

# Dell EMC Networking with Isilon Front-End Deployment and Best Practices Guide

Dell EMC Networking Infrastructure Solutions  
June 2018

## Revisions

Date	Rev	Description	Authors
June 2018	1.0	Initial release	Gerald Myres, Jordan Wilson

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

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

Copyright © 2018 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA 6/25/2018

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

# Table of contents

1	Introduction .....	5
1.1	Typographical conventions .....	7
1.2	Attachments .....	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 Z9100-ON .....	8
2.4	Dell EMC PowerEdge R730xd .....	8
2.5	Dell EMC PowerEdge R640 .....	9
2.6	Isilon X210 .....	9
3	Management .....	10
4	Leaf-spine overview .....	11
4.1	Design considerations .....	11
4.2	Oversubscription .....	12
4.3	Scaling .....	13
4.4	Layer 3 leaf-spine topology .....	14
4.5	Layer 2 leaf-spine topology .....	15
5	Layer 3 Topology preparation .....	16
5.1	BGP ASN configuration .....	16
5.2	Loopback addresses .....	16
5.3	Point-to-point interfaces .....	17
6	Configuration of Layer 3 Topology .....	19
6.1	Configuration of Z9100-ON OS10EE Spine Switches .....	19
6.2	Configuration of S4048-ON OS10EE Leaf Switches .....	23
6.3	Configuration of S4048-ON OS9 Leaf Switches .....	32
7	Configuration of Layer 2 Topology .....	41
7.1	Configuration of Z9100-ON OS10EE Spine Switches .....	41
7.2	Configuration of S4048-ON OS10EE Leaf Switches .....	45
7.3	Configuration of S4048-ON OS9 Leaf Switches .....	50
8	Validation .....	58
8.1	OS10EE Validation commands .....	58
8.2	OS9 Validation commands .....	63

9	Isilon configuration .....	67
9.1	Building the cluster .....	67
9.2	Configuring LACP to each node .....	69
9.3	Configuring SMB share .....	72
9.4	Setting up DNS for SmartConnect.....	73
9.5	Validation .....	74
A	Validated hardware and components.....	76
A.1	Dell EMC Networking Switches .....	76
A.2	Dell EMC Isilon Array .....	76
A.3	Dell EMC PowerEdge Servers .....	76
B	Product Manuals and technical guides .....	77
C	Support and feedback .....	78

# 1 Introduction

Dell EMC Isilon is the leading scale-out network-attached storage (NAS) platform and offers the right blend of performance and capacity to support a wide range of unstructured data workloads including high-performance computing (HPC), file shares, home directories, archives, media content, video surveillance, and in-place data analytics. Isilon offers all-flash, hybrid, and archive storage systems that can be combined into a single cluster. All Isilon models are powered by the Isilon OneFS operating system. With its modular hardware design, Isilon solutions scale easily from tens of terabytes to tens of petabytes in a single cluster.

The OneFS operating system uses a single volume, single namespace, single file system architecture, making Isilon storage systems simple to install, manage and scale. And with automated, policy-based management options including data protection, replication, load-balancing, storage tiering and cloud integration, Isilon solutions remain simple to manage no matter how large the data environment becomes.

All nodes work together as peers in the cluster, leaving no single point of failure. As nodes are added, OneFS expands dynamically and redistributes data, eliminating the work of partitioning disks and creating volumes. Additionally, OneFS ensures that the workloads are dynamically reassigned when a failure occurs. This is achieved using the OneFS SmartConnect feature.

## ISILON SCALE-OUT NAS ARCHITECTURE

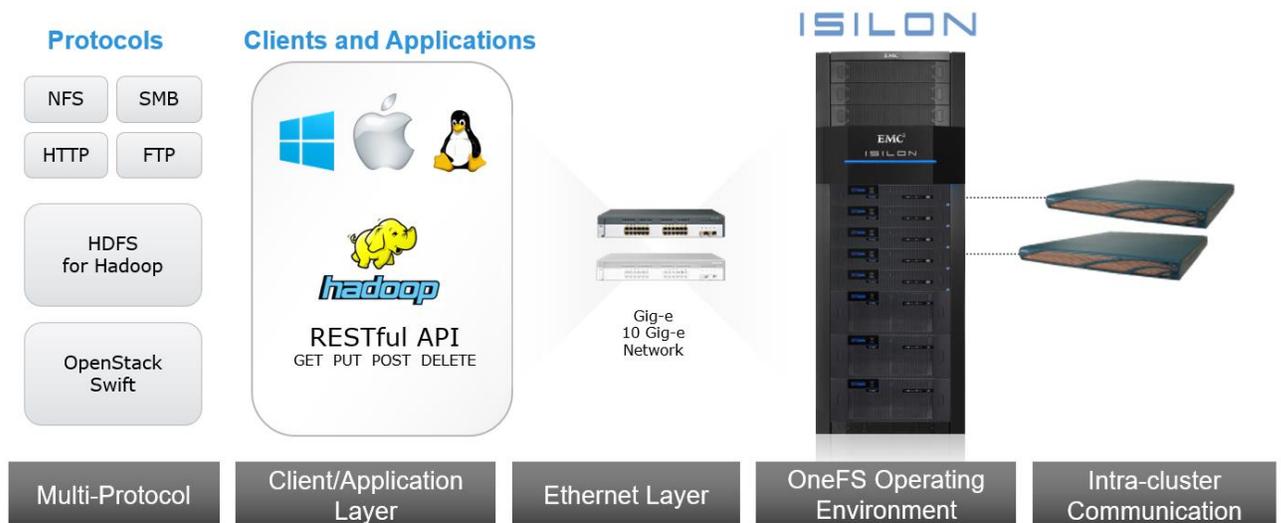


Figure 1 Isilon Scale-Out NAS

The Isilon Scale-out NAS utilizes Dell EMC Networking ethernet switches to provide the network. Dell EMC Networking OS9 and the new OS10 Enterprise Edition (OS10EE) are used in the provided examples. OS10EE is a disaggregated native Linux-based network operating system that has a fully disaggregated software architecture. OS10EE decouples the base software from the Layer 2 and Layer 3 protocol stack and services, which brings forth the ability for open programmability and portability. This allows for greater utilization of Dell EMC's Open Networking, in this guide we will utilize Dell EMC Networking operating systems.

However, Dell EMC Networking's legacy OS9 is still prevalent in the industry and supported on a large cross-section of the currently-shipping portfolio. This document encompasses the use of both operating systems within the same network architecture.

The Dell EMC S4048-ON will be used as leaf switches, and the Dell EMC Z9100-ON will be used in this guide as a spine switch. The Dell EMC S3048-ON will be used as a management switch for the iDRAC connections, out-of-band switch management, as well as the external Isilon connections. Z9100-ON and S4048-ON leaf pair in rack 1 will utilize Dell EMC Networking OS10EE whereas the S3048-ON and S4048-ON leaf pair in rack 2 will utilize OS9.

This guide will demonstrate how to utilize Dell EMC Networking, Dell EMC PowerEdge R730xd servers, and the flexibility of Dell EMC Isilon OneFS in two separate topologies. The examples provided will be using the Isilon X210. However, all Isilon storage systems utilize OneFS. This allows the steps outlined in this guide to be used regardless of the hardware that is available.

The first example will highlight the benefits of using the dynamic routing protocol Border Gateway Protocol (BGP) in a leaf/spine environment. This will be followed up with an example that displays the configuration and benefits of a layer 2 leaf/spine topology.

While the steps in this document were validated using the specified Dell EMC Networking switches and operating system(s), they may be leveraged for other Dell EMC Networking switch models utilizing the same networking OS version or later assuming the switch has the available port numbers, speeds, and types.

**Note:** For more specific details on deploying a spine-leaf architecture using Dell EMC Networking see [Dell EMC Leaf-Spine Deployment Guide](#) and [Dell EMC Networking L3 Design for Leaf-Spine with OS10EE](#) for more information.

In addition to covering the details of the network configuration, specific configurations within OneFS will be discussed. The focus will be on the front-end networking configurations, the back-end network that Isilon utilizes is beyond the scope of this guide. The main configurations discussed within Isilon's OneFS will be as follows:

- Creation of an Isilon cluster
- Adding additional nodes into the cluster
- Creating an SMB share
- Setting up LACP for each node in the cluster
- The configuration of the SmartConnect feature

## 1.1 Typographical conventions

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

Monospace Text	CLI examples
<u>Underlined Monospace Text</u>	CLI examples that wrap the page. This text is entered as a single command.
<i>Italic Monospace Text</i>	Variables in CLI examples
<b>Bold Monospace Text</b>	Commands entered at the CLI prompt
<b>Bold text</b>	GUI fields and information entered in the GUI

## 1.2 Attachments

This .pdf includes switch configuration file attachments. To access attachments in Adobe Acrobat Reader, click the  icon in the left pane halfway down the page, then click the  icon.

## 2 Hardware overview

This section briefly describes the hardware used to validate the deployment example in this guide. [Appendix A](#) contains a complete listing of hardware and components.

### 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 supports management traffic in each rack.



Figure 2 Dell EMC Networking S3048-ON

### 2.2 Dell EMC Networking S4048-ON

The S4048-ON is a 1RU switch with forty-eight 10GbE SFP+ ports and six ports of 40GbE. In this guide, this switch is deployed as a leaf switch performing basic gateway functionality for attached Windows servers and Isilon nodes.



Figure 3 Dell EMC Networking S4048-ON

### 2.3 Dell EMC Networking Z9100-ON

The Dell EMC Networking Z9100-ON is a 1RU, multilayer switch with thirty-two ports supporting 10/25/40/50/100GbE as well as two 10GbE ports. The Z9100-ON is used as a spine in this guide.



Figure 4 Dell EMC Networking Z9100-ON

### 2.4 Dell EMC PowerEdge R730xd

The Dell EMC PowerEdge R730xd is a 2-RU, two-socket server platform. It allows up to 32 x 2.5" SSDs or HDDs with SAS, SATA, and NVMe support. In this guide, two R730xd servers are used to connect to the Isilon storage.



Figure 5 Dell EMC PowerEdge R730xd

## 2.5 Dell EMC PowerEdge R640

This R640 server is a 2-socket, 1 RU server. This server is used in the second rack to connect to the Isilon storage.



Figure 6 Dell EMC PowerEdge R640

## 2.6 Isilon X210

The Dell EMC Isilon X210 is a 2-RU, scale-out NAS system. The Isilon X210 is used to create a storage cluster to support an SMB share in this guide.



Figure 7 Dell EMC Isilon X210 front view



Figure 8 Dell EMC Isilon X210 rear view

### 3 Management

The S3048-ON is used as a Top of Rack (ToR) switch, that aggregates all of the management connections. Each switch has a connection from the out-of-band (OOB) management port to the ToR, as well as each Isilon Node and PowerEdge server. The first Isilon node management connections are shown, each additional node is cabled similarly. This can be seen in Figure 9 below, all equipment in other racks is configured in a similar fashion.

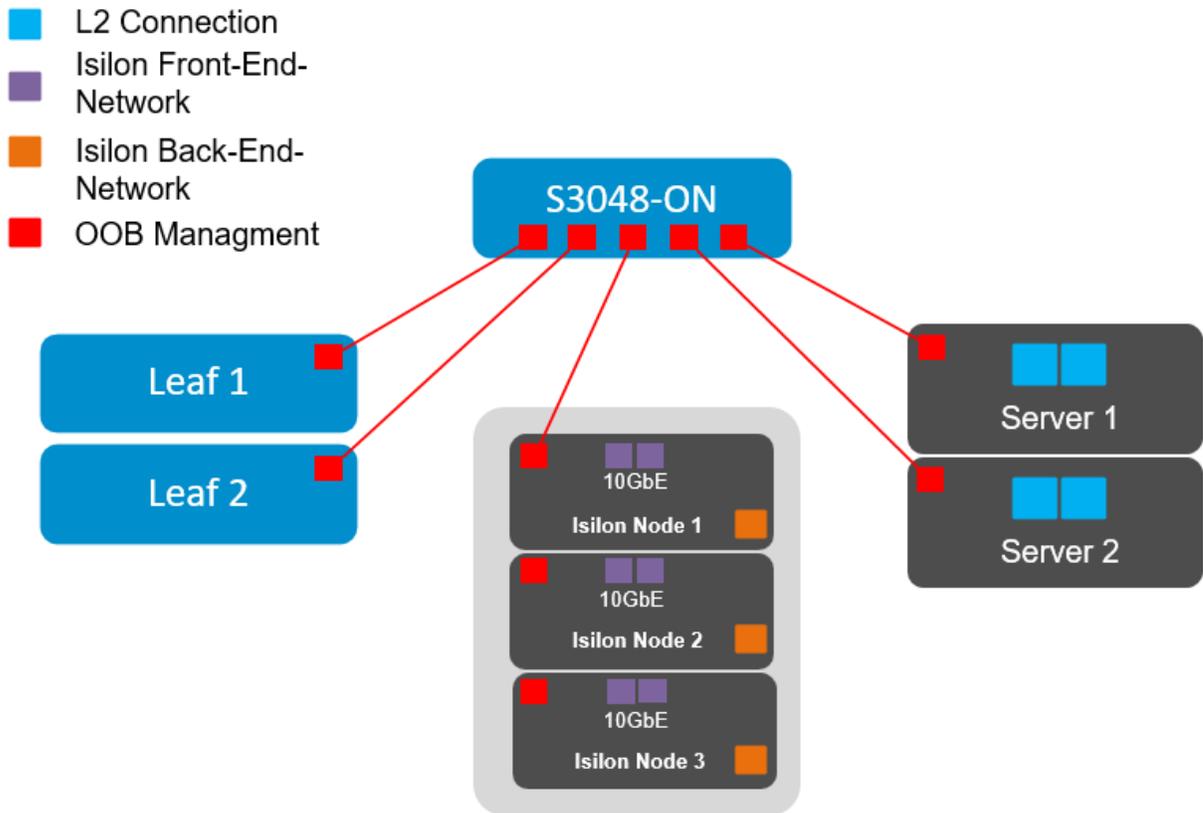


Figure 9 Management configurations for rack 1

## 4 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. The bandwidth of the fabric may be increased by adding connections between the leaf and spine layer as long as the spine layer has the capacity for the additional connections.

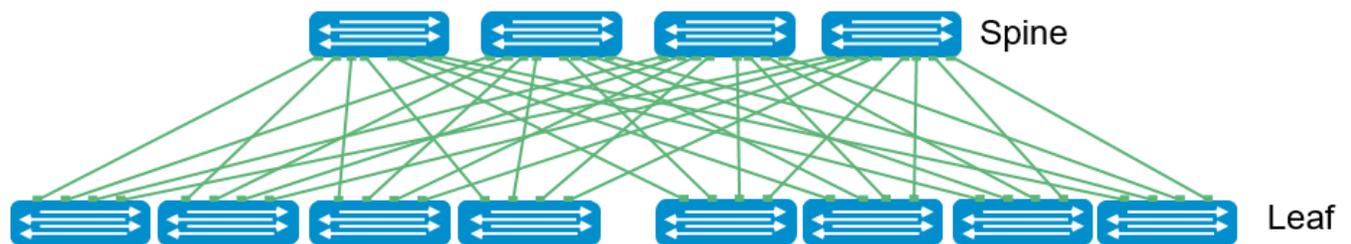


Figure 10 Leaf-Spine architecture

### 4.1 Design considerations

There are many different options regarding the selection of the correct topology that will best fit the needs of the data center. In this section, the different protocols, topologies, and best practices will be covered. The main differentiation will be whether the L2/L3 boundary is located at the spine layer or at the leaf layer. 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.

- 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.

In addition to the considerations for the L2 topology, some options need to be considered in the L3 topology. The primary design choice will be in the dynamic routing protocol that best fits the environment. BGP may be selected for scalability and is well suited for very large networks, while 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. This generally causes more memory and CPU usage than BGP. However, OSPF may offer faster convergence. OSPF is often used in smaller networks.

Design choices that are common in both topologies include:

- Whether to use LACP or non-LACP configurations downstream.
- Configuring each leaf pair in a Virtual Link Trunking (VLT).
- Dell EMC recommends that RSTP always is configured as a best practice. Although configuring VLT will create a loop free topology, RSTP will prevent loops if there is a misconfiguration in the network.

## 4.2 Oversubscription

Oversubscription is equal to the total amount of bandwidth available to all servers connected to a leaf switch divided by the amount of uplink bandwidth. In a leaf-spine network, oversubscription occurs at the leaf layer.

Oversubscription = total bandwidth / uplink bandwidth

Other configurations are using available servers, storage and leaf switches that also could be used to manage the subscription rates. The following are examples of oversubscription ratios based on downlink/uplink bandwidth.

Table 1 Oversubscription ratios based on uplink/downlink availability

NIC Speed	# Servers/Storage interfaces/Leaf	Total Host Bandwidth/Leaf	# Spines	Spine Uplink Speed	Total Uplink Bandwidth	Downlink/Uplink Bandwidth	Oversubscription Ratio
10GbE	19	190GbE	2	40GbE	80GbE	190/80	2.375 : 1
10GbE	19	190GbE	3	40GbE	120GbE	190/120	1.583 : 1
10GbE	19	190GbE	4	40GbE	160GbE	190/160	1.187 : 1
10GbE	38	380GbE	2	40GbE	80GbE	380/80	4.750 : 1
10GbE	38	380GbE	3	40GbE	120GbE	380/120	3.167 : 1
10GbE	38	380GbE	4	40GbE	160GbE	380/160	2.375 : 1

10GbE	38	380GbE	2	100GbE	80GbE	380/200	1.900 : 1
10GbE	38	380GbE	3	100GbE	120GbE	380/300	1.267 : 1
10GbE	38	380GbE	4	100GbE	160GbE	380/400	0.950 : 1

### 4.3 Scaling

An example of scaling this solution in a two-tier leaf-spine is a configuration of up to 16 racks. The Dell/EMC Z9100-ON has thirty-two 40/100GbE interfaces that would support 16 leaf pairs using VLT. This provides one rack that contains WAN-edge connectivity and 15 racks for servers and storage nodes. Each rack of the compute/storage rack holding a combination of up to 19 PowerEdge R730's or Isilon X210's.

This particular example, each R730 has four 10GbE uplinks, and each Isilon node has four 10GbE uplinks with 19 servers/nodes per rack. Additionally, the example architecture has four spine switches to minimize oversubscription.

Table 2 Connections for 16 racks with 4 spine switches

	Server/Storage interfaces	Server connections to leaf switches	Leaf connections to spine switches per rack	Total connections for leaf switches to four spine switches
Connections	4	19 x 4 = 76	4 per leaf switch, 2 leaf switches per rack = 8 links	16 racks * 8 = 128
Speed of Ports	10 GbE	10 GbE	40 GbE	40 GbE
Total theoretical available bandwidth	4 x 10 = 40GbE	76 x 10 = 760GbE	8 * 40GbE per rack = 320GbE	16 * 320GbE = 5120GbE

This example provides for an oversubscription rate of 2.375:1 using 40GbE spine connectivity.

Scaling beyond 16 racks would require a three-tier leaf-spine network. The proof-of-concept scaling that Figure 11 shows allows four 16-rack pods connected using an additional spine layer to scale in excess of 1,000 nodes with the same oversubscription ratio. This scenario requires reducing the number of racks available per pod to accommodate the uplinks required to connect to the super spine layer.

It is important to understand the port-density of switches used and their feature sets' impact on the number of available ports. This directly influences the number of switches necessary for proper scaling.

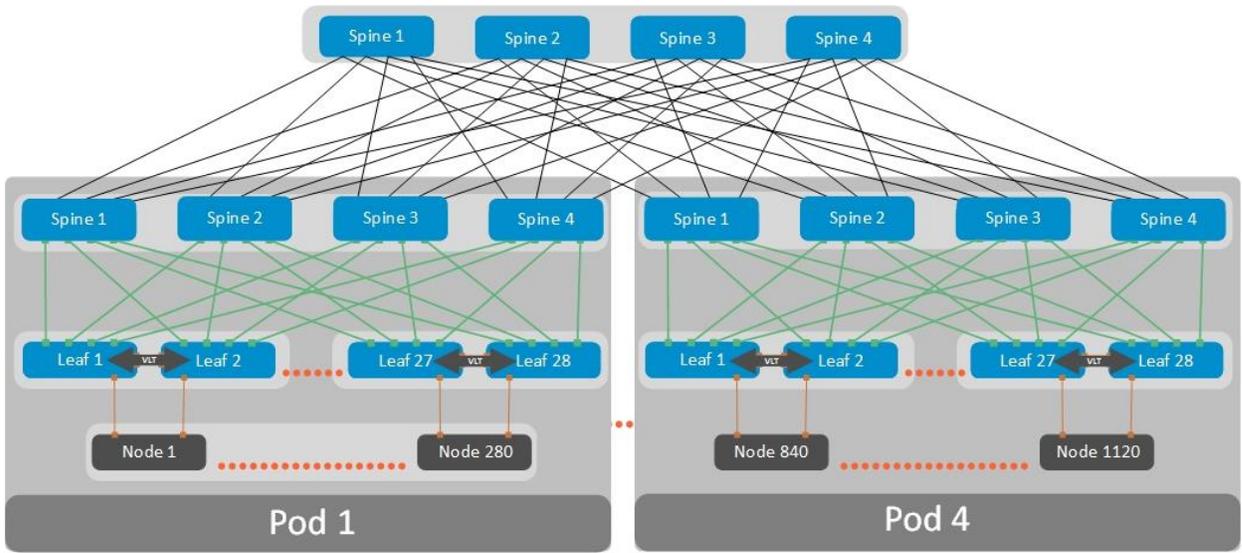


Figure 11 Scaling out the existing networking topology

## 4.4 Layer 3 leaf-spine topology

In a layer 3 leaf-spine network, traffic between leaves and spines is routed. The layer 2/3 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 leaves as shown in Figure 12.

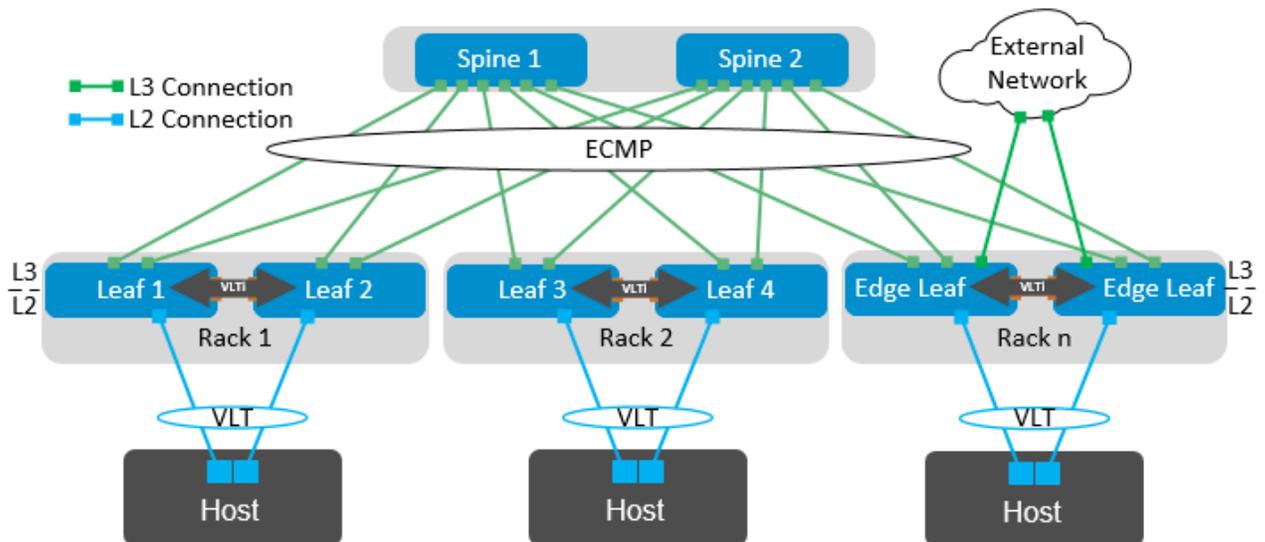


Figure 12 Layer 3 leaf-spine network

## 4.5 Layer 2 leaf-spine topology

In a layer 2 leaf-spine network, traffic between leaves and spines is switched (except for a pair of edge leaves) as shown in Figure 13. 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 leaves.

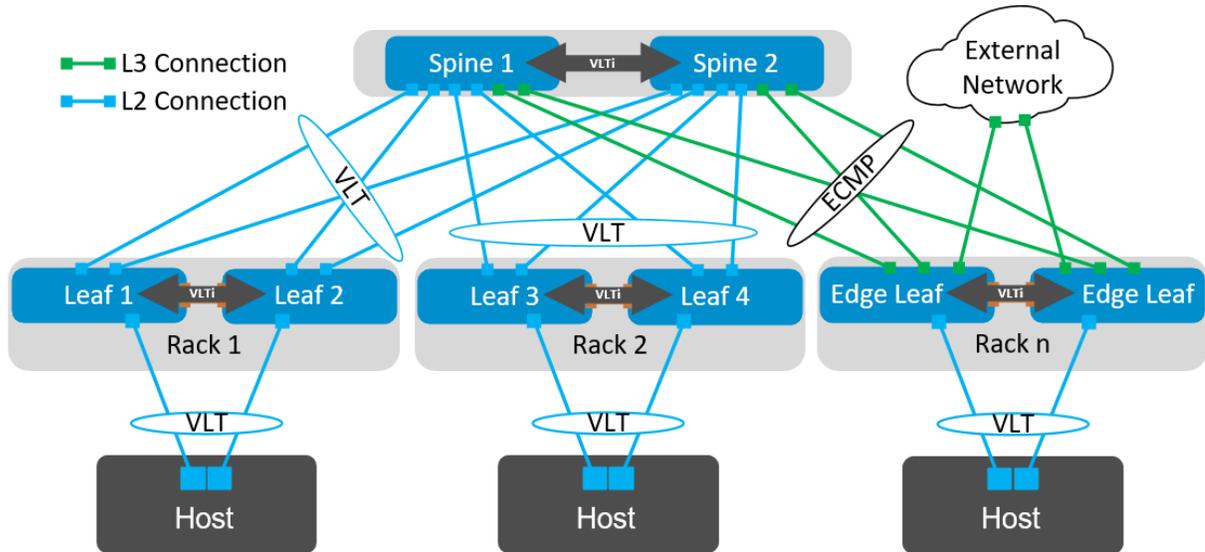


Figure 13 Layer 2 leaf-spine network

## 5 Layer 3 Topology preparation

The layer 3 topology used in this example will use external border gateway protocol (eBGP) as well as ECMP. In order to correctly configure this topology, several things need to be considered and planned.

### 5.1 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 14 shows the ASN assignments used for leaf and spine switches in the BGP examples in this guide.

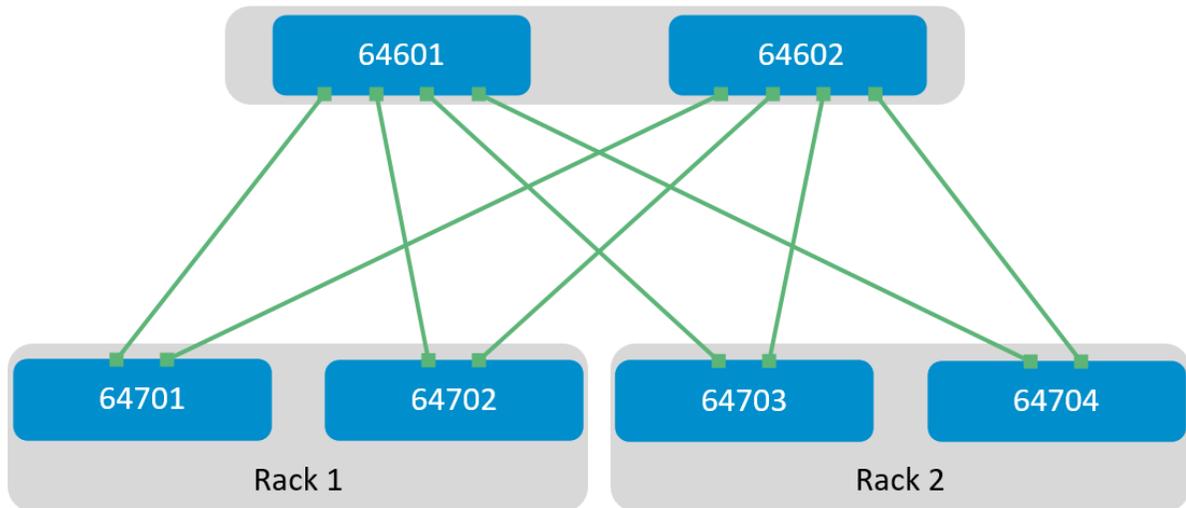


Figure 14 BGP ASN assignments

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., 64601), and an ASN with a “7” in the hundreds place represents a leaf switch (e.g., 64701).

**Note:** The same ASN can be used across all tier-2 spine switches if the growth plans do not require an additional layer of spine switches.

### 5.2 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 15 shows the loopback addresses used as router IDs in the example provided.

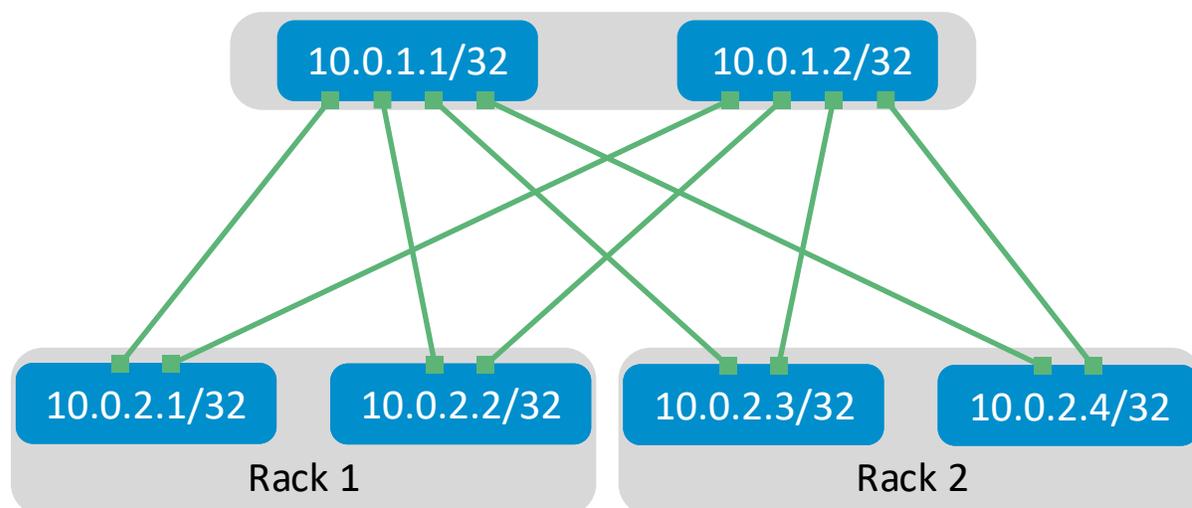


Figure 15 Loopback addressing

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 the spine and “2” for the 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.

## 5.3 Point-to-point interfaces

Table 3 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 16.

Table 3 Interface and IP configuration

Link Label	Source switch	Source interface	Source IP	Network	Destination switch	Destination interface	Destination IP
A	Leaf 1	Eth1/1/49	.1	192.168.1.0/31	Spine 1	Eth1/1/1	.0
B	Leaf 1	Eth1/1/50	.1	192.168.2.0/31	Spine 2	Eth1/1/1	.0
C	Leaf 2	Eth1/1/49	.3	192.168.1.2/31	Spine 1	Eth1/1/2	.2
D	Leaf 2	Eth1/1/50	.3	192.168.2.2/31	Spine 2	Eth1/1/2	.2

E	Leaf 3	Eth1/1/49	.5	192.168.1.4/31	Spine 1	Eth1/1/3	.4
F	Leaf 3	Eth1/1/50	.5	192.168.2.4/31	Spine 2	Eth1/1/3	.4
G	Leaf 4	Eth1/1/49	.7	192.168.1.6/31	Spine 1	Eth1/1/4	.6
H	Leaf 4	Eth1/1/50	.7	192.168.2.6/31	Spine 2	Eth1/1/4	.6

The point-to-point IP addresses used in this guide are shown in Figure 16:

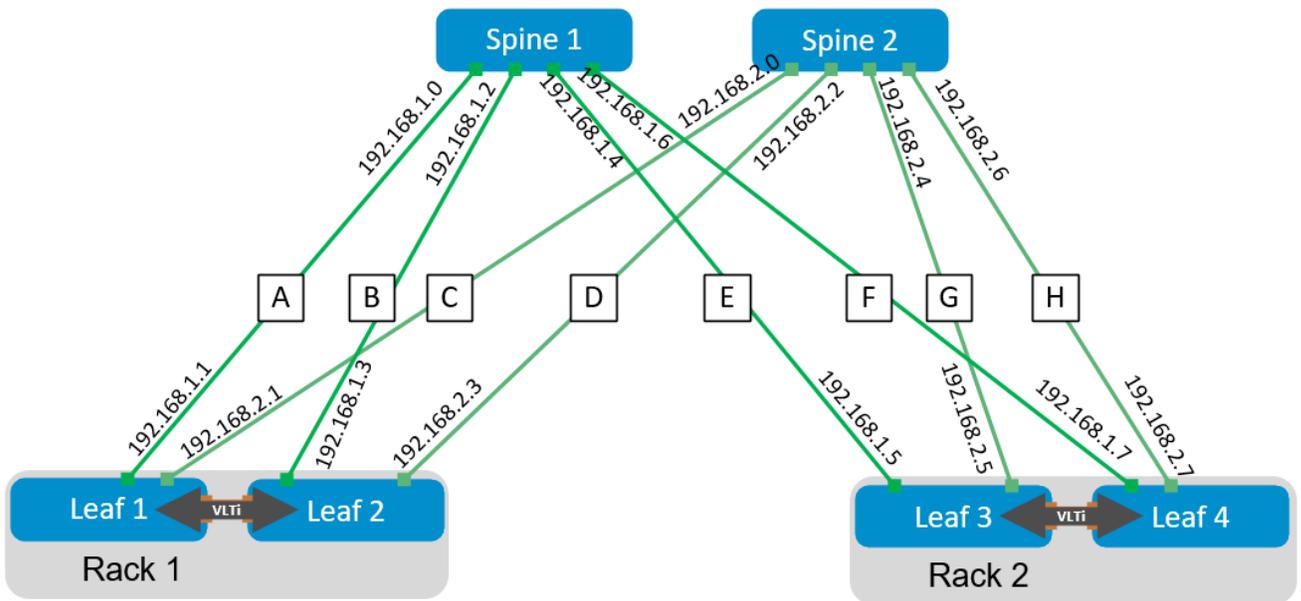


Figure 16 Point-to-point IP addresses

## 6 Configuration of Layer 3 Topology

This section will cover the configuration of the layer 3 topology. In the example provided there are three Isilon nodes connected to each leaf pair as well as two Dell PowerEdge servers that will be consuming the storage. The connections for the first node are shown in the diagram. Each subsequent node is cabled similarly. The configurations for all connections are detailed in the following sections. The Isilon nodes will all connect on the back-end network through two InfiniBand switches, creating one single six node cluster. The Isilon back-end networking is not covered in this guide.

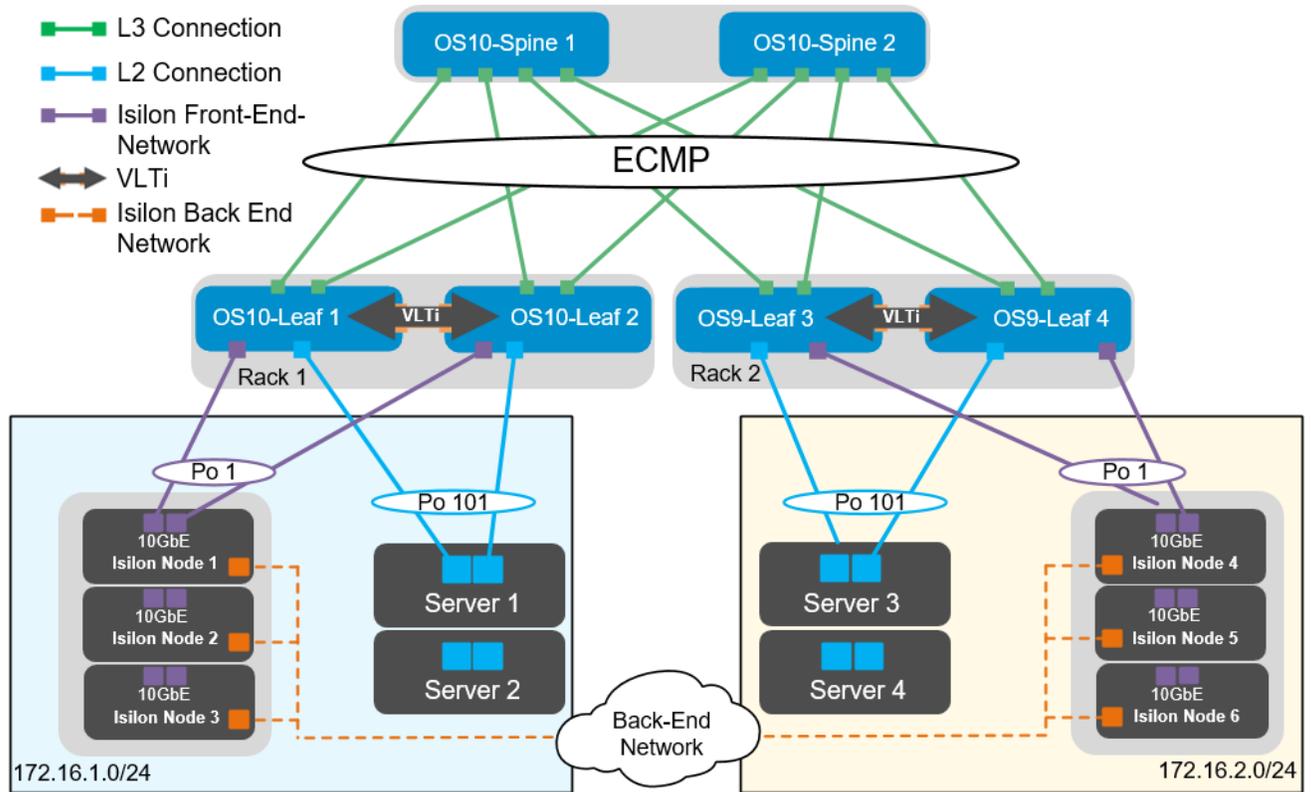


Figure 17 Layer 3 Topology

**Note:** The configuration files for every switch in this topology are listed in the attachments section.

### 6.1 Configuration of Z9100-ON OS10EE Spine Switches

The configuration of the example used in this guide will begin with the two Z9100-ON's Spine1 and Spine2 as seen in Figure 18.

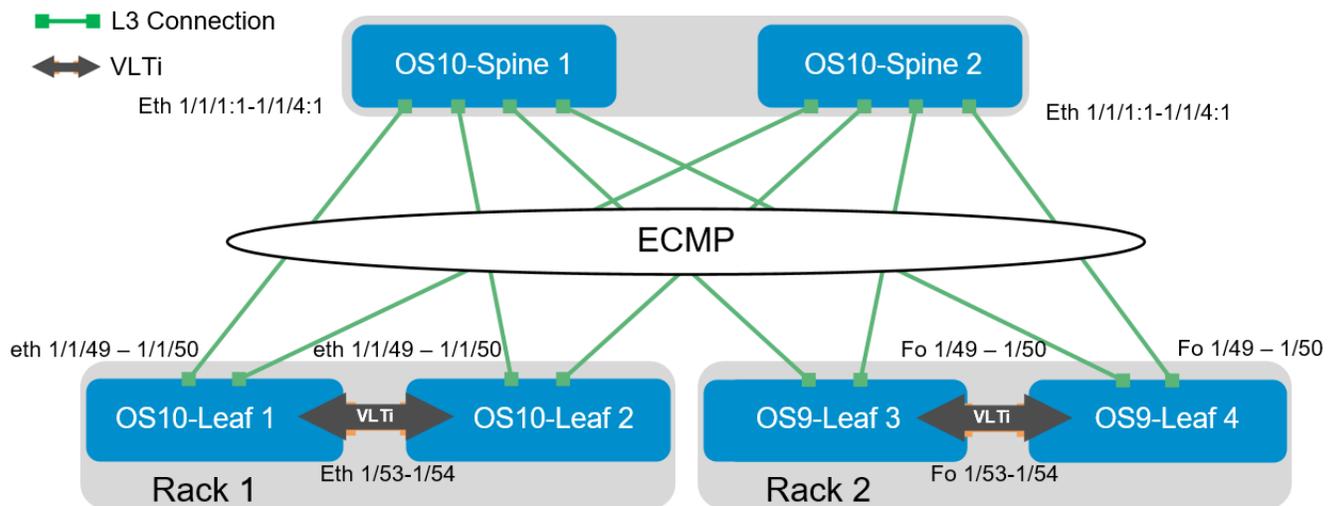


Figure 18 OS10EE Spine

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

Z9100-ON Spine 1	Z9100-ON Spine 2
configure terminal	configure terminal
hostname Z9100-Spine1	hostname Z9100-Spine2
interface mgmt 1/1/1 no ip address dhcp no shutdown ip address 100.67.169.37/24	interface mgmt 1/1/1 no ip address dhcp no shutdown ip address 100.67.169.36/24
management route 0.0.0.0/0 100.67.169.254	management route 0.0.0.0/0 100.67.169.254

Configure the four point-to-point interfaces connected to leaf switches. In this example, each of the connections from the Z9100-ON spine to the S4048-ON leaf switches needs to have the speed set to 40GbE. Next, assign IP addresses per Table 3. Configure a loopback interface to be used as the router ID. Isilon's OneFS supports and recommends the use of jumbo frames. Each interface used below will be configured using jumbo frames.

Z9100-ON Spine 1	Z9100-ON Spine 2
<pre>interface breakout 1/1/1 map 40g-1x interface breakout 1/1/2 map 40g-1x interface breakout 1/1/3 map 40g-1x interface breakout 1/1/4 map 40g-1x  interface ethernet1/1/1:1 description "S4048-Leaf1" no shutdown no switchport mtu 9216 ip address 192.168.1.0/31  interface ethernet1/1/2:1 description "S4048-Leaf2" no shutdown no switchport mtu 9216 ip address 192.168.1.2/31  interface ethernet1/1/3:1 description "S4048-Leaf3" no shutdown no switchport mtu 9216 ip address 192.168.1.4/31  interface ethernet1/1/4:1 description "S4048-Leaf4" no shutdown no switchport mtu 9216 ip address 192.168.1.6/31  interface loopback0 description "Router ID" no shutdown ip address 10.0.1.1/32</pre>	<pre>interface breakout 1/1/1 map 40g-1x interface breakout 1/1/2 map 40g-1x interface breakout 1/1/3 map 40g-1x interface breakout 1/1/4 map 40g-1x  interface ethernet1/1/1:1 description "S4048-Leaf1" no shutdown no switchport mtu 9216 ip address 192.168.2.0/31  interface ethernet1/1/2:1 description "S4048-Leaf2" no shutdown no switchport mtu 9216 ip address 192.168.2.2/31  interface ethernet1/1/3:1 description "S4048-Leaf3" no shutdown no switchport mtu 9216 ip address 192.168.2.4/31  interface ethernet1/1/4:1 description "S4048-Leaf4" no shutdown no switchport mtu 9216 ip address 192.168.2.6/31  interface loopback0 description "Router ID" no shutdown ip address 10.0.1.2/32</pre>

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

The command `seq 10 permit 10.0.1.0/24 ge 32` includes all addresses in the 10.0.1.0/24 address range with a mask greater than or equal to 32. This includes all loopback addresses used as router IDs.

Z9100-ON Spine 1	Z9100-ON Spine 2
<pre>route-map spine-leaf permit 10 match ip address prefix-list spine-leaf</pre>	<pre>route-map spine-leaf permit 10 match ip address prefix-list spine-leaf</pre>
<pre>ip prefix-list spine-leaf seq 10 permit 10.0.1.0/24 ge 32</pre>	<pre>ip prefix-list spine-leaf seq 10 permit 10.0.1.0/24 ge 32</pre>

Use these commands to configure BGP.

First, enable eBGP with the `router bgp ASN` command. The ASN is from Figure 14.

The `bgp bestpath as-path multipath-relax` command 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 Neighbor fall-over is enabled. Graceful restart enables the data plane to continue forwarding traffic for a time if the BGP process fails or quits.

BGP hello and hold down timers are set to three and nine seconds, respectively. Neighbor fall-over will trigger route withdrawal when IP connectivity between BGP peers is lost. The more aggressive timers (default is 60 seconds hello and 180 seconds hold down) are for backup. The advertisement interval is set to one second. This is to prevent BGP speakers from advertising updates immediately upon receipt. Instead, they will advertise them in batched intervals of one second. This delay is to prevent overhead.

Finally, exit configuration mode and save the configuration.

Z9100-ON Spine 1	Z9100-ON Spine 2
<pre> router bgp 64601 bestpath as-path multipath-relax graceful-restart role receiver-only maximum-paths ebgp 2  address-family ipv4 unicast redistribute connected route-map spine-leaf  template spine-leaf advertisement-interval 1 fall-over timers 3 9  neighbor 192.168.1.1 inherit template spine-leaf remote-as 64701 no shutdown  neighbor 192.168.1.3 inherit template spine-leaf remote-as 64702 no shutdown  neighbor 192.168.1.5 inherit template spine-leaf remote-as 64703 no shutdown  neighbor 192.168.1.7 inherit template spine-leaf remote-as 64704 no shutdown  end write memory </pre>	<pre> router bgp 64602 bestpath as-path multipath-relax graceful-restart role receiver-only maximum-paths ebgp 2  address-family ipv4 unicast redistribute connected route-map spine-leaf  template spine-leaf advertisement-interval 1 fall-over timers 3 9  neighbor 192.168.2.1 inherit template spine-leaf remote-as 64701 no shutdown  neighbor 192.168.2.3 inherit template spine-leaf remote-as 64702 no shutdown  neighbor 192.168.2.5 inherit template spine-leaf remote-as 64703 no shutdown  neighbor 192.168.2.7 inherit template spine-leaf remote-as 64704 no shutdown  end write memory </pre>

## 6.2 Configuration of S4048-ON OS10EE Leaf Switches

Configurations of the leaf switches one and two running OS10EE will be shown next as seen in Figure 19.

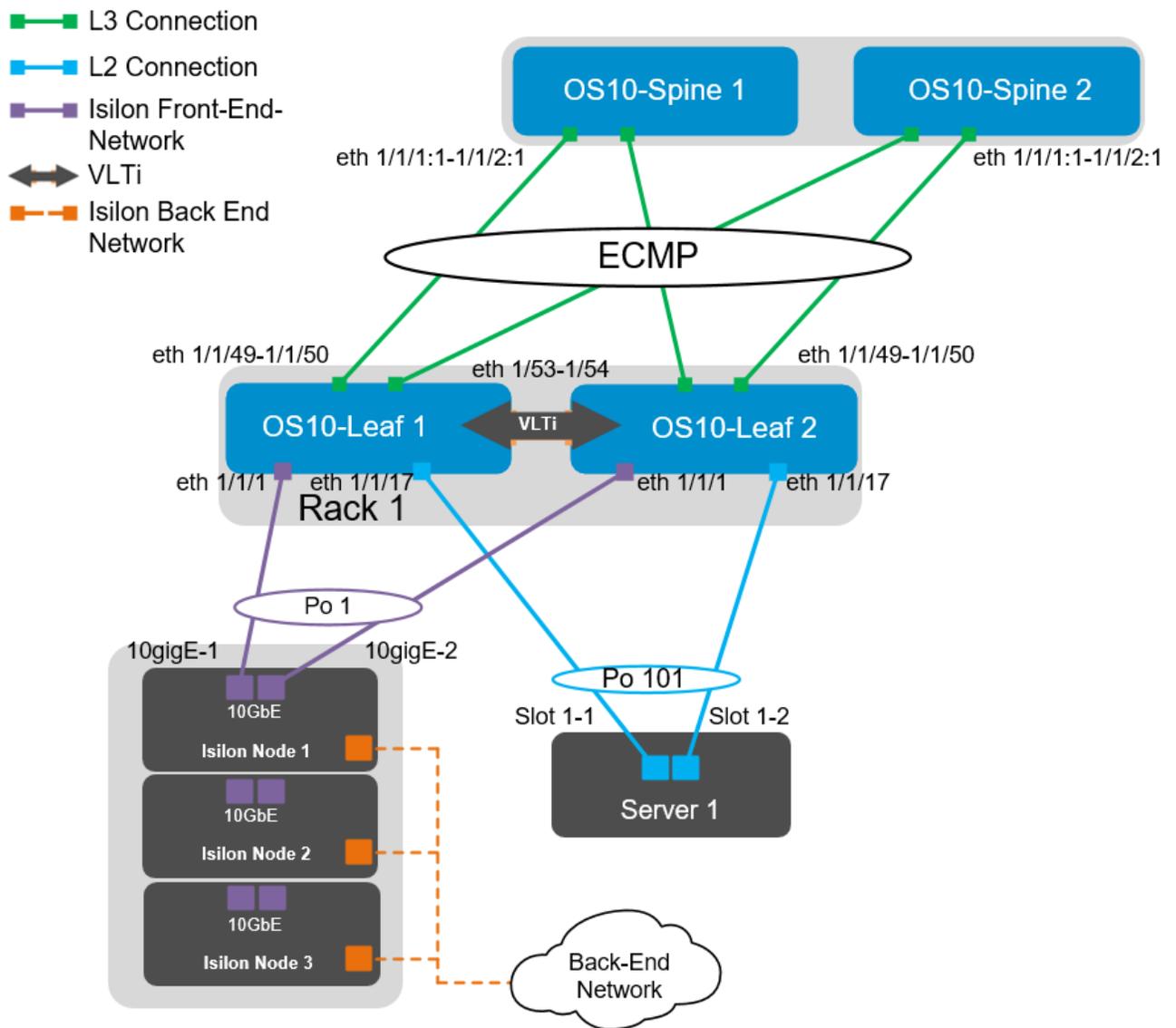


Figure 19 OS10EE Leaf pair

The interfaces for additional Isilon nodes and servers in this example can be found in the following table:

Table 4 Interface enumeration

Server/Isilon Node	Switch	Leaf interface
Node 01-1	Leaf 1	ethernet 1/1/1
	Leaf 2	ethernet 1/1/1
Node 01-2	Leaf 1	ethernet 1/1/3
	Leaf 2	ethernet 1/1/3

Server/Isilon Node	Switch	Leaf interface
Node 01-3	Leaf 1	ethernet 1/1/5
	Leaf 2	ethernet 1/1/5
Server 01	Leaf 1	ethernet 1/1/17
	Leaf 2	ethernet 1/1/17
Server 02	Leaf 1	ethernet 1/1/19
	Leaf 2	ethernet 1/1/19

First, set the hostname, configure the OOB management interface and default gateway. Enable RSTP as a precaution against creating a networking loop. S4048-Leaf 1 is configured as the primary RSTP root bridge using the `bridge-priority 0` command. S4048-Leaf 2 is configured as the secondary RSTP root bridge using the `bridge-priority 4096` command. This ensures that leaf 1 will always be the root bridge. The `no iscsi enable` command is issued to ensure that iSCSI snooping is disabled.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>configure terminal hostname S4048-Leaf1 interface mgmt 1/1/1 no ip address dhcp ip address 100.67.170.15/24 management route 0.0.0.0/0 100.67.170.254 no iscsi enable spanning-tree mode rstp spanning-tree rstp priority 0</pre>	<pre>configure terminal hostname S4048-Leaf2 interface mgmt 1/1/1 no ip address dhcp ip address 100.67.170.14/24 management route 0.0.0.0/0 100.67.171.254 no iscsi enable spanning-tree mode rstp spanning-tree rstp priority 4096</pre>

Configure the VLT interconnect between S4048-Leaf1 and S4048-Leaf2. In this configuration, remove each interface from layer 2 mode with the `no switchport` command for interfaces `eth 1/1/53-1/1/54`. Once this is done, create the VLT domain, add the backup destination, and add the interfaces participating in VLTi. Finally, enable VLT peer routing on each switch. Once VLT is enabled, it will dynamically change the MTU settings for the VLTi to a value of 9216. Isilon's OneFS supports and recommends the use of jumbo frames.

**Note:** Refer to [Isilon Network Design Considerations](#) for more information on specific protocols as they relate to OneFS.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface range ethernet 1/1/53-1/1/54 description VLTi no switchport  vlt-domain 127 backup destination 100.67.170.14 discovery-interface ethernet 1/1/53, 1/1/54</pre>	<pre>interface range ethernet 1/1/53-1/1/54 description VLTi no switchport  vlt-domain 127 backup destination 100.67.170.15 discovery-interface ethernet 1/1/53, 1/1/54</pre>

Next, the VLANs will be configured. There will be one VLAN for each Isilon subnet; this will also include the servers in each rack. Virtual Router Redundancy Protocol (VRRP) will be used to as a secondary from of redundancy. Although VRRP is an active/standby First Hop Redundancy Protocol (FHRP), when combined with VLT it becomes active/active. The priority is assigned to give predictability of VRRP master. The higher priority will be elected as the master. If no priority is given, the higher IP address will be elected as the master.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface vlan 100 ip address 172.16.1.252/24 no shutdown mtu 9216  vrrp-group 100 priority 150 virtual-address 172.16.1.254</pre>	<pre>interface vlan 100 ip address 172.16.1.253/24 no shutdown mtu 9216  vrrp-group 100 priority 100 virtual-address 172.16.1.254</pre>

**Note:** Further information on all protocols used in this document and their use cases can be found in the [Dell EMC Leaf-Spine Deployment Guide](#) and [Dell EMC Networking L3 Design for Leaf-Spine with OS10EE](#).

Configure the necessary port channels. There will be one port channel for each server as well as a single port channel for each of the Isilon nodes. Each port channel is assigned an MTU of 9216 to support jumbo frames, as well configured to allow access to VLAN 100. The `vlt-port-channel` command enables port channels on VLT peers to function as a single port channel.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface port-channel1 description "Isilon Node 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 1 spanning-tree port type edge</pre>	<pre>interface port-channel1 description "Isilon Node 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 1 spanning-tree port type edge</pre>
<pre>interface port-channel3 description "Isilon Node 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 3 spanning-tree port type edge</pre>	<pre>interface port-channel3 description "Isilon Node 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 3 spanning-tree port type edge</pre>
<pre>interface port-channel5 description "Isilon Node 3" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 5 spanning-tree port type edge</pre>	<pre>interface port-channel5 description "Isilon Node 3" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 5 spanning-tree port type edge</pre>
<pre>interface port-channel101 description "Server 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 101 spanning-tree port type edge</pre>	<pre>interface port-channel101 description "Server 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 101 spanning-tree port type edge</pre>
<pre>interface port-channel102 description "Server 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 102 spanning-tree port type edge</pre>	<pre>interface port-channel102 description "Server 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 102 spanning-tree port type edge</pre>

The downstream interfaces will be configured in the next section. Dell recommends asymmetrical flow control (rx on tx off) on S4048-ON leaf switches which have sufficient egress buffers to reduce packet drops and resulting retransmissions. This feature should be applied to all host interfaces that support flow control. In OS10EE this is enabled by default, in OS9 it must be configured. Flow control commands are added below for reference; however, they are not required in OS10EE.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface ethernet1/1/1 description "Isilon Node 1-10gige-1" no shutdown channel-group 1 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/1 description "Isilon Node 1-10gige-2" no shutdown channel-group 1 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/3 description "Isilon Node 2-10gige-1" no shutdown channel-group 3 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/3 description "Isilon Node 2-10gige-2" no shutdown channel-group 3 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/5 description "Isilon Node 3-10gige-1" no shutdown channel-group 5 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/5 description "Isilon Node 3-10gige-2" no shutdown channel-group 5 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/17 description "Server 1-1" no shutdown channel-group 101 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/17 description "Server 1-2" no shutdown channel-group 101 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/19 description "Server 2-1" no shutdown channel-group 102 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/19 description "Server 2-2" no shutdown channel-group 102 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>

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

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface ethernet1/1/49 description "Spine 1" no shutdown no switchport mtu 9216 ip address 192.168.1.1/31  interface ethernet1/1/50 description "Spine 2" no shutdown no switchport mtu 9216 ip address 192.168.2.1/31  interface loopback0 description "Router ID" no shutdown ip address 10.0.2.1/32</pre>	<pre>interface ethernet1/1/49 description "Spine 1" no shutdown no switchport mtu 9216 ip address 192.168.1.3/31  interface ethernet1/1/50 description "Spine 2" no shutdown no switchport mtu 9216 ip address 192.168.2.3/31  interface loopback0 description "Router ID" no shutdown ip address 10.0.2.2/32</pre>

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

The command `ip prefix-list spine-leaf seq 10 permit 10.0.2.0/24 ge 32` includes all addresses in the 10.0.2.0/24 address range with a mask greater than or equal to 32. This includes all loopback addresses used as router IDs.

The command `ip prefix-list spine-leaf seq 20 permit 172.16.0.0/16 ge 24` includes the 172.16.2.0/24 network used on Leaf switches 3 and 4 as shown in Figure 17.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>route-map spine-leaf permit 10 match ip address prefix-list spine-leaf  ip prefix-list spine-leaf seq 10 permit 10.0.2.0/24 ge 32  ip prefix-list spine-leaf seq 20 permit 172.16.0.0/16 ge 24</pre>	<pre>route-map spine-leaf permit 10 match ip address prefix-list spine-leaf  ip prefix-list spine-leaf seq 10 permit 10.0.2.0/24 ge 32  ip prefix-list spine-leaf seq 20 permit 172.16.0.0/16 ge 24</pre>

Uplink Failure Detection (UFD) is configured next. UFD is a feature that shuts down specified downstream interfaces when all specified uplinks become disabled.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre> uplink-state-group 1 enable downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream ethernet1/1/49 upstream ethernet1/1/50 </pre>	<pre> uplink-state-group 1 enable downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream ethernet1/1/49 upstream ethernet1/1/50 </pre>

Use these commands to configure eBGP.

First, enable eBGP with the `router bgp ASN` command. The ASN is from Figure 14.

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 Neighbor fall-over is enabled. Graceful restart enables the data plane to continue forwarding traffic for a time if the BGP process fails or quits.

BGP hello and hold down timers are set to three and nine seconds, respectively. Neighbor fall-over will trigger route withdrawal when IP connectivity between BGP peers is lost. The more aggressive timers (default is 60 seconds hello and 180 seconds hold down) are for backup. The advertisement interval is set to one second. This is to prevent BGP speakers from advertising updates immediately upon receipt. Instead, they will advertise them in batched intervals of one second. This delay is to prevent overhead.

Finally, exit configuration mode and save the configuration.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre> router bgp 64701 bestpath as-path multipath-relax graceful-restart role receiver-only maximum-paths ebgp 2  address-family ipv4 unicast redistribute connected route-map spine-leaf  template spine-leaf advertisement-interval 1 fall-over timers 3 9  neighbor 192.168.1.0 inherit template spine-leaf remote-as 64601 no shutdown  neighbor 192.168.2.0 inherit template spine-leaf remote-as 64602 no shutdown  end write memory </pre>	<pre> router bgp 64702 bestpath as-path multipath-relax graceful-restart role receiver-only maximum-paths ebgp 2  address-family ipv4 unicast redistribute connected route-map spine-leaf  template spine-leaf advertisement-interval 1 fall-over timers 3 9  neighbor 192.168.1.2 inherit template spine-leaf remote-as 64601 no shutdown  neighbor 192.168.2.2 inherit template spine-leaf remote-as 64602 no shutdown  end write memory </pre>

## 6.3 Configuration of S4048-ON OS9 Leaf Switches

Configurations in this section cover leaf switches 3 and 4 running OS9 as seen in Figure 20.

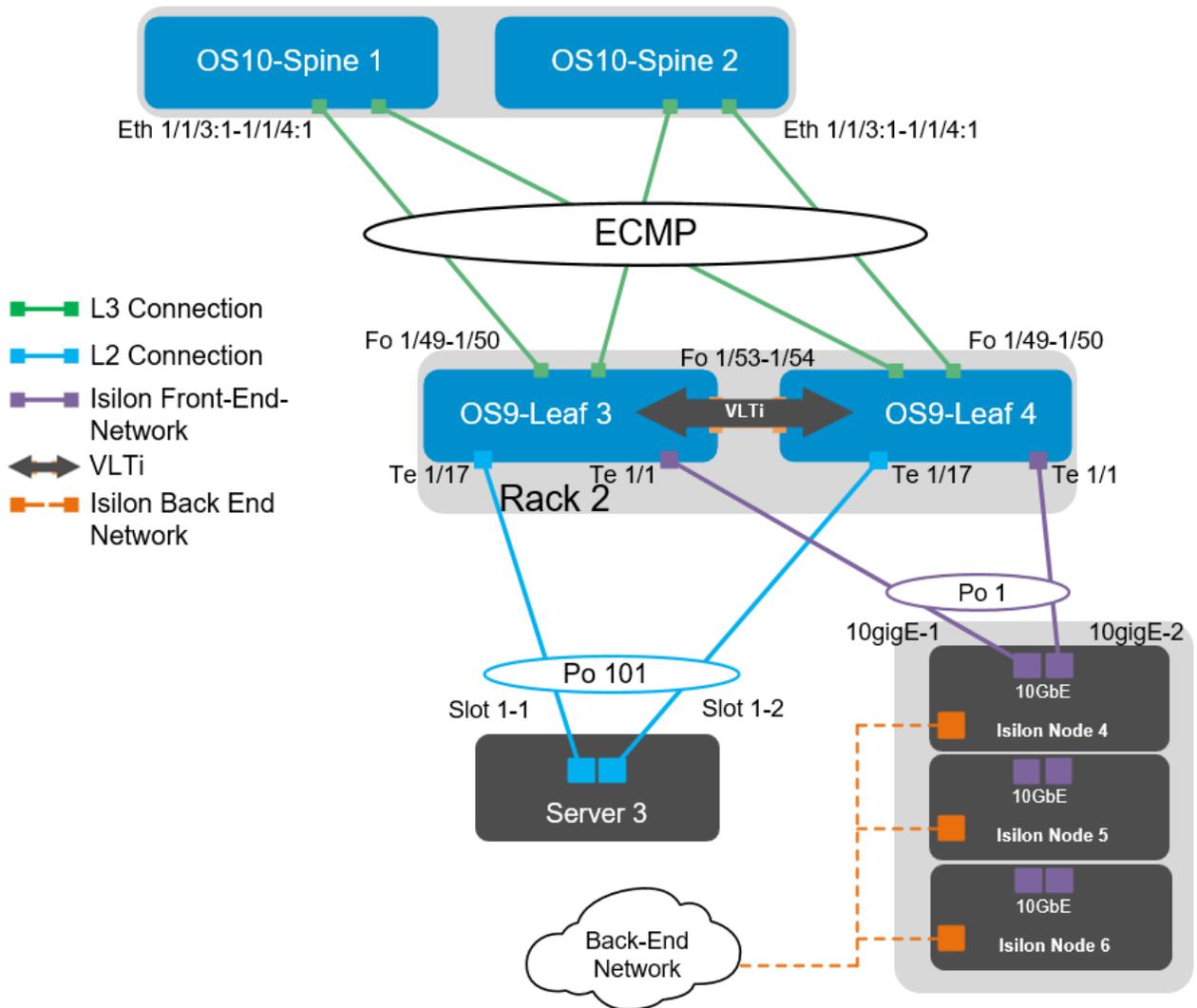


Figure 20 OS9 Leaf pair

The ports for additional Isilon nodes and servers in this example can be found in the following table:

Table 5 Interface enumeration

Server/Isilon Node	Switch	Leaf interface
Node 01-4	Leaf 3	tengigabitethernet 1/1/1
	Leaf 4	tengigabitethernet 1/1/1

Server/Isilon Node	Switch	Leaf interface
Node 01-5	Leaf 3	tengigabitethernet 1/1/3
	Leaf 4	tengigabitethernet 1/1/3
Node 01-6	Leaf 3	tengigabitethernet 1/1/5
	Leaf 4	tengigabitethernet 1/1/5
Server 03	Leaf 3	tengigabitethernet 1/1/17
	Leaf 4	tengigabitethernet 1/1/17
Server 04	Leaf 3	tengigabitethernet 1/1/19
	Leaf 4	tengigabitethernet 1/1/19

First, set the hostname, configure the OOB management interface and default gateway. Next, enable LLDP as well as RSTP. In OS10EE LLDP is enabled by default, in OS9 it must be enabled. S4048-Leaf 3 is configured as the primary RSTP root bridge using the bridge-priority 0 command. S4048-Leaf 4 is configured as the secondary RSTP root bridge using the bridge-priority 4096 command. This ensures that leaf 3 will always be the root bridge.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>configure terminal hostname S4048-Leaf3 interface management 1/1 ip address 100.67.171.15/24 no shutdown management route 0.0.0.0/0 100.67.171.254 protocol lldp   advertise management-tlv management-address system-description system-name   advertise interface-port-desc protocol spanning-tree rstp bridge-priority 0 no disable</pre>	<pre>configure terminal hostname S4048-Leaf4 interface management 1/1 ip address 100.67.171.14/24 no shutdown management route 0.0.0.0/0 100.67.171.254 protocol lldp   advertise management-tlv management-address system-description system-name   advertise interface-port-desc protocol spanning-tree rstp bridge-priority 4096 no disable</pre>

Configure the VLT interconnect between S4048-Leaf3 and S4048-Leaf4. Create a static port channel and add the two VLTi connections to port channel 128. Configure the VLT domain using the IP address assigned to the management interface of the opposing switch as the backup destination. The backup destination is only utilized for the VLT heartbeat. Once VLT is enabled, it will dynamically change the MTU settings for the VLTi to a value of 9216.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre> interface port-channel128 description VLTi channel-member fortyGigE 1/53 channel-member fortyGigE 1/54 no shutdown no switchport  interface fortyGigE 1/53 description VLTi no shutdown no switchport  interface fortyGigE 1/54 description VLTi no shutdown no switchport  vlt domain 127 peer-link port-channel 128 back-up destination 100.67.171.14 unit-id 0 peer-routing </pre>	<pre> interface port-channel128 description VLTi channel-member fortyGigE 1/53 channel-member fortyGigE 1/54 no shutdown no switchport  interface fortyGigE 1/53 description VLTi no shutdown no switchport  interface fortyGigE 1/54 description VLTi no shutdown no switchport  vlt domain 127 peer-link port-channel 128 back-up destination 100.67.171.15 unit-id 1 peer-routing </pre>

The downstream interfaces will be configured in the next section. Dell recommends asymmetrical flow control (rx on tx off) on S4048-ON leaf switches which have sufficient egress buffers to reduce packet drops and resulting retransmissions. This feature should be applied on all host interfaces that support flow control. In OS10EE this is enabled by default, in OS9 it must be configured. Flow Control is configured with the `flowcontrol rx on tx off` command.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface tengigabitethernet 1/1 description "Isilon Node 4-10gige-1" no shutdown port-channel-protocol LACP port-channel 1 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/1 description "Isilon Node 4-10gige-2" no shutdown port-channel-protocol LACP port-channel 1 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/3 description "Isilon Node 5-10gige-1" no shutdown port-channel-protocol LACP port-channel 3 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/3 description "Isilon Node 5-10gige-2" no shutdown port-channel-protocol LACP port-channel 3 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/5 description "Isilon Node 6-10gige-1" no shutdown port-channel-protocol LACP port-channel 5 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/5 description "Isilon Node 6-10gige-2" no shutdown port-channel-protocol LACP port-channel 5 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/17 description "Server 3-1" no shutdown port-channel-protocol LACP port-channel 101 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/17 description "Server 3-2" no shutdown port-channel-protocol LACP port-channel 101 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/19 description "Server 4-1" no shutdown port-channel-protocol LACP port-channel 102 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/19 description "Server 4-2" no shutdown port-channel-protocol LACP port-channel 102 mode active mtu 9216 flowcontrol rx on tx off</pre>

If the S4048-ON running OS9 is used for the first leaf pair the interfaces for the first Isilon node need to be left out of the port channel to facilitate the configuration of the LACP connections in OneFS. In this example that would be tengigabitethernet 1/1 on both switches. The configuration would be as follows for that example:

```
interface tengigabitethernet 1/1
  description "Isilon Node 4-10gige-1"
  switchport
  no shutdown
  mtu 9216
```

```
interface vlan200
  no shutdown
  mtu 9216
  ip address 172.16.2.252/24
  untagged port-channel 3,5,101,102
  untagged interface te 1/1
```

Configure the necessary port channels. There will be one port channel for each server as well as a single port channel for each of the Isilon nodes. Each port channel is assigned an MTU of 9216 to support jumbo frames. The `vlt-port-channel #` command enables port channels on VLT peers to function as a single port channel.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface port-channel1 description "Isilon Node 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 1 spanning-tree rstp edge-port</pre>	<pre>interface port-channel1 description "Isilon Node 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 1 spanning-tree rstp edge-port</pre>
<pre>interface port-channel3 description "Isilon Node 5" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 3 spanning-tree rstp edge-port</pre>	<pre>interface port-channel3 description "Isilon Node 5" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 3 spanning-tree rstp edge-port</pre>
<pre>interface port-channel5 description "Isilon Node 6" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 5 spanning-tree rstp edge-port</pre>	<pre>interface port-channel5 description "Isilon Node 6" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 5 spanning-tree rstp edge-port</pre>
<pre>interface port-channel101 description "Server 3" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 101 spanning-tree rstp edge-port</pre>	<pre>interface port-channel101 description "Server 3" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 101 spanning-tree rstp edge-port</pre>
<pre>interface port-channel102 description "Server 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 102 spanning-tree rstp edge-port</pre>	<pre>interface port-channel102 description "Server 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 102 spanning-tree rstp edge-port</pre>

Next, the VLANs will be created. There will be one VLAN for each Isilon subnet; this will also include the servers in each rack. Virtual Router Redundancy Protocol (VRRP) will be used as a secondary form of redundancy. VRRP becomes active/active with the use of VLT over the standard active/passive. The priority is assigned to give predictability to the VRRP master.

**Note:** In this example, Server 3's NIC is configured as an LACP NIC team. It is assigned the IP address 172.16.2.240/24. Server 3's default gateway is configured as 172.16.2.254/24. The virtual address assigned to the VRRP group will provide fault tolerance if either leaf goes offline.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface vlan200 no shutdown mtu 9216 ip address 172.16.2.252/24 untagged port-channel 1,3,5,101,102  vrrp-group 200 priority 150 virtual-address 172.16.2.254</pre>	<pre>interface vlan200 no shutdown mtu 9216 ip address 172.16.2.253/24 untagged port-channel 1,3,5,101,102  vrrp-group 200 priority 100 virtual-address 172.16.2.254</pre>

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

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface fortyGigE 1/49 description "Spine 1" no shutdown no switchport mtu 9216 ip address 192.168.1.5/31  interface fortyGigE 1/50 description "Spine 2" no shutdown no switchport mtu 9216 ip address 192.168.2.5/31  interface loopback0 description "Router ID" no shutdown ip address 10.0.2.3/32</pre>	<pre>interface fortyGigE 1/49 description "Spine 1" no shutdown no switchport mtu 9216 ip address 192.168.1.7/31  interface fortyGigE 1/50 description "Spine 2" no shutdown no switchport mtu 9216 ip address 192.168.2.7/31  interface loopback0 description "Router ID" no shutdown ip address 10.0.2.4/32</pre>

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

The command `seq 10 permit 10.0.2.0/24 ge 32` includes all addresses in the 10.0.2.0/24 address range with a mask greater than or equal to 32. This includes all loopback addresses used as router IDs.

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-ON Leaf 3	S4048-ON Leaf 4
<pre>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 10.0.2.0/24 ge 32 seq 20 permit 172.16.0.0/16 ge 24</pre>	<pre>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 10.0.2.0/24 ge 32 seq 20 permit 172.16.0.0/16 ge 24</pre>

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-ON Leaf 3	S4048-ON Leaf 4
<pre>ecmp-group 1 interface fortyGigE 1/49 interface fortyGigE 1/50 link-bundle-monitor enable</pre>	<pre>ecmp-group 1 interface fortyGigE 1/49 interface fortyGigE 1/50 link-bundle-monitor enable</pre>

Uplink Failure Detection (UFD) is configured next. UFD shuts down all downstream interfaces if all uplinks fail.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>uplink-state-group 1 downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream fortyGigE 1/49 upstream fortyGigE 1/50</pre>	<pre>uplink-state-group 1 downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream fortyGigE 1/49 upstream fortyGigE 1/50</pre>

Use these commands to configure BGP.

First, enable BGP with the `router bgp ASN` command. The ASN is from Figure 14.

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.

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

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>router bgp 64703   bgp bestpath as-path multipath-relax   bgp graceful-restart   bgp graceful-restart role receiver-only   maximum-paths ebgp 2    redistribute connected route-map spine-leaf    neighbor spine-leaf peer-group   neighbor spine-leaf fall-over   neighbor spine-leaf timers 3 9   neighbor spine-leaf advertisement-interval 1   neighbor spine-leaf no shutdown    neighbor 192.168.1.4 remote-as 64601   neighbor 192.168.1.4 peer-group spine-leaf   neighbor 192.168.1.4 no shutdown    neighbor 192.168.2.4 remote-as 64602   neighbor 192.168.2.4 peer-group spine-leaf   neighbor 192.168.2.4 no shutdown  end write memory</pre>	<pre>router bgp 64704   bgp bestpath as-path multipath-relax   bgp graceful-restart   bgp graceful-restart role receiver-only   maximum-paths ebgp 2    redistribute connected route-map spine-leaf    neighbor spine-leaf peer-group   neighbor spine-leaf fall-over   neighbor spine-leaf timers 3 9   neighbor spine-leaf advertisement-interval 1   neighbor spine-leaf no shutdown    neighbor 192.168.1.6 remote-as 64601   neighbor 192.168.1.6 peer-group spine-leaf   neighbor 192.168.1.6 no shutdown    neighbor 192.168.2.6 remote-as 64602   neighbor 192.168.2.6 peer-group spine-leaf   neighbor 192.168.2.6 no shutdown  end write memory</pre>

## 7 Configuration of Layer 2 Topology

This section will cover the configuration of the layer 2 topology. In the example provided there are three Isilon nodes connected to each leaf pair as well as two Dell PowerEdge servers that will be consuming the storage. The connections from the spine switches to the OS10EE leaf pair will be configured in port channel 49, while the connections to the OS9 leaf pair will be configured in port channel 50. The two spine switches will be configured in a VLT. The benefits and draw backs of a L2 configuration were referenced in [Chapter 4.1](#). The Isilon nodes will all connect to each switch in the leaf pair using an LACP port channel for each node, the connections for the Windows servers will be similarly configured. Also, each Isilon Node will be connected on the back-end network through two InfiniBand switches, creating one single six node cluster.

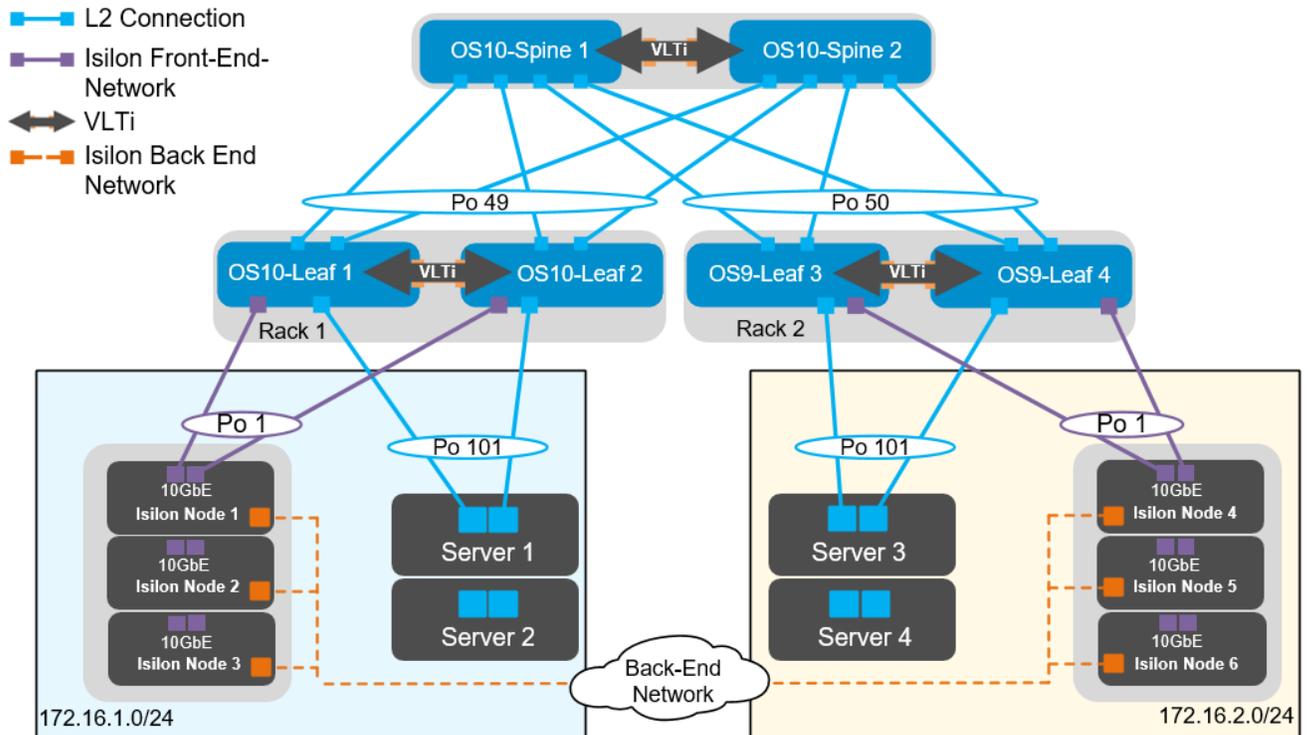


Figure 21 L2 topology

**Note:** The configuration files for every switch in this topology are listed in the attachments section.

### 7.1 Configuration of Z9100-ON OS10EE Spine Switches

The configuration of the example used in this guide will begin with the two Z9100-ON's Spine1 and Spine2 as seen in Figure 22.

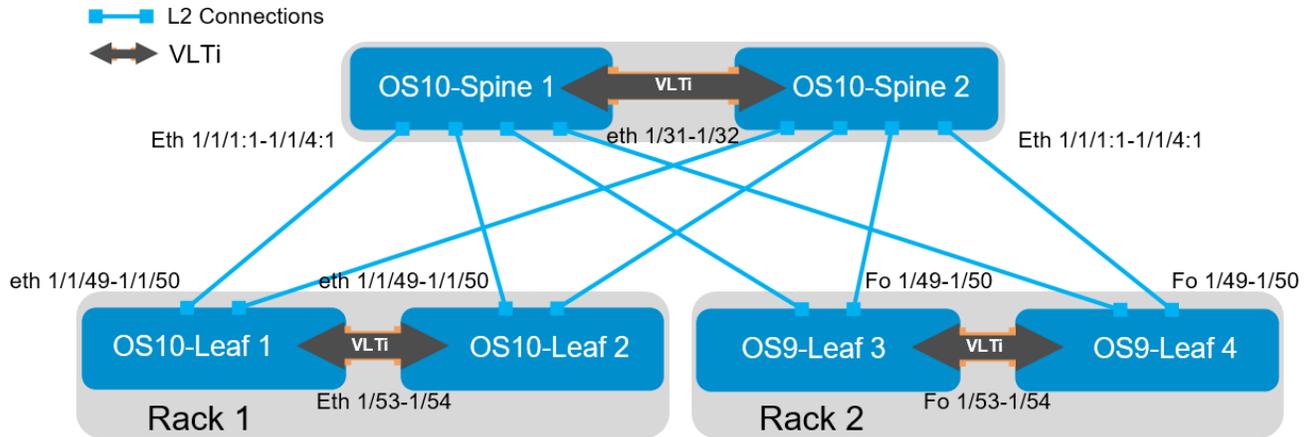


Figure 22 OS10EE Spine

Set the hostname, configure the OOB management interface and default gateway. Additionally, with the spine switches configured in a VLT pair, RSTP will be configured at the spine layer. The priority for RSTP will be set on each spine switch. This ensures that the root bridge will always be one of the spine switches.

Z9100-ON Spine 1	Z9100-ON Spine 2
configure terminal	configure terminal
hostname Z9100-Spine1	hostname Z9100-Spine2
interface mgmt 1/1/1 no ip address dhcp no shutdown ip address 100.67.169.37/24	interface mgmt 1/1/1 no ip address dhcp no shutdown ip address 100.67.169.36/24
management route 0.0.0.0/0 100.67.169.254	management route 0.0.0.0/0 100.67.169.254
spanning-tree mode rstp spanning-tree rstp priority 0	spanning-tree mode rstp spanning-tree rstp priority 4096

First, configure the VLT on each spine switch. Once VLT is enabled, it will dynamically change the MTU settings for the VLTi to a value of 9216.

Z9100-ON Spine 1	Z9100-ON Spine 2
<pre>interface range ethernet1/1/31-1/1/32 description VLTi no shutdown no switchport  vlt-domain 127 backup destination 100.67.169.36 discovery-interface ethernet1/1/31-1/1/32 peer-routing</pre>	<pre>interface range ethernet1/1/31-1/1/32 description VLTi no shutdown no switchport  vlt-domain 127 backup destination 100.67.169.37 discovery-interface ethernet1/1/31-1/1/32 peer-routing</pre>

Next, configure the required VLANs on each spine switch. There will be one VLAN assigned for each of the subnets that will be defined as seen in Figure 21. VRRP will be used to ensure availability in the event of a switch failure. The priority is set to 150 on spine 1 to ensure that it will be the VRRP master.

Z9100-ON Spine 1	Z9100-ON Spine 2
<pre>interface vlan100 no shutdown mtu 9216 ip address 172.16.1.252/24  vrrp-group 100 priority 150 virtual-address 172.16.1.254  interface vlan200 no shutdown mtu 9216 ip address 172.16.2.252/24  vrrp-group 200 priority 150 virtual-address 172.16.2.254</pre>	<pre>interface vlan100 no shutdown mtu 9216 ip address 172.16.1.253/24  vrrp-group 100 priority 100 virtual-address 172.16.1.254  interface vlan200 no shutdown mtu 9216 ip address 172.16.2.253/24  vrrp-group 200 priority 100 virtual-address 172.16.2.254</pre>

Configure the port channels between the spine and leaf switches. There will be four connections from each spine to the leaf switches. Port channel 49 will be configured on the OS10EE leaf pair, and port channel 50 will be configured on the OS9 leaf pair. Also, the downstream switches are S4048-ON switches with 40 GbE uplinks. Each of the connections from the Z9100-ON spines to the S4048-ON leaf switches needs to have the speed set to 40GbE.

Z9100-ON Spine 1	Z9100-ON Spine 2
<pre>interface breakout 1/1/1 map 40g-1x interface breakout 1/1/2 map 40g-1x</pre>	<pre>interface breakout 1/1/1 map 40g-1x interface breakout 1/1/2 map 40g-1x</pre>

Z9100-ON Spine 1	Z9100-ON Spine 2
<pre> interface breakout 1/1/3 map 40g-1x interface breakout 1/1/4 map 40g-1x  interface port-channel49 description R1-Leaf-AandB no shutdown switchport mode trunk switchport trunk allowed vlan 100 mtu 9216 vlt-port-channel 49  interface port-channel50 description R2-Leaf-AandB no shutdown switchport mode trunk switchport trunk allowed vlan 200 mtu 9216 vlt-port-channel 50  interface ethernet1/1/1:1 description R1-Leaf-A no shutdown no switchport channel-group 49 mode active mtu 9216  interface ethernet1/1/2:1 description R1-Leaf-B no shutdown no switchport channel-group 49 mode active mtu 9216  interface ethernet1/1/3:1 description R2-Leaf-A no shutdown no switchport channel-group 50 mode active mtu 9216  interface ethernet1/1/4:1 description R2-Leaf-B no shutdown no switchport channel-group 50 mode active mtu 9216  end write memory </pre>	<pre> interface breakout 1/1/3 map 40g-1x interface breakout 1/1/4 map 40g-1x  interface port-channel49 description R1-Leaf-AandB no shutdown switchport mode trunk switchport trunk allowed vlan 100 mtu 9216 vlt-port-channel 49  interface port-channel50 description R2-Leaf-AandB no shutdown switchport mode trunk switchport trunk allowed vlan 200 mtu 9216 vlt-port-channel 50  interface ethernet1/1/1:1 description R1-Leaf-A no shutdown no switchport channel-group 49 mode active mtu 9216  interface ethernet1/1/2:1 description R1-Leaf-B no shutdown no switchport channel-group 49 mode active mtu 9216  interface ethernet1/1/3:1 description R2-Leaf-A no shutdown no switchport channel-group 50 mode active mtu 9216  interface ethernet1/1/4:1 description R2-Leaf-B no shutdown no switchport channel-group 50 mode active mtu 9216  end write memory </pre>

Finally, exit configuration mode and save the configuration.

## 7.2 Configuration of S4048-ON OS10EE Leaf Switches

Configurations of Leaf 1 and Leaf 2 running OS10EE will be shown next as seen in Figure 23.

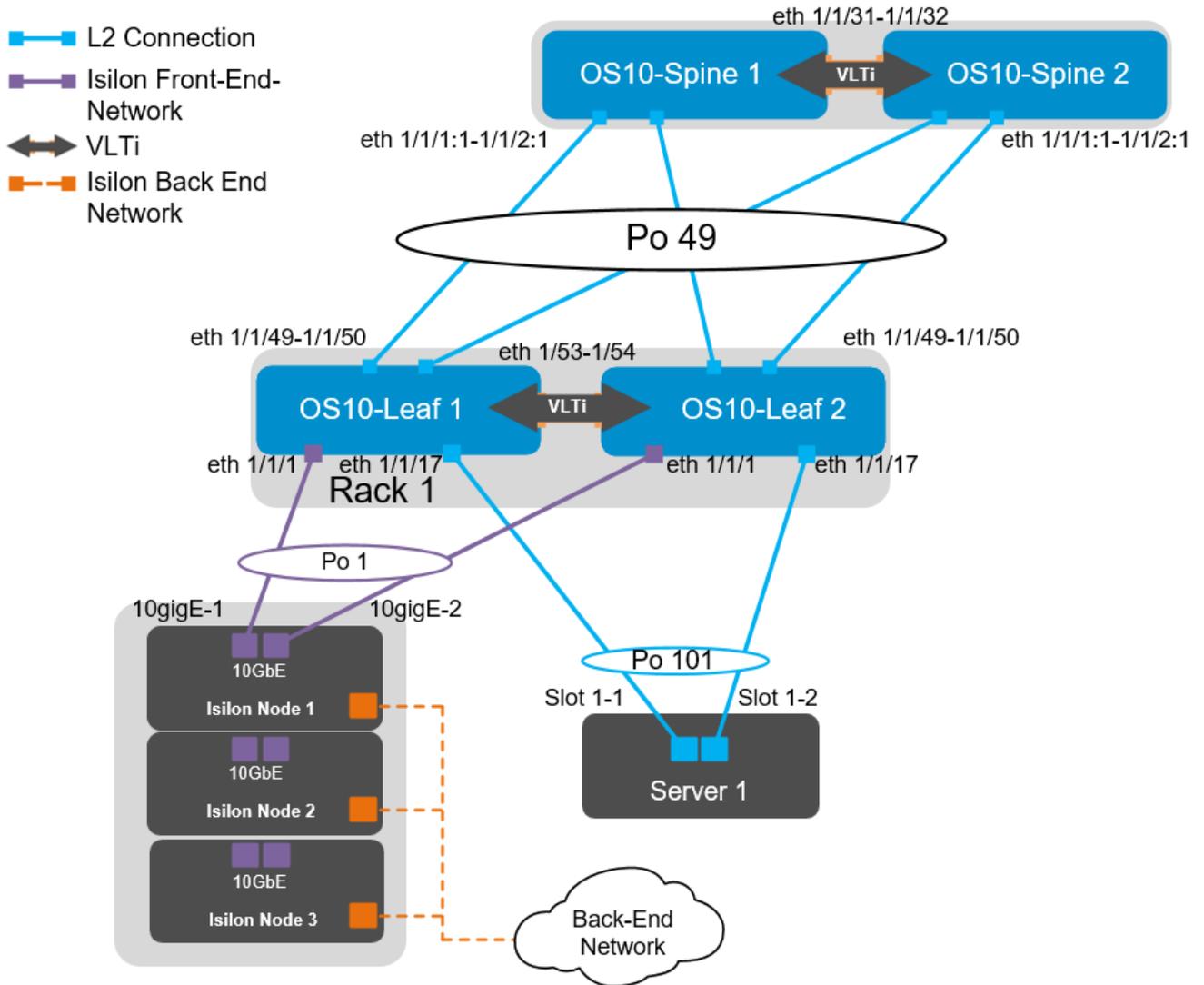


Figure 23 OS10EE Leaf pair

The interfaces for additional Isilon nodes and servers in this example can be found in the following table:

Table 6 Interface enumeration

Server/Isilon Node	Switch	Leaf interface
Node 01-1	Leaf 1	ethernet 1/1/1
	Leaf 2	ethernet 1/1/1
Node 01-2	Leaf 1	ethernet 1/1/3
	Leaf 2	ethernet 1/1/3
Node 01-3	Leaf 1	ethernet 1/1/5
	Leaf 2	ethernet 1/1/5
Server 01	Leaf 1	ethernet 1/1/17
	Leaf 2	ethernet 1/1/17
Server 02	Leaf 1	ethernet 1/1/19
	Leaf 2	ethernet 1/1/19

First, set the hostname, configure the OOB management interface and default gateway. Enable RSTP as a precaution for network loops. The RSTP priority has been set at the spine layer; each leaf switch will be left at the default value for RSTP. The `no iscsi enable` command is issued to ensure that iSCSI snooping is disabled.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>configure terminal hostname S4048-Leaf1 interface mgmt 1/1/1 no ip address dhcp ip address 100.67.170.15/24 management route 0.0.0.0/0 100.67.170.254 no iscsi enable</pre>	<pre>configure terminal hostname S4048-Leaf2 interface mgmt 1/1/1 no ip address dhcp ip address 100.67.170.14/24 management route 0.0.0.0/0 100.67.171.254 no iscsi enable</pre>

Configure the VLT interconnect between S4048-Leaf1 and S4048-Leaf2. In this configuration, remove each interface from layer 2 mode with the `no switchport` command for interfaces eth 1/1/53-1/1/54. Once this is done, enter the VLT domain, add the backup destination, and add the interfaces participating in VLTi. Finally, enable VLT peer routing on each switch. Once VLT is enabled, it will dynamically change the MTU settings for the VLTi to a value of 9216. Isilon's OneFS supports and recommends the use of jumbo frames.

**Note:** Refer to [Isilon Network Design Considerations](#) for more information on specific protocols as they relate to OneFS.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface range ethernet 1/1/53-1/1/54 description VLTi no switchport no shutdown  vlt-domain 127 backup destination 100.67.170.14 discovery-interface ethernet 1/1/53-1/1/54</pre>	<pre>interface range ethernet 1/1/53-1/1/54 description VLTi no switchport no shutdown  vlt-domain 127 backup destination 100.67.170.15 discovery-interface ethernet 1/1/53-1/1/54</pre>

Next, we will create the VLANs. IP addresses and VRRP is configured at the spine layer. On each leaf switch, the respective VLAN will be defined without an IP address or VRRP.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface vlan100 no shutdown</pre>	<pre>interface vlan100 no shutdown</pre>

Configure the downstream port channels. There will be one port channel for each server as well as a single port channel for each of the Isilon nodes. Each port channel is assigned an MTU of 9216 to support jumbo frames, as well configured to allow access to VLAN 100. The `vlt-port-channel #` command enables port channels on VLT peers to function as a single port channel.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface port-channel1 description "Isilon Node 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 1 spanning-tree port type edge</pre>	<pre>interface port-channel1 description "Isilon Node 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 1 spanning-tree port type edge</pre>
<pre>interface port-channel3 description "Isilon Node 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 3 spanning-tree port type edge</pre>	<pre>interface port-channel3 description "Isilon Node 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 3 spanning-tree port type edge</pre>
<pre>interface port-channel5 description "Isilon Node 3" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 5 spanning-tree port type edge</pre>	<pre>interface port-channel5 description "Isilon Node 3" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 5 spanning-tree port type edge</pre>
<pre>interface port-channel101 description "Server 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 101 spanning-tree port type edge</pre>	<pre>interface port-channel101 description "Server 1" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 101 spanning-tree port type edge</pre>
<pre>interface port-channel102 description "Server 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 102 spanning-tree port type edge</pre>	<pre>interface port-channel102 description "Server 2" no shutdown switchport access vlan 100 mtu 9216 vlt-port-channel 102 spanning-tree port type edge</pre>

The downstream interfaces will be configured in the next section. Dell recommends asymmetrical flow control (rx on tx off) on S4048-ON leaf switches which have sufficient egress buffers to reduce packet drops and resulting retransmissions. This feature should be applied on all host interfaces that support flow control. In OS10EE this is enabled by default, in OS9 it must be configured. Flow control commands are added below for reference; however, they are not required in OS10EE.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface ethernet1/1/1 description "Isilon Node 1-10gige-1" no shutdown channel-group 1 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/1 description "Isilon Node 1-10gige-2" no shutdown channel-group 1 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/3 description "Isilon Node 2-10gige-1" no shutdown channel-group 3 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/3 description "Isilon Node 2-10gige-2" no shutdown channel-group 3 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/5 description "Isilon Node 3-10gige-1" no shutdown channel-group 5 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/5 description "Isilon Node 3-10gige-2" no shutdown channel-group 5 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/17 description "Server 1-1" no shutdown channel-group 101 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/17 description "Server 1-2" no shutdown channel-group 101 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>
<pre>interface ethernet1/1/19 description "Server 2-1" no shutdown channel-group 102 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>	<pre>interface ethernet1/1/19 description "Server 2-2" no shutdown channel-group 102 mode active no switchport mtu 9216 flowcontrol receive on flowcontrol transmit off</pre>

Next, Configure the upstream port channels as well as the corresponding interfaces.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>interface port-channel49 description "Spine 1" no shutdown switchport mode trunk switchport trunk allowed vlan 100 mtu 9216 vlt-port-channel 49  interface ethernet1/1/49 description "Spine 1" no shutdown no switchport channel-group 49 mode active mtu 9216  interface ethernet1/1/50 description "Spine 2" no shutdown no switchport channel-group 49 mode active mtu 9216</pre>	<pre>interface port-channel49 description "Spine 1" no shutdown switchport mode trunk switchport trunk allowed vlan 100 mtu 9216 vlt-port-channel 49  interface ethernet1/1/49 description "Spine 1" no shutdown no switchport channel-group 49 mode active mtu 9216  interface ethernet1/1/50 description "Spine 2" no shutdown no switchport channel-group 49 mode active mtu 9216</pre>

Uplink Failure Detection (UFD) is configured next. UFD is a feature that shuts down specified downstream interfaces when all specified uplinks become disabled.

S4048-ON Leaf 1	S4048-ON Leaf 2
<pre>uplink-state-group 1 enable downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream port-channel49  end write memory</pre>	<pre>uplink-state-group 1 enable downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream port-channel49  end write memory</pre>

Finally, exit configuration mode and save the configuration.

## 7.3 Configuration of S4048-ON OS9 Leaf Switches

Configurations in this section cover the leaf switches 3 and 4 running OS9 as seen in Figure 24.

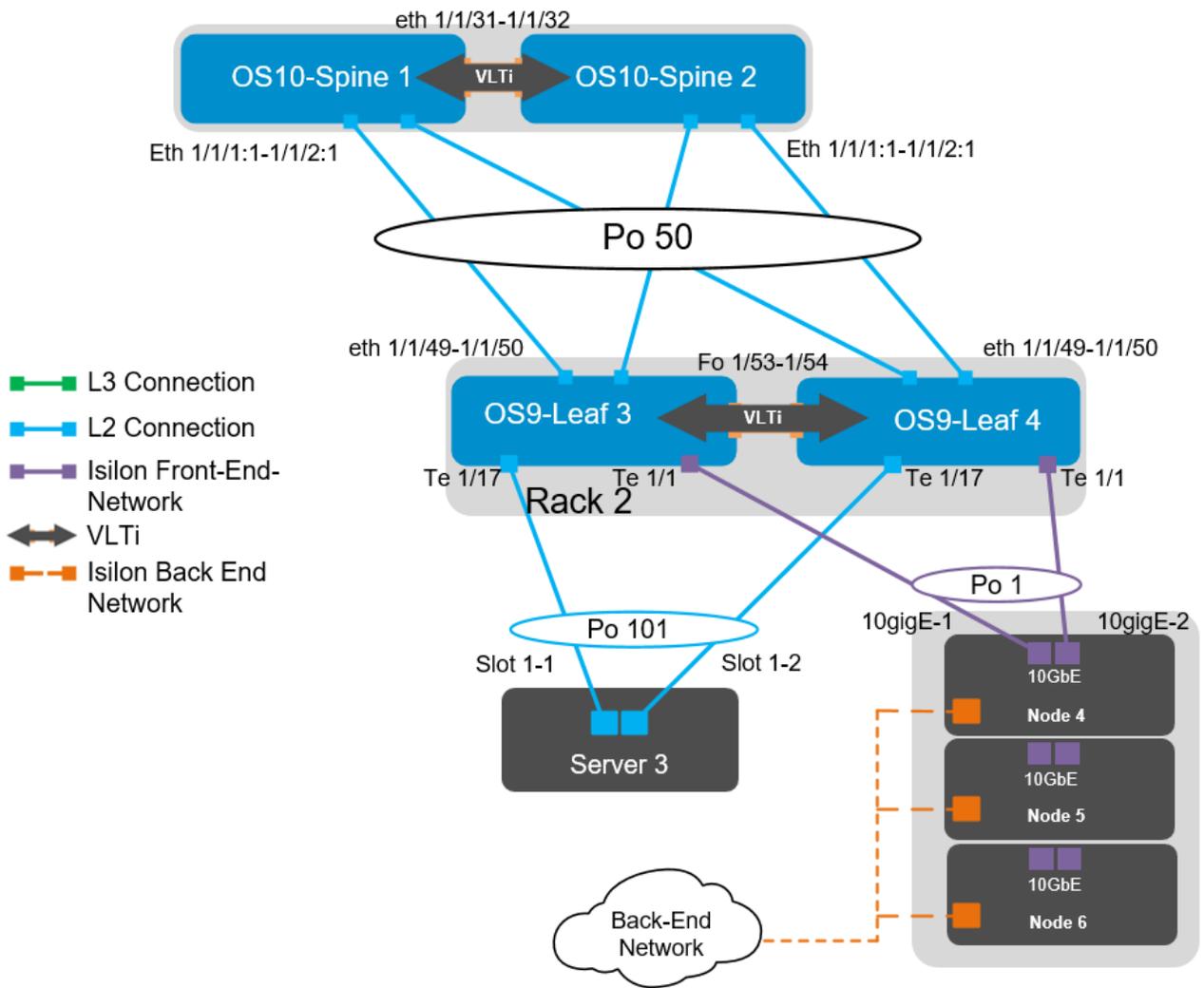


Figure 24 OS9 Leaf pair

The ports for additional Isilon nodes and servers in this example can be found in the following table:

Table 7 Interface enumeration

Server/Isilon Node	Switch	Leaf interface
Node 01-4	Leaf 3	tengigabitethernet 1/1/1
	Leaf 4	tengigabitethernet 1/1/1
Node 01-5	Leaf 3	tengigabitethernet 1/1/3
	Leaf 4	tengigabitethernet 1/1/3
Node 01-6	Leaf 3	tengigabitethernet 1/1/5
	Leaf 4	tengigabitethernet 1/1/5

Server/Isilon Node	Switch	Leaf interface
Server 03	Leaf 3	tengigabitethernet 1/1/17
	Leaf 4	tengigabitethernet 1/1/17
Server 04	Leaf 3	tengigabitethernet 1/1/19
	Leaf 4	tengigabitethernet 1/1/19

First, set the hostname, configure the OOB management interface and default gateway. Next, enable LLDP as well as RSTP. In OS10EE, LLDP is enabled by default, in OS9, it must be enabled.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>configure terminal hostname S4048-Leaf3 interface management 1/1 ip address 100.67.171.15/24 no shutdown management route 0.0.0.0/0 100.67.171.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable</pre>	<pre>configure terminal hostname S4048-Leaf4 interface management 1/1 ip address 100.67.171.15/24 no shutdown management route 0.0.0.0/0 100.67.171.254 protocol lldp advertise management-tlv management- address system-description system-name advertise interface-port-desc protocol spanning-tree rstp no disable</pre>

Configure the VLT interconnect between S4048-Leaf3 and S4048-Leaf4. Create a static port channel and add the two VLTi connections to port channel 128. Configure the VLT domain using the management interface of the peer switch as the backup destination. The backup destination is only utilized for the VLT heartbeat. The `unit-id #` command will designate the primary VLT member. Once VLT is enabled, it will dynamically change the MTU settings for the VLTi to a value of 9216. Isilon's OneFS supports and recommends the use of jumbo frames.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre> interface port-channel128 description VLTi channel-member fortyGigE 1/53 channel-member fortyGigE 1/54 no shutdown no switchport  interface fortyGigE 1/53 description VLTi no shutdown no switchport  interface fortyGigE 1/54 description VLTi no shutdown no switchport  vlt domain 127 peer-link port-channel 128 back-up destination 100.67.171.14 unit-id 0 </pre>	<pre> interface port-channel128 description VLTi channel-member fortyGigE 1/53 channel-member fortyGigE 1/54 no shutdown no switchport  interface fortyGigE 1/53 description VLTi no shutdown no switchport  interface fortyGigE 1/54 description VLTi no shutdown no switchport  vlt domain 127 peer-link port-channel 128 back-up destination 100.67.171.15 unit-id 1 </pre>

**Note:** When port channel 128 is added to the vlt domain the MTU setting is automatically changed to 9216.

The downstream interfaces will be configured in the next section. Dell recommends asymmetrical flow control (rx on tx off) on S4048-ON leaf switches which have sufficient egress buffers to reduce packet drops and resulting retransmissions. This feature should be applied on all host interfaces that support flow control. In OS10EE this is enabled by default, in OS9 it must be configured. Flow Control is configured with the `flowcontrol rx on tx off` command.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface tengigabitethernet 1/1 description "Isilon Node 4-10gige-1" no shutdown port-channel-protocol LACP port-channel 1 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/1 description "Isilon Node 4-10gige-2" no shutdown port-channel-protocol LACP port-channel 1 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/3 description "Isilon Node 5-10gige-1" no shutdown port-channel-protocol LACP port-channel 3 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/3 description "Isilon Node 5-10gige-2" no shutdown port-channel-protocol LACP port-channel 3 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/5 description "Isilon Node 6-10gige-1" no shutdown port-channel-protocol LACP port-channel 5 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/5 description "Isilon Node 6-10gige-2" no shutdown port-channel-protocol LACP port-channel 5 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/17 description "Server 3-1" no shutdown port-channel-protocol LACP port-channel 101 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/17 description "Server 3-2" no shutdown port-channel-protocol LACP port-channel 101 mode active mtu 9216 flowcontrol rx on tx off</pre>
<pre>interface tengigabitethernet 1/19 description "Server 4-1" no shutdown port-channel-protocol LACP port-channel 102 mode active mtu 9216 flowcontrol rx on tx off</pre>	<pre>interface tengigabitethernet 1/19 description "Server 4-2" no shutdown port-channel-protocol LACP port-channel 102 mode active mtu 9216 flowcontrol rx on tx off</pre>

Configure the necessary downstream port channels. There will be one port channel for each server as well as a single port channel for each of the Isilon nodes. Each port channel is assigned an MTU of 9216 to support jumbo frames. The `vlt-peer-lag-port-channel #` command ensures that each switch in the VLT is aware of all port channels and can act accordingly in the event of a switch failure.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface port-channel1 description "Isilon Node 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 1 spanning-tree rstp edge-port</pre>	<pre>interface port-channel1 description "Isilon Node 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 1 spanning-tree rstp edge-port</pre>
<pre>interface port-channel3 description "Isilon Node 5" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 3 spanning-tree rstp edge-port</pre>	<pre>interface port-channel3 description "Isilon Node 5" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 3 spanning-tree rstp edge-port</pre>
<pre>interface port-channel5 description "Isilon Node 6" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 5 spanning-tree rstp edge-port</pre>	<pre>interface port-channel5 description "Isilon Node 6" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 5 spanning-tree rstp edge-port</pre>
<pre>interface port-channel101 description "Server 3" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 101 spanning-tree rstp edge-port</pre>	<pre>interface port-channel101 description "Server 3" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 101 spanning-tree rstp edge-port</pre>
<pre>interface port-channel102 description "Server 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 102 spanning-tree rstp edge-port</pre>	<pre>interface port-channel102 description "Server 4" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 102 spanning-tree rstp edge-port</pre>

The upstream interfaces and port channels will be configured as follows:

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface port-channel50 description "Spine 1 and 2" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 50  interface fo 1/49 description "Spine 1" no shutdown port-channel-protocol LACP port-channel 50 mode active mtu 9216  interface fo 1/50 description "Spine 2" no shutdown port-channel-protocol LACP port-channel 50 mode active mtu 9216</pre>	<pre>interface port-channel50 description "Spine 1 and 2" no shutdown switchport mtu 9216 vlt-peer-lag port-channel 50  interface fo 1/49 description "Spine 1" no shutdown port-channel-protocol LACP port-channel 50 mode active mtu 9216  interface fo 1/50 description "Spine 2" no shutdown port-channel-protocol LACP port-channel 50 mode active mtu 9216</pre>

Next, we will create the VLANs. There will be one VLAN for each Isilon subnet; this will also include the servers in each rack.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre>interface vlan200 no shutdown untagged port-channel 1,3,5,101,102 tagged port-channel 50</pre>	<pre>interface vlan200 no shutdown untagged port-channel 1,3,5,101,102 tagged port-channel 50</pre>

Uplink Failure Detection (UFD) is configured next. UFD shuts down all downstream interfaces if all uplinks fail.

S4048-ON Leaf 3	S4048-ON Leaf 4
<pre> uplink-state-group 1 downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream port-channel 50  end write memory </pre>	<pre> uplink-state-group 1 downstream port-channel1 downstream port-channel3 downstream port-channel5 downstream port-channel101 downstream port-channel102 upstream port-channel 50  end write memory </pre>

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

## 8 Validation

In addition to sending traffic between hosts, the configuration shown in both topologies can be validated with the commands shown in this section. The validation for the layer 2 topology will only reference sections [8.1.3](#) – [8.1.9](#) for OS10EE and sections [8.2.3](#) – [8.2.8](#) for OS9.

**Note:** For more information on commands and output, see the Command Line Reference Guide for the applicable switch (links to documentation are provided in [Appendix B](#)).

Command and output examples are provided for one spine and one leaf. The OS9 section will only cover the validation of a single leaf as the spine layer is running OS10EE. Command output on other switches is similar.

### 8.1 OS10EE Validation commands

This section will cover the validation of the S4048-ON leaf switches that are running OS10EE as well the Z9100-ON spine switches.

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

#### 8.1.1 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).

##### Z9100-Spine1# show ip bgp summary

```
BGP router identifier 10.0.1.1 local AS number 64601
Neighbor      AS      MsgRcvd   MsgSent   Up/Down   State/Pfx
192.168.1.1   64701   2119     2146     00:55:27  3
192.168.1.3   64702   2101     2113     00:54:38  3
192.168.1.5   64703   2291     2322     01:37:49  3
192.168.1.7   64704   2290     2310     01:37:36  3
```

##### S4048-Leaf1#show ip bgp summary

```
BGP router identifier 10.0.2.1 local AS number 64701
Neighbor      AS      MsgRcvd   MsgSent   Up/Down   State/Pfx
192.168.1.0   64601   131660    131649    3d:21:49:06  10
192.168.2.0   64602   131685    131625    3d:21:49:04  10
```

#### 8.1.2 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.

S4048-Leaf1 has two paths to all other leafs and two paths to Server 3's network, 172.16.2.0. 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.

**Z9100-Spine1# show ip route bgp**

Codes: C - connected

S - static

B - BGP, IN - internal BGP, EX - external BGP

O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,

N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,

E2 - OSPF external type 2, \* - candidate default,

+ - summary route, > - non-active route

Gateway of last resort is not set

Destination	Gateway	Dist/Metric	Last Change
B EX 10.0.1.2/32	via 192.168.1.1 via 192.168.1.3 via 192.168.1.5 via 192.168.1.7	20/0	00:54:54
B EX 10.0.2.1/32	via 192.168.1.1	20/0	00:55:44
B EX 10.0.2.2/32	via 192.168.1.3	20/0	00:54:54
B EX 10.0.2.3/32	via 192.168.1.5	20/0	01:38:06
B EX 10.0.2.4/32	via 192.168.1.7	20/0	01:37:52
B EX 172.16.1.0/24	via 192.168.1.1 via 192.168.1.3	20/0	00:54:54
B EX 172.16.2.0/24	via 192.168.1.5 via 192.168.1.7	20/0	01:37:49

**S4048-Leaf1# show ip route bgp**

Codes: C - connected

S - static

B - BGP, IN - internal BGP, EX - external BGP

O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,

N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,

E2 - OSPF external type 2, \* - candidate default,

+ - summary route, > - non-active route

Gateway of last resort is not set

Destination	Gateway	Dist/Metric	Last Change
B EX 10.0.1.1/32	via 192.168.1.0	20/0	01:00:13
B EX 10.0.1.2/32	via 192.168.2.0	20/0	01:00:11
B EX 10.0.2.2/32	via 192.168.1.0 via 192.168.2.0	20/0	00:59:23
B EX 10.0.2.3/32	via 192.168.1.0 via 192.168.2.0	20/0	01:00:11
B EX 10.0.2.4/32	via 192.168.1.0 via 192.168.2.0	20/0	01:00:11

```

B   EX 172.16.2.0/24   via 192.168.1.0           20/0           01:00:11
                        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.

### 8.1.3 Show VLT

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

```
S4048-Leaf1# show vlt 127
```

```

Domain ID           : 127
Unit ID            : 2
Role               : secondary
Version           : 1.0
Local System MAC address : 64:00:6a:e6:b6:14
VLT MAC address    : f4:8e:38:20:44:29
IP address         : fda5:74c8:b79e:1::2
Delay-Restore timer : 90 seconds
Peer-Routing       : Enabled
Peer-Routing-Timeout timer : 0 seconds
VLTi Link Status
  port-channel1000 : up

```

VLT Peer Unit ID	System MAC Address	Status	IP Address	Version
1	f4:8e:38:20:44:29	up	fda5:74c8:b79e:1::1	1.0

### 8.1.4 Show VLT backup-link

This command is used to verify that the back-up destination is set correctly and the peer heartbeat is up.

```
S4048-Leaf1# show vlt 127 backup-link
```

```

VLT Backup Link
-----
Destination: 100.67.170.14
Peer Heartbeat status: Up

```

### 8.1.5 Show vlt-port-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. The local and peer status must both be up.

```
S4048-Leaf1# show vlt 127 vlt-port-detail
```

```

vlt-port-channel ID : 1
VLT Unit ID      Port-Channel      Status      Configured ports      Active ports

```

```

-----
  1          port-channel1    up      1      1
* 2          port-channel1    up      1      1
vlt-port-channel ID : 3
VLT Unit ID  Port-Channel    Status  Configured ports  Active ports
-----
  1          port-channel3    up      1      1
* 2          port-channel3    up      1      1
vlt-port-channel ID : 5
VLT Unit ID  Port-Channel    Status  Configured ports  Active ports
-----
  1          port-channel5    up      1      1
* 2          port-channel5    up      1      1
vlt-port-channel ID : 101
VLT Unit ID  Port-Channel    Status  Configured ports  Active ports
-----
  1          port-channel101   up      1      1
* 2          port-channel101   up      1      1
vlt-port-channel ID : 102
VLT Unit ID  Port-Channel    Status  Configured ports  Active ports
-----
  1          port-channel102   up      1      1
* 2          port-channel102   up      1      1

```

### 8.1.6 Show VLT mismatch

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

```
S4048-Leaf1# show vlt 127 mismatch
```

```
Peer-routing mismatch:
No mismatch
```

```
VLAN mismatch:
No mismatch
```

```
VLT VLAN mismatch:
No mismatch
```

### 8.1.7 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 1
```

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
Upstream Interfaces: ethernet1/1/49(down) ethernet1/1/50(down)
Downstream Interfaces: port-channel1(disabled) port-channel101(disabled) port-
channel102(disabled) port-channel3(disabled) port-channel5(disabled)
```

**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-Leaf1# show interface port-channel 1
Port-channel 1 is up, line protocol is down(error-disabled[UFD])
--- Output is truncated ---
```

### 8.1.8 Show spanning-tree brief

This command validates STP 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 1,3,5,101,102 in this example) are edge ports.

```
S4048-Leaf1# show spanning-tree brief
Spanning tree enabled protocol rstp with force-version rstp
Executing IEEE compatible Spanning Tree Protocol
Root ID      Priority 0, Address 6400.6ae6.b614
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 0, Address 6400.6ae6.b614
We are the root
Configured hello time 2, max age 20, forward delay 15
Flush Interval 200 centi-sec, Flush Invocations 73
Flush Indication threshold 0 (MAC flush optimization is disabled)
Interface
Name          PortID  Prio Cost      Sts  Cost  Bridge ID          PortID
-----
port-channel1 128.2517 128 1000      FWD  0     0     6400.6ae6.b614    128.2517
port-channel3 128.2519 128 1000      FWD  0     0     6400.6ae6.b614    128.2519
port-channel5 128.2521 128 1000      FWD  0     0     6400.6ae6.b614    128.2521
port-channel101 128.2617 128 1000      FWD  0     0     6400.6ae6.b614    128.2617
port-channel102 128.2618 128 1000      FWD  0     0     6400.6ae6.b614    128.2618
--- Output is truncated ---
Interface
Name          Role    PortID  Prio  Cost      Sts  Cost  Link-type  Edge
-----
port-channel1 Desg    128.2517 128  1000      FWD  0     AUTO      Yes
port-channel3 Desg    128.2519 128  1000      FWD  0     AUTO      Yes
port-channel5 Desg    128.2521 128  1000      FWD  0     AUTO      Yes
```

```
port-channel101 Desg    128.2617 128    1000      FWD  0      AUTO      Yes
port-channel102 Desg    128.2618 128    1000      FWD  0      AUTO      Yes
--- Output is truncated ---
```

### 8.1.9 Show vrrp brief

This command is used to verify that VRRP is correctly configured and both switches in the VLT have the same virtual address.

```
S4048-Leaf1# show vrrp brief
```

Interface	Group	Priority	Preempt	State	Version	Master addr(s)	Virtual addr
vlan100	IPv4	100	150 true	master-state	2	172.16.1.252	172.16.1.254

## 8.2 OS9 Validation commands

### 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). The 2 spine neighbors are shown below.

```
S4048-Leaf3#show ip bgp summary
```

```
BGP router identifier 10.0.2.3, local AS number 64703
BGP local RIB : Routes to be Added 0, Replaced 0, Withdrawn 0
16 network entrie(s) using 1216 bytes of memory
30 paths using 3240 bytes of memory
BGP-RIB over all using 3270 bytes of memory
19 BGP path attribute entrie(s) using 3456 bytes of memory
17 BGP AS-PATH entrie(s) using 170 bytes of memory
2 neighbor(s) using 16384 bytes of memory
```

Neighbor	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/Pfx
192.168.1.4	64601	28655	128086	0	0	0	3d:19:43:08	14
192.168.2.4	64602	128677	128092	0	0	0	3d:19:43:06	14

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

S4048-Leaf3 has two paths to all other leafs and two paths to Server 1's network, 172.16.1.0. 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-Leaf3#show ip route bgp**

Destination	Gateway	Dist/Metric	Last Change
B EX 10.0.1.1/32	via 192.168.1.4	20/0	00:21:22
B EX 10.0.1.2/32	via 192.168.2.4	20/0	00:21:22
B EX 10.0.2.1/32	via 192.168.1.4	20/0	00:21:22
	via 192.168.2.4		
B EX 10.0.2.2/32	via 192.168.1.4	20/0	00:21:22
	via 192.168.2.4		
B EX 10.0.2.4/32	via 192.168.1.4	20/0	00:20:27
	via 192.168.2.4		
B EX 172.16.1.0/24	via 192.168.1.4	20/0	00:21:22
	via 192.168.2.4		

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

### 8.2.3 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 of one switch in the VLT pair is primary (not shown), and its peer switch is assigned the secondary role. Ensure peer routing is shown as enabled.

**S4048-Leaf3#show vlt brief**

```
VLT Domain Brief
-----
Domain ID:                127
Role:                     Secondary
Role Priority:            32768
ICL Link Status:         Up
HeartBeat Status:        Up
VLT Peer Status:         Up
Local Unit Id:           0
Version:                  6(8)
Local System MAC address: f4:8e:38:2e:a0:f7
Remote System MAC address: f4:8e:38:2e:76:f7
Remote system version:   6(8)
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
```

## 8.2.4 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. The local and peer status must both be up.

```
S4048-Leaf3#show vlt detail
```

Local LAG Id	Peer LAG Id	Local Status	Peer Status	Active VLANs
1	1	UP	UP	200
3	3	UP	UP	200
5	5	UP	UP	200
101	101	UP	UP	200
102	102	UP	UP	200

## 8.2.5 Show VLT mismatch

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

```
S4048-Leaf3#show vlt mismatch
```

```
S4048-Leaf3#
```

## 8.2.6 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-Leaf3#show uplink-state-group 1
```

```
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
```

When an interface or port-channel has been disabled by UFD, the show port-channel value command for affected interfaces indicates it is error-disabled as follows:

```
S4048-Leaf3#show interface port-channel 1
```

```
Port-channel 1 is up, line protocol is down(error-disabled[UFD])
```

```
-- Output truncated --
```

## 8.2.7 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 are edge ports.

### S4048-Leaf3#show spanning-tree rstp brief

```
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 0, Address f48e.382e.a0f7
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID  Priority 0, Address f48e.382e.a0f7
We are the root
Configured hello time 2, max age 20, forward delay 15
```

Interface Name	PortID	Prio	Cost	Sts	Cost	Designated Bridge ID	PortID
Po 1	128.2	128	1800	FWD(vlt)	0	f48e.382e.a0f7	128.2
Po 3	128.4	128	1800	FWD(vlt)	0	f48e.382e.a0f7	128.4
Po 5	128.6	128	1800	FWD(vlt)	0	f48e.382e.a0f7	128.6
Po 101	128.102	128	1800	FWD(vlt)	0	f48e.382e.a0f7	128.102
Po 102	128.103	128	1800	FWD(vlt)	0	f48e.382e.a0f7	128.103
Po 128	128.129	128	600	FWD(vltI)	0	f48e.382e.a0f7	128.129

Interface Name	Role	PortID	Prio	Cost	Sts	Cost	Link-type	Edge
Po 1	Desg	128.2	128	1800	FWD	0	(vlt) P2P	Yes
Po 3	Desg	128.4	128	1800	FWD	0	(vlt) P2P	Yes
Po 5	Desq	128.6	128	1800	FWD	0	(vlt) P2P	Yes
Po 101	Desg	128.102	128	1800	FWD	0	(vlt) P2P	Yes
Po 102	Desg	128.103	128	1800	FWD	0	(vlt) P2P	Yes
Po 128	Desg	128.129	128	600	FWD	0	(vltI) P2P	No

## 8.2.8 Show vrrp brief

This command is used to verify that VRRP is correctly configured and both switches in the VLT have the same virtual address.

### S4048-Leaf3#show vrrp brief

Interface Description	Group	Pri	Pre	State	Master addr	Virtual addr(s)
Vl 200	IPv4 200	150	Y	Master	172.16.2.252	172.16.2.254

## 9 Isilon configuration

The following section outlines the steps necessary to add the Isilon X210 nodes into a cluster, setup a functioning SMB share, designate a secondary subnet, and configure the SmartConnect feature in OneFS.

### 9.1 Building the cluster

The first step in configuring the Isilon array is building the cluster. Once the array is powered on, a serial connection can be made to each node. Connect to the first node and use the cluster creation wizard to build the cluster. Table 8 shows the values that were used in this example. If any additional information is needed on the setup and configuration of the Isilon Array cluster, refer to the [Isilon X210 Installation Guide](#).

```
Welcome to the Isilon IQ configuration wizard.
Copyright (c) 2001-2017 EMC Corporation. All Rights Reserved.
Enter 'help' at any prompt for additional information on that step.
Enter 'back' at any prompt to return to previous step.
Enter 'quit' at any prompt to discard all changes.
Enter 'manual' at any prompt to switch to manual mode (console).

Node build: Isilon OneFS v8.1.0.1 B_MR_8_1_0_1_157(RELEASE)
Node serial number: JWXER164900015

Select an option:
[ 1] Create a new cluster
[ 2] Join an existing cluster
[ 3] Exit wizard and configure manually
[ 4] Reboot into SmartLock Compliance mode
Wizard >>>
Please select an option.
```

Figure 25 Initial configuration wizard

The following table outlines the configurations that were made in this example. The following configurations will be made in this section:

- The cluster name is defined
- IP ranges are assigned for the back-end-network
- Configuration of the management 1GbE ext-1 connections are made using out-of-band management IP range 100.67.170.0/24

The 10gige-1 and 10gige-2 interfaces for user connectivity will be configured in the next section.

Table 8 Isilon configuration parameters

Parameter	Value
Option	[ 1] Create a new cluster
Configure root/admin passwords	"password"
Configuration Name	Cluster-01
Cluster Encoding	[ 8] UTF-8

Parameter	Value
Interface int-a netmask	255.255.255.0
Interface int-a IP Range	10.10.10.1-253
Interface int-b netmask	255.255.255.0
Interface int-b IP Range	10.10.20.1-253
Failover IP range	10.10.30.1-253
Interface ext-1 netmask	255.255.255.0
MTU	1500
Interface ext-1 IP pool	100.67.170.140-148
ext-1 default gateway	100.67.170.254
SmartConnect Zone name	Mgmtzone1
SmartConnect service IP	100.67.170.149
DNS servers	100.67.10.1
Search domain	dell.local
Cluster date and time	Default
Cluster join mode	Default
Commit changes	Yes

**Note:** Additional information and configuration for SmartConnect will be provided in [Section 9.5](#).

When the first node is configured, and the cluster has been created, the other nodes can then be added into the cluster. All nodes will be able to discover the cluster over the back-end network. The assorted options for back-end networking and Isilon are not covered in this document, for more information on the back-end network refer to Isilon documentation.

1. At the login screen of node two, select **[ 2 ] Join existing cluster**.
2. The cluster that was created will be displayed, select **1** to add node two to Cluster-01.
3. Repeat for all nodes that need to be added to the cluster. Each node will be dynamically assigned an IP address out of the IP pool that was designated in the cluster creation.

```

Node build: Isilon OneFS v8.1.0.1 B_MR_8_1_0_1_157 (RELEASE)
Node serial number: JWXER164900015

Select an option:
[ 1] Create a new cluster
[ 2] Join an existing cluster
[ 3] Exit wizard and configure manually
[ 4] Reboot into SmartLock Compliance mode
Wizard >>> 2

Select the cluster you want to join
Index Name      Version      Status
-----
1      Cluster-01 B_MR_8_1_0_1_157R available
[Enter] Refresh the list
Join cluster >>> 1

```

Figure 26 Join cluster

## 9.2 Configuring LACP to each node

Once the cluster has been created, and each node has been added to the cluster, the rest of the configurations can be done through the OneFS web GUI. This section will cover the creation of two subnets and their associated IP pools, as well as the addition of static routes to ensure end to end connectivity. This will include ensuring that the LACP option is selected and the appropriate interfaces are selected.

1. Navigate to a web page and use the first IP address in the range defined for the ext-1 interface, <https://100.67.170.140:8080>. This will bring you to the first node in the cluster.

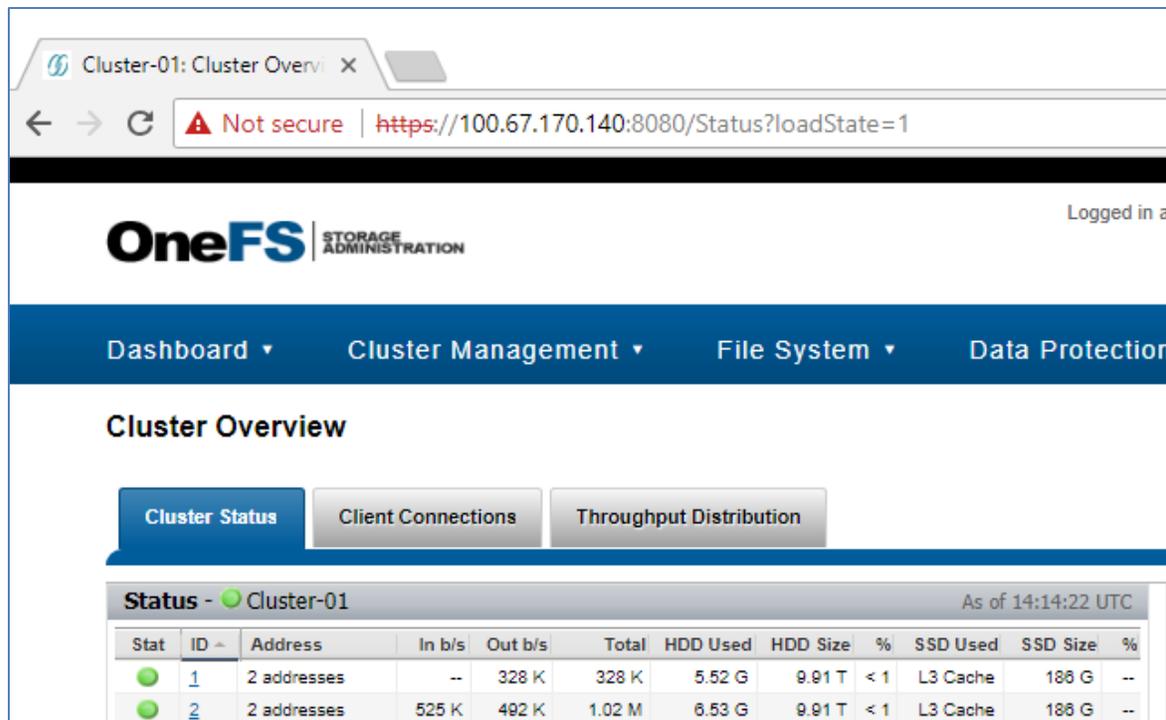


Figure 27 OneFS Web GUI

2. After successfully logging into the OneFS web GUI, navigate to **Cluster Management > Network Configuration**. This will allow changes to the network configuration.
3. Select the **More** option next to groupnet0, and **Add subnet**. Define values for the next two subnets. Refer to Table 9.

Table 9 Subnet Configurations

Parameter	Value
Subnet Name	subnet1
Netmask	255.255.255.0
Gateway Address	172.16.1.254
MTU	9000(Jumbo Frame)
SmartConnect Service IP	172.16.1.249
SmartConnect Service Name	sczone1.dell.local
Subnet Name	subnet2
Netmask	255.255.255.0
Gateway Address	172.16.2.254
MTU	9000(Jumbo Frame)
SmartConnect Service IP	172.16.2.249
SmartConnect Service Name	sczone2.dell.local

4. Select the **More** option next to subnet1, and **Add pool**.
5. Define the pool name, **pool1**.
6. Set the IP range, **172.16.1.1-50**.
7. **Add** in the **10gige-agg-1** interfaces for nodes 1-3 into the pool.
8. Set the SmartConnect Zone name, **sczone1**.
9. Set the SmartConnect Service subnet, **subnet1**.
10. Ensure that the **Aggregation Mode** is set to **LACP**.
11. Complete steps 4-10 to create **pool2** in **subnet2**.

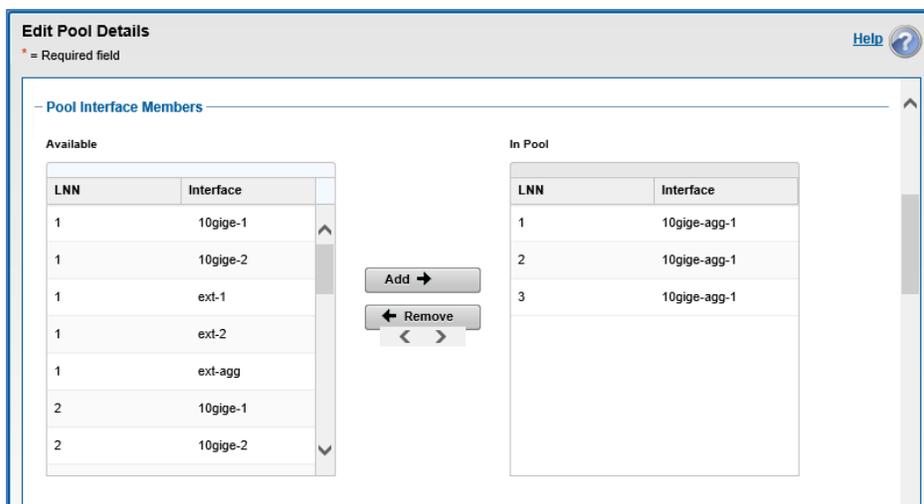


Figure 28 LACP configuration

The next section will cover the configuration of static routes within OneFS. This is necessary to enable traffic to cross between the two subnets.

1. Navigate **Network configurations>subnet1>pool1**, **Add static route** to subnet 2.

Parameter	Value
Subnet	172.16.2.0
Netmask	255.255.255.0
Gateway	172.16.1.254

2. Navigate **Network configurations>subnet2>pool2**, **Add static route** to subnet 1.

Parameter	Value
Subnet	172.16.1.0
Netmask	255.255.255.0
Gateway	172.16.2.254

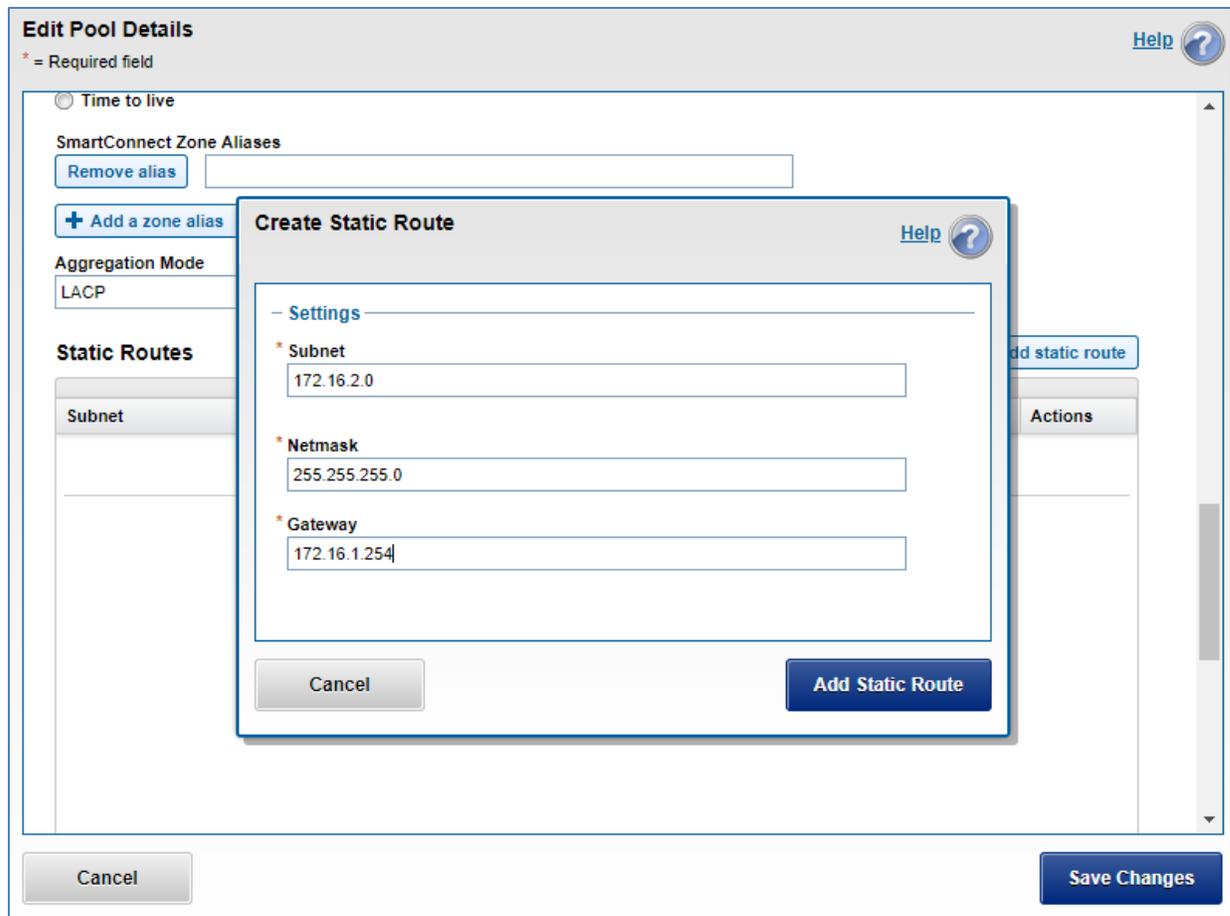


Figure 29 Static Routes

In addition to the Isilon nodes, each Windows server is configured in an LACP NIC team that utilizes the address hash setting. The specific configuration steps for a Windows NIC team can be found in Microsoft's documentation, [Create a New NIC Team on a Host Computer or VM](#).

### 9.3 Configuring SMB share

This section will cover the creation of a folder structure that will then be configured under OneFS Protocols Windows Sharing SMB. Once the folders are added to the Windows Sharing SMB section and the network is properly configured, the folder will be accessible via a Windows server.

1. From the OneFS GUI select **File System > File System Explorer**.
2. Select **Create Directory**.
3. Define the directory name, "**SMB**," as well as the user, group, and permission settings as shown.

**Create a Directory**

\* = Required field

**Settings**

Containing Path  
/ifs

\* Directory Name  
SMB

\* User  
root

\* Group  
wheel

**Permissions**

Type	Read	Write	Execute
User	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Group	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Others	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Figure 30 Directory creation

4. Once the folder has been created, click the **SMB** folder.
5. Once again create a directory and assign the name, "**Engineering**," user, group, and permissions settings.
6. When both folders have been created, navigate to the **Protocols > Windows Sharing (SMB)**.
7. Select **Create an SMB Share**.
8. Assign a name, "**Engineering SMB**," use the **Browse** button to navigate to the **Engineering** folder that was created previously.
9. Select the member's window and assign access permissions.
10. All other selections were left at the default values.
11. Click **Create Share**.

## 9.4 Setting up DNS for SmartConnect

One of the features of the OneFS file system is SmartConnect. SmartConnect allows the workload to be dynamically reassigned to other nodes in the cluster in the event of a node failure. SmartConnect is a feature that must be configured. There is also a DNS dependency for SmartConnect to function properly. Configuring DNS host A records and Delegation zones is necessary. Once this is done, SmartConnect will assign the workload to each node using the round robin method. Using the FQDN that was configured previously for the SmartConnect zones is required. This section will cover the completion of the SmartConnect configuration.

1. Login to the DNS server.

**Note:** Refer to [Microsoft's DNS Resource Record Management](#) for the specific steps in creating a DNS server.

2. Navigate to **Server Manager > Tools > DNS** and configure the following:

Table 10 DNS parameters

Parameter	Value
Forward Lookup Zone	dell.local
Reverse Lookup Zones	172.16.1, 172.16.2
Host A record for sczone1	sczonea, 172.16.1.249
Host A record for sczone2	sczoneb, 172.16.2.249
Delegation zone for sczone1	sczone1.dell.local, 172.16.1.249
Delegation zone for sczone2	sczone2.dell.local, 172.16.2.249

It is important that the DNS Host A record have a unique name that does not match the name that was assigned to the SmartConnect zone within the OneFS file system. The Delegation zone, however, needs to have a name that matches the SmartConnect zone. This will ensure that a connection to sczone1.dell.local will reference the SmartConnect Zone 1 IP of 172.16.1.249. When this IP is referenced, it will round robin through each node IP, assigning the workload to each node in turn, ensuring a balanced workload.

## 9.5 Validation

At this point in the configuration, the file system and the network can be validated. This can be accomplished by selecting a server on one leaf pair and connecting to the file system assigned to the opposite subnet.

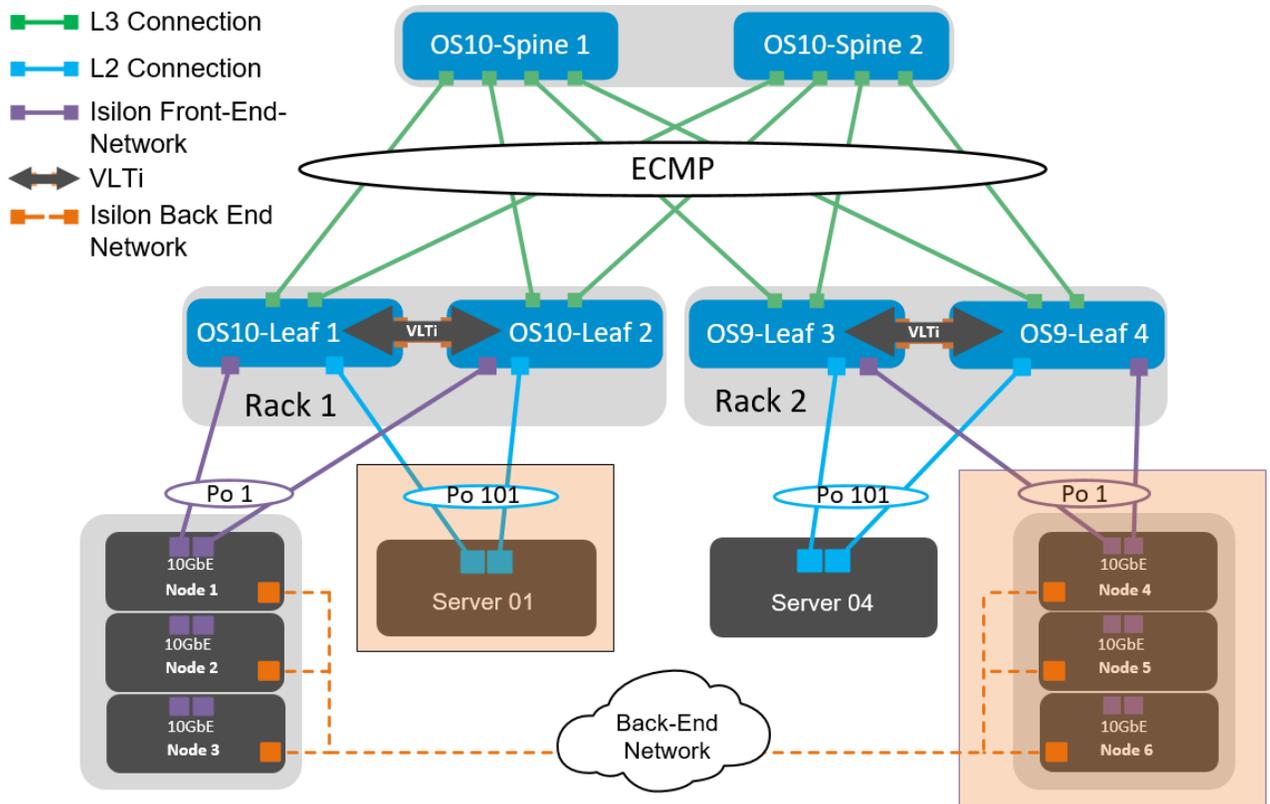


Figure 31 Validation

1. Login to Server 01 and issue a Run command of `\\sczone2.dell.local`. This will bring up the folder structure that was created in the OneFS file system.
2. Create a file or a folder in order to verify read/write permissions.
3. Login to Server04 and issue the Run command `\\sczone1.dell.local`. This will display the same folder, and the previously created files will be present.

In addition to accessing the files on the opposite subnet, a test should be done to ensure that each server can access the files through the local subnet. Example: **Server01** to `\\sczone1.dell.local`.

## A Validated hardware and components

The following tables list the hardware and components used to configure and validate the example configurations in this guide.

### A.1 Dell EMC Networking Switches

Qty	Item	OS/Firmware version	
2	Z9100-ON	OS	10.4.0 R3
2	S4048-ON	OS	10.4.0 R3
2	S4048-ON	OS	9.13 (0.0)
1	S3048-ON	OS	9.13 (0.0)

### A.2 Dell EMC Isilon Array

Qty	Item	OS/Firmware version
6	Isilon X210	8.1.0.1

### A.3 Dell EMC PowerEdge Servers

Qty	Item	OS/Firmware version	
2	R730xd	OS	Windows Server 2016
		BIOS	2.5.5
		iDRAC	2.50.50.50
		Mellanox ConnectX-4 LX 25GbE SFP adapter	14.20.18.20
		Inter®Ethernet 10G X520	18.3.6
		Intel® Ethernet 10G X710 rNDC	18.3.6
2	R640	OS	Windows Server 2016
		BIOS	1.2.11
		iDRAC	3.15.17.15
		Intel® Ethernet 10G X710 rNDC	18.3.6

## B Product Manuals and technical guides

[Manuals and documentation for Dell Networking Z9100-ON](#)

[Manuals and documentation for Dell Networking S4048-ON](#)

[Manuals and documentation for Dell Networking S3048-ON](#)

[Manuals and documentation for PowerEdge R730xd](#)

[Manuals and documentation for PowerEdge R640](#)

[Isilon OneFS](#)

[Isilon X210 Installation Guide](#)

[Create a New NIC Team on a Host Computer or VM](#)

[Microsoft's DNS Resource Record Management](#)

[Dell EMC Leaf-Spine Deployment Guide](#)

[Dell EMC Networking L3 Design for Leaf-Spine with OS10EE](#)

## C 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](mailto:Dell_Networking_Solutions@Dell.com).