths using 1404 bytes of memory
BGP-RIB over all using 1417 bytes of memory
BFD is enabled, Interval 100 Min_rx 100 Multiplier 3 Role Active
29 BGP path attribute entrie(s) using 4816 bytes of memory
27 BGP AS-PATH entrie(s) using 270 bytes of memory
4 neighbor(s) using 32768 bytes of memory

Neighbor      AS       MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.1.1   64701    3014     3014          0    0     0 1d:19:31:07 3
192.168.1.3   64702    3013     3011          0    0     0 1d:19:31:11 3
192.168.1.5   64703    3014     3012          0    0     0 1d:19:30:59 3
192.168.1.7   64704    3014     3012          0    0     0 1d:19:31:06 3

S4048-Leaf-1#show ip bgp summary
BGP router identifier 10.0.2.1, local AS number 64701
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
8 network entrie(s) using 608 bytes of memory
12 paths using 1296 bytes of memory
BGP-RIB over all using 1308 bytes of memory
BFD is enabled, Interval 100 Min_rx 100 Multiplier 3 Role Active
17 BGP path attribute entrie(s) using 2752 bytes of memory
15 BGP AS-PATH entrie(s) using 150 bytes of memory
2 neighbor(s) using 16384 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down
State/Pfx
192.168.1.0     64601            15       17          0    0     0 00:03:41 5
192.168.2.0     64602            13       13          0    0     0 00:03:41 5
```

7.3.2 show ip ospf neighbor

When OSPF is configured, this command shows the state of all connected OSPF neighbors. In this configuration, each spine has four neighbors (the four leafs) and each leaf has two neighbors (the two spines).

```
Z9100-Spine-1#sh ip ospf neighbor
```

```
Neighbor ID    Pri      State         Dead Time Address         Interface   Area
10.0.2.1         1      FULL/DR       00:00:32  192.168.1.1     Fo 1/1/1    0
10.0.2.2         1      FULL/DR       00:00:34  192.168.1.3     Fo 1/2/1    0
10.0.2.3         1      FULL/DR       00:00:35  192.168.1.5     Fo 1/3/1    0
10.0.2.4         1      FULL/DR       00:00:35  192.168.1.7     Fo 1/4/1    0

S4048-Leaf-1#sh ip ospf neighbor

Neighbor ID    Pri      State         Dead Time Address         Interface   Area
10.0.1.1         1      FULL/BDR      00:00:38  192.168.1.0     Fo 1/49     0
10.0.1.2         1      FULL/BDR      00:00:39  192.168.2.0     Fo 1/50     0
```

7.3.3 show ip route bgp

On switches with BGP configured, this command is used to verify the BGP entries in the Routing Information Base (RIB). Entries with multiple paths shown are used with ECMP. The two server networks in this example, 10.60.1.0 and 172.16.1.0, each have two paths from Z9100-Spine1, one through each leaf.

The first set of routes with a subnet mask of /32 are the IPs configured for router IDs.

```
Z9100-Spine1#show ip route bgp
       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  B EX 10.0.1.2/32        via 192.168.1.1                    20/0    00:00:37
                          via 192.168.1.3
  B EX 10.0.2.1/32        via 192.168.1.1                    20/0    00:00:37
  B EX 10.0.2.2/32        via 192.168.1.3                    20/0    00:03:37
  B EX 10.0.2.3/32        via 192.168.1.5                    20/0    00:03:31
  B EX 10.0.2.4/32        via 192.168.1.7                    20/0    00:03:23
  B EX 10.60.1.0/24       via 192.168.1.5                    20/0    00:03:19
                          via 192.168.1.7
  B EX 172.16.1.0/24      via 192.168.1.1                    20/0    00:00:37
                          via 192.168.1.3
```

S4048-Leaf1 has two paths to all other leafs and two paths to Server 2's network, 10.60.1.0. There is one path through each spine. If all paths do not appear, make sure the `maximum-paths` statement in the BGP configuration is equal to or greater than the number of spines in the topology.

```
S4048-Leaf1#show ip route bgp
       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  B EX 10.0.1.1/32        via 192.168.1.0                    20/0    00:03:56
  B EX 10.0.1.2/32        via 192.168.2.0                    20/0    00:07:02
  B EX 10.0.2.2/32        via 192.168.1.0                    20/0    00:03:56
                          via 192.168.2.0
  B EX 10.0.2.3/32        via 192.168.1.0                    20/0    00:03:56
```

```
                          via 192.168.2.0
  B EX 10.0.2.4/32        via 192.168.1.0                    20/0    00:03:56
                          via 192.168.2.0
  B EX 10.60.1.0/24       via 192.168.1.0                    20/0    00:03:56
                          via 192.168.2.0
```

Note: The command `show ip route <cr>` can also be used to verify the information above as well as static routes and direct connections.

7.3.4 show ip route ospf

On switches with OSPF configured, this command is used to verify the OSPF entries in the Routing Information Base (RIB). Entries with multiple paths shown are used with ECMP. The two server networks in this example, 10.60.1.0 and 172.16.1.0, each have two paths from Z9100-Spine1, one through each leaf.

The first set of routes with a subnet mask of /32 are the IPs configured for router IDs.

```
Z9100-Spine1#show ip route ospf
       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  O E2 10.0.1.2/32        via 192.168.1.3, Fo 1/2/1         110/20    16:46:28
                          via 192.168.1.5, Fo 1/3/1
  O E2 10.0.2.1/32        via 192.168.1.1, Fo 1/1/1         110/20    17:20:59
  O E2 10.0.2.2/32        via 192.168.1.3, Fo 1/2/1         110/20    17:20:59
  O E2 10.0.2.3/32        via 192.168.1.5, Fo 1/3/1         110/20    17:20:59
  O E2 10.0.2.4/32        via 192.168.1.7, Fo 1/4/1         110/20    17:20:59
  O E2 10.60.1.0/24       via 192.168.1.5, Fo 1/3/1         110/20    16:46:28
                          via 192.168.1.7, Fo 1/4/1
  O E2 172.16.1.0/24      via 192.168.1.1, Fo 1/1/1         110/20    16:46:28
                          via 192.168.1.3, Fo 1/2/1
  O    192.168.2.0/31     via 192.168.1.1, Fo 1/1/1          110/2    17:20:59
  O    192.168.2.2/31     via 192.168.1.3, Fo 1/2/1          110/2    17:20:59
  O    192.168.2.4/31     via 192.168.1.5, Fo 1/3/1          110/2    17:20:59
  O    192.168.2.6/31     via 192.168.1.7, Fo 1/4/1          110/2    17:20:59
```

S4048-Leaf1 has two paths to all other leafs and two paths to Server 2's network, 10.60.1.0. There is one path through each spine. If all paths do not appear, make sure the `maximum-paths` statement in the OSPF configuration is equal to or greater than the number of spines in the topology.

```
S4048-Leaf1#show ip route ospf
       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  O E2 10.0.1.1/32        via 192.168.1.0, Fo 1/49          110/20    17:30:11
  O E2 10.0.1.2/32        via 192.168.2.0, Fo 1/50          110/20    18:18:43
  O E2 10.0.2.2/32        via 192.168.1.0, Fo 1/49          110/20    17:30:11
                          via 192.168.2.0, Fo 1/50
  O E2 10.0.2.3/32        via 192.168.1.0, Fo 1/49          110/20    17:30:11
```

```
                          via 192.168.2.0, Fo 1/50
  O E2 10.0.2.4/32        via 192.168.1.0, Fo 1/49          110/20    17:30:11
                          via 192.168.2.0, Fo 1/50
  O E2 10.60.1.0/24       via 192.168.1.0, Fo 1/49          110/20    17:30:11
                          via 192.168.2.0, Fo 1/50
  O    192.168.1.2/31     via 192.168.1.0, Fo 1/49           110/2    17:30:11
  O    192.168.1.4/31     via 192.168.1.0, Fo 1/49           110/2    17:30:11
  O    192.168.1.6/31     via 192.168.1.0, Fo 1/49           110/2    17:30:11
  O    192.168.2.2/31     via 192.168.2.0, Fo 1/50           110/2    18:18:43
  O    192.168.2.4/31     via 192.168.2.0, Fo 1/50           110/2    18:18:43
  O    192.168.2.6/31     via 192.168.2.0, Fo 1/50           110/2    18:18:43
```

Note: The command `show ip route <cr>` can also be used to verify the information above as well as static routes and direct connections.

7.3.5 show bfd neighbors

This command may be used to verify BFD is properly configured and sessions are established as indicated by `Up` in the `State` column.

Note: The output shown below is for BGP configurations as indicated by a `B` in the `Clients` column. On OSPF configurations, the output is identical except there is an `O` in the `Clients` column.

```
Z9100-Spine-1#show bfd neighbors

*       - Active session role
B       - BGP
O       - OSPF

  LocalAddr      RemoteAddr          Interface  State Rx-int Tx-int Mult Clients
* 192.168.1.0    192.168.1.1         Fo 1/1/1   Up    100    100    3    B
* 192.168.1.2    192.168.1.3         Fo 1/2/1   Up    100    100    3    B
* 192.168.1.4    192.168.1.5         Fo 1/3/1   Up    100    100    3    B
* 192.168.1.6    192.168.1.7         Fo 1/4/1   Up    100    100    3    B

S4048-Leaf-1#show bfd neighbors

*       - Active session role
B       - BGP
O       - OSPF

  LocalAddr       RemoteAddr          Interface  State Rx-int Tx-int Mult Clients
* 192.168.1.1     192.168.1.0         Fo 1/49    Up    100    100    3    B
* 192.168.2.1     192.168.2.0         Fo 1/50    Up    100    100    3    B
```

7.3.6 show vlt brief

This command is used to validate VLT configuration status on leaf switches in this topology. The Inter-chassis link (ICL) Link Status, Heart Beat Status and VLT Peer Status must all be up. The role for one switch in the VLT pair is primary and its peer switch (not shown) is assigned the secondary role. Ensure Peer-Routing is shown as enabled.

```
S4048-Leaf-1#show vlt brief
 VLT Domain Brief
------------------
 Domain ID:                    127
 Role:                         Primary
 Role Priority:                32768
 ICL Link Status:              Up
 HeartBeat Status:             Up
 VLT Peer Status:              Up
 Local Unit Id:                0
 Version:                      6(7)
 Local System MAC address:     f4:8e:38:20:37:29
 Remote System MAC address:    f4:8e:38:20:54:29
 Remote system version:        6(7)
 Delay-Restore timer:          90 seconds
 Delay-Restore Abort Threshold: 60 seconds
 Peer-Routing :                Enabled
 Peer-Routing-Timeout timer:   0 seconds
 Multicast peer-routing timeout: 150 seconds
```

7.3.7 show vlt detail

This command is used to validate VLT LAG status on leaf switches in this topology. This command shows the status and active VLANs of all VLT LAGs (Port channel 1 in this example). The local and peer status must both be up.

```
S4048-Leaf-1#show vlt detail
Local LAG Id  Peer LAG Id  Local Status  Peer Status  Active VLANs
------------  -----------  ------------  -----------  -------------
1             1            UP            UP            50
```

7.3.8 show vlt mismatch

This command highlights configuration issues between VLT peers. Mismatch examples include incompatible VLT configuration settings, VLAN differences, different switch operating system versions and spanning-tree inconsistencies. There should be no output to this command on any switch configured for VLT. If there is, resolve the mismatch.

```
S4048-Leaf1#show vlt mismatch
S4048-Leaf1#
```

7.3.9 show uplink-state-group

This command is used to validate the UFD status on leaf switches in this topology. `Status: Enabled, Up` indicates UFD is enabled and no interfaces are currently disabled by UFD.

S4048-Leaf1#**show uplink-state-group**

```
Uplink State Group: 1   Status: Enabled, Up
```

If an interface happens to be disabled by UFD, the `show uplink-state-group` command output will appear as follows:

```
Uplink State Group: 1   Status: Enabled, Down
```

Note: When an interface has been disabled by UFD, the `show interfaces` *`interface`* command for affected interfaces indicates it is error-disabled as follows:

```
S4048-Leaf-1#show interfaces te 1/4
TenGigabitEthernet 1/4 is up, line protocol is down(error-disabled[UFD])
-- Output truncated –-
```

7.3.10 show spanning-tree rstp brief

This command validates spanning tree is enabled on the leaf switches. All interfaces are forwarding (`Sts` column shows `FWD`). One of the leaf switches (S4048-Leaf1 in this example) is the root bridge and sever-facing interfaces (Po 1 in this example) are edge ports.

```
S4048-Leaf1#show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 0, Address f48e.3820.3729
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 0, Address f48e.3820.3729
We are the root
Configured hello time 2, max age 20, forward delay 15

Interface                                                  Designated
 Name      PortID   Prio Cost    Sts         Cost        Bridge ID        PortID
---------- -------- ---- ------- ----------- ----- -------------------- --------
Po 1       128.2    128  1800    FWD(vlt)    0     0     f48e.3820.3729 128.2
Po 127     128.128  128  600     FWD(vltI)   0     0     f48e.3820.3729 128.128

Interface
 Name      Role   PortID   Prio Cost    Sts         Cost    Link-type Edge
---------- ------ -------- ---- ------- ----------- ------- --------- ----
Po 1       Desg   128.2    128  1800    FWD         0       (vlt) P2P Yes
Po 127     Desg   128.128  128  600     FWD         0       (vltI)P2P No
```

8 Example 2: Layer 3 with Dell EMC leaf and Cisco Nexus spine switches

In this section, the Dell EMC Networking Z9100-ON spines used in the previous example are replaced with Cisco Nexus 5600 series spines as shown in Figure 15. BGP and OSPF configuration examples are included. S4048-ON leaf switch configuration is identical to that covered in Section 7.1 and is not repeated in this section.

Note: The BGP ASNs and IP addresses defined in Section 6 are used here.

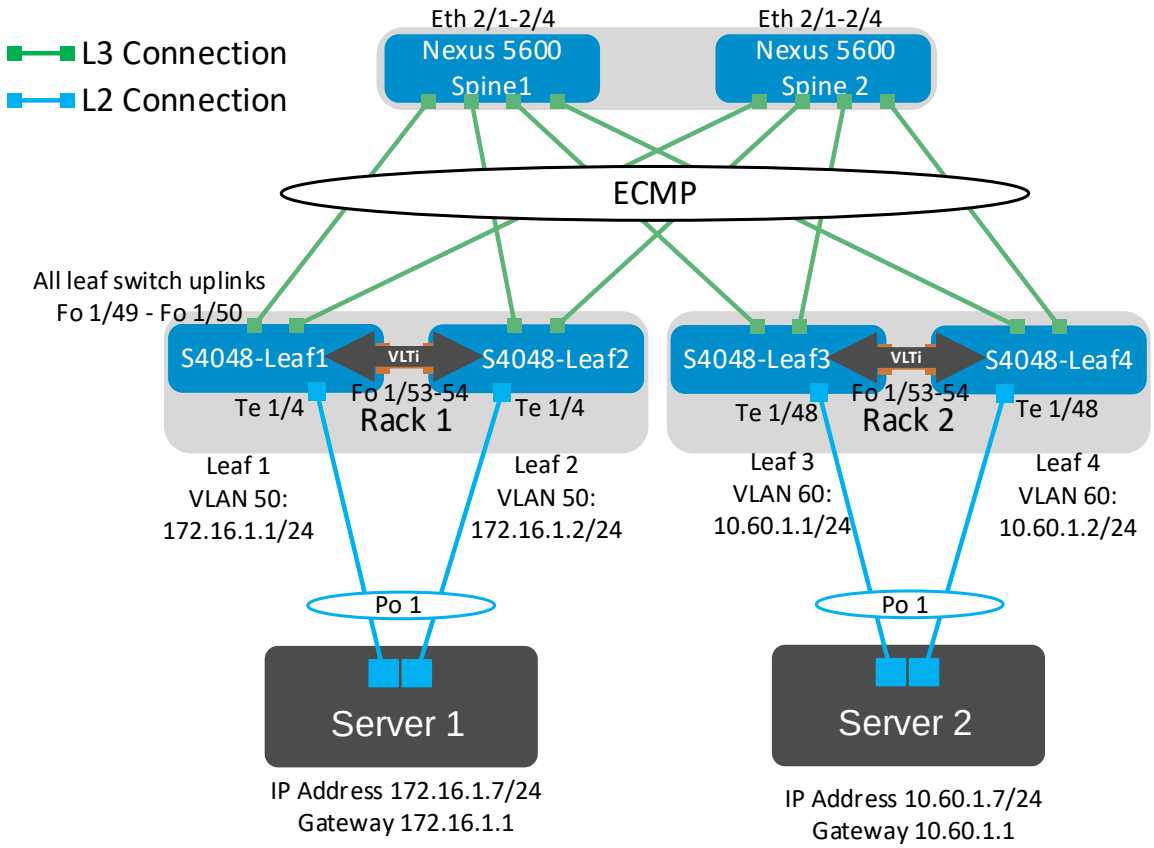

Figure 15 Example 2: Layer 3 leaf-spine topology with Dell EMC leaf and Cisco Nexus spine switches

Note: All switch configuration files for the topology in Figure 15 are contained in the attachment named **Example2_config_files.pdf.** The files may be edited as needed in a plain text editor and commands pasted directly into switch consoles.
Dell EMC Networking switches start at their factory default settings per Appendix A.
Cisco Nexus switches in this example were reset to their factory default configurations by running `write erase` followed by `reload`. After reload, "Power on Auto Provisioning" was not used, the admin password was configured and the Nexus “basic configuration dialog” was not used. Refer to your Nexus system documentation for more information.

8.1 Nexus 5600 series spine switch configuration

The following configuration details are for Nexus5600-Spine1 and Nexus5600-Spine2 in Figure 15.

First, set the hostname, enable LLDP and disable switchport as the default port type. Configure the management interface and default management route.

Nexus5600-Spine1	Nexus5600-Spine2

`enable`
`configure`
 	
`hostname Nexus5600-Spine1`
`feature lldp`
 	
`no system default switchport`
 	
`interface mgmt0`
`vrf member management`
`ip address 100.67.219.34/24`
 	
`vrf context management`
`ip route 0.0.0.0/0 100.67.219.254`
	`enable`
	`configure`
	
	`hostname Nexus5600-Spine2`
	`feature lldp`
	
	`no system default switchport`
	
	`interface mgmt0`
	`vrf member management`
	`ip address 100.67.219.33/24`
	
	`vrf context management`
	`ip route 0.0.0.0/0 100.67.219.254`

Configure the four point-to-point interfaces connected to leaf switches. Assign IP addresses per Table 1 (replacing destination interfaces Fo 1/1/1-1/4/1 in Table 1 with Nexus interfaces Ethernet 2/1-2/4). Configure a loopback interface to be used as the router ID. This is used with BGP or OSPF.

Finally, exit configuration mode and save the configuration with the `end` and `copy running-config startup-config` commands.

Nexus5600-Spine1	Nexus5600-Spine2

`interface ethernet 2/1`
`description Leaf 1 fo1/49`
`ip address 192.168.1.0/31`
`no shutdown`
 	
`interface ethernet 2/2`
`description Leaf 2 fo1/49`
`ip address 192.168.1.2/31`
`no shutdown`
 	
`interface ethernet 2/3`
`description Leaf 3 fo1/49`
`ip address 192.168.1.4/31`
`no shutdown`
 	
`interface ethernet 2/4`
`description Leaf 4 fo1/49`
`ip address 192.168.1.6/31`
`no shutdown`
	`interface ethernet 2/1`
	`description Leaf 1 fo1/50`
	`ip address 192.168.2.0/31`
	`no shutdown`
	
	`interface ethernet 2/2`
	`description Leaf 2 fo1/50`
	`ip address 192.168.2.2/31`
	`no shutdown`
	
	`interface ethernet 2/3`
	`description Leaf 3 fo1/50`
	`ip address 192.168.2.4/31`
	`no shutdown`
	
	`interface ethernet 2/4`
	`description Leaf 4 fo1/50`
	`ip address 192.168.2.6/31`
	`no shutdown`

Nexus5600-Spine1	Nexus5600-Spine2
interface loopback 0 description Router ID ip address 10.0.1.1/32 no shutdown end copy running-config startup-config	interface loopback 0 description Router ID ip address 10.0.1.2/32 no shutdown end copy running-config startup-config

8.1.1 Nexus 5600 series BGP configuration

Use these commands to configure BGP and BFD. If OSPF is used, skip to Section 8.1.2.

First, enable the BGP and BFD features.

Note: After running the command `feature bfd`, the message `Please disable the ICMP redirects on all interfaces running BFD sessions using the command 'no ip redirects'` may be displayed. This is done in the subsequent commands.

Run the `no ip redirects` command on the interfaces that will run BFD.

BGP is enabled with the `router bgp` *ASN* command. The ASN is from Figure 11.

The `bestpath as-path multipath-relax` command enables ECMP. The `maximum-paths 2` command specifies the maximum number of parallel paths to a destination to add to the routing table. In this topology, there are two equal cost best paths from a spine to a host, one to each leaf that the host is connected.

BGP neighbors are configured and BFD is enabled for each neighbor connection.

Finally, exit configuration mode and save the configuration with the `end` and `copy running-config startup-config` commands.

Note: On Nexus 5600 series switches, BGP graceful-restart, fast-external-fallover and BFD interval commands are configured by default.

Nexus5600-Spine1	Nexus5600-Spine2
enable configure feature bgp feature bfd interface ethernet 2/1-4 no ip redirects router bgp 64601 bestpath as-path multipath-relax address-family ipv4 unicast maximum-paths 2	enable configure feature bgp feature bfd interface ethernet 2/1-4 no ip redirects router bgp 64602 bestpath as-path multipath-relax address-family ipv4 unicast maximum-paths 2

Nexus5600-Spine1	Nexus5600-Spine2
neighbor 192.168.1.1 remote-as 64701 address-family ipv4 unicast bfd neighbor 192.168.1.3 remote-as 64702 address-family ipv4 unicast bfd neighbor 192.168.1.5 remote-as 64703 address-family ipv4 unicast bfd neighbor 192.168.1.7 remote-as 64704 address-family ipv4 unicast bfd end copy running-config startup-config	neighbor 192.168.2.1 remote-as 64701 address-family ipv4 unicast bfd neighbor 192.168.2.3 remote-as 64702 address-family ipv4 unicast bfd neighbor 192.168.2.5 remote-as 64703 address-family ipv4 unicast bfd neighbor 192.168.2.7 remote-as 64704 address-family ipv4 unicast bfd end copy running-config startup-config

8.1.2 Nexus 5600 series OSPF configuration

Use these commands to configure OSPF and BFD. Skip this section if BGP is used.

First, enable the OSPF and BFD features.

Note: After running the `feature bfd` command, the following message may be displayed: `Please disable the ICMP redirects on all interfaces running BFD sessions using the command 'no ip redirects'`. This is done in the subsequent commands.

OSPF is enabled with the `router ospf` *`process-tag`* command.

The maximum`-paths 2` command enables ECMP and specifies the maximum number of parallel paths to a destination to add to the routing table. In this topology, there are two equal cost best paths from a spine to a host, one to each leaf that the host is connected.

Run the `no ip redirects` command on the interfaces that will run BFD. Add the interfaces connected to the leaf switches to OSPF area 0. Enable BFD on the interfaces.

Finally, exit configuration mode and save the configuration with the `end` and `copy running-config startup-config` commands.

Nexus5600-Spine1	Nexus5600-Spine2
enable configure feature ospf feature bfd router ospf 1	enable configure feature ospf feature bfd router ospf 1

Nexus5600-Spine1	Nexus5600-Spine2
log-adjacency-changes maximum-paths 2 interface ethernet 2/1-4 no ip redirects ip router ospf 1 area 0 ip ospf bfd end copy running-config startup-config	log-adjacency-changes maximum-paths 2 interface ethernet 2/1-4 no ip redirects ip router ospf 1 area 0 ip ospf bfd end copy running-config startup-config

8.2 Example 2 validation

In addition to sending traffic between hosts, the configuration shown in Figure 15 can be validated with the commands shown in this section. For more information on commands and output, see the Command Line Reference Guide for the applicable switch (links to documentation are provided in Appendix C).

Command and output examples are provided for one spine and one leaf. Command output on other switches is similar.

8.2.1 show ip bgp summary

When BGP is configured, this command shows the status of all BGP connections. Each spine has four neighbors (the four leafs) and each leaf has two neighbors (the two spines). On Dell EMC switches, this command also confirms BFD is enabled on the 6th line of output.

```
Nexus5600-Spine1# show ip bgp summary
BGP summary information for VRF default, address family IPv4 Unicast
BGP router identifier 10.0.1.1, local AS number 64601
BGP table version is 59, IPv4 Unicast config peers 4, capable peers 4
6 network entries and 8 paths using 1024 bytes of memory
BGP attribute entries [4/576], BGP AS path entries [4/24]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.1.1     4 64701      89      91       59    0    0 00:46:58 2
192.168.1.3     4 64702      90      91       59    0    0 00:51:12 2
192.168.1.5     4 64703      91      94       59    0    0 00:47:07 2
192.168.1.7     4 64704      92      91       59    0    0 00:51:07 2

S4048-Leaf1#show ip bgp summary
BGP router identifier 10.0.2.1, local AS number 64701
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
6 network entrie(s) using 456 bytes of memory
11 paths using 1188 bytes of memory
BGP-RIB over all using 1199 bytes of memory
BFD is enabled, Interval 100 Min_rx 100 Multiplier 3 Role Active
13 BGP path attribute entrie(s) using 2064 bytes of memory
```

```
11 BGP AS-PATH entrie(s) using 110 bytes of memory
2 neighbor(s) using 16384 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down
State/Pfx
192.168.1.0     64601              54       58          0    0     0 00:47:30 5
192.168.2.0     64602              59       66          0    0     0 00:33:00 4
```

8.2.2 show ip ospf neighbor

When OSPF is configured, this command shows the state of all connected OSPF neighbors. In this configuration, each spine has four neighbors (the four leafs) and each leaf has two neighbors (the two spines).

```
Nexus5600-Spine1# show ip ospf neighbor
 OSPF Process ID 1 VRF default
 Total number of neighbors: 4
 Neighbor ID     Pri State            Up Time  Address         Interface
 10.0.2.1          1 FULL/DR          00:22:25 192.168.1.1     Eth2/1
 10.0.2.2          1 FULL/DR          00:22:05 192.168.1.3     Eth2/2
 10.0.2.3          1 FULL/DR          00:21:56 192.168.1.5     Eth2/3
 10.0.2.4          1 FULL/DR          00:21:47 192.168.1.7     Eth2/4

S4048-Leaf1#show ip ospf neighbor

Neighbor ID    Pri      State        Dead Time Address        Interface   Area
10.0.1.1         1      FULL/BDR     00:00:33  192.168.1.0    Fo 1/49     0
10.0.1.2         1      FULL/BDR     00:00:39  192.168.2.0    Fo 1/50     0
```

Note: All neighbor states should be FULL. If a neighbor is stuck in EXSTART or EXCHANGE, there may be an MTU setting mistmatch between the two connected interfaces.

8.2.3 show ip route bgp

On switches with BGP configured, this command is used to verify the BGP entries in the Routing Information Base (RIB). Entries with multiple paths shown are used with ECMP. The two server networks in this example, 10.60.1.0 and 172.16.1.0, each have two best paths from Nexus5600-Spine1, one through each leaf.

Note: The first set of routes with a subnet mask of /32 are the IPs configured for router IDs.

```
Nexus5600-Spine1# show ip route bgp-64601
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>
```

```
10.0.2.1/32, ubest/mbest: 1/0
    *via 192.168.1.1, [20/0], 00:51:59, bgp-64601, external, tag 64701,
10.0.2.2/32, ubest/mbest: 1/0
    *via 192.168.1.3, [20/0], 00:56:12, bgp-64601, external, tag 64702,
10.0.2.3/32, ubest/mbest: 1/0
    *via 192.168.1.5, [20/0], 00:52:07, bgp-64601, external, tag 64703,
10.0.2.4/32, ubest/mbest: 1/0
    *via 192.168.1.7, [20/0], 00:56:08, bgp-64601, external, tag 64704,
10.60.1.0/24, ubest/mbest: 2/0
    *via 192.168.1.5, [20/0], 00:52:07, bgp-64601, external, tag 64703,
    *via 192.168.1.7, [20/0], 00:52:23, bgp-64601, external, tag 64704,
172.16.1.0/24, ubest/mbest: 2/0
    *via 192.168.1.1, [20/0], 00:51:59, bgp-64601, external, tag 64701,
    *via 192.168.1.3, [20/0], 00:53:42, bgp-64601, external, tag 64702,
```

S4048-Leaf1 has two paths to all other leafs and two paths to Server 2's network, 10.60.1.0. There is one path through each spine. If all paths do not appear, make sure the `maximum-paths` statement in the BGP configuration is equal to or greater than the number of spines in the topology.

```
S4048-Leaf1#show ip route bgp
       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  B EX 10.0.2.2/32        via 192.168.1.0                     20/0    00:39:04
                          via 192.168.2.0
  B EX 10.0.2.3/32        via 192.168.1.0                     20/0    00:39:04
                          via 192.168.2.0
  B EX 10.0.2.4/32        via 192.168.1.0                     20/0    00:39:03
                          via 192.168.2.0
  B EX 10.60.1.0/24       via 192.168.1.0                     20/0    00:39:04
                          via 192.168.2.0
```

8.2.4 show ip route ospf

On switches with OSPF configured, this command is used to verify the OSPF entries in the Routing Information Base (RIB). Entries with multiple paths shown are used with ECMP. The two server networks in this example, 10.60.1.0 and 172.16.1.0, each have two best paths from Nexus5600-Spine1, one through each leaf.

The first set of routes with a subnet mask of /32 are the IPs configured for router IDs.

```
Nexus5600-Spine1# show ip route ospf
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>
```

```
10.0.2.1/32, ubest/mbest: 1/0
    *via 192.168.1.1, Eth2/1, [110/20], 00:32:09, ospf-1, type-2
10.0.2.2/32, ubest/mbest: 1/0
    *via 192.168.1.3, Eth2/2, [110/20], 00:31:49, ospf-1, type-2
10.0.2.3/32, ubest/mbest: 1/0
    *via 192.168.1.5, Eth2/3, [110/20], 00:31:42, ospf-1, type-2
10.0.2.4/32, ubest/mbest: 1/0
    *via 192.168.1.7, Eth2/4, [110/20], 00:31:30, ospf-1, type-2
10.60.1.0/24, ubest/mbest: 2/0
    *via 192.168.1.5, Eth2/3, [110/20], 00:31:30, ospf-1, type-2
    *via 192.168.1.7, Eth2/4, [110/20], 00:31:30, ospf-1, type-2
172.16.1.0/24, ubest/mbest: 2/0
    *via 192.168.1.1, Eth2/1, [110/20], 00:31:49, ospf-1, type-2
    *via 192.168.1.3, Eth2/2, [110/20], 00:31:49, ospf-1, type-2
192.168.2.0/31, ubest/mbest: 1/0
    *via 192.168.1.1, Eth2/1, [110/2], 00:32:09, ospf-1, intra
192.168.2.2/31, ubest/mbest: 1/0
    *via 192.168.1.3, Eth2/2, [110/2], 00:31:49, ospf-1, intra
192.168.2.4/31, ubest/mbest: 1/0
    *via 192.168.1.5, Eth2/3, [110/2], 00:31:42, ospf-1, intra
192.168.2.6/31, ubest/mbest: 1/0
    *via 192.168.1.7, Eth2/4, [110/2], 00:31:30, ospf-1, intra
```

S4048-Leaf1 has two paths to all other leafs and two paths to Server 2's network, 10.60.1.0. There is one path through each spine. If all paths do not appear, make sure the `maximum-paths` statement in the OSPF configuration is equal to or greater than the number of spines in the topology.

```
S4048-Leaf1#show ip route ospf
       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  O E2 10.0.2.2/32        via 192.168.1.0, Fo 1/49          110/20    00:34:45
                          via 192.168.2.0, Fo 1/50
  O E2 10.0.2.3/32        via 192.168.1.0, Fo 1/49          110/20    00:34:45
                          via 192.168.2.0, Fo 1/50
  O E2 10.0.2.4/32        via 192.168.1.0, Fo 1/49          110/20    00:34:35
                          via 192.168.2.0, Fo 1/50
  O E2 10.60.1.0/24       via 192.168.1.0, Fo 1/49          110/20    00:34:45
                          via 192.168.2.0, Fo 1/50
  O    192.168.1.2/31     via 192.168.1.0, Fo 1/49           110/2    00:35:16
  O    192.168.1.4/31     via 192.168.1.0, Fo 1/49           110/2    00:35:16
  O    192.168.1.6/31     via 192.168.1.0, Fo 1/49           110/2    00:35:16
  O    192.168.2.2/31     via 192.168.2.0, Fo 1/50           110/2    00:34:45
  O    192.168.2.4/31     via 192.168.2.0, Fo 1/50           110/2    00:35:16
  O    192.168.2.6/31     via 192.168.2.0, Fo 1/50           110/2    00:35:16
```

8.2.5 show bfd neighbors

This command may be used to verify BFD is properly configured and sessions are established as indicated by `Up` in the `RH` (Remote Heard) and `State` columns on the Nexus spine and `Up` in the `State` column on the Dell EMC leaf.

Note: The output for S4048-Leaf1 shown is for BGP configurations as indicated by a `B` in the `Clients` column. On OSPF configurations, the output is identical except there is an `O` in the `Clients` column. Nexus spine output is the same for either protocol.

```
Nexus5600-Spine1# show bfd neighbors

OurAddr       NeighAddr       RH/RS    Holdown(mult)   State   Int       Vrf
192.168.1.2   192.168.1.3     Up       219(3)          Up      Eth2/2    default
192.168.1.4   192.168.1.5     Up       211(3)          Up      Eth2/3    default
192.168.1.0   192.168.1.1     Up       276(3)          Up      Eth2/1    default
192.168.1.6   192.168.1.7     Up       202(3)          Up      Eth2/4    default

S4048-Leaf1#show bfd neighbors

*       - Active session role
B       - BGP
O       - OSPF

  LocalAddr       RemoteAddr         Interface  State Rx-int Tx-int Mult Clients
* 192.168.1.1     192.168.1.0        Fo 1/49    Up    100    100    3    B
* 192.168.2.1     192.168.2.0        Fo 1/50    Up    100    100    3    B
```

8.2.6 Dell EMC Networking leaf validation commands previously covered

The following commands previously covered may be run to validate the Dell EMC Networking leaf switches for this configuration. The output is the same or similar to that shown in the referenced sections.

`show vlt brief` – see Section 7.3.6

`show vlt detail` – see Section 7.3.7

`show vlt mismatch` – see Section 7.3.8

`show uplink-state-group` – see Section 7.3.9

`show spanning-tree rstp brief` – see Section 7.3.10

9 Example 3: Layer 2 with Dell EMC leaf and spine switches

This section provides configuration information to build the layer 2 leaf-spine topology shown in Figure 16. Dell EMC Networking S4048-ON switches are used at the leaf layer and Dell EMC Networking S6010-ON switches are used at the spine layer.

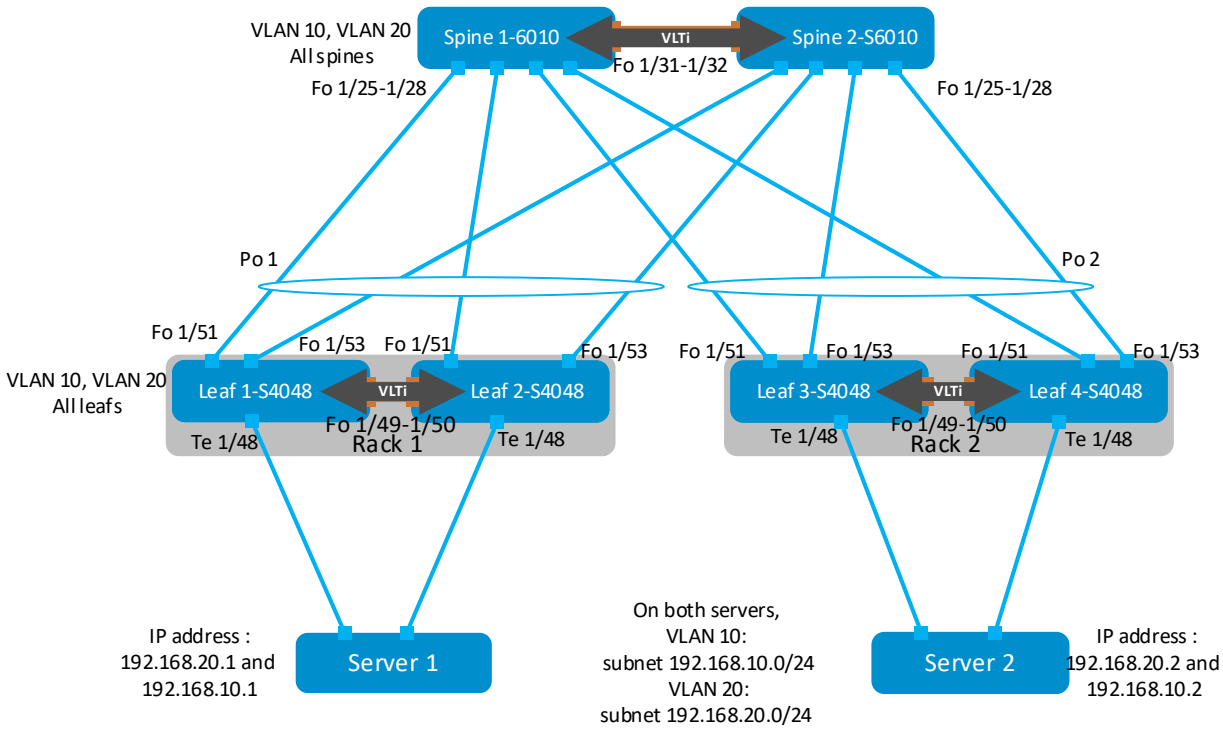

Figure 16 Example 3: Layer 2 leaf-spine topology with Dell EMC leaf and spine switches

Note: All switch configuration files for the topology in Figure 16 are contained in the attachment named **Example3_config_files.pdf.** The files may be edited as needed in a plain text editor and commands pasted directly into switch consoles. Dell EMC Networking switches start at their factory default settings per Appendix A.

9.1 S4048-ON leaf switch configuration

The following sections outline the configuration commands issued to the S4048-ON leaf switches to build the topology in. The commands detailed below are for L2-Leaf1-S4048 and L2-Leaf2-S4048. The configuration commands for L2-Leaf3-S4048 and L2-Leaf4-S4048 are similar and are provided in the attachments.

Note: On S4048-ON switches, Telnet is enabled and SSH is disabled by default. Both services require the creation of a non-root user account to login. If needed, it is a best practice to use SSH instead of Telnet for security. SSH can optionally be enabled with the command: `(conf)#ip ssh server enable.`
A user account can be created to access the switch via SSH with the command
`(conf)#username ssh_user sha256-password ssh_password`

First, configure the serial console enable password and disable Telnet.

L2-Leaf1-S4048	L2-Leaf2-S4048
enable configure enable sha256-password *enable_password* no ip telnet server enable	enable configure enable sha256-password *enable_password* no ip telnet server enable

Set the hostname, configure the OOB management interface and default gateway. Enable LLDP. Enable RSTP as a precaution.

Note: In this layer 2 topology, the RSTP root bridge is configured at the spine level.

L2-Leaf1-S4048	L2-Leaf2-S4048
hostname L2-Leaf1-S4048 interface ManagementEthernet 1/1 ip address 100.67.194.1/24 no shutdown management route 0.0.0.0/0 100.67.194.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable	hostname L2-Leaf2-S4048 interface ManagementEthernet 1/1 ip address 100.67.194.2/24 no shutdown management route 0.0.0.0/0 100.67.194.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable

Configure the VLT interconnect between Leaf1 and Leaf2. In this configuration, add interfaces fortyGigE 1/49-50 to static port channel 127 for the VLT interconnect. The backup destination is the management IP address of the VLT peer switch.

Note: Dell EMC recommends that the VLTi is configured as a static LAG (without LACP) per the commands shown below.

L2-Leaf1-S4048	L2-Leaf2-S4048
interface Port-channel 127 description VLTi Port-Channel no ip address channel-member fortyGigE 1/49,1/50 no shutdown interface range fortyGigE 1/49 - 1/50 description VLTi no ip address no shutdown	interface Port-channel 127 description VLTi Port-Channel no ip address channel-member fortyGigE 1/49,1/50 no shutdown interface range fortyGigE 1/49 - 1/50 description VLTi no ip address no shutdown

L2-Leaf1-S4048	L2-Leaf2-S4048
vlt domain 127 peer-link port-channel 127 back-up destination 100.67.194.2 unit-id 0	vlt domain 127 peer-link port-channel 127 back-up destination 100.67.194.1 unit-id 1

Interface Te 1/48 connects downstream to Server 1 and is configured as an RSTP edge port. Interfaces Fo 1/51 and Fo 1/53 connect to the spines upstream and are configured in LACP port channel 1. The port channel is configured for VLT.

L2-Leaf1-S4048	L2-Leaf2-S4048
interface TenGigabitEthernet 1/48 description Server 1 no ip address portmode hybrid switchport spanning-tree rstp edge-port no shutdown interface fortyGigE 1/51 description Spine1-Port25 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/53 description Spine2-Port25 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface Port-channel 1 description To Spines no ip address portmode hybrid switchport vlt-peer-lag port-channel 1 no shutdown	interface TenGigabitEthernet 1/48 description Server 1 no ip address portmode hybrid switchport spanning-tree rstp edge-port no shutdown interface fortyGigE 1/51 description Spine1-Port26 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/53 description Spine2-Port26 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface Port-channel 1 description To Spines no ip address portmode hybrid switchport vlt-peer-lag port-channel 1 no shutdown

VLANs 10 and 20 are configured on each switch. Port-channel 1 is tagged in both VLANs.

Note: The `shutdown/no shutdown` commands on a VLAN have no effect unless the VLAN is assigned an IP address (configured as an SVI).

L2-Leaf1-S4048	L2-Leaf2-S4048
interface Vlan 10 no ip address tagged TenGigabitEthernet 1/48 tagged Port-channel 1	interface Vlan 10 no ip address tagged TenGigabitEthernet 1/48 tagged Port-channel 1

L2-Leaf1-S4048	L2-Leaf2-S4048
shutdown interface Vlan 20 no ip address tagged TenGigabitEthernet 1/48 tagged Port-channel 1 shutdown	shutdown interface Vlan 20 no ip address tagged TenGigabitEthernet 1/48 tagged Port-channel 1 shutdown

Configure UFD. This shuts the downstream interfaces if all uplinks fail. The hosts attached to the switch use the remaining LACP port member to continue sending traffic across the fabric.

Finally, exit configuration mode and save the configuration with the `end` and `write` commands.

L2-Leaf1-S4048	L2-Leaf2-S4048
uplink-state-group 1 description Disable all edge ports in event all spines uplinks fail downstream TenGigabitEthernet 1/1-1/48 upstream Port-channel 1 end write	uplink-state-group 1 description Disable all edge ports in event all spines uplinks fail downstream TenGigabitEthernet 1/1-1/48 upstream Port-channel 1 end write

9.2 S6010-ON spine configuration

The following sections outline the configuration commands issued to the S6010-ON spine switches to build the topology in Figure 16.

Note: On S6010-ON switches, Telnet is enabled and SSH is disabled by default. Both services require the creation of a non-root user account to login. If needed, it is a best practice to use SSH instead of Telnet for security. SSH can optionally be enabled with the command: `(conf)#`**`ip ssh server enable.`**
A user account can be created to access the switch via SSH with the command
`(conf)#`**`username`** ***`ssh_user`*** **`sha256-password`** ***`ssh_password`***

First, configure the serial console enable password and disable Telnet.

L2-Spine1-S6010	L2-Spine2-S6010
enable configure enable sha256-password *enable_password* no ip telnet server enable	enable configure enable sha256-password *enable_password* no ip telnet server enable

Set the hostname, configure the OOB management interface and default gateway. Enable LLDP.

Enable RSTP as a precaution. L2-Spine1-S6010 is configured as the primary RSTP root bridge using the `bridge-priority 0` command. L2-Spine2-S6010 is configured as the secondary RSTP root bridge using the `bridge-priority 4096` command.

L2-Spine1-S6010	L2-Spine2-S6010
hostname L2-Spine1-S6010 interface ManagementEthernet 1/1 ip address 100.67.194.15/24 no shutdown management route 0.0.0.0/0 100.67.194.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable bridge-priority 0	hostname L2-Spine2-S6010 interface ManagementEthernet 1/1 ip address 100.67.194.16/24 no shutdown management route 0.0.0.0/0 100.67.194.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable bridge-priority 4096

Configure the VLT interconnect between Spine1 and Spine2. In this configuration, add interfaces fortyGigE 1/31-32 to static port channel 127 for the VLT interconnect. The backup destination is the management IP address of the VLT peer switch.

Note: Dell EMC recommends that the VLTi is configured as a static LAG (without LACP) per the commands shown below.

L2-Spine1-S6010	L2-Spine2-S6010
interface Port-channel 127 description VLTi Port-Channel no ip address channel-member fortyGigE 1/31,1/32 no shutdown interface range fortyGigE 1/31 - 1/32 description VLTi no ip address no shutdown vlt domain 127 peer-link port-channel 127 back-up destination 100.67.194.16 unit-id 0	interface Port-channel 127 description VLTi Port-Channel no ip address channel-member fortyGigE 1/31,1/32 no shutdown interface range fortyGigE 1/31 - 1/32 description VLTi no ip address no shutdown vlt domain 127 peer-link port-channel 127 back-up destination 100.67.194.15 unit-id 1

Interfaces Fo 1/25-28 connect to the leaf switches downstream via LACP port channels. Port channel 1 has members Fo 1/25 and Fo 1/26 and port channel 2 has members Fo 1/27 and Fo 1/28. The port channels are configured for VLT.

L2-Spine1-S6010	L2-Spine2-S6010
interface fortyGigE 1/25 description Leaf1-Port51 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/26 description Leaf2-Port51 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/27 description Leaf3-Port51 no ip address port-channel-protocol LACP port-channel 2 mode active no shutdown interface fortyGigE 1/28 description Leaf4-Port51 no ip address port-channel-protocol LACP port-channel 2 mode active no shutdown interface Port-channel 1 description Leaf 1 & 2 no ip address portmode hybrid switchport vlt-peer-lag port-channel 1 no shutdown interface Port-channel 2 description Leaf 3 & 4 no ip address portmode hybrid switchport vlt-peer-lag port-channel 2 no shutdown	interface fortyGigE 1/25 description Leaf1-Port53 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/26 description Leaf2-Port53 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/27 description Leaf3-Port53 no ip address port-channel-protocol LACP port-channel 2 mode active no shutdown interface fortyGigE 1/28 description Leaf4-Port53 no ip address port-channel-protocol LACP port-channel 2 mode active no shutdown interface Port-channel 1 description Leaf 1 & 2 no ip address portmode hybrid switchport vlt-peer-lag port-channel 1 no shutdown interface Port-channel 2 description Leaf 3 & 4 no ip address portmode hybrid switchport vlt-peer-lag port-channel 2 no shutdown

VLANs 10 and 20 are configured on each switch. Port-channels 1 and 2 are tagged in both VLANs.

Note: The `shutdown/no shutdown` commands on a VLAN have no effect unless the VLAN is assigned an IP address (configured as an SVI).

Finally, exit configuration mode and save the configuration with the `end` and `write` commands.

L2-Spine1-S6010	L2-Spine2-S6010
interface Vlan 10 no ip address tagged Port-channel 1-2 shutdown interface Vlan 20 no ip address tagged Port-channel 1-2 shutdown end write	interface Vlan 10 no ip address tagged Port-channel 1-2 shutdown interface Vlan 20 no ip address tagged Port-channel 1-2 shutdown end write

9.3 Example 3 validation

In addition to sending traffic between hosts, the configuration shown in Figure 16 can be validated with the commands shown in this section. For more information on commands and output, see the Command Line Reference Guide for the applicable switch (links to documentation are provided in Appendix C).

Command and output examples are provided for one spine and one leaf. Command output on other switches is similar.

9.3.1 show vlt brief

The Inter-chassis link (ICL) Link Status, Heart Beat Status and VLT Peer Status must all be up. The role for one switch in the VLT pair is primary and its peer switch (not shown) is assigned the secondary role.

```
L2-Spine1-S6010#show vlt brief
 VLT Domain Brief
------------------
 Domain ID:                    127
 Role:                         Primary
 Role Priority:                32768
 ICL Link Status:              Up
 HeartBeat Status:             Up
 VLT Peer Status:              Up
 Local Unit Id:                0
 Version:                      6(7)
 Local System MAC address:     f4:8e:38:2b:08:69
 Remote System MAC address:    f4:8e:38:2b:36:e9
 Remote system version:        6(7)
```

```
 Delay-Restore timer:            90 seconds
 Delay-Restore Abort Threshold:  60 seconds
 Peer-Routing :                  Disabled
 Peer-Routing-Timeout timer:     0 seconds
 Multicast peer-routing timeout: 150 seconds

L2-Leaf1-S4048#show vlt brief
 VLT Domain Brief
------------------
 Domain ID:                      127
 Role:                           Primary
 Role Priority:                  32768
 ICL Link Status:                Up
 HeartBeat Status:               Up
 VLT Peer Status:                Up
 Local Unit Id:                  0
 Version:                        6(7)
 Local System MAC address:       f4:8e:38:20:c5:29
 Remote System MAC address:      64:00:6a:e6:cc:14
 Remote system version:          6(7)
 Delay-Restore timer:            90 seconds
 Delay-Restore Abort Threshold:  60 seconds
 Peer-Routing :                  Disabled
 Peer-Routing-Timeout timer:     0 seconds
 Multicast peer-routing timeout: 150 seconds
```

9.3.2 show vlt detail

This command shows the status and active VLANs of all VLT LAGs (Port channels 1 and 2 in this example). The local and peer status must both be up.

```
L2-Spine1-S6010#show vlt detail
Local LAG Id  Peer LAG Id  Local Status  Peer Status  Active VLANs
------------  ----------  ------------  -----------  -------------
1             1           UP            UP            1, 10, 20
2             2           UP            UP            1, 10, 20

L2-Leaf1-S4048#show vlt detail
Local LAG Id  Peer LAG Id  Local Status  Peer Status  Active VLANs
------------  ----------  ------------  -----------  -------------
1             1           UP            UP            1, 10, 20
```

9.3.3 show vlt mismatch

This command highlights configuration issues between VLT peers. Mismatch examples include incompatible VLT configuration settings, VLAN differences, different switch operating system versions and spanning-tree inconsistencies. There should be no output to this command on any switch configured for VLT. If there is, resolve the mismatch.

```
L2-Spine1-S6010#show vlt mismatch
L2-Spine1-S6010#

L2-Leaf1-S4048#show vlt mismatch
L2-Leaf1-S4048#
```

9.3.4 show uplink-state-group

This command is used to validate the UFD status on leaf switches in this topology. `Status: Enabled, Up` indicates UFD is enabled and no interfaces are currently disabled by UFD.

S4048-Leaf1#**show uplink-state-group**

```
Uplink State Group: 1   Status: Enabled, Up
```

If an interface happens to be disabled by UFD, the `show uplink-state-group` command output will appear as follows:

```
Uplink State Group: 1   Status: Enabled, Down
```

Note: When an interface has been disabled by UFD, the `show interfaces` *`interface`* command for affected interfaces indicates it is error-disabled as follows:

```
S4048-Leaf-1#show interfaces te 1/48
TenGigabitEthernet 1/48 is up, line protocol is down(error-disabled[UFD])
-- Output truncated –-
```

9.3.5 show spanning-tree rstp brief

This command validates spanning tree is enabled. All interfaces are forwarding (`Sts` column shows `FWD`) because VLT is configured at the leaf and spine layers, eliminating the need for blocked ports. One of the spine switches (L2-Spine1-S6010 in this example) is the root bridge. Sever-facing interfaces on leaf switches (L2-Leaf1-S4048 interface Te 1/48 in this example) are edge ports.

```
L2-Spine1-S6010#show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 0, Address f48e.382b.0869
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 0, Address f48e.382b.0869
We are the root
Configured hello time 2, max age 20, forward delay 15

Interface                                                 Designated
 Name      PortID   Prio Cost    Sts         Cost        Bridge ID         PortID
---------- -------- ---- ------- ----------- ----- ------------------- --------
Po 1       128.2    128  188     FWD(vlt)    0     0     f48e.382b.0869 128.2
Po 2       128.3    128  188     FWD(vlt)    0     0     f48e.382b.0869 128.3
```

```
Po 127     128.128  128  600     FWD(vltI)   0     0     f48e.382b.0869 128.128

Interface
 Name      Role   PortID   Prio Cost    Sts          Cost    Link-type Edge
---------- ------ -------- ---- ------- ----------- ------- --------- ----
Po 1       Desg   128.2    128  188     FWD          0       (vlt) P2P No
Po 2       Desg   128.3    128  188     FWD          0       (vlt) P2P No
Po 127     Desg   128.128  128  600     FWD          0       (vltI)P2P No

L2-Leaf1-S4048#show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 0, Address f48e.382b.0869
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 32768, Address f48e.3820.c529
Configured hello time 2, max age 20, forward delay 15

Interface                                                  Designated
 Name      PortID   Prio Cost    Sts          Cost        Bridge ID        PortID
---------- -------- ---- ------- ----------- ----- ------------------- --------
Po 1       128.2    128  188     FWD(vlt)    788   32768 f48e.3820.c529 128.2
Po 127     128.128  128  600     FWD(vltI)   788   32768 6400.6ae6.cc14 128.128
Te 1/48    128.249  128  2000    FWD         788   32768 f48e.3820.c529 128.249

Interface
 Name      Role   PortID   Prio Cost    Sts          Cost    Link-type Edge
---------- ------ -------- ---- ------- ----------- ------- --------- ----
Po 1       Root   128.2    128  188     FWD          788     (vlt) P2P No
Po 127     Root   128.128  128  600     FWD          788     (vltI)P2P No
Te 1/48    Desg   128.249  128  2000    FWD          788     P2P       Yes
```

10 Example 4: Layer 2 with Dell EMC leaf and Cisco Nexus spine switches

This section provides configuration information to build the layer 2 leaf-spine topology shown in Figure 17. Dell EMC Networking Z9100-ON switches are used at the leaf layer and Cisco Nexus 7000 series switches are used at the spine layer.

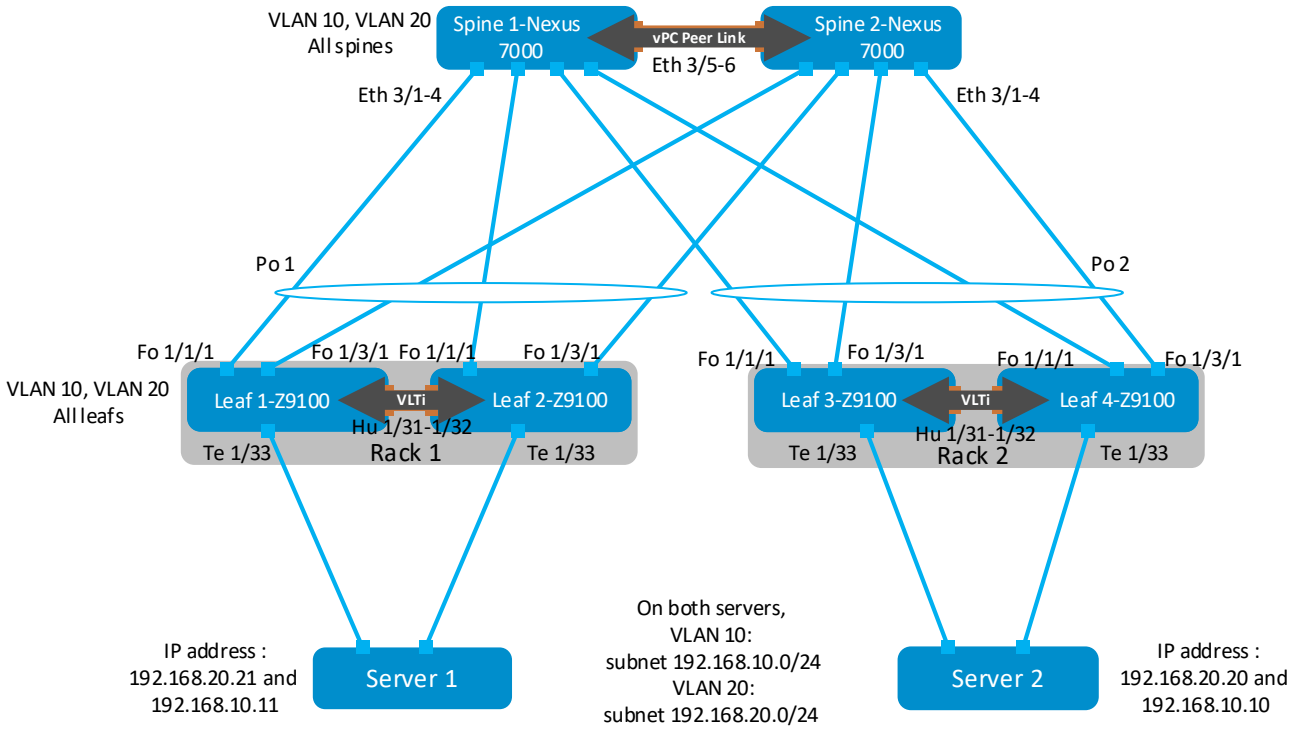

Figure 17 Example 4: Layer 2 leaf-spine topology with Dell EMC leaf and Cisco Nexus spine switches

Note: All switch configuration files for the topology in Figure 17 are contained in the attachment named **Example4_config_files.pdf.** The files may be edited as needed in a plain text editor and commands pasted directly into switch consoles.
Dell EMC Networking switches start at their factory default settings per Appendix A.
Cisco Nexus switches in this example were reset to their factory default configurations by running `write erase` followed by `reload`. After reload, "Power on Auto Provisioning" was not used, the admin password was configured and the Nexus “basic configuration dialog” was not used. Refer to your Nexus system documentation for more information.

10.1 Z9100-ON leaf switch configuration

The following section outlines the configuration commands issued to the Z9100-ON leaf switches to build the topology in Figure 17. The commands detailed below are for L2-Leaf1-Z9100 and L2-Leaf2-Z9100. The configuration commands for L2-Leaf3-Z9100 and L2-Leaf4-Z9100 are similar and are provided in the attachments.

Note: On Z9100-ON switches, Telnet is enabled and SSH is disabled by default. Both services require the creation of a non-root user account to login. If needed, it is a best practice to use SSH instead of Telnet for security. SSH can optionally be enabled with the command: `(conf)#ip ssh server enable.`
A user account can be created to access the switch via SSH with the command
`(conf)#username ssh_user sha256-password ssh_password`

First, configure the serial console enable password and disable Telnet.

L2-Leaf1-Z9100	L2-Leaf2-Z9100
enable configure enable sha256-password *enable_password* no ip telnet server enable	enable configure enable sha256-password *enable_password* no ip telnet server enable

Set the hostname, configure the OOB management interface and default gateway. Enable LLDP. Enable RSTP as a precaution.

Note: In this layer 2 topology, the RSTP root bridge is configured at the spine level.

L2-Leaf1-Z9100	L2-Leaf2-Z9100
hostname L2-Leaf1-Z9100 interface ManagementEthernet 1/1 ip address 100.67.194.5/24 no shutdown management route 0.0.0.0/0 100.67.194.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable	hostname L2-Leaf2-Z9100 interface ManagementEthernet 1/1 ip address 100.67.194.6/24 no shutdown management route 0.0.0.0/0 100.67.194.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable

Convert interfaces connected to the Nexus 5600 spines from their default speed of 100GbE to 40GbE.

L2-Leaf1-Z9100	L2-Leaf2-Z9100
stack-unit 1 port 1 portmode single speed 40G no-confirm stack-unit 1 port 3 portmode single speed 40G no-confirm	stack-unit 1 port 1 portmode single speed 40G no-confirm stack-unit 1 port 3 portmode single speed 40G no-confirm

Configure the VLT interconnect between Leaf1 and Leaf2. In this configuration, add interfaces hundredGigE 1/31 – 1/32 to static port channel 127 for the VLT interconnect. The backup destination is the management IP address of the VLT peer switch.

L2-Leaf1-Z9100	L2-Leaf2-Z9100
interface Port-channel 127 description VLTi Port-Channel no ip address channel-member hundredGigE 1/31,1/32 no shutdown interface range hundredGigE 1/31-1/32 description VLTi no ip address no shutdown vlt domain 127 peer-link port-channel 127 back-up destination 100.67.194.6 unit-id 0	interface Port-channel 127 description VLTi Port-Channel no ip address channel-member hundredGigE 1/31,1/32 no shutdown interface range hundredGigE 1/31-1/32 description VLTi no ip address no shutdown vlt domain 127 peer-link port-channel 127 back-up destination 100.67.194.5 unit-id 1

Interface Te 1/33 connects downstream to Server 1 and is configured as an RSTP edge port. Interfaces Fo 1/1/1 and Fo 1/3/1 connect to the spines upstream and are configured in LACP port channel 1. The port channel is configured for VLT.

L2-Leaf1-Z9100	L2-Leaf2-Z9100
interface TenGigabitEthernet 1/33 description Server-1 no ip address portmode hybrid switchport spanning-tree rstp edge-port no shutdown interface fortyGigE 1/1/1 description Spine1-Port1 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/3/1 description Spine2-Port1	interface TenGigabitEthernet 1/33 description Server-1 no ip address portmode hybrid switchport spanning-tree rstp edge-port no shutdown interface fortyGigE 1/1/1 description Spine1-Port2 no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface fortyGigE 1/3/1 description Spine2-Port2

no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface Port-channel 1 description To Spines no ip address portmode hybrid switchport vlt-peer-lag port-channel 1 no shutdown	no ip address port-channel-protocol LACP port-channel 1 mode active no shutdown interface Port-channel 1 description To Spines no ip address portmode hybrid switchport vlt-peer-lag port-channel 1 no shutdown

VLANs 10 and 20 are configured on each switch. Port-channel 1 is tagged in both VLANs.

Note: The `shutdown/no shutdown` commands on a VLAN have no effect unless the VLAN is assigned an IP address (configured as an SVI).

L2-Leaf1-Z9100	L2-Leaf2-Z9100
interface Vlan 10 no ip address tagged TenGigabitEthernet 1/33 tagged Port-channel 1 shutdown interface Vlan 20 no ip address tagged TenGigabitEthernet 1/33 tagged Port-channel 1 shutdown	interface Vlan 10 no ip address tagged TenGigabitEthernet 1/33 tagged Port-channel 1 shutdown interface Vlan 20 no ip address tagged TenGigabitEthernet 1/33 tagged Port-channel 1 shutdown

Configure UFD. This shuts the downstream interfaces if all uplinks fail. The hosts attached to the switch use the remaining LACP port member to continue sending traffic across the fabric.

Finally, exit configuration mode and save the configuration with the `end` and `write` commands.

L2-Leaf1-Z9100	L2-Leaf2-Z9100
uplink-state-group 1 description Disable edge port in event all spine uplinks fail downstream TenGigabitEthernet 1/33 upstream Port-channel 1 end write	uplink-state-group 1 description Disable edge port in event all spine uplinks fail downstream TenGigabitEthernet 1/33 upstream Port-channel 1 end write

10.2 Nexus 7000 series spine switch configuration

The following sections outline the configuration commands issued to the Nexus 7000 series switches to build the topology in Figure 17.

First, enable the LACP and virtual port channel (vPC) features. Configure the hostname, management IP address and default management route.

Note: Cisco enables Rapid Per VLAN Spanning Tree Plus (RPVST+), its implementation of RSTP, on Nexus 7000 series switches by default.

L2-Spine1-Nexus7K	L2-Spine2-Nexus7K
enable configure feature lacp feature vpc hostname L2-Spine1-Nexus7K interface mgmt0 vrf member management ip address 100.67.184.21/24 no shutdown vrf context management ip route 0.0.0.0/0 100.67.184.254	enable configure feature lacp feature vpc hostname L2-Spine2-Nexus7K interface mgmt0 vrf member management ip address 100.67.184.28/24 no shutdown vrf context management ip route 0.0.0.0/0 100.67.184.254

Create VLAN 10 and 20. All VLANs are added to RPVST+ as a precaution against loops. L2-Spine1-Nexus7K is configured as the primary spanning tree root bridge using the `spanning tree vlan` *`vlan_numbers`* `priority 0` command. L2-Spine1-Nexus7K is configured as the secondary spanning tree root bridge using the `spanning tree vlan` *`vlan_numbela`* `priority 4096` command.

L2-Spine1-Nexus7K	L2-Spine2-Nexus7K
vlan 10 vlan 20 spanning-tree vlan 1,10,20 spanning-tree vlan 1,10,20 priority 0	vlan 10 vlan 20 spanning-tree vlan 1,10,20 spanning-tree vlan 1,10,20 priority 4096

Create a vPC domain and vPC peer link between the two spine switches. On spine 1, assign a role priority of 1 to make it the vPC primary.

Specify the management IP address of the vPC peer as the vPC peer-keepalive destination. In this example, interfaces Ethernet 3/5 and 3/6 are used to create the vPC peer link. Interfaces are configured as trunk ports and allow applicable VLANs.

L2-Spine1-Nexus7K	L2-Spine2-Nexus7K
vpc domain 1 role priority 1 peer-keepalive destination 100.67.184.28 source 100.67.184.21 auto-recovery interface port-channel 20 switchport switchport mode trunk switchport trunk allowed vlan 1,10,20 spanning-tree port type network vpc peer-link interface Ethernet3/5 switchport switchport mode trunk switchport trunk allowed vlan 1,10,20 channel-group 20 mode active no shutdown interface Ethernet3/6 switchport switchport mode trunk switchport trunk allowed vlan 1,10,20 channel-group 20 mode active no shutdown	vpc domain 1 role priority 65535 peer-keepalive destination 100.67.184.21 source 100.67.184.28 auto-recovery interface port-channel 20 switchport switchport mode trunk switchport trunk allowed vlan 1,10,20 spanning-tree port type network vpc peer-link interface Ethernet3/5 switchport switchport mode trunk switchport trunk allowed vlan 1,10,20 channel-group 20 mode active no shutdown interface Ethernet3/6 switchport switchport mode trunk switchport trunk allowed vlan 1,10,20 channel-group 20 mode active no shutdown

Configure port channels and member ports for downstream connectivity to the leaf switches.

Finally, exit configuration mode and save the configuration with the `end` and `copy running-config startup-config` commands.

L2-Spine1-Nexus7K	L2-Spine2-Nexus7K
interface port-channel1 switchport switchport mode trunk vpc 1 interface port-channel2 switchport switchport mode trunk vpc 2 interface Ethernet3/1 switchport switchport mode trunk channel-group 1 mode active no shutdown interface Ethernet3/2 switchport switchport mode trunk channel-group 1 mode active no shutdown interface Ethernet3/3 switchport switchport mode trunk channel-group 2 mode active no shutdown interface Ethernet3/4 switchport switchport mode trunk channel-group 2 mode active no shutdown end copy running-config startup-config	interface port-channel1 switchport switchport mode trunk vpc 1 interface port-channel2 switchport switchport mode trunk vpc 2 interface Ethernet3/1 switchport switchport mode trunk channel-group 1 mode active no shutdown interface Ethernet3/2 switchport switchport mode trunk channel-group 1 mode active no shutdown interface Ethernet3/3 switchport switchport mode trunk channel-group 2 mode active no shutdown interface Ethernet3/4 switchport switchport mode trunk channel-group 2 mode active no shutdown end copy running-config startup-config

10.3 Example 4 validation

In addition to sending traffic between hosts, the configuration shown in Figure 17 can be validated with the commands shown in this section. For more information on commands and output, see the Command Line Reference Guide for the applicable switch (links to documentation are provided in Appendix C).

10.3.1 show vpc

This command displays the vPC status on the Nexus spine switches. Peer status and vPC keep-alive status must be as shown. The consistency status fields should all show "success". If not, see the `show vpc consistency-parameters` command.

```
L2-Spine1-Nexus7K# show vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 1
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 inconsistency reason       : Consistency Check Not Performed
vPC role                          : primary
Number of vPCs configured         : 2
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Enabled (timeout = 240 seconds)

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po20   up     1,10,20

vPC status
----------------------------------------------------------------------------
id   Port      Status Consistency Reason                Active vlans
--   ----      ------ ----------- ------                ------------
1    Po1       up     success     success                  1,10,20
2    Po2       up     success     success                  1,10,20
```

10.3.2 show vpc consistency-parameters

This command pinpoints inconsistencies between vPC peers on the Nexus spine switches. Depending on the severity of the misconfiguration, vPC may either warn the user (Type-2 misconfiguration) or suspend the port channel (Type-1 misconfiguration). In the specific case of a VLAN mismatch, only the VLAN that differs between the vPC member ports is suspended on the port channels.

```
L2-Spine1-Nexus7K# show vpc consistency-parameters ?
  global     Global Parameters
  interface  Specify interface
  vlans      Vlans
  vpc        Virtual Port Channel configuration

L2-Spine1-Nexus7K# show vpc consistency-parameters global

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Mode                    1     Rapid-PVST             Rapid-PVST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Allowed VLANs               -     1,10,20                1,10,20
Local error VLANs           -     -                      -
```

10.3.3 show spanning-tree

This command validates spanning tree is enabled on all VLANs on the Nexus spine switches and all interfaces are forwarding (`Sts` column shows `FWD`). One of the spine switches (L2-Spine1-Nexus7K in this example) is the root bridge on all VLANs.

```
L2-Spine1-Nexus7K# show spanning-tree

VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    1
             Address     8478.ac11.e341
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    1      (priority 0 sys-id-ext 1)
             Address     8478.ac11.e341
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Desg FWD 1         128.4096 (vPC) P2p
Po2              Desg FWD 1         128.4097 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC peer-link) Network P2p


VLAN0010
  Spanning tree enabled protocol rstp
  Root ID    Priority    10
             Address     8478.ac11.e341
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    10     (priority 0 sys-id-ext 10)
             Address     8478.ac11.e341
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Desg FWD 1         128.4096 (vPC) P2p
Po2              Desg FWD 1         128.4097 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC peer-link) Network P2p


VLAN0020
  Spanning tree enabled protocol rstp
  Root ID    Priority    20
```

```
             Address     8478.ac11.e341
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    20     (priority 0 sys-id-ext 20)
             Address     8478.ac11.e341
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Desg FWD 1         128.4096 (vPC) P2p
Po2              Desg FWD 1         128.4097 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC peer-link) Network P2p
```

10.3.4 Dell EMC Networking leaf validation commands previously covered

The following commands previously covered may be run to validate the Dell EMC Networking leaf switches for this configuration. The output is the same or similar to that shown in the referenced sections.

`show vlt brief` – see Section 9.3.1

`show vlt detail` – see Section 9.3.2

`show vlt mismatch` – see Section 9.3.3

`show uplink-state-group` – see Section 9.3.4

`show spanning-tree rstp brief` – see Section 9.3.5

A Dell EMC Networking switch factory default settings

All Dell EMC Networking switches in this guide can be reset to factory defaults as follows:

```
Dell#restore factory-defaults stack-unit unit# clear-all
Proceed with factory settings? Confirm [yes/no]:yes
```

Factory settings are restored and the switch reloads. After reload, enter **A** at the [A/C/L/S] prompt as shown below to exit Bare Metal Provisioning mode.

```
This device is in Bare Metal Provisioning (BMP) mode.
To continue with the standard manual interactive mode, it is necessary to
abort BMP.

Press A to abort BMP now.
Press C to continue with BMP.
Press L to toggle BMP syslog and console messages.
Press S to display the BMP status.
[A/C/L/S]:A

% Warning: The bmp process will stop ...

Dell>
```

The switch is now ready for configuration.

B Validated hardware and operating systems

The following table includes the hardware and operating systems used to validate the examples in this guide.

Table 2 Switches and operating systems used in this guide

Switch	OS / Version
Dell EMC Networking S3048-ON	DNOS 9.11.2.0 P0
Dell EMC Networking S4048-ON	DNOS 9.11.2.0 P0
Dell EMC Networking S6010-ON	DNOS 9.11.2.0 P0
Dell EMC Networking Z9100-ON	DNOS 9.11.2.0 P0
Cisco Nexus 7000 series (validated with Nexus 7004)	NX-OS 6.2(16)
Cisco Nexus 5600 series (validated with Nexus 5672UP)	NX-OS 7.3(2)N1(1)

C Technical support and resources

Dell EMC TechCenter is an online technical community where IT professionals have access to numerous resources for Dell EMC software, hardware and services.

Dell EMC TechCenter Networking Guides

Manuals and documentation for Dell EMC Networking S3048-ON

Manuals and documentation for Dell EMC Networking S4048-ON

Manuals and documentation for Dell EMC Networking S6010-ON

Manuals and documentation for Dell EMC Networking Z9100-ON

Cisco Nexus 5000 Series Switches Documentation

Cisco Nexus 7000 Series Switches Documentation

D Support and Feedback

Contacting Technical Support

Support Contact Information

Web: http://support.dell.com/

Telephone: USA: 1-800-945-3355

Feedback for this document

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