TECHNICAL BRIEF

Pluribus Adaptive Cloud Fabric
Reference Deployment Guide for VxRail Network

- Simple smart networking fabric: consistent and automatic provisioning with single point of management without error-prone box-by-box operation
- Single graphical dashboard for provisioning and monitoring with optional flow analytics module
- Automated network configuration from vCenter for tasks such as VLAN, LAG and virtual-chassis LAG provisioning
- Fully interoperable with existing third-party networks using standard networking protocols: allows inserting VxRail into a legacy network without rip-and-replace
- Elastically scales from several VxRail nodes to hundreds of nodes with linear performance within the same data center or across data centers for active-active data protection over geographically dispersed data centers and disaster recovery (DR) solutions

The Dell EMC VxRail™ appliance family provides a turn-key, easy-to-deploy and optimized Hyper-Converged Infrastructure (HCI) solution that seamless integrates into VMware ecosystems.

A modern network realized by the Pluribus Adaptive Cloud Fabric™ running on Dell EMC Open Networking (ON) switches is essential to supply the network agility needed to support VxRail with simplified management, elasticity, visibility and adaptability.

The Pluribus Adaptive Cloud Fabric is easy to deploy, and provides a simple centralized, single point-of-control to provision and manage all devices across the fabric regardless of the fabric size and the location of the switches. Integration with VMware vCenter allows one-touch provisioning for network, compute and storage to administer HCI services and the network in unison.

Introduction

This Deployment and Configuration Guide offers a reference network design example based on a pair of Dell EMC ON Top of Rack (ToR) switches that enables a VxRail cluster solution. The intended audience of this guide is a group of network or IT architects and engineers involved in the design and deployment of a VxRail solution.

About Pluribus Networks

Pluribus Networks offers a unique and highly differentiated approach to software-defined networking and is driving the revolution of the networking market to a more open operating environment. The Adaptive Cloud Fabric architecture enables organizations to build scalable private and public clouds that improve service velocity, performance, and reliability. The company's innovative Netvisor® ONE software virtualizes open networking hardware to build a holistic, distributed network that is more intelligent, automated, and resilient. The company's Insight Analytics™ platform leverages embedded telemetry and other data sources to enable pervasive visibility across the network to reveal network and application performance that speeds troubleshooting and improves operational and security intelligence.

Reference Topology and Solution Components

The diagram in Figure 1 represents in their physical topology the components required to deploy a VxRail cluster of up to 48 nodes using a pair of Dell ON switches, like S4148-ON, S4128-ON or S4048-ON, equipped with Pluribus Netvisor ONE software and the license enabling Adaptive Cloud Fabric.

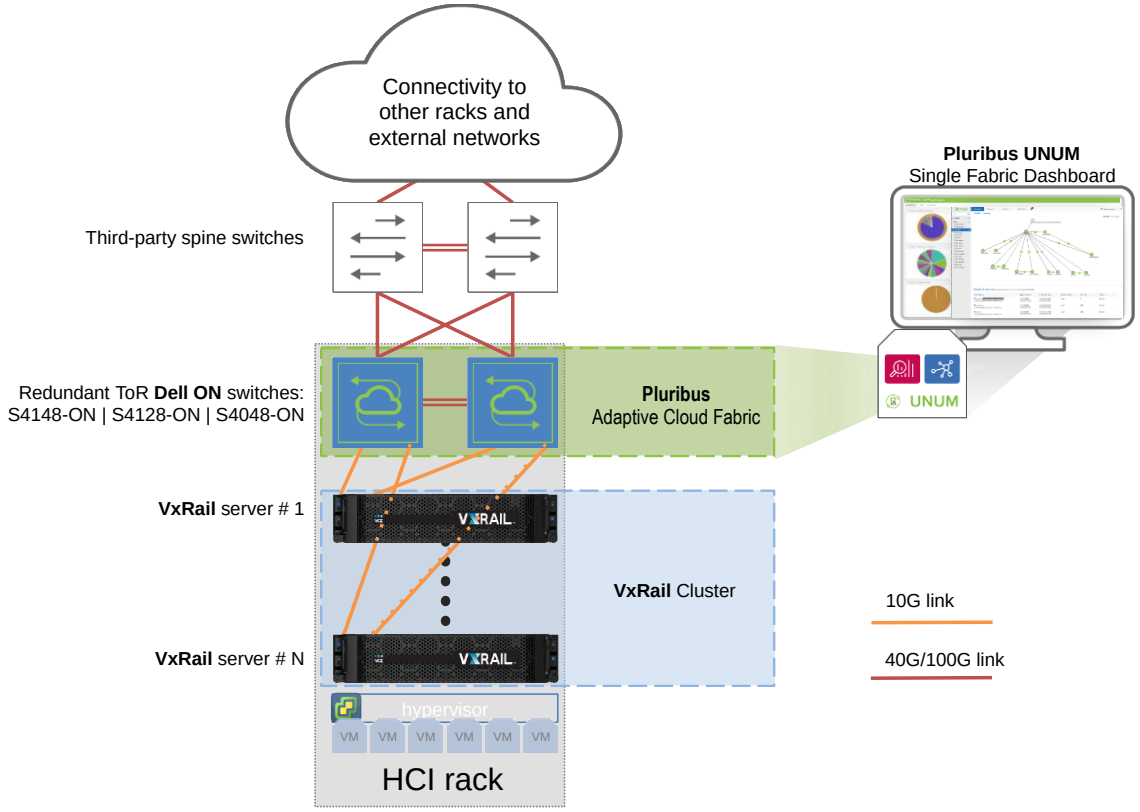


Figure 1. Reference topology used in this guide: VxRail nodes interconnect to a pair of Dell EMC ON switches running Netvisor software

Dell-Pluribus switches provide redundant 10Gbs links to each VxRail server for node-to-node interconnectivity. Netvisor software enables key functions like vSAN and vMotion, and offers a seamless integration with the existing network infrastructure, using standard layer-2 and layer-3 protocols. This example assumes a typical design where the ToR aggregation employs a pair of third-party spine switches, which interconnect to other ToRs and to external networks. Depending on the switch models and required capacity, switch-to-switch connectivity can use either 40 Gbs or 100 Gbs links.

The Adaptive Cloud Fabric infrastructure offers an integrated automation framework that simplifies the deployment of common network constructs like: virtual-chassis emulation, Link Aggregation Group (LAG), virtual-chassis LAG (vLAG), VLANs. The Fabric provides a unified view of the entire network as a single entity, allowing the provisioning of network constructs as single objects, which are consistently implemented in the switches that are part of the Fabric. This software-defined architectural approach achieves an unsurpassed agility compared to traditional network stacks that require box-by-box provisioning, which is slow and error-prone.

Other than offering a Command Line Interface (CLI) and a fully equivalent set of RESTful APIs, Netvisor switches can be easily managed using a graphical tool: Pluribus UNUM™ is a unified management platform that integrates a comprehensive range of advanced

management capabilities to enhance the intrinsic automation of the Adaptive Cloud Fabric architecture with workflow automation, topology visualization, network diagnostics and integrated performance analytics.

UNUM software is available in a virtual machine form factor and provides the following capabilities:

- Device and fabric management and provisioning
- Workflow automation with zero-touch provisioning
- Dynamic multi-vendor topology mapping
- Multi-vendor device and network diagnostics
- Real-time multi-vendor network-wide monitoring with advanced analytics[1]

[1] Advanced Analytics requires the optional Insight Analytics module

Provisioning the Fabric as a Highly Available ToR for VxRail

The initial Fabric provisioning is executed with an automated zero-touch procedure of three simple steps. Being the switches connected by two inter-switch links, as in Figure 1 above, they will also be configured in a high-availability mode called Netvisor **cluster**, which emulates a virtual-chassis.

As shown in Figure 2, the mandatory general information requested in **Step 1** is:

- The *Name* of the fabric
- The *User Name* and *Password* credentials used to connect to the switches

As in Figure 2, **Step 2** sets up the switches that are part of the fabric, which can be added using the *Add Switch* button and removed using the *trash-icon* button:

- The out-of-band management *IP address* of each switch in the topology
- The role of the switch: in our example, still referring to Figure 1, there are two Dell-Pluribus switches acting as *leaf* elements (an equivalent name for ToR) and two *third-party* switches.

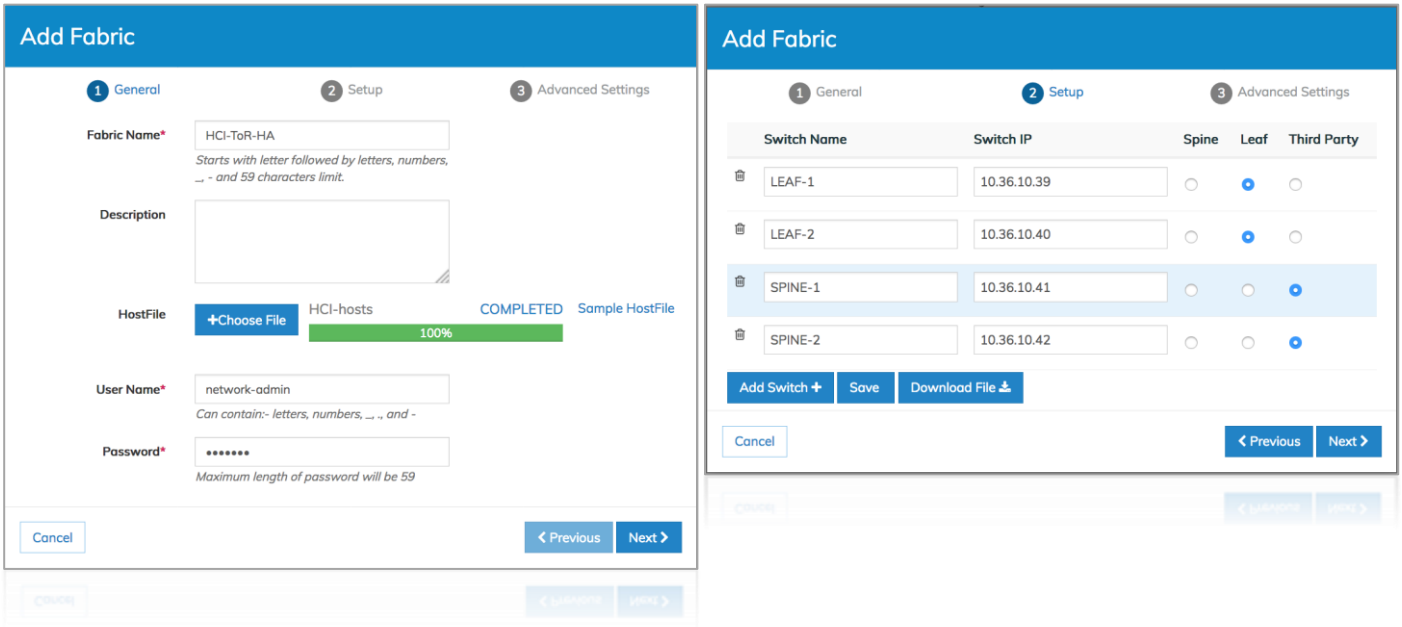


Figure 2. Fabric Zero-Touch-Provisioning: Step 1 and Step 2

Note also that the switch information entered at Step 2, can be optionally imported during Step 1 using a text based *HostFile*.

Advanced settings can be entered during Step 3, as represented in Figure 3. An explanation of the parameters and their recommended value is provided below:

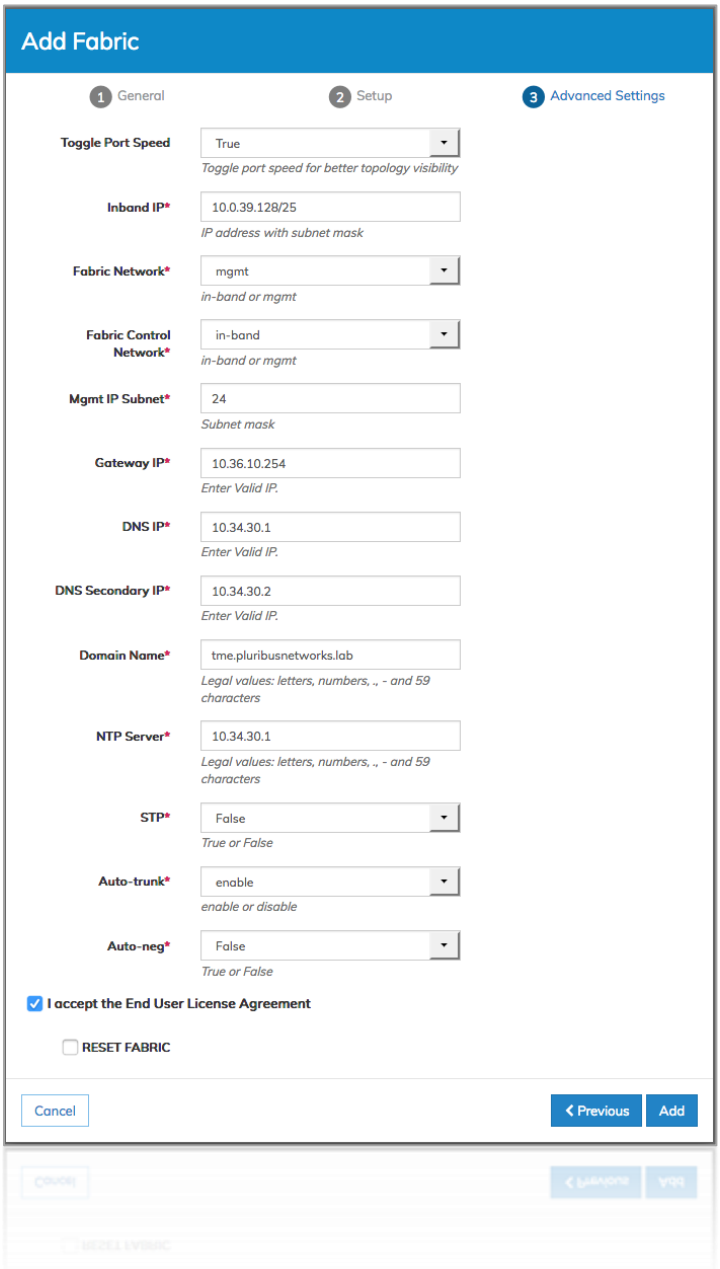


Figure 3. Fabric Zero-Touch-Provisioning: Step 3

- *Toggle Port Speed*: improves the physical topology discovery effectiveness by configuring the speed of 40G/100G ports to 10G/25G
- The *in-band* management *IP* address of each switch in the topology. Primarily used for in-band Fabric communication
- *Fabric Network*: determines if the Fabric communication uses the out-of-band (mgmt) or the in-band network. For the simple design example in this document, mgmt is recommended
- *Fabric Control Network*: selects the transport network for the cluster control plane; in-band is recommended
- *Management IP Subnet* mask value for out-of-band network
- *Gateway IP* address for out-of-band network
- Domain Network System (*DNS*) Primary *IP* address
- *DNS Secondary IP* address
- DNS *Domain Name*
- Network Time Protocol (*NTP*) *Server* hostname or IP address
- Enable or disable the Spanning Tree Protocol (*STP*) on the switches as recommended by the network administrator
- *Auto-trunk*: self-provision LAGs (trunks) when detecting two or more ports connecting to the same device
- Enable or disable port *auto-negotiation* as recommended by the network administrator
- *Accept End User License Agreement* (EULA) on each switch: mandatorily yes
- *Reset the fabric*: used to erase any previous fabric configuration on the switches

The execution of the Fabric zero-touch-provisioning process will take few minutes per switch and can be monitored on the topology view, as shown in Figure 4: switches being configured are represented with bigger icons, like switch LEAF-2 in the picture. Detailed steps of the zero-touch-provisioning can be viewed by clicking on the Playbook Logs button.

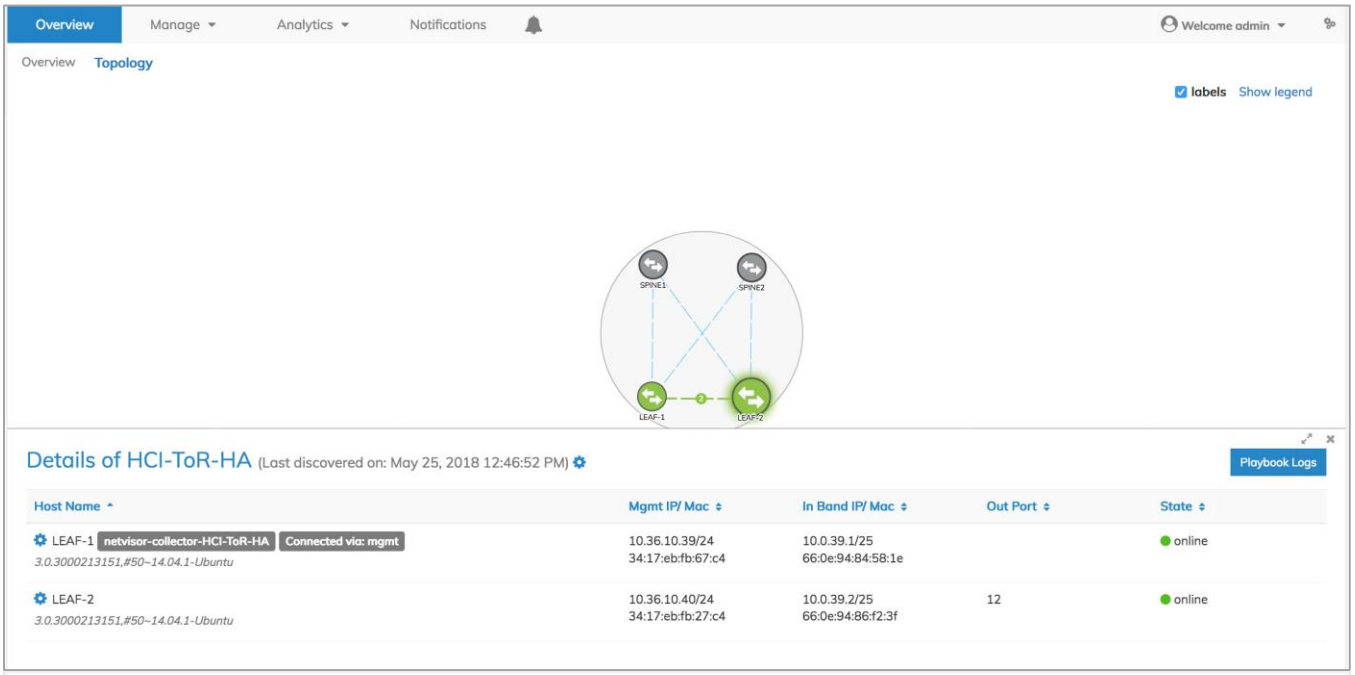


Figure 4. Fabric Zero-Touch-Provisioning: topology view of the fabric at process completion

Integration with Third-Party Networks

The typical design discussed in this guide assumes that the ToR aggregation layer consists of a pair of third-party spine switches, configured as a virtual-chassis for high-availability.

Pluribus cluster can interoperate with most virtual-chassis implementation, like Cisco Systems vPC, Arista MLAG and Dell VLT.

With this design approach, as shown in Figure 5, the ports facing the spine virtual-chassis on each Dell-Pluribus switch are configured as a LAG. The two resulting LAGs are configured as a virtual-chassis LAG (vLAG).

Pluribus UNUM provides many configuration wizards for creating layer-2 constructs like LAG, called *trunk*, and *vLAG*. These wizards can be activated by selecting a switch element in the leftmost fabric widget and by selecting the *Manage→Layer2* menu on top.

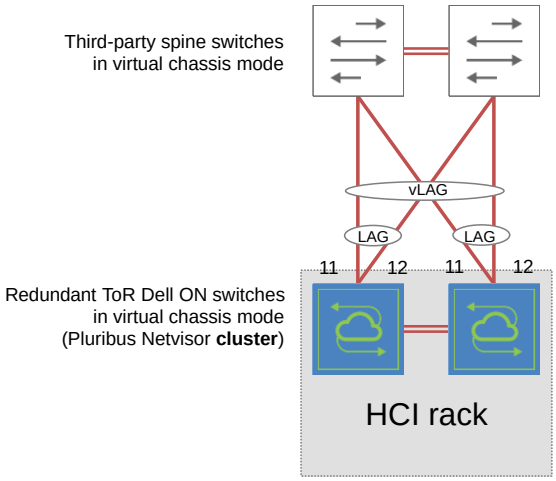


Figure 5. Integration with third-party virtual chassis

The following steps explain how to provision LAGs and vLAG to interconnect the Dell-Pluribus cluster to the third party virtual chassis. Using Figure 5 as reference, provision the LAG that interconnects LEAF-1 to the spine pair by selecting LEAF-1 on the left, activating the top menu Manage → Layer2 → Trunks and clicking on *Create a trunk configuration for link aggregation* button, as shown in Figure 6.

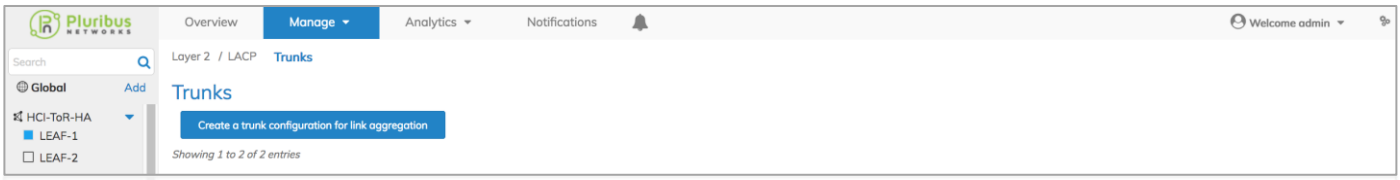


Figure 6. Creating a trunk (LAG) object on switch LEAF-1 using Manage wizard: Step 1

In the trunk configuration wizard, enter the name that identifies the LAG and, based on the physical interface information of Figure 5 above, enter the local ports that will be part of the LAG, and select active as Link Aggregation Control Protocol (LACP) mode, as shown in Figure 7. Leave the other parameters with their default value and click the *Save* button.

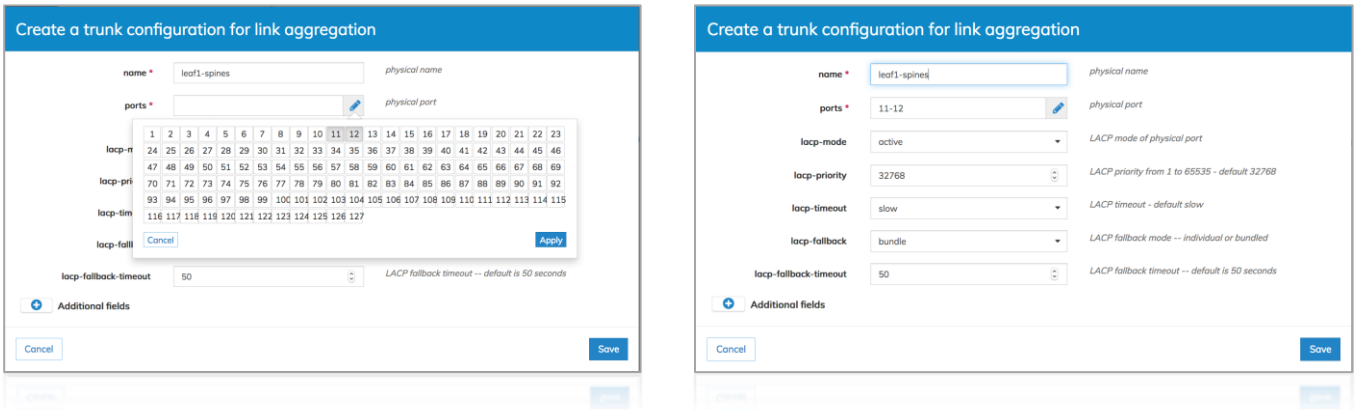


Figure 7. Creating a trunk (LAG) object on switch LEAF-1 using Manage wizard: Step 2

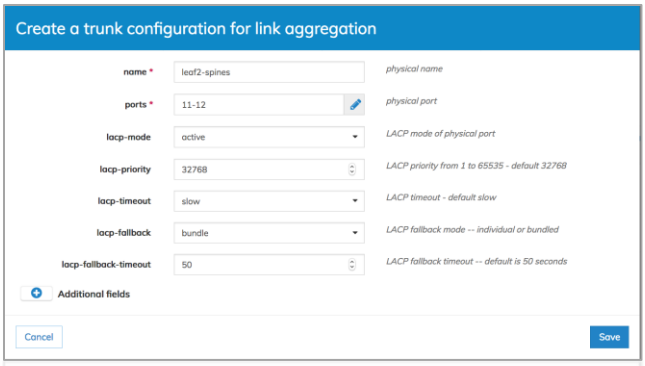


Figure 8. Creating a trunk (LAG) object on switch LEAF-2 using Manage wizard

Create the trunk object on LEAF-2, by selecting LEAF-2 in the left side widget bar and following a symmetric procedure of previous step, as shown in Figure 8.

Leave the other parameters with their default value and click the *Save* button.

The resulting LAGs can be grouped into a vLAG with the following procedure.

Select LEAF-1 or LEAF-2 on the left side bar, activate the top menu Manage → Layer2 →Manage VLAG; start the vLAG configuration wizard by clicking on *Create virtual link aggregation* button, as shown in Figure 9.

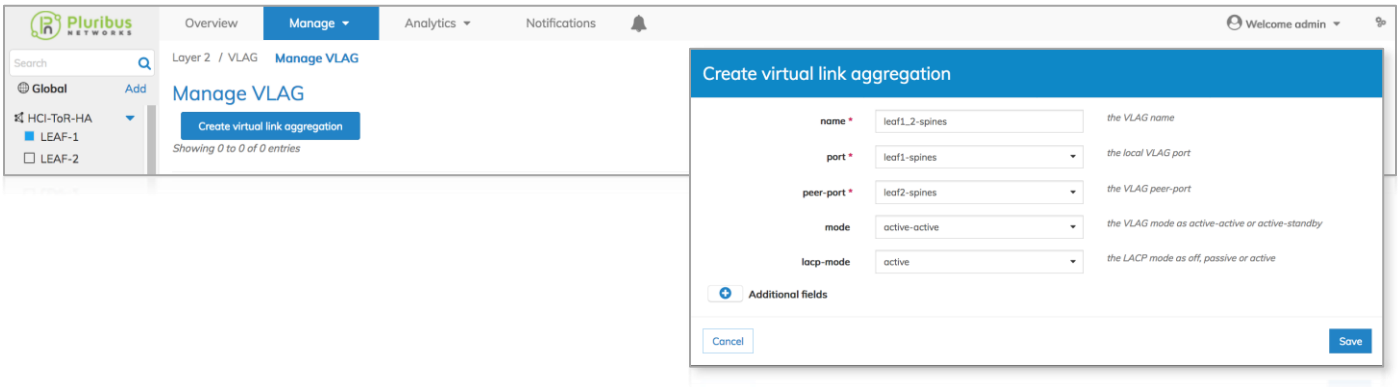


Figure 9. Creating a vLAG object on the switch cluster LEAF-1 & LEAF-2 using Manage wizard

In the wizard, enter the name that identifies the vLAG object, and then select the name of the trunk objects provisioned on LEAF-1 and LEAF-2 at previous steps. Specify high-availability *mode* active-active and *LACP mode* active. Press the Save button.

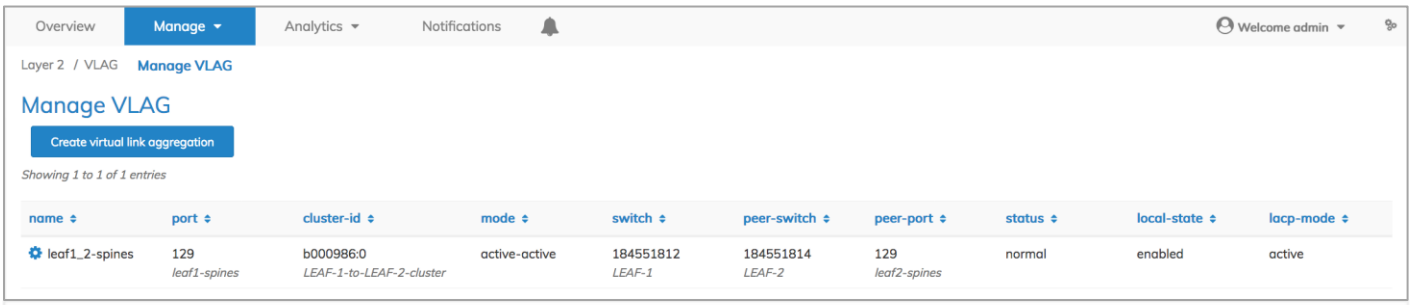


Figure 10. vLAG object on the switch cluster LEAF-1 & LEAF-2

The resulting vLAG object is visible in Manage → Layer2 →Manage VLAG, as shown in Figure 10.

Other Integration options with Third-Party Network

This chapter provides a reference to other network design options whose provisioning is also fully automated by UNUM. These options are more advanced and hence not discussed in detail.

As an alternative to the typical layer-2 design described in previous chapters, Dell-Pluribus ToR cluster can be also configured as a pair or redundant routers that can form a routing adjacency with third-party spine switches using a dynamic routing protocol like Open Shortest Path First (OSPF) or Border Gateway Protocol (BGP).

Figure 11 represents a corresponding network topology example using two spine switches, but it can expand to a higher number of spine switches.

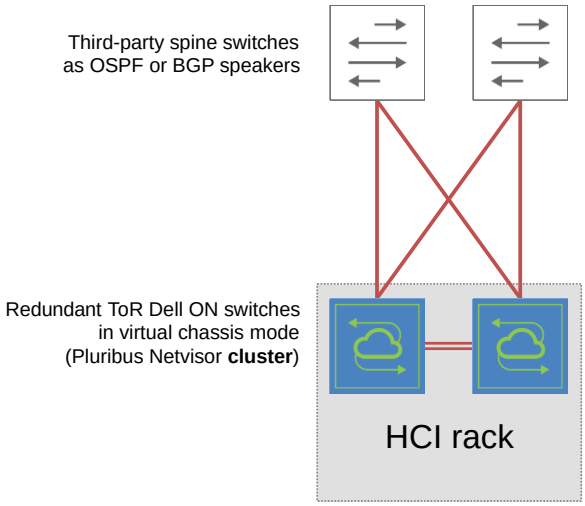


Figure 11. Integration with third-party routers

UNUM provides a set of six *Playbooks* that can be executed at the end of the Fabric zero-touch-provisioning process, in order to automate the layer-2 and layer-3 provisioning for the entire fabric. In the context of the topology in Figure 11, the relevant playbooks are shown in Figure 12 and 13.

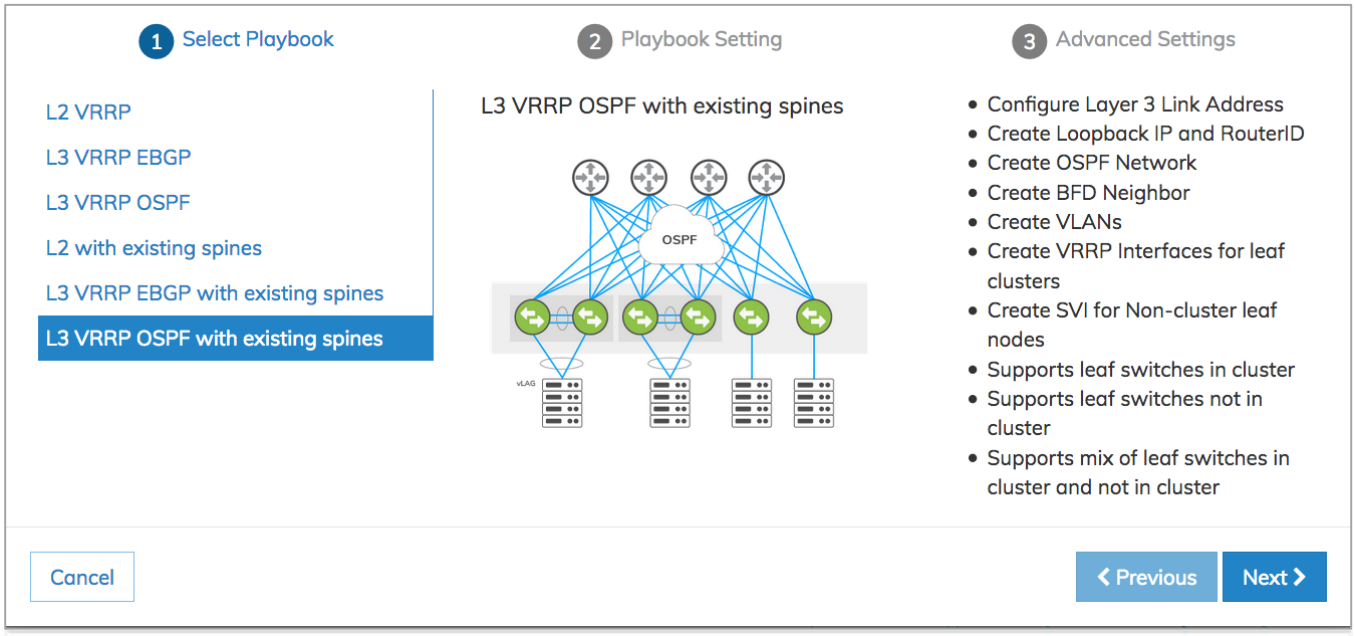


Figure 12. Automating the integration with third-party OSPF switch spines using UNUM Playbook

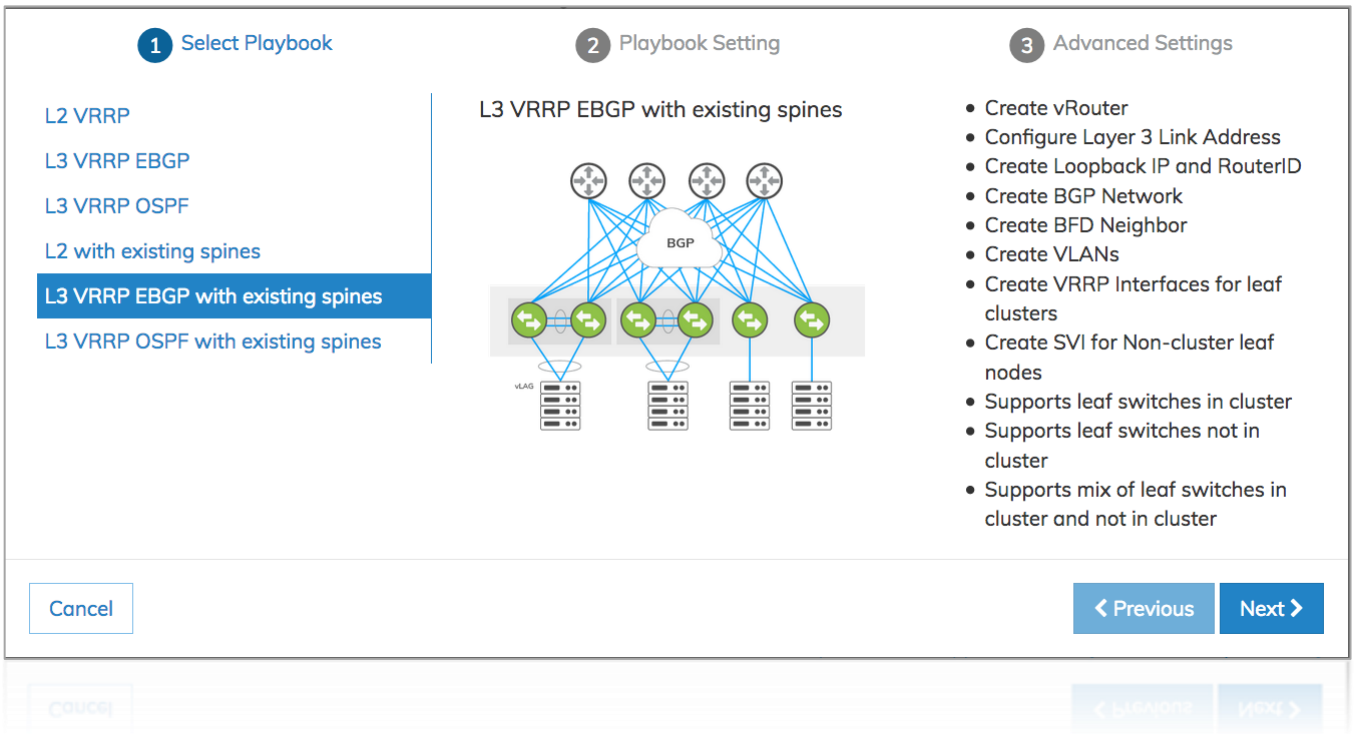


Figure 13. Automating the integration with third-party BGP switch spines using UNUM Playbook

Enabling VxRail services in the network from UNUM

This chapter discusses individual workflow implementation in UNUM for various network services required by VxRail. Please refer to the next chapter for a fully automated approach for provisioning the same services from vCenter.

Enabling vSAN in the network

VxRail is based on a VMware vSAN technology that traditionally used to require a multicast enabled network. Layer 2 Multicast is frequently misconfigured or not enabled at all due to complexity, resulting in uncontrolled large flooding domains. The Pluribus solution aims at simplifying the provisioning of this IP storage service into the network by detecting vSAN VMkernel interfaces and by automatically enabling Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) snooping.

Regardless of the vSAN software version and the use of a multicast or unicast based control plane, a fundamental requirement of the ToR is to offer a transport network for vSAN services. This chapter describes the simple steps required to instantiate the vSAN Virtual Local Area Network (VLAN) in the Pluribus fabric, according to the specific VLAN ID value used for VxRail configuration.

VLAN provisioning can be done via the VLAN configuration wizard: select LEAF-1 or LEAF-2 on the left side bar, activate the top menu Manage → Layer2 →Manage VLAN; start the wizard by clicking on *Create VLANs by range* button, as shown in Figure 14 on the left side.

The wizard allows the creation of one or multiple VLAN objects on a chosen switch. By selecting *scope* Fabric, the switch that receives the command synchronizes the VLAN database of the entire fabric and instantiates the VLAN object on each switch: this removes the operational burden of defining the same VLAN ID on each individual switch. The wizard allows also the association of the created VLAN(s) with any specified ports on the selected switch. In the example shown, the VLAN is created on LEAF-1 and associated with ports 1-2, which connect to locally attached VxRail nodes, and 65 and 69, which connect to the cluster peer, LEAF-2. Press the button *Save* to complete.

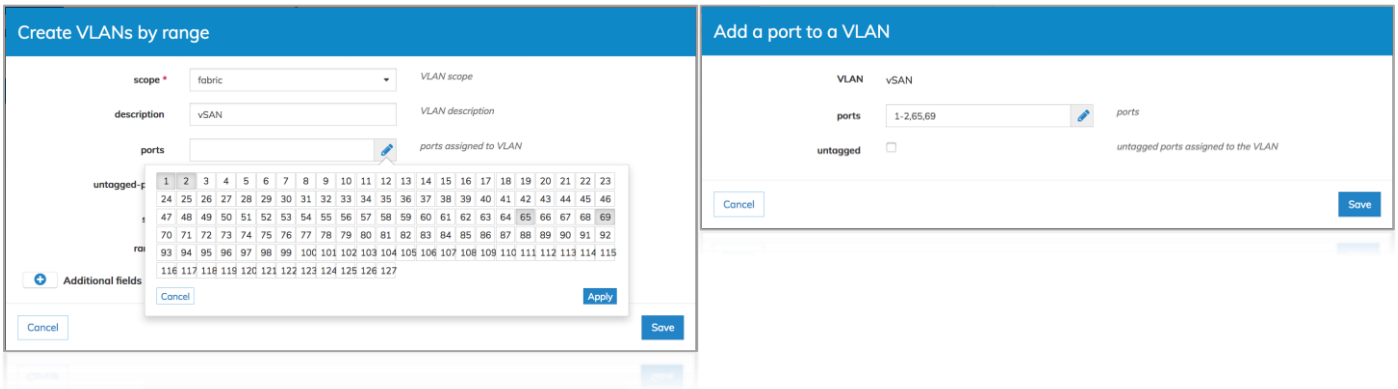
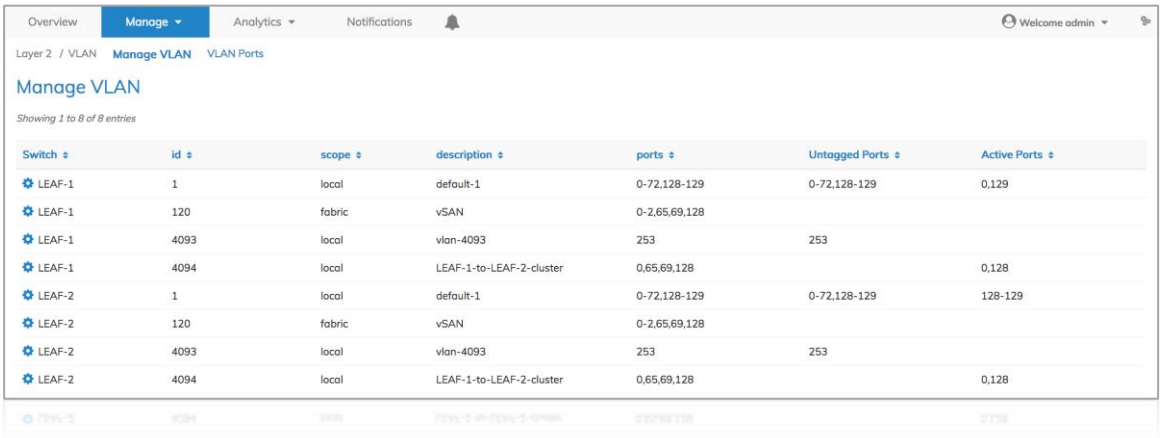


Figure 14. Provisioning vSAN VLAN: defining the VLAN (left) and assigning the VLAN to physical ports (left and right).

As a second step, associate the vSAN VLAN with physical interfaces on switch LEAF-2: activate the top menu *Manage → Layer2 →VLAN ports*; select the *VLAN* vSAN and start the wizard by pressing the *Add a port to a VLAN* button. On the wizard, as shown on the right side of Figure 14, enter the required port range to include physical interfaces connecting to the VxRail nodes and to the peer cluster switch, LEAF-2. Press the button *Save* to complete. The obtained configuration can be verified from the top menu Manage → Layer2 →Manage VLAN, as in Figure 15.



Switch	id	scope	description	ports	Untagged Ports	Active Ports
LEAF-1	1	local	default-1	0-72,128-129	0-72,128-129	0,129
LEAF-1	120	fabric	vSAN	0-2,65,69,128		
LEAF-1	4093	local	vlan-4093	253	253	
LEAF-1	4094	local	LEAF-1-to-LEAF-2-cluster	0,65,69,128		0,128
LEAF-2	1	local	default-1	0-72,128-129	0-72,128-129	128-129
LEAF-2	120	fabric	vSAN	0-2,65,69,128		
LEAF-2	4093	local	vlan-4093	253	253	
LEAF-2	4094	local	LEAF-1-to-LEAF-2-cluster	0,65,69,128		0,128

Figure 15. vSAN VLAN provisioned on LEAF-1 and LEAF-2 with VLAN ID 120 as a single fabric scope object

Enabling vMotion in the network

According to VxRail recommended design, the vMotion service requires layer-2 reachability between cluster nodes. Similarly to provisioning vSAN transport in the previous chapter, vMotion can be enabled by creating a Fabric VLAN object with the proper ID value, as configured in VxRail, and assigning it to the ports connecting to the servers and to inter-cluster links on both LEAF-1 and LEAF-2.

Enabling other services

Additional VLANs can be manually provisioned as required in order to provide connectivity between guest Virtual Machines (VMs). The configuration steps can simply follow the procedure described for enabling vSAN service.

It is also possible that certain ESXi servers are configured with a Network Interface Card (NIC) teaming strategy that requires provisioning of LAG and/or vLAG objects in the network. In this case, it is possible to follow the procedure used in previous sections of the document for interconnecting the leaf cluster to the spine virtual chassis.

Automating VxRail network provisioning using vCenter

All workflows described in previous chapter can also be fully automated by instantiating the vCenter Connection Service (vCCS) on Dell-Pluribus fabric. This service can be enabled using Netvisor Command Line Interface (CLI) or RESTful Application Programming Interface (API). For example, using the CLI via an SSH session:

```
network-admin@LEAF-1> vcenter-connection-create name vCCS host 10.37.30.100 user
administrator@vsphere.local vlans 3000-3999 network-provisioning l2-underlay
```

In the above example, the following parameters are used:

- name: string that uniquely identifies the vCCS service
- host: hostname or IP address of vCenter server
- user: vCenter administrator credentials
- vlans: comma/hyphen separated list of VLAN IDs, whose provisioning is delegated to the vCenter administrator
- network-provisioning: instructs VLAN provisioning to follow a layer-2 underlay design or a layer-3 underlay design. In the context of this guide, the underlay is layer-2.

As a result of enabling vCCS with the above command, the following three configuration workflows are automated:

ESXi host discovery and NIC teaming automation

Once ESXi servers are connected to the Fabric, leveraging dynamic discovery protocols on both servers and the network, the switches automatically create LAGs or vLAGs to match the teaming configuration on the server side. This completely automates the addition of new ESXi hosts in the network without any manual operation on the Adaptive Cloud Fabric.

VM network provisioning

When a new VM is created or associated to a standard or distributed port-group, the VLAN IDs configured at the host level are automatically defined within the Adaptive Cloud Fabric. Moreover, for security reasons and resources optimization, the Adaptive Cloud Fabric adds VLANs to downlinks towards servers only where required, and prunes automatically VLANs when they are not used on the servers by VMs. As a further security check, the Adaptive Cloud Fabric allows the network administrator the ability to restrict the dynamic usage of VLAN IDs that vCenter can automatically provision in the fabric according to the value of the parameter *vlans*. The provisioning is also extended to ToR uplinks and, according to the specified *network-provisioning* parameter value, it is performed using IEEE 802.1q encapsulation for *l2-underlay* or using Virtual extensible LAN (VxLAN) for *l3-underlay*.

VM kernel services like vSAN and vMotion

Pluribus vCCS fully automates the discovery of vSAN and vMotion services and dynamically creates and provisions the required VLAN IDs on physical ports, and instantiates any required multicast service.

VMware Endpoint Visibility

One of the key components of the Pluribus Adaptive Cloud Fabric is the ability to build a database of all endpoints or virtual Ports (vPorts) connected to the network and to provide an easy way to display this information to the network administrator. By using the vPort database as a central source of information for all devices connected to the fabric, physical or virtual, the network admin can considerably reduce the time to identify an endpoint and use the vPort details as a tool to simplify troubleshooting.

By default, the Adaptive Cloud Fabric provides the following information to the admin for every vPort (non-exhaustive list):

- Switch
- Port
- Tenant
- MAC address
- VLAN
- VXLAN Segment
- IP address
- Status

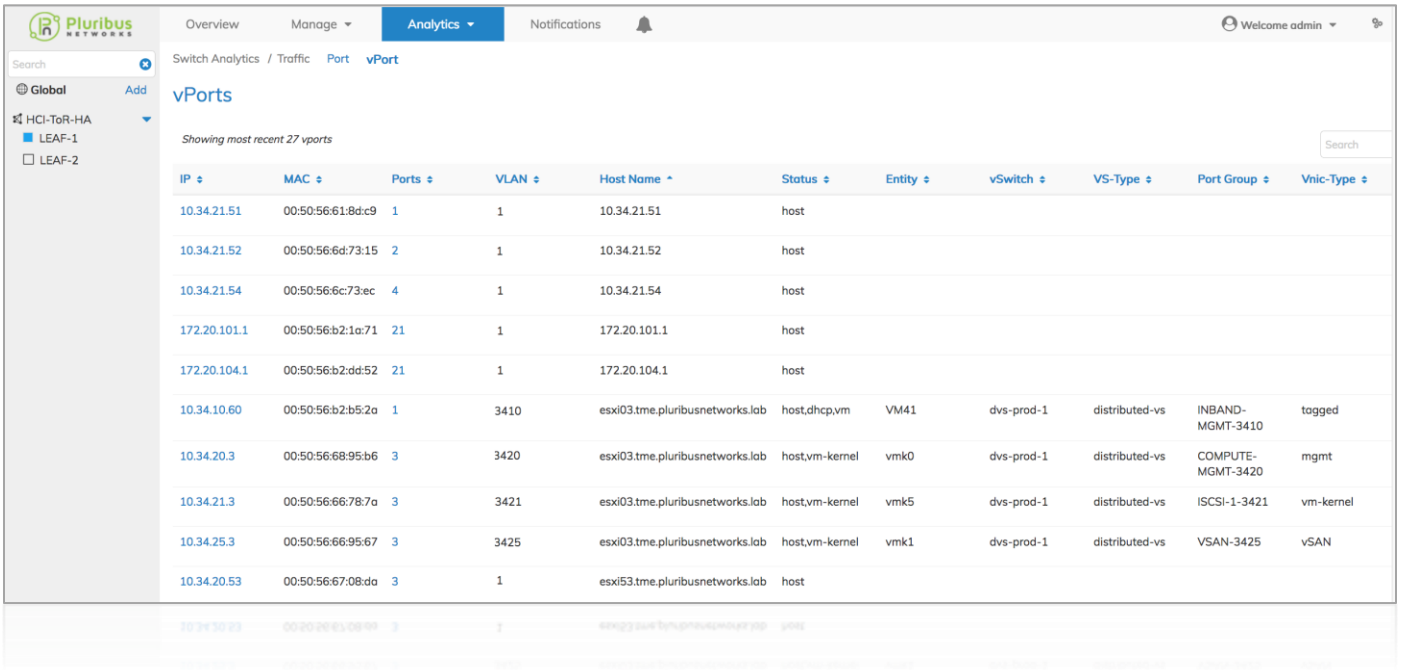
With vCCS, the fabric can provide a tremendous amount of information in addition to what already exists for each vPort.

For each vPort associated to a vCenter, the following attributes will be added to the vPort metadata:

- VM Name
- Host
- vNIC (name, type)
- VMKernel (name, service)
- vSwitch name and type (standard or distributed)
- Port group name and configuration (VLAN config, VLANs, teaming, security, QoS)
- VM Guest OS type
- VMware Tag

These attributes can also be used by the optional UNUM Insight Analytics module to display comprehensive and customized dashboards to report flow analytics of any application running on VMware vSphere.

An example of vCenter-aware vPort database, available on UNUM, can be seen in Figure 16.



IP	MAC	Ports	VLAN	Host Name	Status	Entity	vSwitch	VS-Type	Port Group	Vnic-Type
10.34.21.51	00:50:56:61:8d:c9	1	1	10.34.21.51	host					
10.34.21.52	00:50:56:6d:73:15	2	1	10.34.21.52	host					
10.34.21.54	00:50:56:6c:73:ec	4	1	10.34.21.54	host					
172.20.101.1	00:50:56:b2:1a:71	21	1	172.20.101.1	host					
172.20.104.1	00:50:56:b2:dd:52	21	1	172.20.104.1	host					
10.34.10.60	00:50:56:b2:b5:2a	1	3410	esxi03.tme.pluribusnetworks.lab	host,dhcp,vm	VM41	dvs-prod-1	distributed-vs	INBAND-MGMT-3410	tagged
10.34.20.3	00:50:56:68:95:b6	3	3420	esxi03.tme.pluribusnetworks.lab	host,vm-kernel	vmk0	dvs-prod-1	distributed-vs	COMPUTE-MGMT-3420	mgmt
10.34.21.3	00:50:56:66:78:7a	3	3421	esxi03.tme.pluribusnetworks.lab	host,vm-kernel	vmk5	dvs-prod-1	distributed-vs	ISCSI-1-3421	vm-kernel
10.34.25.3	00:50:56:66:95:67	3	3425	esxi03.tme.pluribusnetworks.lab	host,vm-kernel	vmk1	dvs-prod-1	distributed-vs	VSAN-3425	vSAN
10.34.20.53	00:50:56:67:08:da	3	1	esxi53.tme.pluribusnetworks.lab	host					

Figure 16. vPorts active on fabric switch LEAF-1, inclusive of vSphere attributes

Conclusion

This guide demonstrates how deploying VxRail in any network can be achieved with simple, consistent and automated tasks using Pluribus Adaptive Cloud Fabric on Dell ON switches.

References

1. Open Netvisor for Linux ONVL Installation Guide
2. ONVL Getting Started Guide
3. UNUM Installation Guide

Copyright © 2018 Pluribus Networks, Inc. All Rights Reserved. Netvisor is a registered trademark, and the Pluribus Networks logo, Pluribus Networks, Adaptive Cloud Fabric, UNUM and Insight Analytics are trademarks of Pluribus Networks, Inc. The Dell EMC Logo, Dell and EMC, are registered trademarks, and VxRail is a trademark of Dell Inc., or its subsidiaries. All other brands and product names are registered and unregistered trademarks of their respective owners. This is a joint solution of Dell EMC and Pluribus Networks and requires multiple products to enable the capabilities referenced, and functionality will vary by physical device deployed. Pluribus Networks and Dell EMC reserve the right to make changes to its technical information and specifications at any time, without notice.

DE/PN-TB-VXR-072218

Pluribus Networks, Inc.
6001 America Center Dr., Suite 450, San Jose, CA 95002

+1-650-289-4717 | +1-855-438-8638
www.pluribusnetworks.com