

Dell Networking with Dell-VMware EVO:RAIL

A Dell Reference Architecture

May 2015



Revisions

| Date | Description | Authors |
|----------|---------------------------|-------------------------------|
| 05/27/15 | Version 1.2 | Humair Ahmed, Dell Networking |
| 05/27/15 | Version 1.1 | Humair Ahmed, Dell Networking |
| 05/22/15 | Version 1 Initial release | Humair Ahmed, Dell Networking |

©2015 Dell Inc., All rights reserved.

Except as stated below, no part of this document may be reproduced, distributed or transmitted in any form or by any means, without express permission of Dell.

You may distribute this document within your company or organization only, without alteration of its contents.

THIS DOCUMENT IS PROVIDED "AS-IS", AND WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED. IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE SPECIFICALLY DISCLAIMED. PRODUCT WARRANTIES APPLICABLE TO THE DELL PRODUCTS DESCRIBED IN THIS DOCUMENT MAY BE FOUND AT:

http://www.dell.com/learn/us/en/19/terms-of-sale-commercial-and-public-sector Performance of network reference architectures discussed in this document may vary with differing deployment conditions, network loads, and the like. Third party products may be included in reference architectures for the convenience of the reader. Inclusion of such third party products does not necessarily constitute Dell's recommendation of those products. Please consult your Dell representative for additional information.

Trademarks used in this text:

Dell™, the Dell logo, Dell Boomi™, Dell Precision™, OptiPlex™, Latitude™, PowerEdge™, PowerVault™, PowerConnect™, OpenManage™, EqualLogic™, Compellent™, KACE™, FlexAddress™, Force10™ and Vostro™ are trademarks of Dell Inc. Other Dell trademarks may be used in this document. Cisco Nexus®, Cisco MDS®, Cisco NX-0S®, and other Cisco Catalyst® are registered trademarks of Cisco System Inc. EMC VNX®, and EMC Unisphere® are registered trademarks of EMC Corporation. Intel®, Pentium®, Xeon®, Core® and Celeron® are registered trademarks of Intel Corporation in the U.S. and other countries. AMD®is a registered trademark and AMD Opteron™, AMD Phenom™ and AMD Sempron™ are trademarks of Advanced Micro Devices, Inc. Microsoft®, Windows®, Windows Server®, Internet Explorer®, MS-DOS®, Windows Vista® and Active Directory® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Red Hat® and Red Hat® Enterprise Linux® are registered trademarks of Red Hat, Inc. in the United States and/or other countries. Novell® and SUSE® are registered trademarks of Novell Inc. in the United States and other countries. Oracle® is a registered trademark of Oracle Corporation and/or its affiliates. Citrix®, Xen®, XenServer® and XenMotion® are either registered trademarks or trademarks of Citrix Systems, Inc. in the United States and/or other countries. VMware®, Virtual SMP®, vMotion®, vCenter® and vSphere® are registered trademarks or trademarks of VMware, Inc. in the United States or other countries. IBM® is a registered trademark of International Business Machines Corporation. Broadcom® and NetXtreme® are registered trademarks of Broadcom Corporation. Qlogic is a registered trademark of QLogic Corporation. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and/or names or their products and are the property of their respective owners. Dell disclaims proprietary interest in the marks and names of others.



Table of contents

| Re | vision | S | 2 |
|----|--------|---|----|
| 1 | | view | |
| 2 | Supp | ported Dell Networking ToR Switches and Requirements | 5 |
| | 2.1 | Dell Networking Requirements for EVO:RAIL | 5 |
| | 2.2 | Recommended Dell Networking Switches | 6 |
| 3 | Unde | erstanding EVO:RAIL | 7 |
| | 3.1 | Dell-VMware EVO:RAIL Hardware Specs | 7 |
| | 3.2 | Dell-VMware EVO:RAIL Auto-configuration | 7 |
| 4 | Dell- | -VMware EVO:RAIL Network Architecture and Configuration | 9 |
| | 4.1 | Dell Infrastructure Physical Architecture | 9 |
| | 4.2 | Understanding EVO:RAIL Network Configuration | 15 |
| | 4.3 | Dell Networking Switch Configuration | 20 |
| | 4.4 | EVO:RAIL Configuration | 29 |
| 5 | Dell- | -VMware EVO:RAIL Setup Validation | 46 |
| 6 | Con | clusion | 51 |



1 Overview

The Dell-VMware EVO:RAIL solution is a highly automated, hyper-converged infrastructure appliance. The word **EVO** is used to define a new family of **evolutionary** hyper-converged infrastructure offerings from VMware. The word **RAIL** represents the rail mount attached to the 2RU appliance with four server nodes which can easily be mounted within a datacenter rack; it is also the first product within the EVO family. The Dell-VMware EVO:RAIL solution is EVO:RAIL customized and running on reliable Dell hardware providing a hyper-converged, cost-effective virtualization platform.



Figure 1 Dell-VMware EVO:RAIL Front View

The Dell-VMware EVO:RAIL solution includes the industry leading VMware vSphere server virtualization platform and VMware VSAN storage virtualization technology installed on a Dell 2RU four server node hardware appliance. The Dell EVO:RAIL 2RU chassis has 2 x 2.10 GHz 6 core Intel

CPUs per node, 192 GB of physical memory per node, 2×10 GbE connectivity per server node, and 13.1 TB of usable VSAN storage per appliance. It can scale up to eight appliances combined for a total of thirty-two server nodes and 104.8 TB usable VSAN storage.

Customers can now get up and running with a complete prepackaged virtualization solution running on reliable Dell hardware in minutes. A pre-installed VMware management engine automates the discovery of server nodes and configuration of the hyper-converged appliance.

Each server node has either 2 x 10GbE fiber or 2 x 10GbE copper ports for a total

Engineered Solutions optimized for VMware EVO: RAIL

Dell 2U/4 nodes of shared local storage
Single support contact for hardware/software
Value-added software, services and support

VMware EVO: RAIL Software Bundle
EVO: RAIL Rapid deployment, configuration, and management engine
Compute, network & storage virtualization with vSphere & vSAN

Virtual Infrastructure Edition
Ideal for virtual infrastructure and general-purpose workload deployments

VMware Horizon Edition
Ideal for virtual desktop infrastructure (VDI deployments based on VMware Horizon)

of 8 x 10GbE ports per Dell-VMware EVO:RAIL appliance. There is also a Baseboard Management Controller (BMC) interface on each node for remote management. The EVO:RAIL appliance can be connected to a pair of ToR Dell Networking switches for network connectivity and high availability (HA).



Figure 2 Dell-VMware EVO:RAIL Back View

The Dell-VMware EVO:RAIL appliance allows for fast, automated, and less error-prone configuration of a virtualized environment. It is bundled with the following components:

- VMware ESXi 5.5U2 on each server node (vSphere Enterprise Plus)
- VMware vCenter Server Appliance (VCSA) 5.5U2
- VMware VSAN 5.5
- VMware vRealize Log Insight 2.05



When looking at the EVO-Rail appliance from the back, as shown in Figure 2, the server nodes are positioned as follows:

- Node 1: top right
- Node 2: bottom right
- Node 3: top left
- Node 4: bottom left

The VMware vCenter Server Appliance and VMware vRealize Log Insight both run on server node 1, and VSAN is distributed across all nodes.

This whitepaper discusses the details of the network connectivity at the ToR for the Dell-VMware EVO:RAIL appliance, reference architecture and respective network configuration required, and scale-out scenarios.

2 Supported Dell Networking ToR Switches and Requirements

2.1 Dell Networking Requirements for EVO:RAIL

All 10/40GbE Dell Networking switches that support the basic networking requirements below are supported with Dell-VMware EVO:RAIL deployments.

- o IPv4 multicast support for VSAN.
- o IPv6 multicast support for EVO:RAIL auto-discovery/management software.
- IGMP Snooping and Querier are recommended but not required for VSAN. All Dell Networking N, S, and Z-series switches support IGMP Snooping and Querier which is used for IPv4 multicast.
- o Although a flat L2 network can be used, for performance and scalability, it is highly recommended to configure VLANs for vMotion, VSAN, management, and compute (VM).
- o If switches will be used as gateways, L3 capabilities and a HA gateway protocol such as VRRP or Routed VLT with Proxy ARP are required.
- Switch interconnectivity via stacking, VLTi, or port-channel is highly recommended for high availability.

Please note the following important points:



- o IPv4 and IPv6 multicast must be enabled on either all ports or specifically the VSAN and Management VLAN ports on the Dell ToR switch that are connected to EVO:RAIL.
- When using multiple Dell ToR switches, switch interconnectivity via stacking, VLTi, or portchannel is required so multicast traffic can communicate between the switches. This is supported by default on all Dell Networking N, S, and Z-series switches.
- LACP must not be used to connect the Dell Networking switches to the EVO:RAIL appliance. Specific VMware NIC teaming options are used instead.
- o The VSAN and Management VLAN must allow for pass-through of multicast traffic. This is supported by default on all Dell Networking N, S, and Z-series switches.
- o Two options exist for handling IPv4 multicast traffic on the ToR switches.
 - 1) Enable IGMP Snooping on ToR switches AND enable IGMP Querier. This is supported on all Dell Networking N, S, and Z-series switches. Note, IGMP Snooping is enabled by default on N-series switches.
 - 2) Disable IGMP Snooping and Querier on ToR switches. This option may lead to additional multicast traffic on the network. This is the default on all Dell Networking S and Z-series switches. N-Series switches have IGMP Snooping enabled by default.
- MLD Snooping and Querier for IPv6 multicast is not required on the Dell ToR switch due to the extremely low volume of IPv6 multicast that is used for EVO:RAIL autoconfiguration/management purposes. Dell Networking S-Series and Z-Series do not support MLD Snooping and Querier. N-Series switches support MLD Snooping and Querier.

2.2 Recommended Dell Networking Switches

VMware EVO:RAIL requires 10GbE switches for connectivity. All Dell Networking 10/40 GbE N, S, and Z-series switches are supported by EVO:RAIL. However, as each EVO:RAIL server node within the EVO:RAIL appliance has either 2 x 10GbE fiber or 2 x 10GbE copper ports, the following Dell Networking switches are recommended for ToR connectivity to the EVO:RAIL appliance:

- o Dell S4810
- o Dell S4820T
- o Dell S4048-ON
- o Dell S5000
- o Dell S6000/S6000-ON
- o Dell N4032/N4064
- Dell N4032F/N4064F



3 Understanding EVO:RAIL

3.1 Dell-VMware EVO:RAIL Hardware Specs

The Dell-VMware EVO:RAIL hardware specs are outlined in Figure 3 below for both a single node, and all 4 x nodes as an aggregate within a single EVO:RAIL appliance. An EVO:RAIL cluster can scale up to 8 x EVO:RAIL appliances.

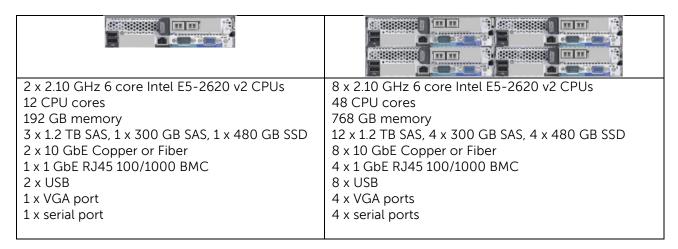


Figure 3 Dell-VMware EVO:RAIL Specs

Each EVO:RAIL node has 5 x hard drives for a total of 20 hard drives per Dell-VMware EVO:RAIL appliance. Each ESXi node has 3×1.2 TB SAS, 1×300 GB SAS, and 1×480 GB SSD drives. The 1×300 GB SAS drive in drive slot 1 on each node is used as the boot drive for ESXi, the 3×1.2 TB SAS drives in drive slots 2-4 are used for VSAN data, and the 1×480 GB SSD drive in drive slot 5 is used for VSAN read/write cache.

3.2 Dell-VMware EVO:RAIL Auto-configuration

EVO:RAIL automatically starts two main daemons when booted: 1) Loudmouth, or VMware's Zeroconf implementation for auto-configuration (on all nodes and VCSA), and 2) MARVIN, or the EVO:RAIL management daemon (on VCSA).

EVO:RAIL leverages VMware's own Zeroconf daemon called VMware Loudmouth. Zeroconf is a set of RFC-backed technologies that leverage each other to provide automated network configuration for services and devices. VMware Loudmouth is based on Zeroconf auto-discovery



Figure 4 Dell-VMware EVO:RAIL Nodes

capabilities and allows for the auto-configuration of the appliance. A Zeroconf/Loudmouth daemon resides on each of the ESXi nodes and inside the VCSA instance; this allows the EVO:RAIL management engine to discover all nodes and respectively automate the configuration.



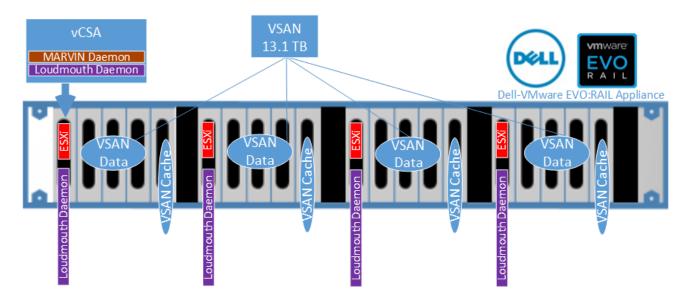


Figure 5 Dell-VMware EVO:RAIL

The Loudmouth service is designed to work over a layer 2-only domain. A new EVO:RAIL appliance will advertise itself on the network using the VMware Loudmouth service. The first EVO:RAIL appliance in a cluster will use a pre-installed instance of vSphere vCenter Server Appliance. Additional EVO:RAIL appliances detected and added to the cluster will use this same vCenter Server Appliance and delete their own local preinstalled vSphere vCenter Server Appliance and vRealize Log Insight. Eight EVO:RAIL appliances are supported within a cluster, which accounts for a total of thirty-two ESXi nodes.

Loudmouth also requires IPv6 multicast as it leverages the innate link-local and auto-configuration capabilities of IPv6. IPv6 multicast with link-local IPv6 addresses is used by EVO:RAIL for auto-discovery of all nodes.

An IPv4 address is 32 bits where-as an IPv6 address is 128 bits long. The IP address has two parts: a subnet prefix representing the network to which the interface is connected and a local identifier.

The IPv6 link local addresses are made from the first 64-bit reservation (FE80::/64) and remaining bits are taken from the MAC address of the interface. Since MAC addresses are only 48 bits, **FF:FE** is added inbetween the MAC address of the related interface to form the complete 128 bit IPv6 link local address. Also the 7th bit (from left) in the MAC address is flipped.

Most IPv4 hosts use link-local addressing only as a last resort when a DHCP server is unavailable. An IPv4 host otherwise uses its DHCP-assigned address for all communications, global or link-local. However, every IPv6 host is required to configure a link-local address even when global addresses are available. IPv6



hosts may additionally self-configure additional addresses on receipt of router advertisement messages, thus eliminating the need for a DHCP server.

The Dell-VMware EVO:RAIL appliance uses IPv6 link local addresses to enable auto-configuration without any communication with the switch. As each one of the four ESXi nodes and the vCenter Server Agent (vCSA) pictured in Figure 5 above comes up, it is assigned an IPv6 link local address starting with FE80::/64.

4 Dell-VMware EVO:RAIL Network Architecture and Configuration

4.1 Dell Infrastructure Physical Architecture

All nodes in the EVO:RAIL solution must be on the same L2 subnet; if scaling up to the max of eight appliances, all appliances must be on the same subnet. Placing appliances of the same EVO:RAIL cluster/solution on different subnets is not supported. In most all cases, all EVO:RAIL appliances that are part of the same cluster/solution will connect to the same pair of Dell Networking ToR switches as shown below in Figure 6. Dell S4810s are shown in this example, but other supported Dell Networking switches like S4820T, S4048-ON, S6000/S6000-ON, or N4000 Series switches could also be used.

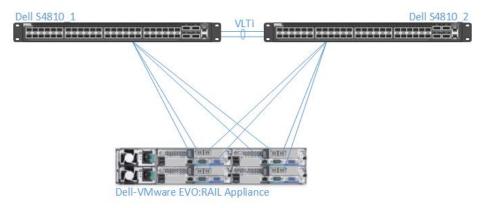


Figure 6 Dell-VMware EVO:RAIL Connected to 2 x S4810s for HA

In this example, two Dell S4810 switches are connected together with a VLTi port-channel. VLT is Dell Networking's layer 2 multipathing technology which allows for running a loop-free L2 multipath network without any blocked ports. The VLTi syncs layer 2 table information between the switches and makes the switch appear as one logical unit. Typically, a server is connected via LACP LAG to the pair of VLT switches, however, in this case no LACP LAG is used on the server side as the default VMware NIC teaming options are used. Other non-EVO:RAIL devices can utilize LACP LAG to the pair of VLT switches, for an active/active L2



multipathing scenario. For, EVO:RAIL, the VLTi provides connectivity between the switches and also allows for better failover.

As 8 x EVO:RAIL appliances are supported per EVO:RAIL cluster, such a basic topology can easily scale up to the max of thirty-two server nodes while utilizing only $64 \times 10 \text{GbE}$ ToR ports ($32 \times 10 \text{GbE}$ ToR ports per switch) for connecting to the server nodes. The $4 \times 40 \text{GbE}$ ports on the S4810 or $6 \times 40 \text{GbE}$ ports on the S4048-ON can easily be used for upstream connectivity for either a 3:1 or 2:1 oversubscription ratio.

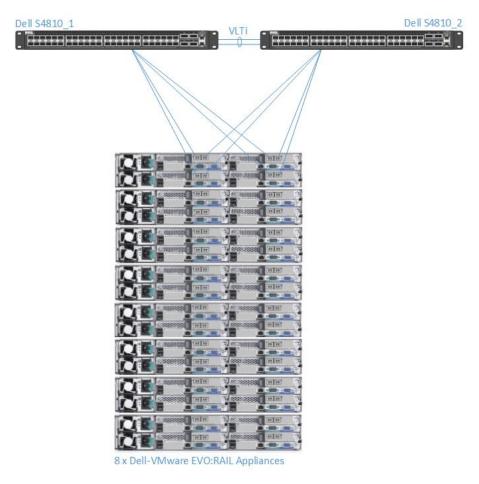


Figure 7 8 x Dell-VMware EVO:RAIL Appliances in One Cluster

Multiple EVO:RAIL clusters can also be deployed utilizing the same Dell Networking ToR switches. Since a max of 8 x EVO:RAIL appliances is supported in one cluster, such a deployment will utilize 32 x 10GbE ToR ports per switch connected to servers.

The amount of appliances to include in a cluster will depend on customer requirements. If four appliances are installed per cluster instead of the max of eight appliances per cluster, EVO:RAIL will utilize 16×10 GbE ports on each ToR switch in a pair for server connectivity. In such a setup, a max of 3×16 node EVO:RAIL clusters can be supported with 2×54810 Dell Networking switches which supplies 48×10 GbE ports and 4×40 GbE



ports. The 40 GbE ports can also be broken out into 4 x 10GbE ports and utilized with a 40GbE to 4 x 10GbE breakout cable.

Additionally, scale-out can be achieved horizontally by using another pair of switches like the Dell S6000s as an aggregate within a collapsed core spine-leaf type architecture as shown in Figure 7 below. An additional pair of S4810s or S4840-ONs connecting to a pair of S6000s at the spine/core provide additional connectivity for more EVO:RAIL clusters.

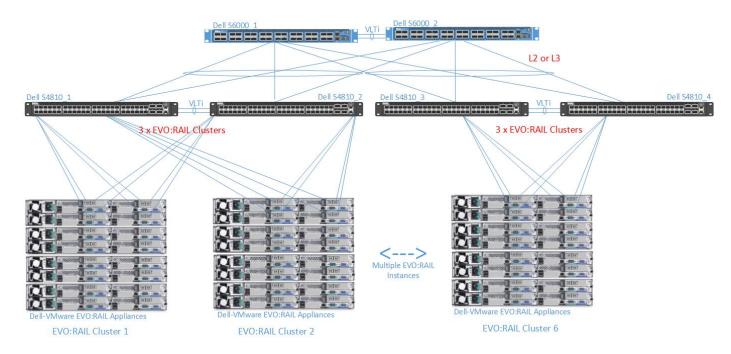


Figure 8 Multiple Dell-VMware EVO:RAIL Clusters – Horizontal Scale-Out (4 x appliances / cluster)

If needed, the Dell S4048-ON, S6000/S6000-ON, or Z9500 switches can be used to provide even higher density vertically. For example, if deploying multiple clusters with the max of 8×10000 x appliances per cluster (32 server nodes per cluster), Dell S6000 would be a better fit at ToR as it can supply 104×1000 x 10GbE ports; this would allow for 3×1000 x EVO:RAIL clusters with the max of 8×1000 x appliances per cluster as shown below in Figure 9.



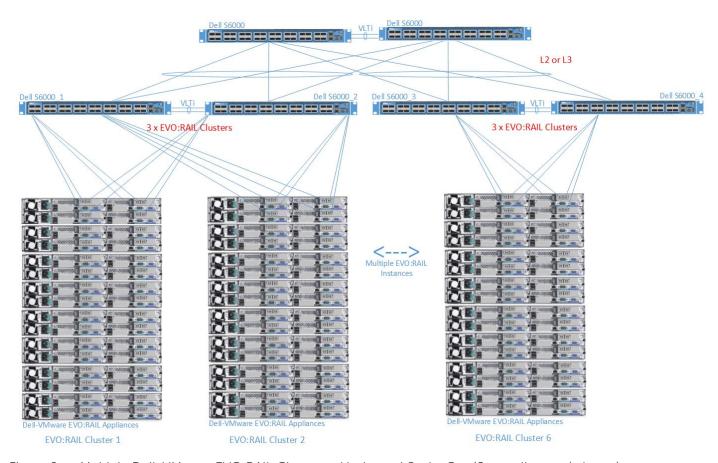


Figure 9 Multiple Dell-VMware EVO:RAIL Clusters - Horizontal Scale-Out (8 x appliances / cluster)

EVO:RAIL auto-discovery and deployment is not supported over L3 boundaries, meaning different EVO:RAIL appliances belonging to the same EVO:RAIL cluster cannot be spread across different subnets. The L2/L3 boundary where gateways reside for compute can be at the spine/core S6000 switches, creating a larger L2 network, or at the ToR S4810 switches, limiting the L2 domain. However, since an EVO:RAIL cluster must have all nodes in the same L2 domain and is limited to eight appliances, there is no need to span L2 across all clusters. L3 ECMP fabric is recommended between spine/core and ToR switches. In such a deployment all nodes of an EVO:RAIL cluster are attached to the same ToR switches.

If a complete L2 design is used, it is important to note that different VLANs should be used for the respective clusters to provide for better performance and scalability.

Figure 10 below shows a typical setup with one EVO:RAIL appliance and Dell S4810s at ToR. All Dell Networking S and Z-series 10GbE/40 Gbe switches will have the same capabilities.



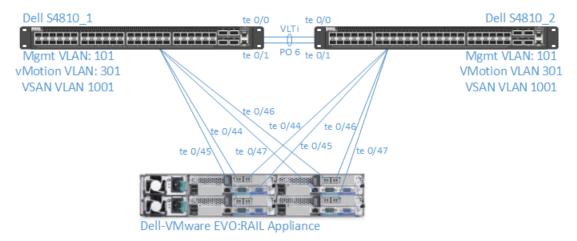


Figure 10 Dell-VMware EVO:RAIL Connected to 2 x S4810 for HA

A VLTi port-channel is being employed between the S4810s to sync layer 2 state and make the S4810s seem as one logical unit; the syncing of the layer 2 tables enabled by the VLTi link also enables better failover. VLTi is the preferred HA option, however, another option here could be to deploy stacking; up to six S4810s can be stacked to appear and be managed as one logical unit. The requirement is that traffic, including multicast, must be able to traverse between the switches. VLTi port-channel is recommended as it meets the requirements and allows for upgrading of switches without downtime as would be required by stacking.

Also, although LACP VLT LAGs are not created and connected to the EVO:RAIL node servers from the S4810 switches, other hosts outside of the EVO:RAIL solution can leverage VLT multipathing technology while at the same time the pair of S4810 switches are being used for EVO:RAIL and providing HA. In this case 2 x 10GbE ports are used on each switch for the VLTi port-channel; with the current setup, if all traffic on one switch with only one node attached has to traverse the VLTi, there is a 2:1 oversubscription. Additional ports can be added to the VLTi link if needed. In most cases, the ToR switches will aggregate to another pair of switches as shown in Figure 10 below; this also enables additional scale-out.



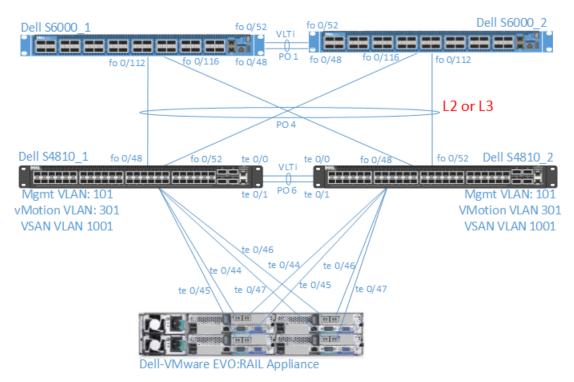


Figure 11 Dell-VMware EVO:RAIL Connected to 2 x S4810 for HA

In the setup shown in Figure 11, the L2/L3 boundary could sit on the spine/core or the ToRs. VLT can be used on the S6000s to connect to the ToR S4810s enabling a non-blocked L2 multipathing solution. VRRP or Proxy ARP can then be deployed on the S6000s to provide for gateway redundancy for compute networks, or, if L2/L3 boundary is on the S4810s, VRRP can be used. Best practice is to isolate VSAN VMkernel multicast traffic to a layer 2 non-routable VLAN.

The S6000s are simply aggregate switches one may already have deployed; this enables a scale-out model for additional EVO:RAIL clusters. The management VLAN, VLAN 101, is trunked down to the S4810s while the L2/L3 boundary for compute is set at the ToR. If desired, for large-scale environments, one can also route the management traffic. Since an EVO:RAIL cluster must have all nodes in the same L2 domain and is limited to eight appliances, there is no need to span L2 across all clusters. The L2/L3 boundary can be moved down to the S4810s and VRRP can be deployed on S4810s. As L2 is only needed to the ToR, L3 ECMP fabric between spine/core and ToR switches is recommended.

If Dell N-Series switches are being used for the ToR switches instead of S-Series switches, note that VLT is not supported, and MLAG does not meet the criteria of passing multicast traffic. Thus, the recommendation is to deploy stacking with a ring topology as shown in Figure 12 below. User port stacking is supported with both Dell S4810 and Dell N4064. In the below setup, Dell N4064, is used and 4 x 10GbE ports per switch are configured for stacking. Anywhere from one to eight 10GbE or 40GbE ports can be utilized for stacking up to 12×10 Dell N4064s.



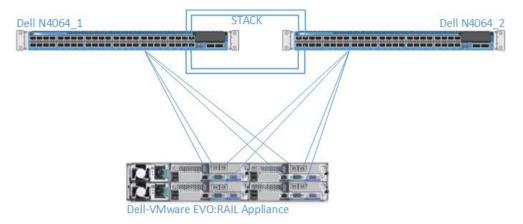


Figure 12 Dell-VMware EVO:RAIL Connected to 2 x N4064s Stacked for HA

4.2 Understanding EVO:RAIL Network Configuration

The EVO:RAIL appliance will handle four types of traffic, and these traffic types will be segregated via VLANs on the ToR switch. As shown in Figure 11, VLANs for management, vMotion, and VSAN are created on the ToR switches. Additionally, any compute VLANs as defined by application/design requirements will also need to be created. The below table outlines the VLANs created on the ToR S4810 switches for the setup shown in Figure 11.

| Traffic | VLANs | Use |
|--------------|---------------------|----------------------------|
| Management | Untagged (VLAN 101) | vSphere management traffic |
| vMotion | 301 | vMotion traffic |
| VSAN | 1001 | VSAN traffic |
| Compute (VM) | 110, 120 | Compute/workload traffic |

Table 1 Dell-VMware EVO:RAIL Traffic and VLANs

The respective vSwitch on server node 1 of the EVO:RAIL appliance as shown below in Figure 13 reflects the information in Table 1.



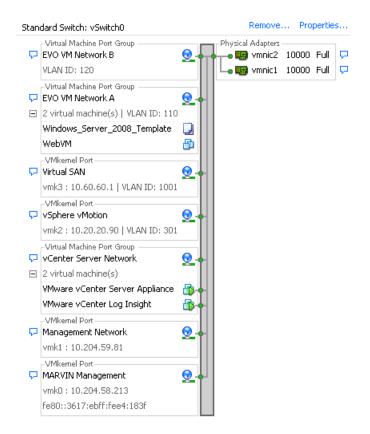


Figure 13 vSwitch on Dell-VMware EVO:RAIL node 1 server

Note, Figure 13 shows an example where EVO:RAIL has already been setup. EVO:RAIL utilizes the standard vSwitch, and the VMware vCenter Server Appliance and VMware vCenter Log Insight sit in the vCenter Server Network, which is untagged. On the ToR S4810 switches the untagged traffic is then placed into VLAN 101, which is the management VLAN of the network. Two default compute VLANs, VLAN 110 and VLAN 120, have also been created for EVO VM Network A and EVO VM Network B respectively. After setup, a custom VM and template was created in EVO VM Network A; the WebVM running Windows Server 2008 Enterprise was created, and a template of the VM, Windows_Server_2008_Template, was also instantiated.

Figure 13 also shows that 2 x 10GbE NICs are utilized by the vSwitch on server node 1. Each server node utilizes 2 x 10GbE NICs in a standard vSwitch configuration. Each server node will have the same management, vMotion, VSAN, and compute networks configured.

The default NIC teaming option for the vSwitch is **Route based on the originating virtual port ID**, and both NICs, **vmnic1** and **vmnic2** are active. This implies that each VM will be linked dynamically to one of the two uplinks. However, as is shown further below in Figure 14, the port group NIC teaming configuration overrides the vSwitch NIC teaming configuration.



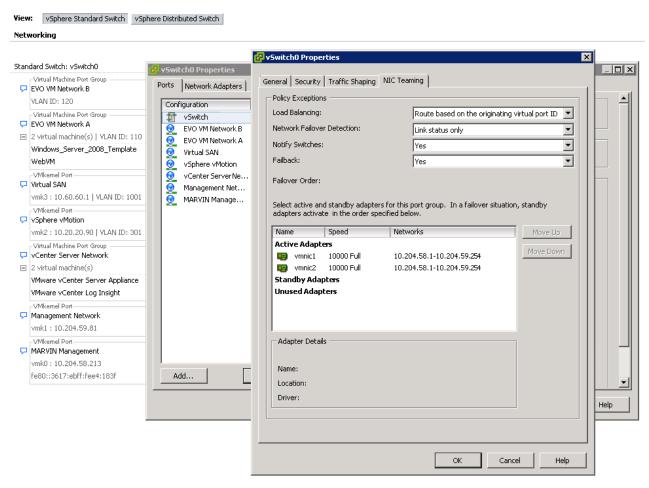


Figure 14 vSwitch on Dell-VMware EVO:RAIL Server Node 1

The port group NIC teaming for each network/traffic type overrides the default vSwitch NIC teaming configuration. The NIC teaming defaults for each network is outlined in Table 2 below.

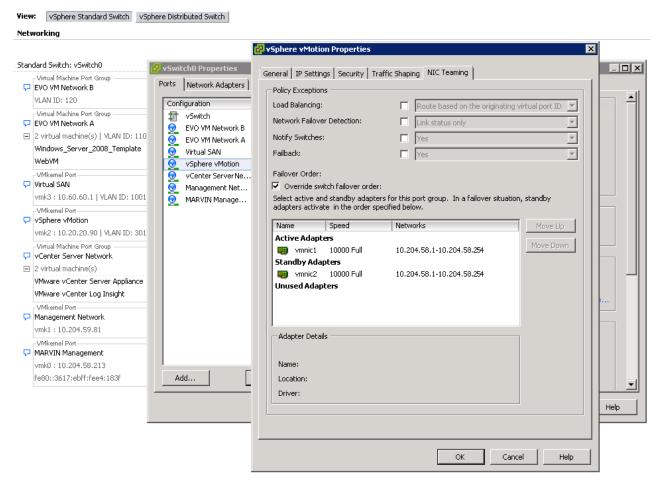
| Network | VMNIC1 | VMNIC2 |
|---|---------|---------|
| Management (vCenter, ESXi, Log InSight) | Active | Standby |
| vMotion | Active | Standby |
| VSAN | Standby | Active |
| Compute (VMs) | Active | Standby |
| EVO:RAIL Management (MARVIN) | Active | Standby |

Table 2 Dell-VMware EVO:RAIL NIC Teaming Configuration by Network/Port Group

Each network/traffic type utilizes an Active/Standby NIC teaming option. All networks, except for VSAN, utilize **VMNIC1** as the active NIC and only utilize **VMNIC2** upon failover. The VSAN network utilizes **VMNIC2** as the active NIC and only utilizes **VMNIC1** upon failover.



NIC teaming settings for the vMotion network is shown below in Figure 15. All other networks/port groups, except for VSAN, will have identical NIC teaming configuration.



vMotion Port Group NIC Teaming Configuration Figure 15

To prevent vMotion traffic from consuming all available bandwidth on the 10GbE port, EVO:RAIL limits vMotion traffic to 4Gbps. This can be seen on the Traffic Shaping tab within the vSphere vMotion port group properties as shown below in Figure 16.



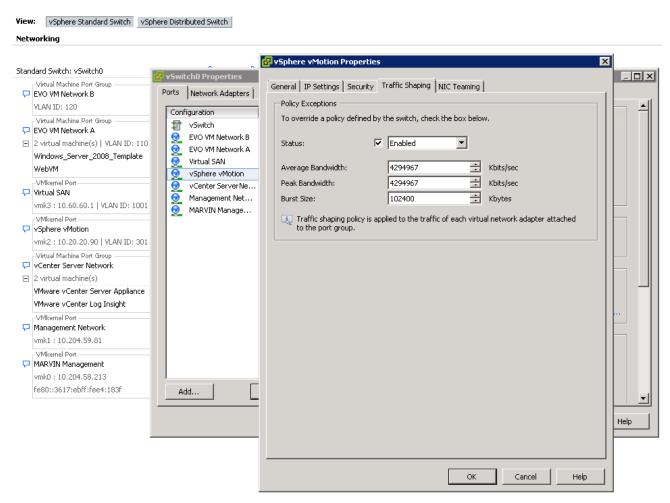


Figure 16 Traffic Shaping for vMotion Traffic

NIC teaming settings for the VSAN network is shown below in Figure 17. Notably different here from the other networks/port groups is the fact that **VMNIC2** is used as the active adapter and **VMNIC1** is the standby.



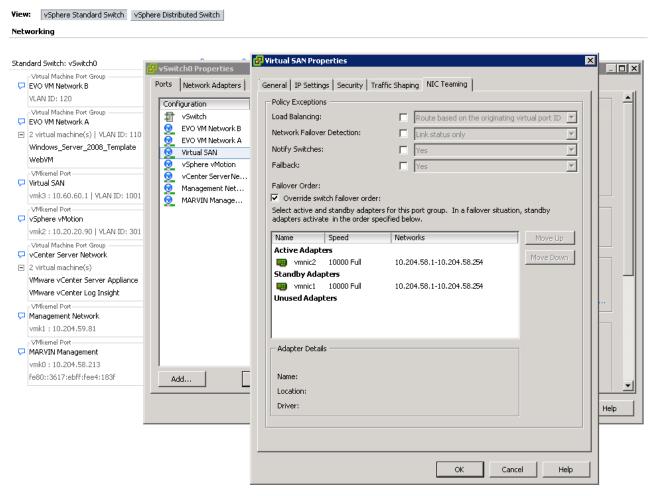


Figure 17 vSAN Port Group NIC Teaming Configuration

4.3 Dell Networking Switch Configuration

Before turning on the Dell-VMware EVO:RAIL appliance, the Dell Networking switches must be configured properly. For initial configuration, the requirement is that the server nodes be able to talk to each other over the management VLAN for auto-configuration. Also, the engineer will connect via a browser to a preconfigured default IP address for EVO:RAIL; this is done over the management network. This IP address should not be manually changed; the EVO:RAIL configuration wizard will step the user through the process of assigning a different IP address.

Since the EVO:RAIL nodes are not tagged on any VLAN, it is easiest to simply configure the management VLAN as untagged on the ToR switch ports that connect to EVO:RAIL. Other types of traffic will be tagged on their respective VLANs during EVO:RAIL configuration, and respective VLANs should already be configured on the switch as described below.

Below is the network diagram and respective configuration on the Dell Networking S4810 ToR switches, before EVO:RAIL is turned on.



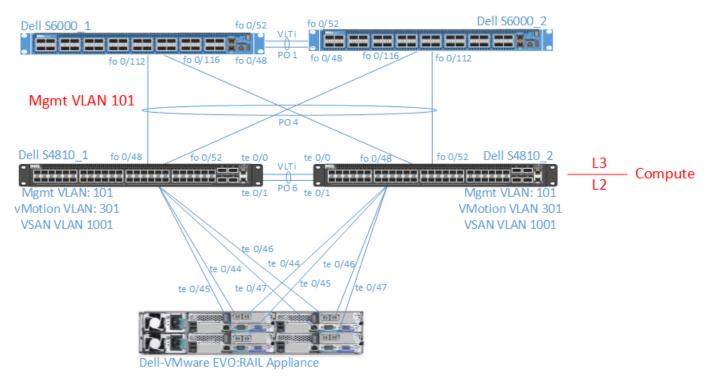


Figure 18 EVO:RAIL Network Setup with S4810s at ToR

In this setup, the Dell S4810 switches are the ToR switches and could be deployed by themselves with the EVO:RAIL appliance if desired. Here, the S6000s are simply aggregate switches one may already have deployed; this enables a scale-out model for many EVO:RAIL clusters. The management VLAN, VLAN 101, is trunked down to the S4810s while the L2/L3 boundary for compute is set at the ToR. Since an EVO:RAIL cluster must have all nodes in the same L2 domain and is limited to eight appliances, there is no need to span L2 across all clusters. Also, best practice is to isolate VSAN VMkernel multicast traffic to a layer 2 non-routable VLAN.

The S4810s provide the L3 gateway for the compute networks and also provide access to the management network via VLAN 101. One can also route to the management network if desired for large-scale environments.

VLANs for management, vMotion, VSAN, and compute need to be created on the S4810 ToRs if they don't already exist. Although not required, it is recommended to enable IGMP Snooping to reduce multicast traffic in the network used by VSAN for metadata operations and to discover new hosts. Also, since there is no Querier for multicast when the multicast source and receivers are in the same VLAN, and therefore not routed, it is recommended to configure the switch as a Querier for multicast for the VSAN VLAN; this allows the switch to send queries for new nodes.

Note, EVO:RAIL does not automatically configure jumbo frames for VSAN. Jumbo frames can provide small additional performance benefit, but can add complexity if there is a frame size mismatch within the topology. If jumbo frames are enabled on the vSwitch and respective VSAN port group, jumbo frames must also be



enabled on the respective physical switch ports. Jumbo frames can be enabled on the physical switch beforehand, and will be utilized if end nodes are later configured for jumbo frames. It is recommended to use the default configuration of EVO:RAIL.

VLTi is configured to provide connectivity between the switches and better failover. In this scenario, and most brown-field deployments, although not required, S4810 ToR switches will typically connect to another pair of aggregate switches. The configuration for the Dell S4810 switches in Figure 18 is shown below.

Note, for conciseness and to focus on the ToR configuration and connectivity to EVO:RAIL, the configuration for the L3 routing between the spine/core S6000s and ToR S4810s is not shown here. If L2/L3 boundary is at the ToR, typically a dynamic routing protocol such as OSPF would be configured on S6000s and S4810s. If L2/L3 boundary is on S6000s, VLANs would be trunked up to S6000s. In this case, the L2/L3 boundary is on the ToR S4810s with the exception of the management VLAN which is trunked up to the S6000s.



Dell S4810_1 Configuration:

```
enable
configure
interface managementethernet //management port
ip address 10.204.59.147/23
no shutdown
exit
interface range tengigabitethernet 0/0-1 //port used for VLTi port-channel
mtu 12000
no shut
exit
interface port-channel 6 //VLTi port-channel
mtu 12000
channel-member TenGigabitEthernet 0/0-1
no shutdown
exit
vlt domain 50 //VLT used as a port-channel and for non-EVO:RAIL hosts
peer-link port-channel 6
back-up destination 10.204.59.148 //IP address of management port on $4810_2
primary-priority 1
system-mac mac-address 00:0b:00:0b:00:0b //unique MAC address from VLT domain on $6000 pair
unit-id 0
exit
interface range fortyGigE 0/48 - 52 //ports connecting to spine/core S6000 switches
mtu 12000
port-channel-protocol LACP
port-channel 4 mode active
no shutdown
interface port-channel 4 //LACP port-channel connecting to spine/core $6000 switches
portmode hybrid
mtu 12000
switchport
vlt-peer-lag port-channel 4
interface range tengigabitethernet 0/44 – 47 //ports connecting to EVO:RAIL
portmode hybrid
mtu 12000
switchport
no shutdown
exit
```



```
interface vlan 101 //management VLAN
mtu 12000
untagged tengigabitethernet 0/44 - 47
exit
interface vlan 301 //vMotion VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
exit
interface vlan 1001 //VSAN VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
exit
interface vlan 110 //compute VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
ip-address 10.10.1.254/24
vrrp-group 1 //setup gateway redundancy for EVO VM Network A
virtual-address 10.10.1.254
exit
interface vlan 120 //compute VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
ip-address 10.20.1.254/24
vrrp-group 1 //setup gateway redundancy for EVO VM Network B
virtual-address 10.20.1.254
exit
ip igmp snooping enable //enable IGMP snooping to limit multicast traffic
int vlan 1001
ip address 10.60.60.200/24 //lower IP than other switch so will be elected IGMP snooping querier
ip igmp snooping querier //needed for IGMP snooping querier
write
```



Dell S4810_2 Configuration:

```
enable
configure
interface managementethernet //management port
ip address 10.204.59.148/23
no shutdown
interface range tengigabitethernet 0/0-1 //port used for VLTi port-channel
mtu 12000
no shutdown
exit
interface port-channel 6 //VLTi port-channel
mtu 12000
no shutdown
channel-member TenGigabitEthernet 0/0-1
exit
vlt domain 50 //VLT used as a port-channel and for non-EVO:RAIL hosts
peer-link port-channel 6
back-up destination 10.204.59.148 //IP address of management port on S4810_2
primary-priority 2
system-mac mac-address 00:0b:00:0b:00:0b //unique MAC address from VLT domain on $6000 pair
unit-id 1
exit
interface range fortyGigE 0/48 - 52 //ports connecting to spine/core $6000 switches
mtu 12000
port-channel-protocol LACP
port-channel 4 mode active
no shutdown
interface port-channel 4 //LACP port-channel connecting to spine/core $6000 switches
portmode hybrid
mtu 12000
switchport
vlt-peer-lag port-channel 4
interface range tengigabitethernet 0/44 - 47 //ports connecting to EVO:RAIL
portmode hybrid
mtu 12000
switchport
no shutdown
exit
```



```
interface vlan 101 //management VLAN
mtu 12000
untagged tengigabitethernet 0/44 - 47
exit
interface vlan 301 //vMotion VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
exit
interface vlan 1001 //VSAN VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
exit
interface vlan 110 //compute VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
ip-address 10.10.1.253/24
vrrp-group 1 //setup gateway redundancy for EVO VM Network A
virtual-address 10.10.1.254
exit
interface vlan 120 //compute VLAN
mtu 12000
tagged tengigabitethernet 0/44 - 47
ip-address 10.20.1.253/24
vrrp-group 1 //setup gateway redundancy for EVO VM Network B
virtual-address 10.20.1.254
exit
ip igmp snooping enable //enable IGMP snooping to limit multicast traffic
int vlan 1001
ip address 10.60.60.201/24 //higher IP than other switch so will not be elected IGMP snooping querier
ip igmp snooping querier //needed for IGMP snooping querier
write
```



It is required to use the following steps for initial configuration and not attempt to change the default initial VLAN or IP address settings. The VLAN and IP address information should only be changed through the EVO:RAIL configuration wizard.

The default IP address configured on the EVO:RAIL appliance is **192.168.10.200/24** and the gateway is set to the default of **192.168.10.254/24**. This is actually the default IP address of the vSphere vCenter Server appliance. It's ok if this is not the same subnet as your management network. You must physically connect a computer/laptop to the ToR switch and change the IP address on the laptop so the laptop is in the same subnet. Then you can navigate the browser to the **192.168.10.200/24** IP address, which is the IP address of the vCenter Server Appliance sitting on server node 1, to start initial EVO:RAIL configuration.

Since the EVO:RAIL management network in untagged, you will need to make sure the port you connect the laptop/PC to is untagged on the management VLAN being used on the network. In this case, interface tengigabitethernet 0/10 on switch **S4810_1** is used to connect to the laptop. In this setup, the management VLAN is VLAN 101, so port tengigabitethernet 0/10 on **Dell S4810_1** is untagged in VLAN 101 using the below commands, and then the laptop/PC NIC is connected to the port.

<u>Dell S4810_1 Configuration of Port Used to Connect to Laptop/PC for Initial Configuration:</u>

enable
configure
interface range tengigabitethernet 0/10
switchport
no shutdown
interface vlan 101 //management VLAN
untagged tengigabitethernet 0/10
write

If a complete L2 topology is used, any aggregate switches that S4810s are connecting to should also be configured with the appropriate VLANs, and all VLANs will need to be trunked to the aggregate switches which should also be providing any L3 gateways for compute.

Dell Networking N-Series Configuration

The above section discussed in detail the network design and configuration using Dell Networking Data Center S-Series switches. For Dell Networking N-Series switches, it is recommended to use stacking for HA as shown below in Figure 19.



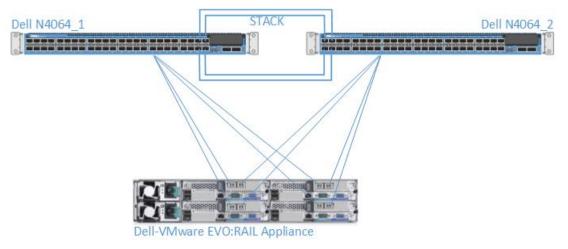


Figure 19 Dell-VMware EVO:RAIL Connected to 2 x N4064s Stacked for HA

The network requirements and VLAN setup will remain the same as that discussed on the S-Series switches, although the configuration commands/syntax will be different. For reference, VLAN and multicast configuration CLI commands for Dell Networking N-Series switches are listed below. See the Dell Networking N-Series configuration/CLI guide for more information on respective CLI commands.

Setting up VLANs (interface configuration mode):

switchport mode general switchport general allowed vlan add <VLANID1>,<VLANID2> tagged

IPV4 multicast (config mode):

ip igmp snooping (configured globally by default) ip igmp snooping querier vlan < vsan vlan id>

IPV6 multicast (config mode):

Note, IPv6 multicast traffic is not significant; therefore, the below commands for IPv6 MLD Snooping and Querier is optional for limiting multicast traffic on the network, and is not required. Dell Networking Sseries and Zseries switches do not support IPv6 MLD Snooping and Querier.

ipv6 mld snooping

ipv6 mld snooping querier vlan <management vlan id>



4.4 EVO:RAIL Configuration

Now that the backend networking is setup and your laptop is connected to the ToR S4810 switch for initial configuration, you can navigate the browser to the default IP address and URL of https://192.168.10.200:7443. You should see the below screen and can now step through the setup wizard. Click the Yes, Let's Go! button to get started.

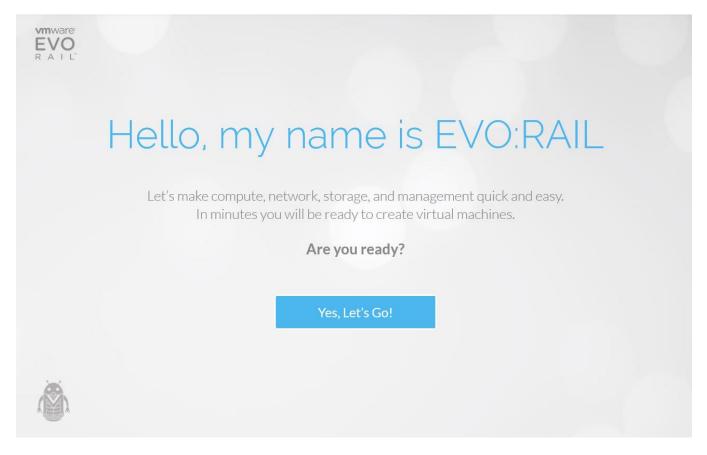


Figure 20 EVO:RAIL Setup Wizard



Click the Yes, I do button to proceed.



Figure 21 EVO:RAIL License Agreement



Figure 22 shows an option available to provide customization which is highly recommended to allow for customization based on requirements and for optimized deployments. Click the **Customize Me!** button.

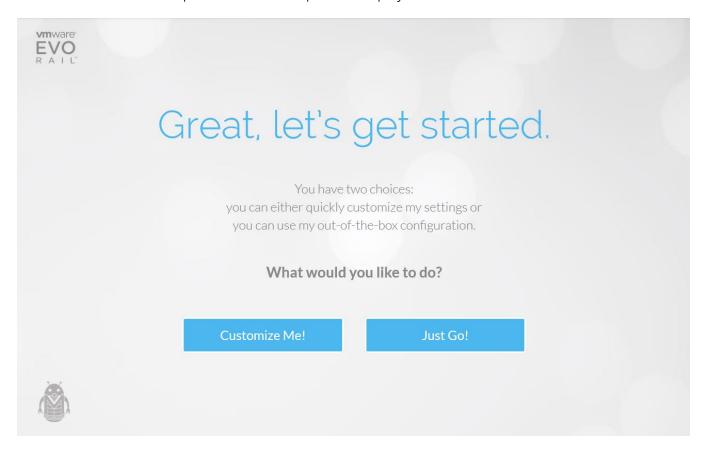


Figure 22 EVO:RAIL Setup Wizard Allowing for Customization



Enter the ESXi host naming scheme you wish to employ and the top-level domain (TLD) you wish to use; the TLD will be what has been setup via your local DNS server. Also, enter the hostname you want to assign to vCenter Server. When done, click the **Networking** tab on the left.

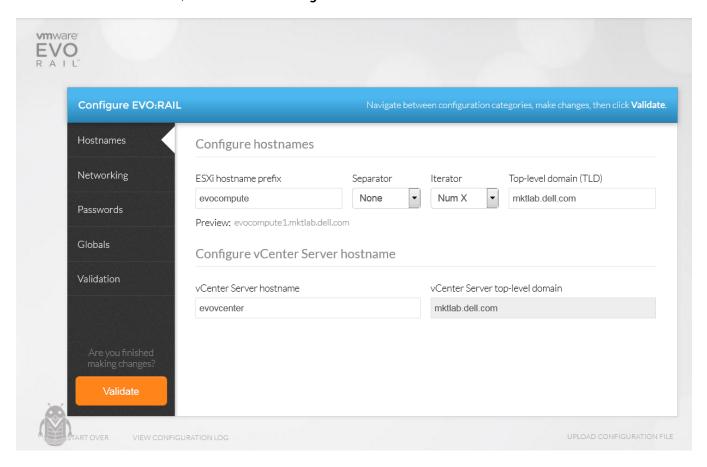


Figure 23 EVO:RAIL: ESXi Host Name Configuration



Enter the IP address information you wish to use for the ESXi hosts. The IP address information should be from the management subnet. Also, make sure to provision extra IP addresses depending on how many future nodes may be added. Below, eight IP addresses are provisioned as another EVO:RAIL appliance is planned to be added to the cluster. Once complete, click the **vMotion** tab on the left inner-menu.

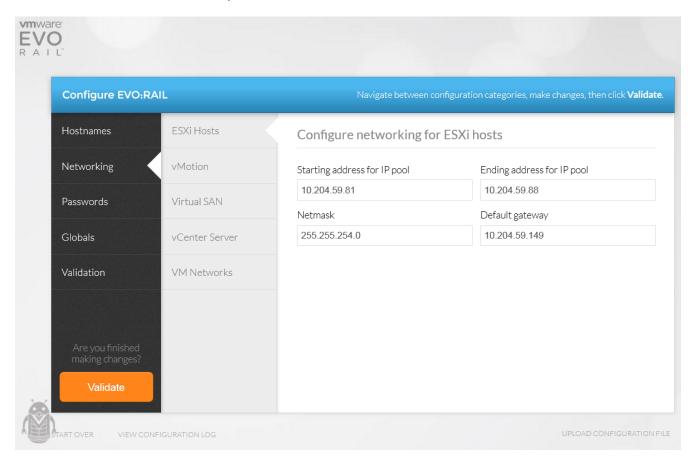


Figure 24 EVO:RAIL: ESXi Host IP Address Configuration



Enter the VLAN and IP address information you wish to use for the vMotion network. The IP address information should be from the vMotion subnet. Also, make sure to provision extra IP addresses depending on how many future nodes may be added. Below, eight IP addresses are provisioned as another EVO:RAIL appliance is planned to be added to the cluster. Once complete, click the **Virtual SAN** tab on the left innermenu.

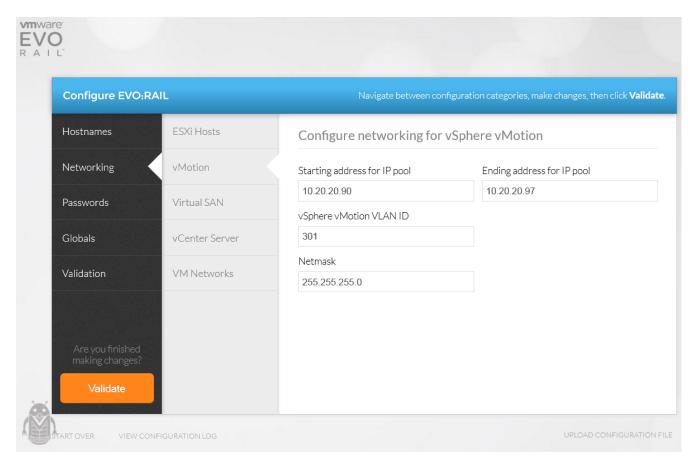


Figure 25 EVO:RAIL: vMotion VLAN and IP Address Configuration



Enter the VLAN and IP address information you wish to use for the VSAN network. The IP address information should be from the VSAN subnet. Also, make sure to provision extra IP addresses depending on how many future nodes may be added. Below, eight IP addresses are provisioned as another EVO:RAIL appliance is planned to be added to the cluster. Once complete, click the **vCenter Server** tab on the left inner-menu.

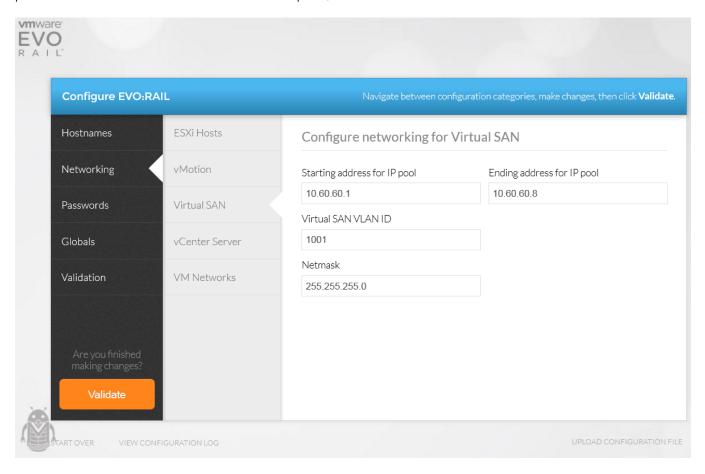


Figure 26 EVO:RAIL: VSAN VLAN and IP Address Configuration



Enter the IP address you would like to assign to the vCenter Server. The IP address should be from the management network. The **Netmask** and **Default Gateway** values are automatically set based on the IP address information provided for the EVO:RAIL ESXi hosts network configuration.

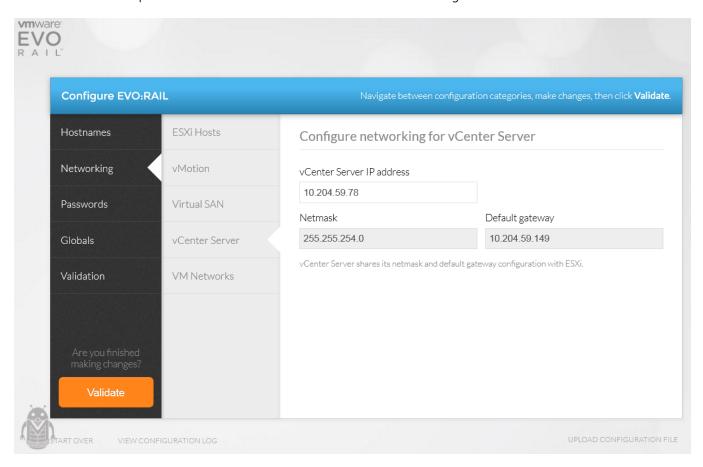


Figure 27 EVO:RAIL: vCenter IP Address Configuration



Enter the name and VLAN information you wish to use for the two default compute/VM networks. If you wish to add additional compute/VM networks, click the **Add a VM network** button and enter the respective information. Once complete, click the **Passwords** tab on the left menu.

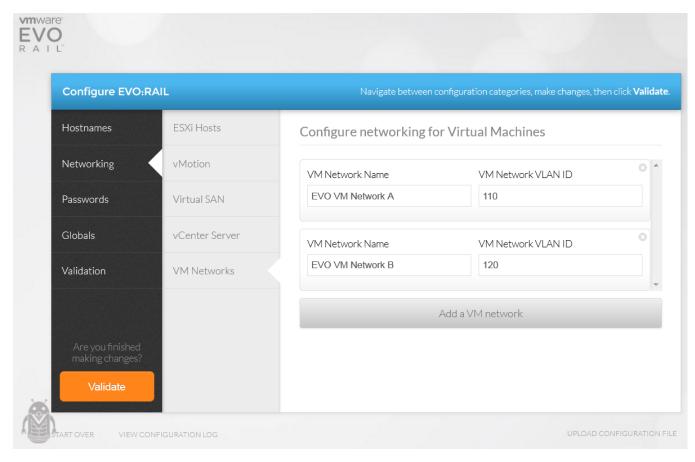


Figure 28 EVO:RAIL: Compute (VM) VLAN Configuration



Enter the ESXi host and vCenter Server passwords you wish to set. Once complete, click the **Globals** tab on the left menu.

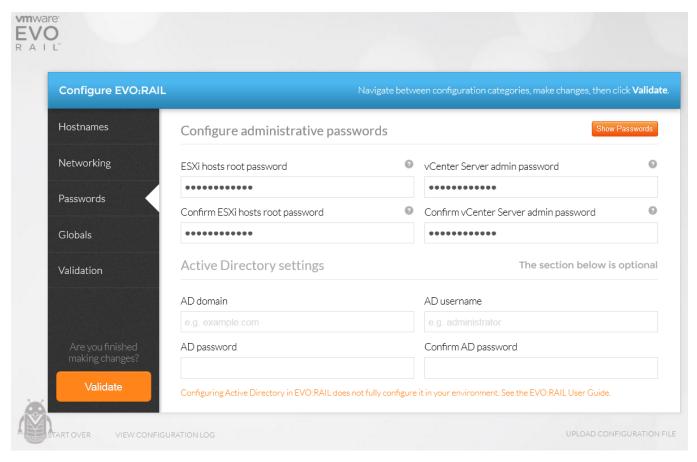


Figure 29 EVO:RAIL: ESXi host and vCenter Server Password Setup



Enter the IP addresses for existing NTP and DNS servers. Choose **vCenter Log Insight** for the logging method and enter the respective hostname and IP address you would like to give it. Note, ensure the hostnames for the ESXi hosts, vSphere vCenter Virtual Appliance, and Log Insight, resolve to the correct IP addresses you've assigned. For example, in this setup, **evologinsight.mktlab.dell.com** should resolve to **10.204.59.79**. Ensure your DNS is setup and properly resolving all hostnames by attempting to ping the hosts by hostname. Make sure the DNS is working correctly before building the appliance in the last step of this initial process. Once complete, click the **Validation** tab on the left menu.

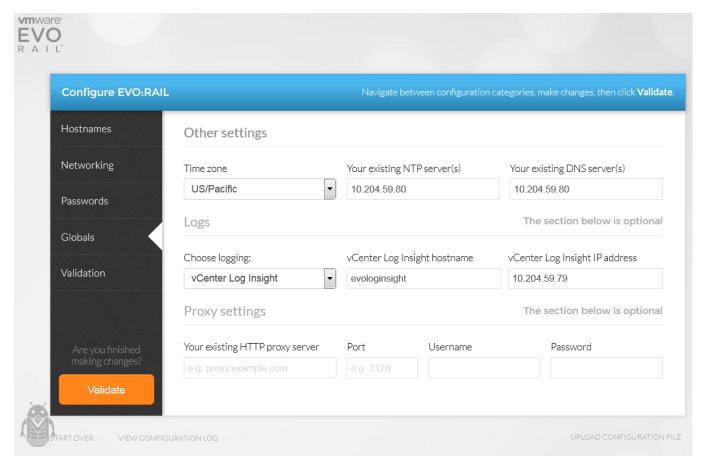


Figure 30 EVO:RAIL: NTP Server, DNS Server, and Log Insight Configuration



EVO:RAIL will first run a validation check to make sure all required information is supplied and there are no parameter issues or errors. If validation is successful, you should see the below, otherwise, you will be provided with some details on missing, incorrect, or error information, so you can take respective action. Once complete, click the **Build Appliance** button.

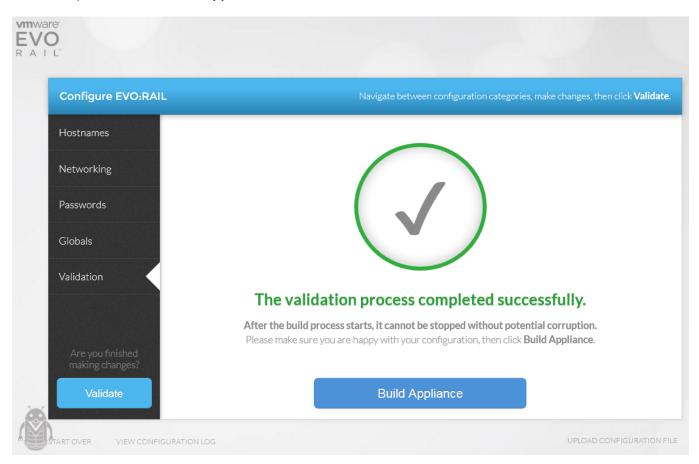


Figure 31 EVO:RAIL Configuration and Validation Successful



EVO:RAIL will start building the appliance and you will see status progress by percentage as shown below.

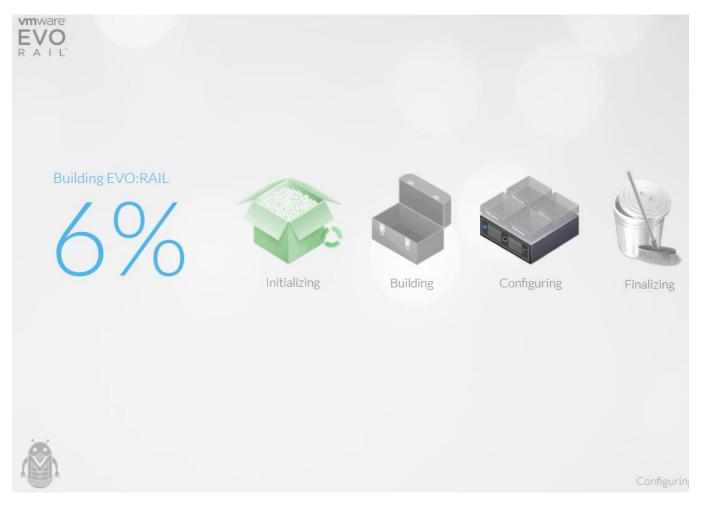


Figure 32 EVO:RAIL Build Appliance Status Progress by Percentage



Once the EVO:RAIL appliance configuration has completed, you should see something similar to the below; the IP address will likely be different depending on the IP address that was used during the EVO:RAIL configuration via the wizard. In this setup, the https://10.204.59.78:7743 URL is what is used via web browser to reach the EVO:RAIL management web client.



Figure 33 EVO:RAIL Appliance Configuration Completed



Navigating via web browser to https//10.204.59.78:7743 or the custom IP address you assigned, displays the below login page. Enter the vSphere vCenter credentials you configured for EVO:RAIL management during EVO:RAIL setup and click the AUTHENTICATE button.



Figure 34 EVO:RAIL Login



Upon logging-in, you should see something similar to the below. From this web client, you can console into, clone, create, start, and stop VMs. In addition, you can monitor the health of EVO:RAIL and also monitor some health/performance metrics such as storage IOPS, storage utilization, CPU consumption, and memory consumption. Click the **HEALTH** menu item on the left to see the health/performance statistics.

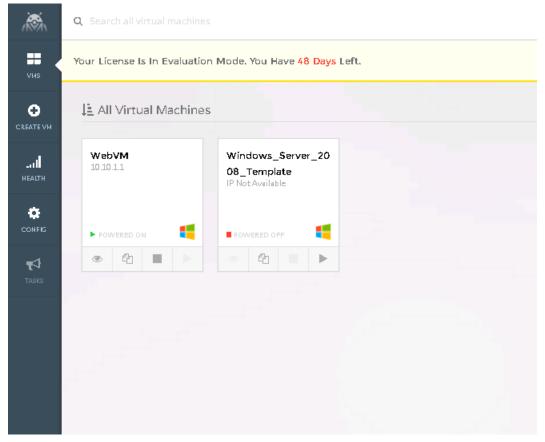


Figure 35 EVO:RAIL Management Web Client



You should see something similar to the below health/performance statistics. Log information can also be obtained from the **CONFIG** menu item.

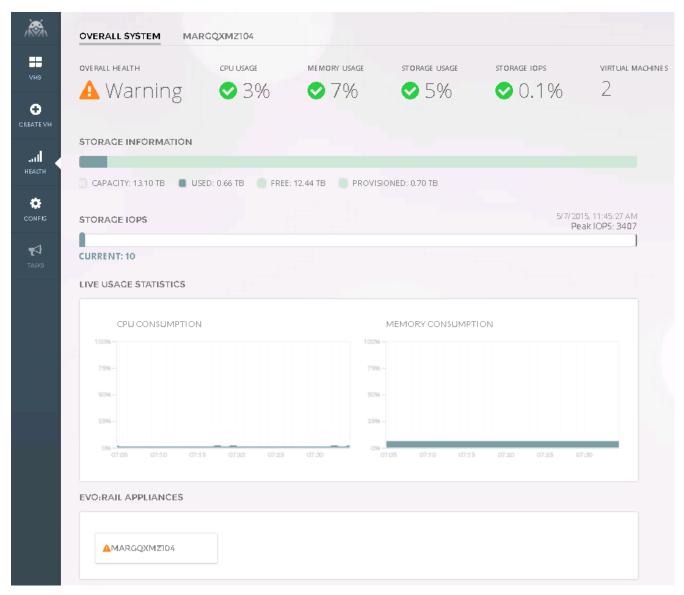


Figure 36 EVO:RAIL Management Web Client: Health/Performance Statistics



Navigating via web browser to the https://10.204.59.78:9443/vsphere-client/ URL provides access to vSphere vCenter Server Appliance and the ESXi hosts via vSphere Web Client.

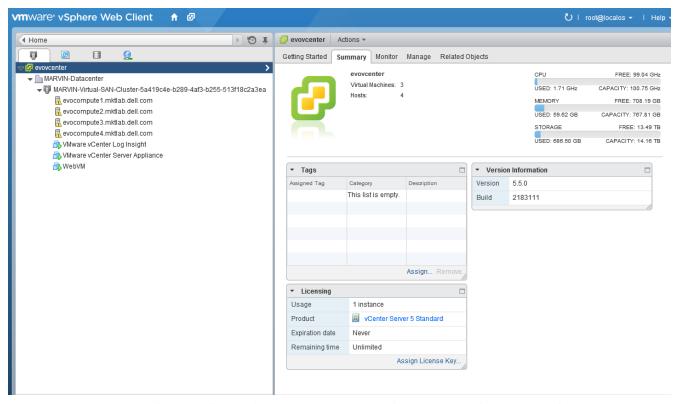


Figure 37 Accessing vSphere vCenter Server Appliance and ESXi Hosts via vSphere Web Client

5 Dell-VMware EVO:RAIL Setup Validation

- 1. First, log into the EVO:RAIL management web client and check the health status: https://<configured-vCenterIPAddress>:7743
- 2. Second, log into vSphere Web Client at https:// <configured-vCenterlPAddress>:/vsphere-client/, navigate to "Home->Hosts and Clusters" and ensure all ESXi nodes, vSphere vCenter Server Appliance, and VMware vRealize Log Insight is accessible. You should not see any red critical alerts.
- 3. Navigate to "Home->Hosts and Clusters->Storage tab" and under the MARVIN-Datacenter data center label, select the datastore starting with "MARVIN-Virtual-SAN-Datastore". This is the VSAN data store spanning all hosts in the EVO:RAIL cluster. Ensure the hosts # listed is equivalent to the number of server nodes in the EVO:RAIL cluster. Also ensure the storage capacity is correct; this will depend



on how many EVO:RAIL appliances are installed. One EVO:RAIL appliance with 4 x server nodes should have a capacity of 13.10 TB listed; see example screenshot below.

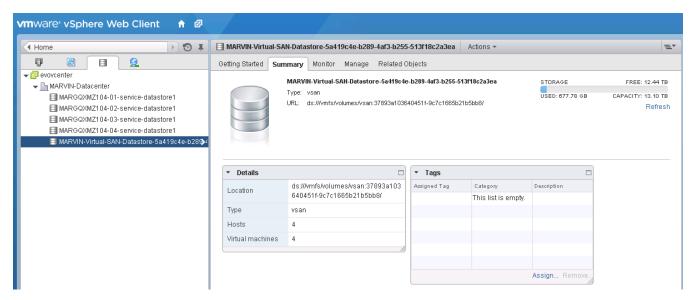


Figure 38 Validating VSAN Setup in EVO:RAIL

4. Console into one of the VMs and confirm you can ping the gateway configured on the Dell Networking switch. Below, in Figure 39, it can be seen the WebVM VM created in the EVO VM Network A network can ping the virtual IP address of the VRRP gateway configured on the Dell Networking S4810 ToR switches.



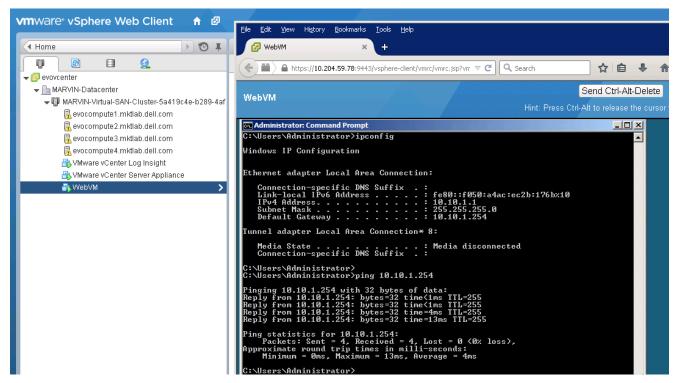


Figure 39 Validating VM Can Reach Gateway Configured on Dell Networking S4810 ToR Switches

5. Console, telnet, or SSH into the second ToR switch in the pair; in this case **Dell S4810_2**. This switch is receiving the SAN traffic since **VMNIC2** (connecting to **Dell S4810_2**) is active and **VMNIC1** (connecting to **Dell S4810_1**) is standby for the VSAN port group as discussed earlier. Verify IGMP snooping is working with the **show ip igmp snooping groups** command. Note, the switch with the lowest IP address on the respective VSAN VLAN is elected the IGMP snooping querier per configuration; this can be verified with the **show ip igmp snooping interface vlan 1001** command.

```
Dell S4810 2#show ip igmp snooping groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address
                 Interface
                                           Mode
                                                         Uptime
                                                                    Expires
                                                                              Last Reporter
224.1.2.3
                 Vlan 1001
                                           IGMPv2
                                                         1d6h
                                                                    00:02:07
                                                                              10.60.60.2
 Member Ports: Te 0/44, Te 0/47
                 Vlan 1001
                                           IGMPv2
                                                         1d6h
                                                                    00:02:07 10.60.60.2
224.2.3.4
 Member Ports: Te 0/44, Te 0/45, Te 0/46, Te 0/47
Dell S4810 2#
```

Figure 40 Validating Multicast and IGMP Snooping

 Navigate to "Home->Hosts and Clusters" and select the VMware vCenter Server Appliance. Click the Summary tab and then click the View all 2 IP addresses link. You should see something similar to the below in Figure 41. Note the IPv6 link-local address starting with fe80.



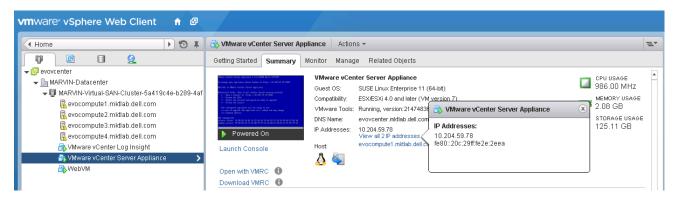


Figure 41 Verifying vSphere vCenter Server Appliance IPv6 Address

Navigate to "Home->Hosts and Clusters" and select the first ESXi server node. Click the Networking tab and then click the Virtual switches item in the inner menu. Within the MARVIN Management port group, click the i symbol connecting to the green port as shown below in Figure 42. Note the IPv6 link-local address for the ESXi host. Repeat this same process for other ESXi hosts within the cluster and note their IPv6 link-local addresses.



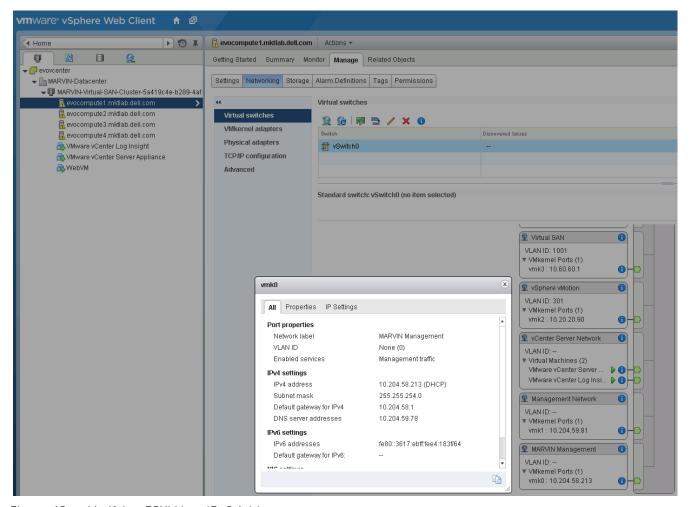


Figure 42 Verifying ESXi Host IPv6 Address

SSH into one of the ESXi hosts and run the command **ping -6 -I vmk0 ff02::1** is the IPv6 link-local scope all-nodes multicast address. All IPv6 entities including the respective EVO:RAIL ESXi nodes should reply back. You will see many replies and replies with **DUP!** posted. Confirm within the list, the IPv6 addresses of the vSphere vCenter Server Appliance and respective ESXi hosts is present. This confirms the required IPv6 and IPv6 multicast is working on the switch.

Note, if SSH is not working, ensure SSH is enabled on the ESXi hosts by logging directly into the Direct Console User Interface (DCUI) by either directly connecting a monitor and keyboard to the respective server node or by connecting via vSphere Web Client to the respective ESXi host.



6 Conclusion

In this whitepaper, we looked at the benefits of the Dell-VMware EVO:RAIL solution. As a highly automated, hyper-converged infrastructure appliance, the Dell-VMware EVO:RAIL solution provides cost savings and a complete prepackaged virtualization solution running on reliable Dell hardware. Additionally, the Dell-VMware EVO:RAIL solution provides flexibility in terms of starting small and adding additional nodes as required. The Dell-VMware EVO:RAIL solution can easily scale up to 8 x EVO:RAIL appliances providing 32 x ESXi server nodes and 104.8 TB usable VSAN storage.

