Hewlett Packard Enterprise

HPE Synergy and Cisco ACI Networking Interop Guide

Version 1.2, January. 2021

Technical white paper

Contents

Executive Summary

With the increasing deployment of HPE Synergy and Cisco ACI fabric, this white paper aims to help customers to understand Synergy/ACI networking interop solutions in more details.

Target Audience

This guide is intended for network architects, administrators and engineers who needs to interconnect HPE Synergy with Cisco ACI fabric.

Equipment

DEVICE	VERSION	QTY
HPE Synergy managed by OneView	5.20.01 and HPE Synergy SPP 2020.07.02	
HPE Virtual Connect SE 40GB F8 Module for Synergy	1.6.1.1002	2
Cisco ACI APIC	3.2(7f)	
Cisco Nexus 93180YC-EX	13.2(7f)	2

Note: All discussions in the white paper also apply to current latest OneView version 5.50 and HPE Synergy Virtual Connect SE 100Gb F32 module.

Overview

The key networking component of HPE Synergy Composable Infrastructure is HPE Synergy Virtual Connect SE 100Gb F32 Module. Its disaggregated, rack-scale design uses a master/satellite architecture to consolidate data center network connections, reduce hardware and scales network bandwidth across multiple HPE Synergy Frames. Each Synergy Virtual Connect SE 100Gb F32 module provides up to six uplinks using QSFP+ interfaces, up to 24 Ethernet/Fibre Channel uplinks are available using splitter cables. The Virtual Connect SE 100Gb F32 Modules reduce the number of components required compared to traditional and other converged network solutions by eliminating the need for separate Ethernet and Fibre Channel switches and cables.

For ACI Interop networking, users could just focus on the 6 QSFP+ uplink ports showing in the following diagram. For each QSFP+ port, users have the option to run single 100G/40G uplink to ACI or use splitter cables to run 4x25G/10G uplink.

6 x 100Gb QSFP28 uplink ports

Eth/FCOE: 100Gb, 40Gb, 4x25Gb or 4x10Gb

FC: 4x32/16/8Gb

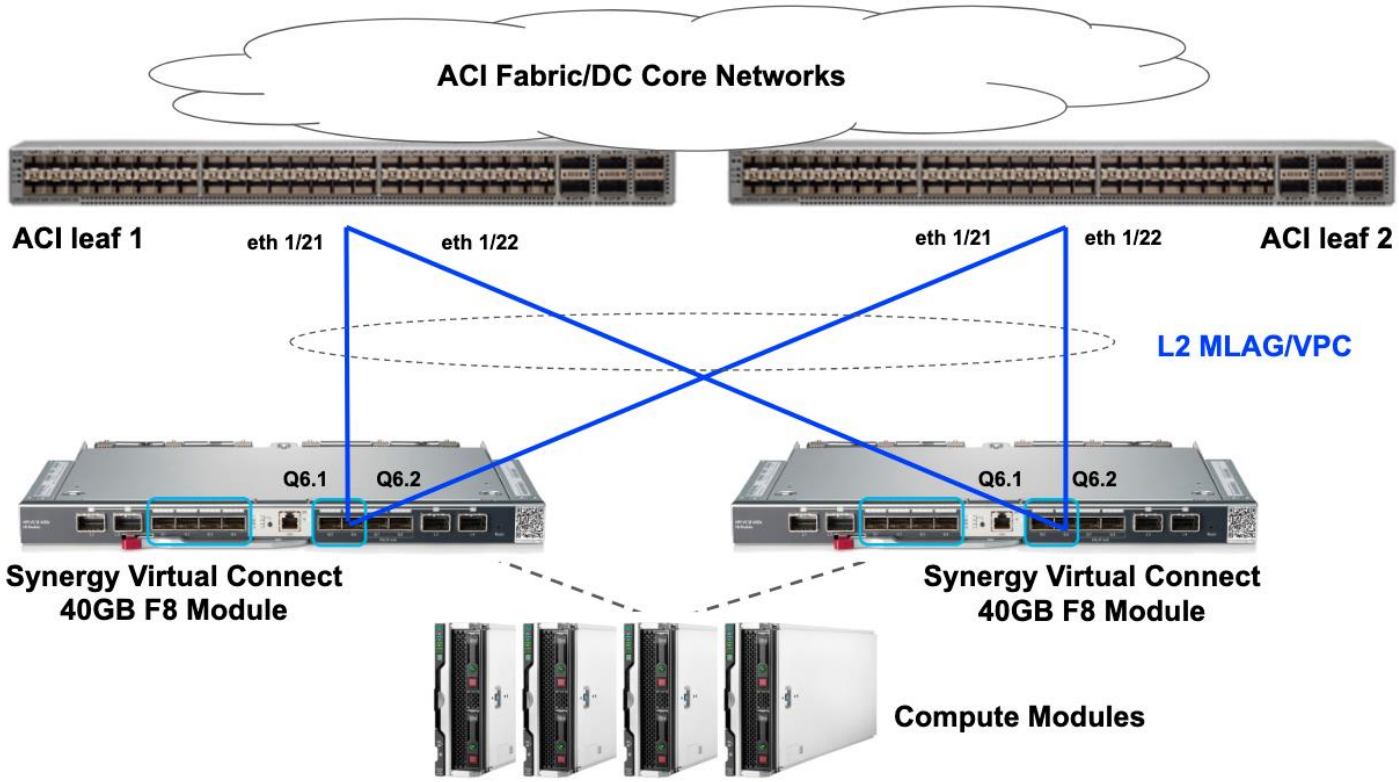
The earlier version of HPE Synergy Virtual Connect SE 100Gb F32 Module is HPE Synergy Virtual Connect SE 40Gb F8 Module show below. All discussion in this white paper apply to both modules.

6x 40Gb uplink ports

Q1-Q6: 40Gb, 4x10Gb Ethernet/FCoE, or 4x8Gb Fibre Channel

Networking Topology

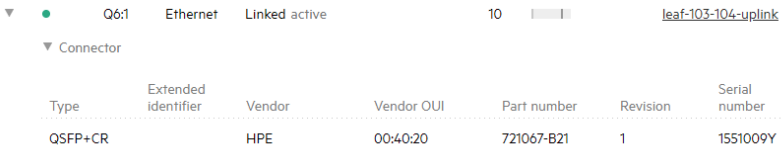
The networking topology used in this white paper consists of 4x10G links across two ACI 93180YC-EX leaf switches and two HPE Synergy Virtual Connect 40GB F8 Modules. On ACI side, VPC port-channel is formed across two leaf nodes while on Synergy side, Multi-chassis Link Aggregation group(MLAG) is formed across two VC modules in separate Synergy frames to provide maximum HA.



Cabling

HPE Synergy Virtual Connect Quickspecs include official support cables and transceivers. 40G QSFP+ to 4x10G SFP+ cables like the following are used in this white paper.

User can show current used cable details in Synergy OneView -> Interconnect menu.



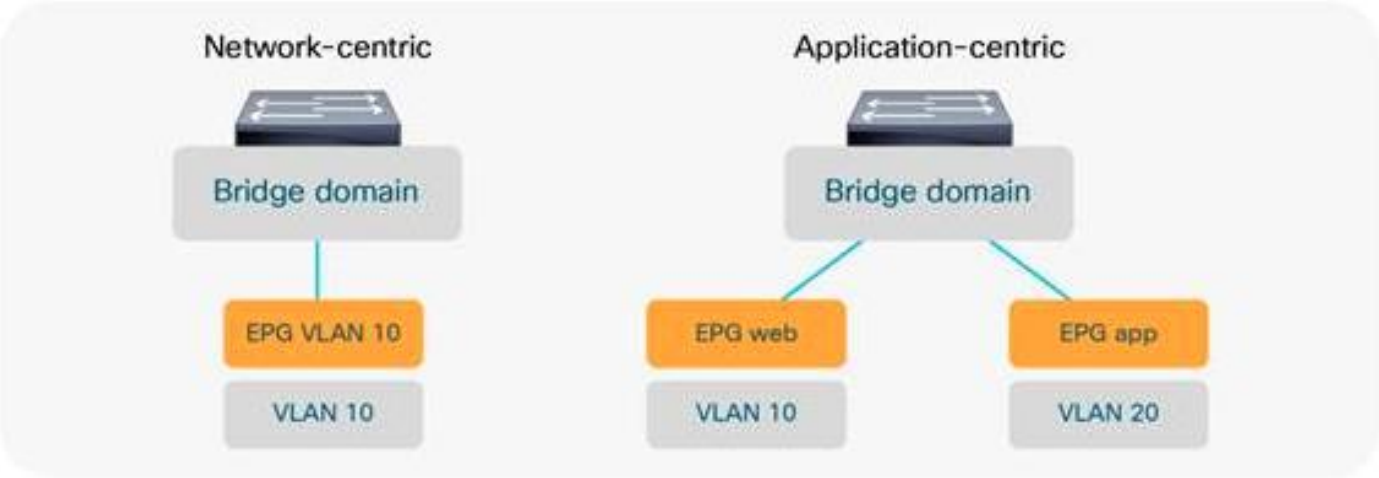
This is the part number shown on Cisco ACI leaf switch side.

```
Ethernet1/21
    transceiver is present
    type is SFP-H10GB-CU5M
```

ACI Design Terms

There are two commonly used terms in regard to ACI designs. The first one is called network-centric design. In this design, users create one or more bridge domains and EPGs and use the mapping 1 bridge domain = 1 EPG = 1 VLAN. Network centric mode suits the use cases where admins want to transition from traditional VLAN-based networking to ACI fabric and they'd like to keep leveraging existing network design logic and IP address scheme.

The other term is referred as application-centric design. In the application-centric design, the bridge domain is divided into EPGs that represent application tiers and between EPGs, contracts are applied to define ACL filtering and security zones. The following diagram from Cisco ACI design white papers highlight the differences between two designs.



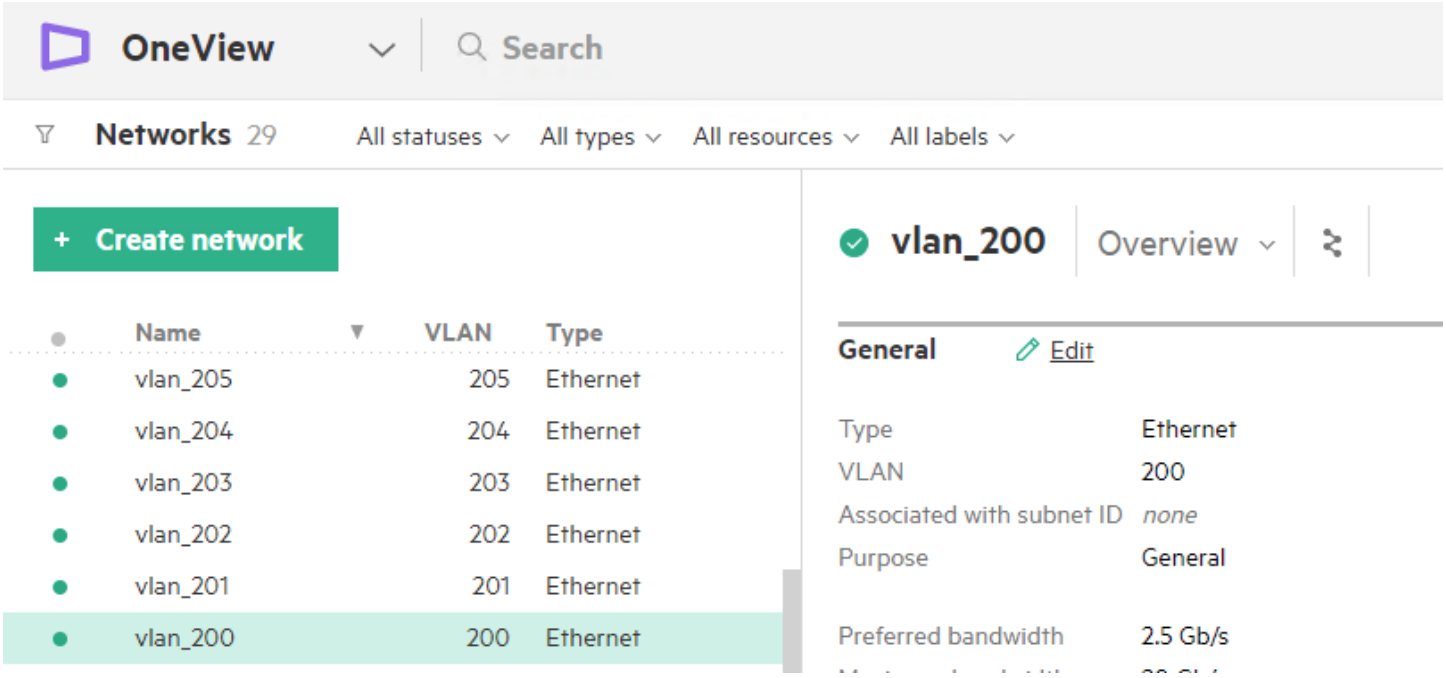
In regard to Synergy interop, network-centric behaves exactly like traditional networking forwarding behavior for network edge devices like Synergy Virtual connect. Therefore, ACI/Synergy VC will interop well without additional notes. ACI app-centric mode provides some unique operation like inter-VLAN bridging and this white paper will also discuss in more details on how this mode interops with Synergy tunnel networks.

Synergy Virtual Connect Networking

Mapped vs Tunnel Networks

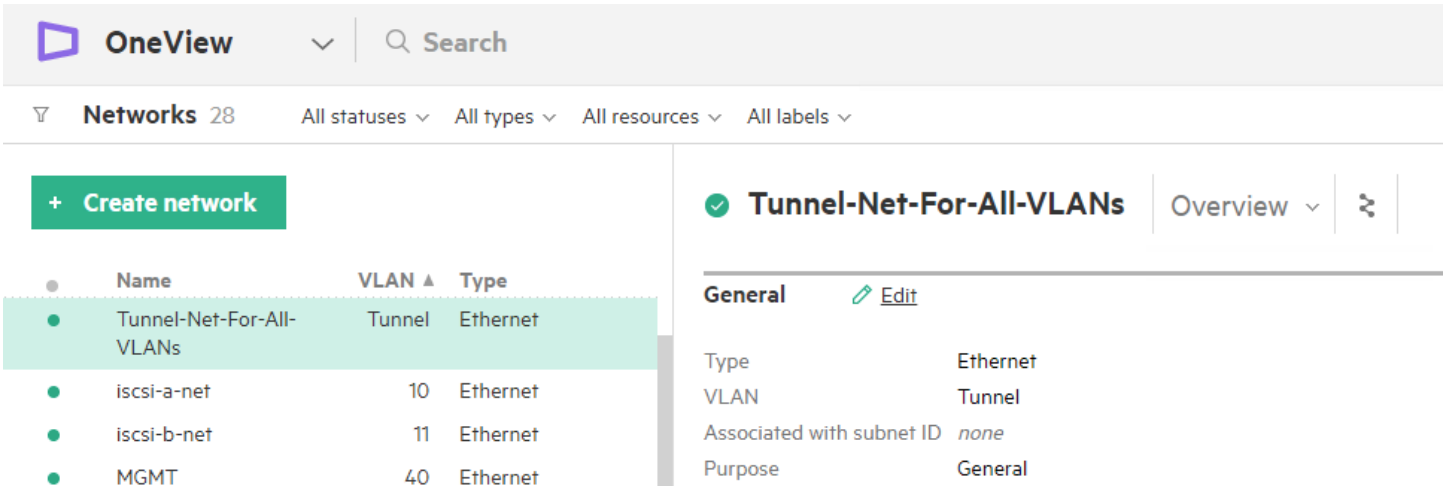
Synergy Virtual Connect has two primary methods of aggregating Ethernet networks internally. The first is "Tagged", aka "Mapped" network and the second is "Tunnel" network.

"Tagged"/"Mapped' networks configuration matches typical networking concept where you'll need to define all VC networks mapping to encapsulation VLANs on ACI leaf switches. For example, if users want to pass VLAN 200-205 from ACI leaf nodes to VMware ESXi hosts, users need to create six separate tagged networks in Synergy matching these six VLANs and then assign these networks to Synergy uplink set(facing ACI leaf nodes) and server profiles(facing ESXi hosts VS/DVS).



"Tunnel" networks configuration uses the network "pseudo-wire" concept. Using the same configuration example, if users want to pass VLAN 200-205 from ACI leaf nodes to VMware ESXi hosts, instead of requiring users to define six separate tagged VC networks, users can just define one single "Tunnel" VC network(like a single pseudo-wire) and assign the network to uplink set(facing ACI leaf nodes) and server profiles(facing ESXi hosts VS/DVS).

Note: In the following picture of tunnel network configuration, users don't need(and can't) specify user VLAN ids like VLAN 200 for this tunnel network. The tunnel/pseudo-wire behavior is implemented internally by Synergy VC and not based on individual user VLANs. In fact, for tunnel network, Synergy VC won't even look into user VLAN IDs when it pass the traffic between ACI leaf nodes and hypervisors/servers.



Tagged/mapped network has advantages of matching typical networking VLAN configuration concepts and users have flexibility to assign different networks to different server NICs if needed. Tagged/mapped networks can scale up to hundreds of VLANs to server NICs. Also, when users make VLAN changes on ACI leaf nodes and hypervisor VS/DVS port-groups, users need to make corresponding changes in Synergy networks to keep "matching" the VLAN id/tagging on both sides.

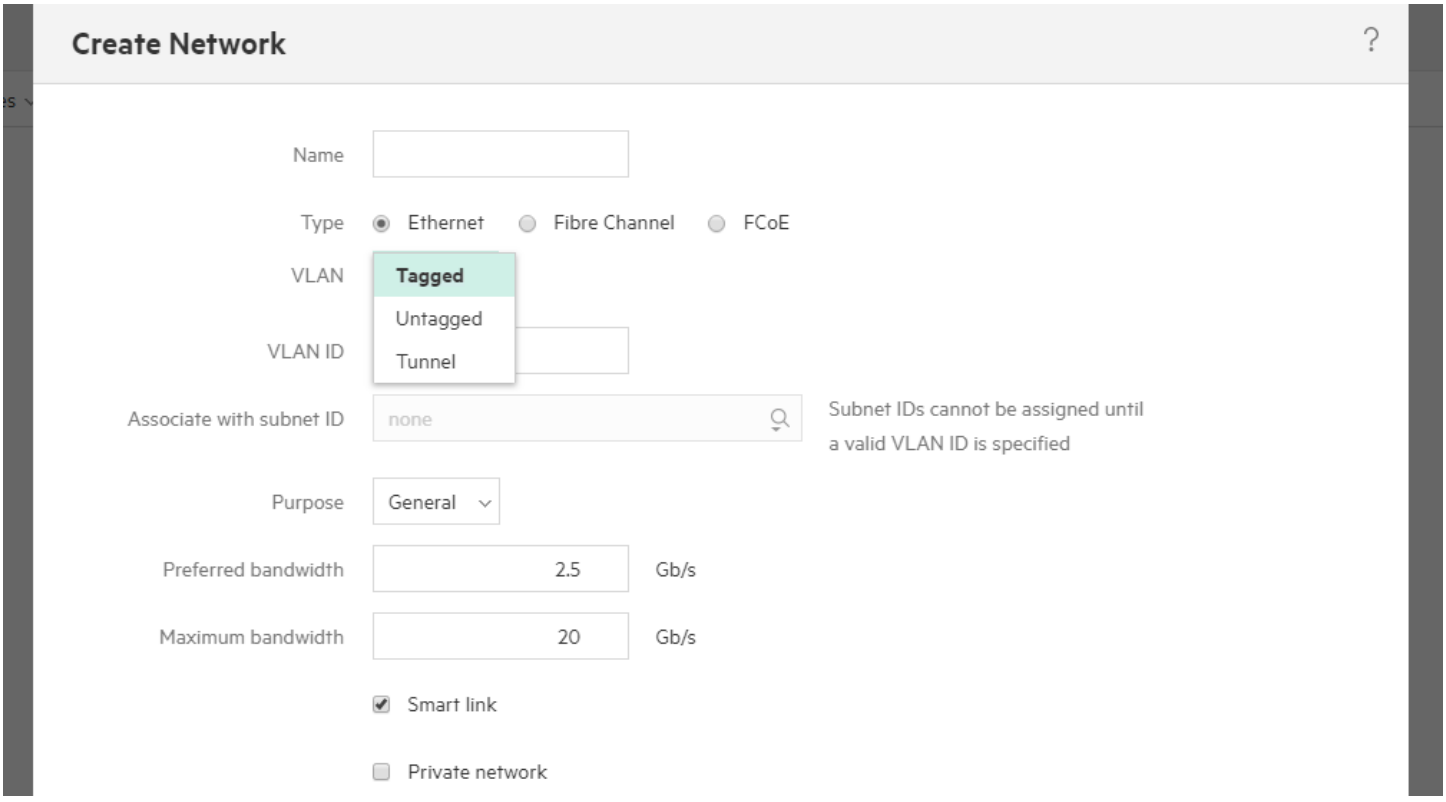
Tunnel network is a unique feature for Synergy VC. It may take some time to get used to this pseudo-wire concept at server edge. Also, VC tunnel network loses some flexibility VC mapped networks have. For example, you can't assign VLAN 200 to server NIC1 and VLAN 201 to server NIC2 when these VLANs are inside a VC tunnel network. You will need to assign the single VC tunnel network to the server NIC and have server VS/DVS port-group differentiate traffic based on VLANs.

However, in return VC tunnel network provides the powerful feature that users can set it and leave it in regard to network management. The moment users create a tunnel network to tunnel all networks between ACI leaf nodes and hypervisor, users don't need to login Synergy to make further changes to this tunnel network even if later they need to add/edit/delete networks on ACI and hypervisor side. The same VC tunnel network operates the same way when it tunnels one user VLAN or 4K user VLANs(which means VC tunnel network doesn't impose any VLAN scaling limits).

VC tunnel networks can also greatly simplify ACI VMM integration where ACI can create DVS port-groups from APIC dynamic VLAN pools. VC tunnel network configuration remains the same in regard to which dynamic VLAN id was chose by APIC to deploy on DVS.

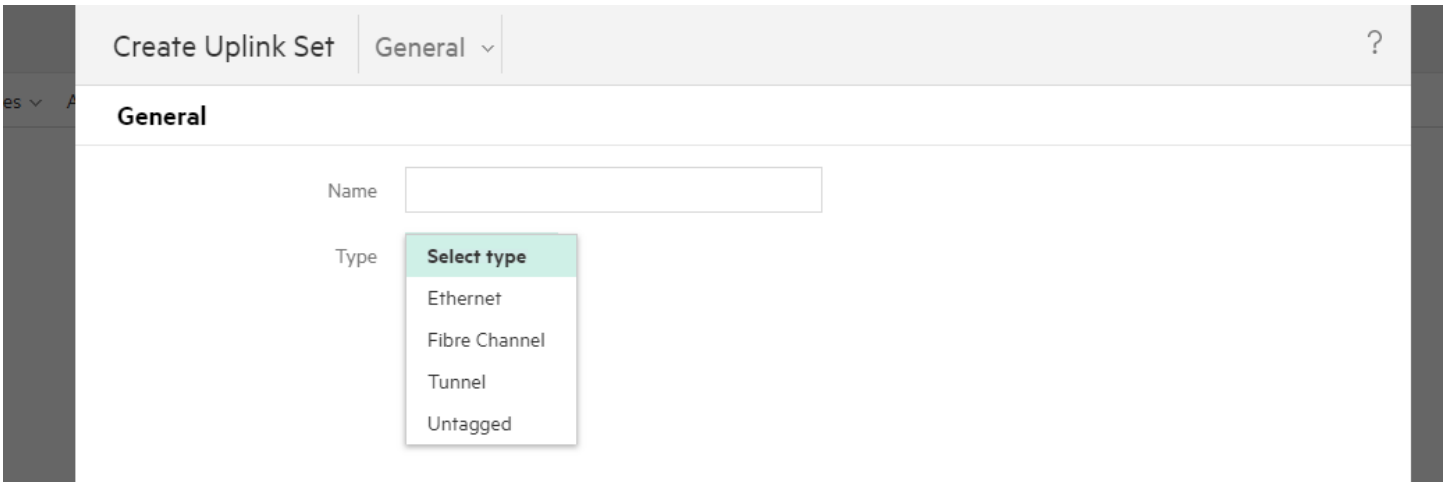
Networks and Uplink Set Configuration

When users first create a VC network, they have to choose network type as the following picture. The third method "Untagged" is also available but since this kind of network can't carry any VLAN tags, it's not used as often as Tagged/Tunnel networks so we'll focus on Mapped/Tagged and Tunnel networks in this white paper.

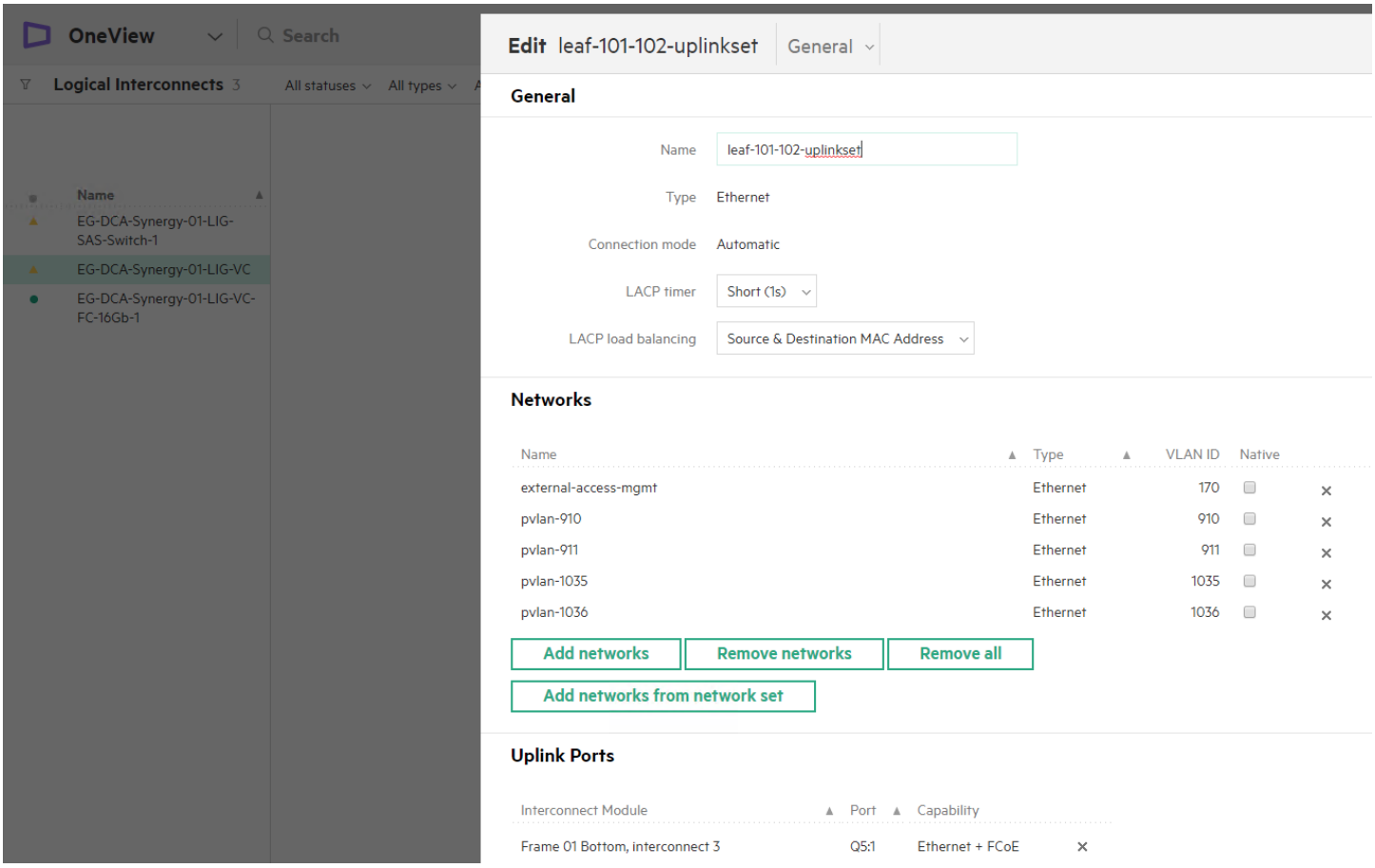


Creating VC network is the first step to connect user server VLANs to ACI fabric. The next step is to create Uplink Set facing Cisco ACI leaf switches. From Cisco admins perspective, an uplink set is equivalent to Cisco port-channel in which you define allowed VLANs and which physical ports will be LACP members of this uplink set.

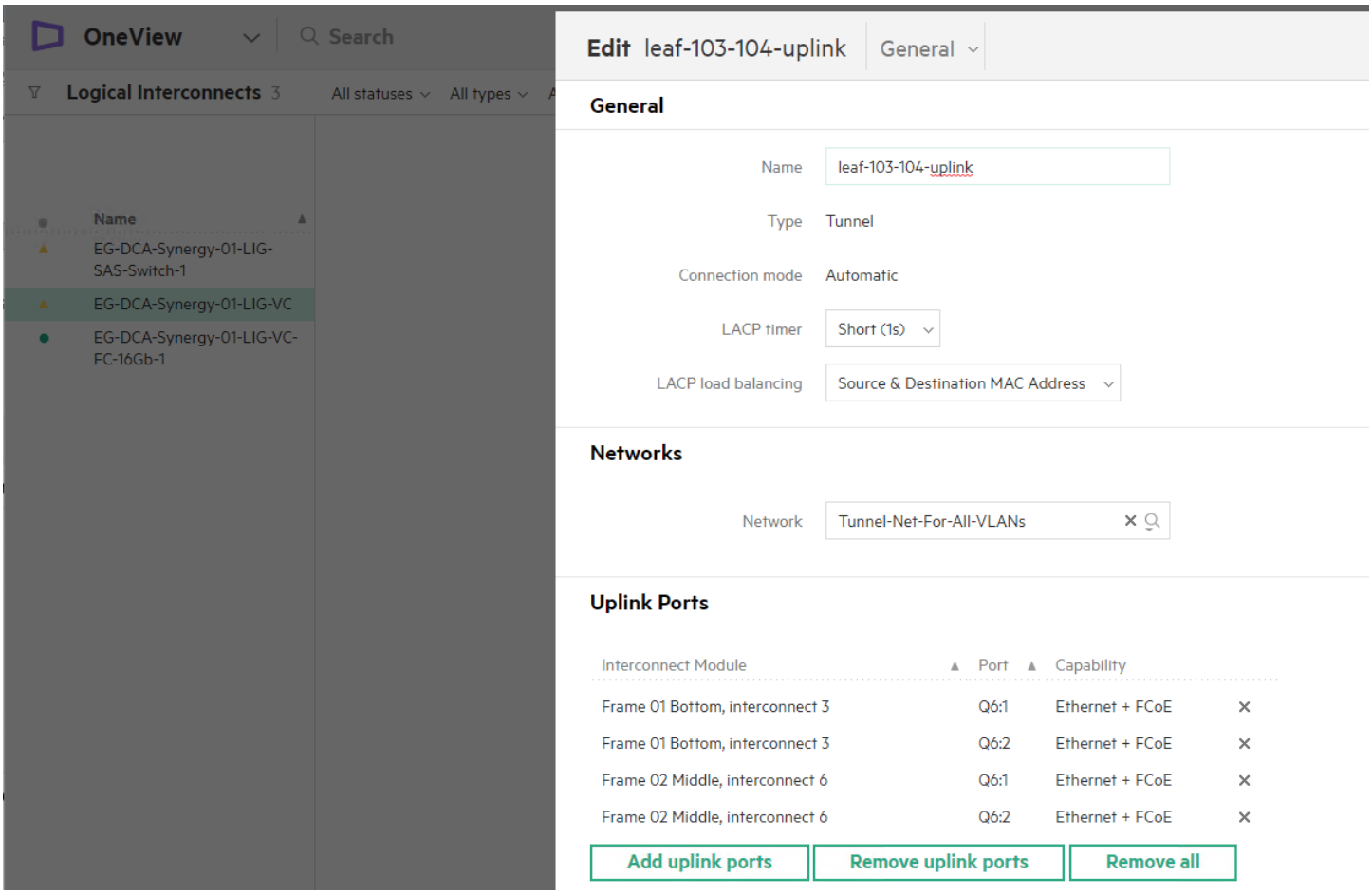
You have the option to choose “Ethernet” or “Tunnel” option for a uplink set. An “Ethernet” uplink set can only include “tagged/mapped” networks and a “Tunnel” uplink set can only include one single “tunnel” network.



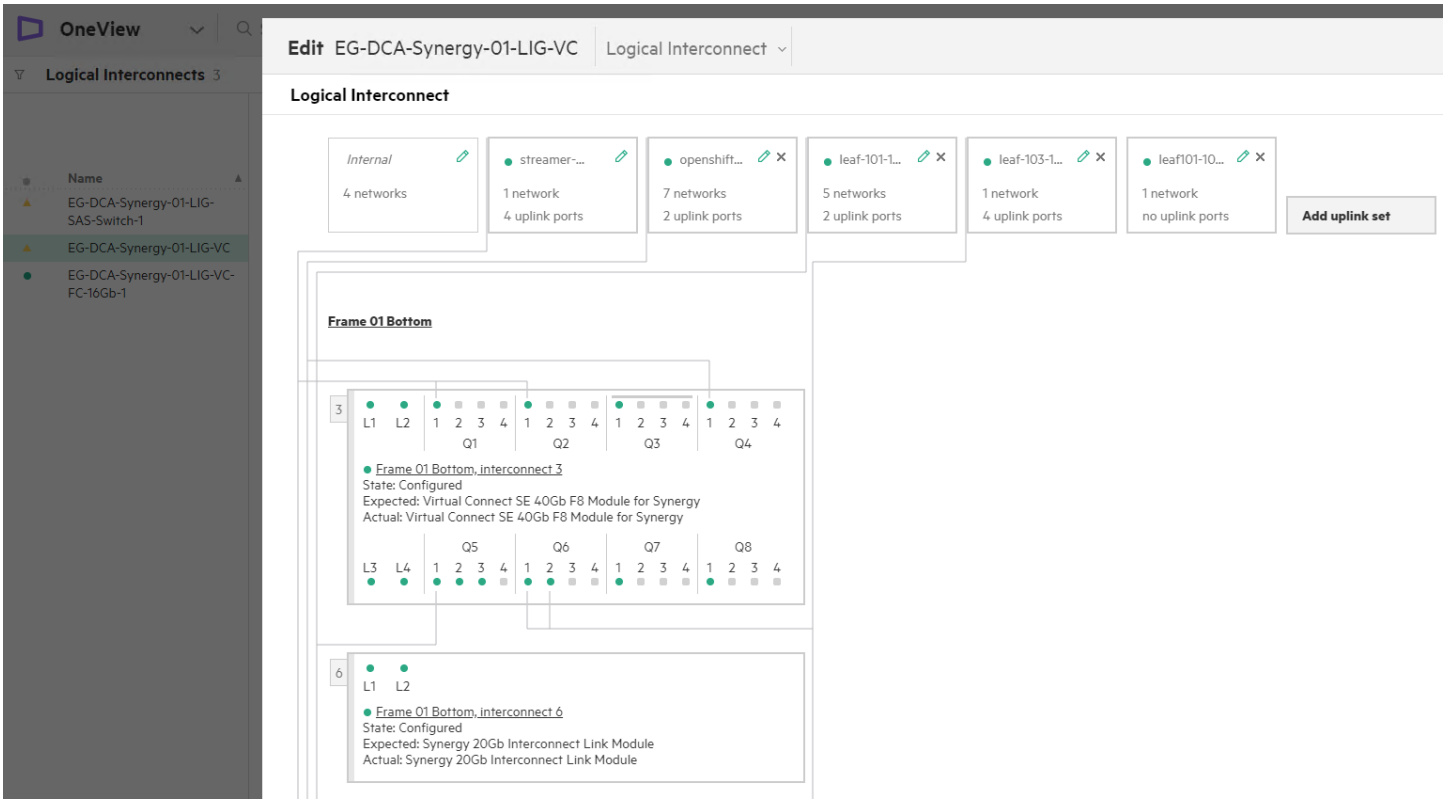
The following capture shows a typical “Ethernet” uplink set including multiple “tagged/mapped” networks and its physical ports



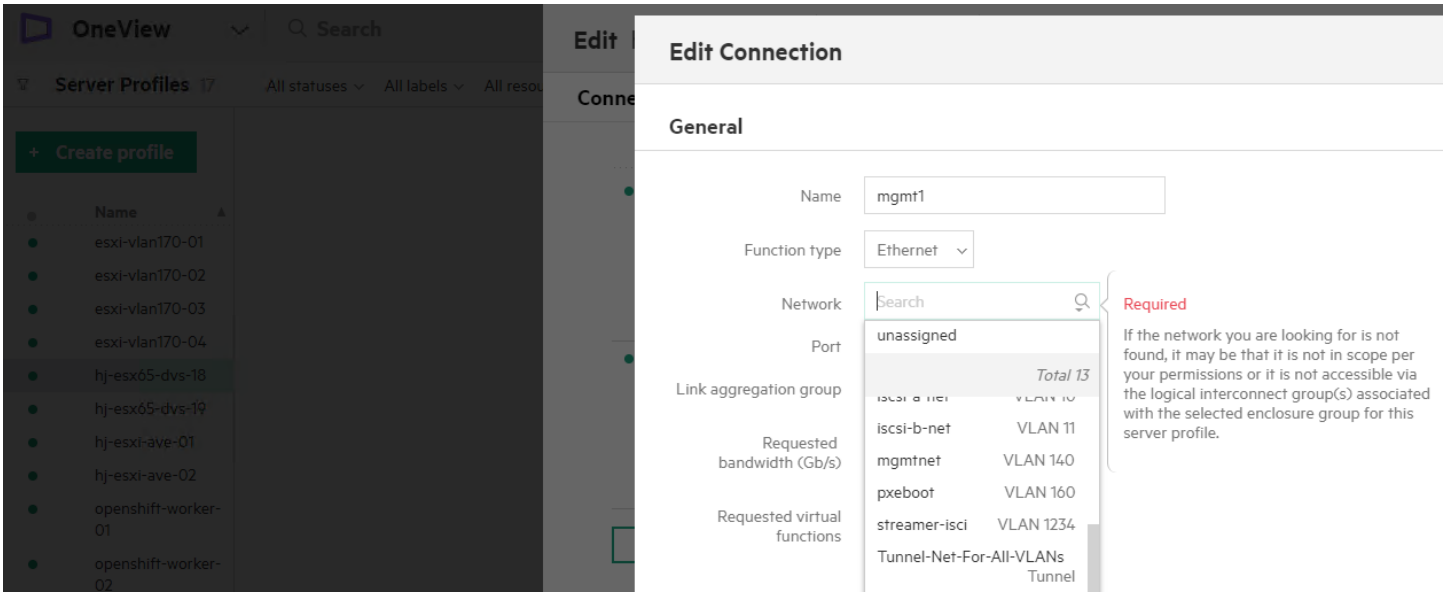
The following capture shows a “tunnel” uplink including a single “tunnel” network and its physical ports



Synergy Virtual Connect modules can have multiple uplink sets to different switch ports or different switches just like a Cisco switch or ACI leaf node can have multiple vPCs. Each uplink set operates independent from other uplink sets and you can have “Ethernet” uplink sets together with “Tunnel” uplink sets on Synergy VC modules. The following capture shows multiple “Ethernet” and “Tunnel” uplink sets connecting to different ACI leaf nodes and other Cisco switches.



For Synergy virtual connect downlinks facing hypervisors and compute modules, users can create multiple connections(physical NICs on servers) and for a giving connection, users can assign one single VLAN network or multiple VLAN networks(using “network-set” if mapped networks or one single tunnel network).



Cisco ACI and HPE Synergy Interop Configuration

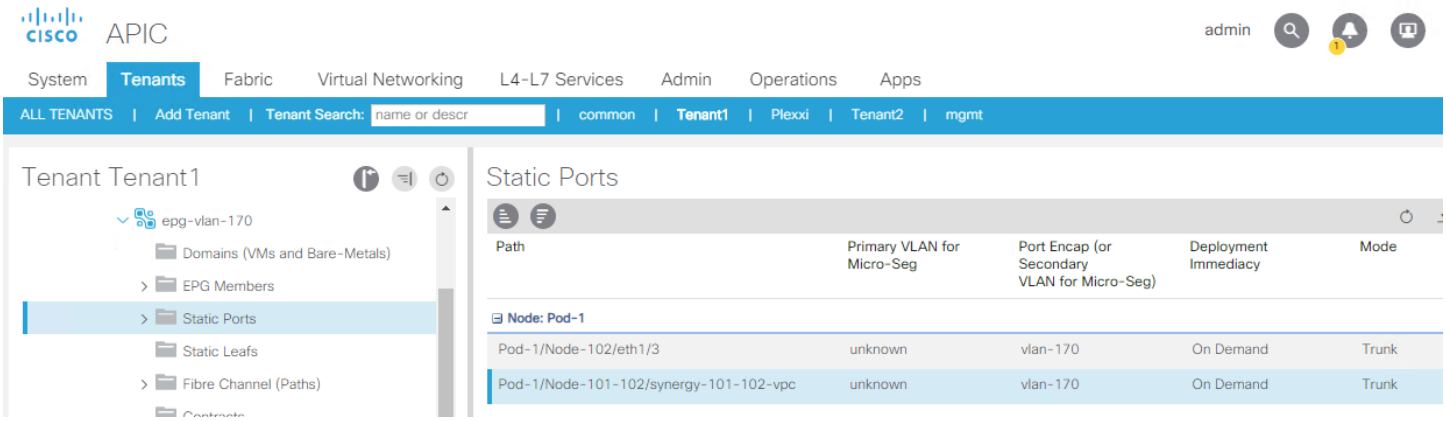
The interop configuration between ACI and Synergy is pretty straightforward. When passing traffic, Cisco ACI doesn’t know it’s talking with HPE synergy and vice versa, HPE Synergy doesn’t know it’s connected to ACI leaf nodes so for

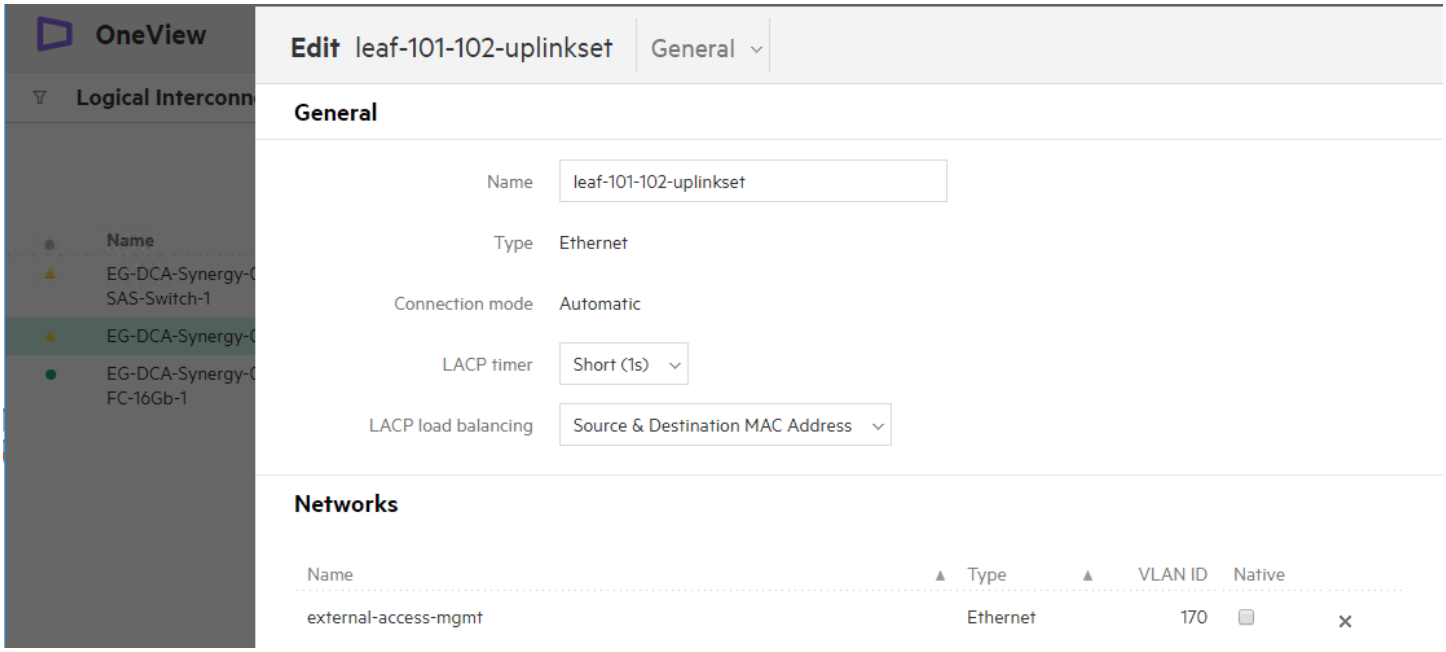
most use cases, it's standard configuration on both sides. ACI is designed to connect to any edge devices and Synergy is designed to connect to any core fabric.

Cisco ACI consists of many building blocks such as Tenants, Application Profiles, Contracts and VRFs. When interop with edge devices like HPE Synergy, the most important concept is Endpoint Groups(EPGs) and especially EPG port encapsulation VLANs.

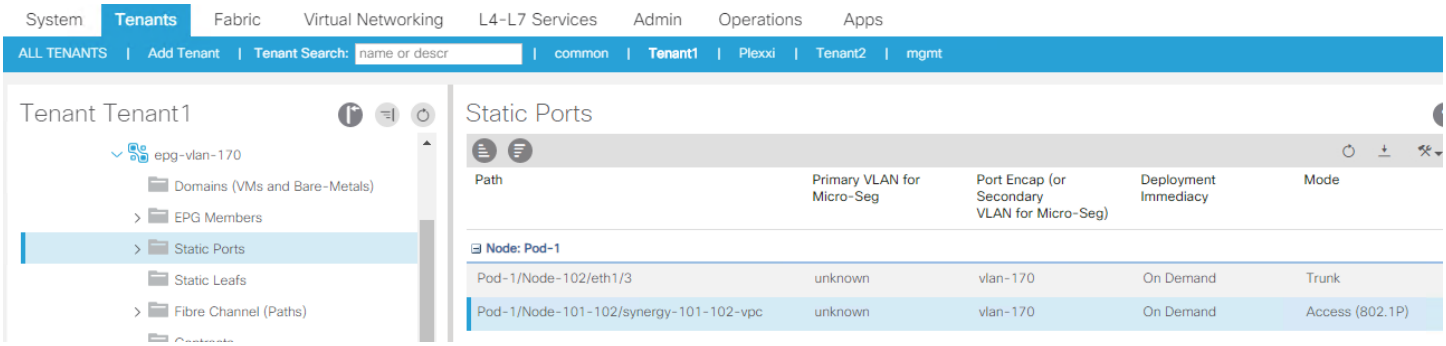
ACI EPG Encapsulation VLAN configuration with Synergy Mapped Networks

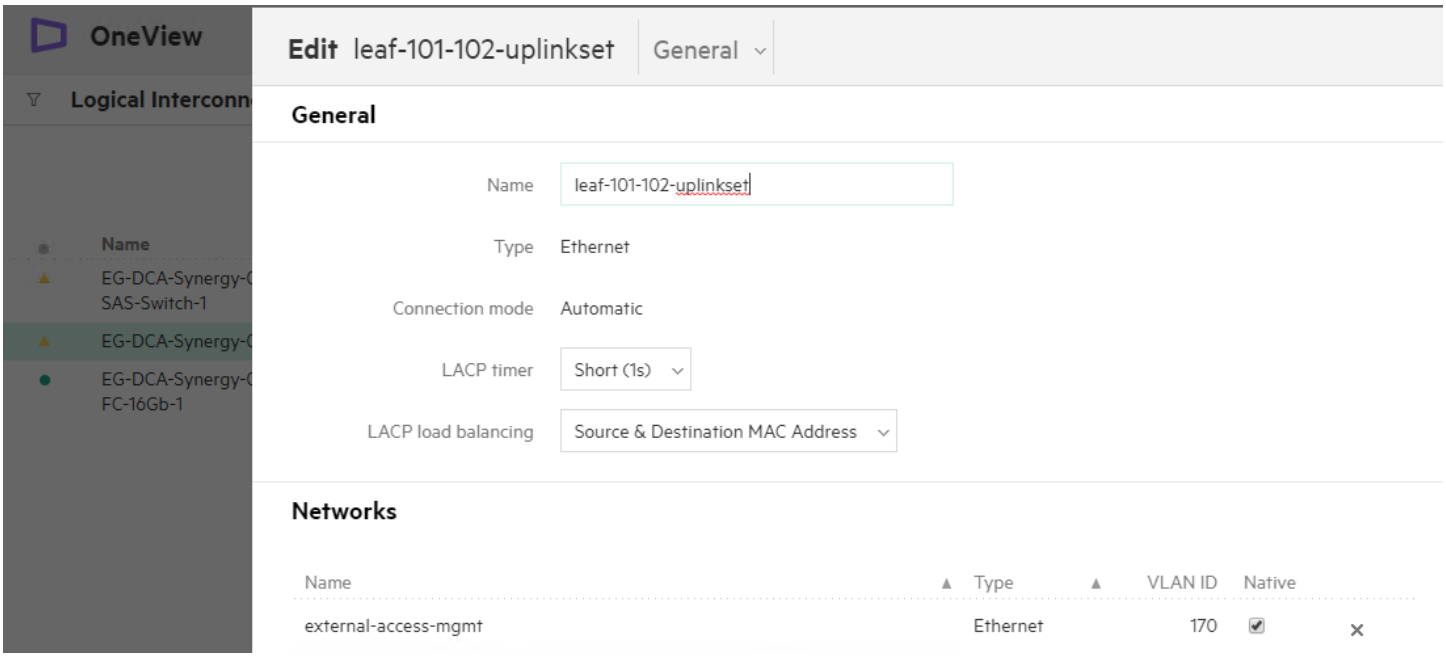
For a EPG encapsulation VLAN using VLAN tag, on ACI side, it'll be configured as mode "Trunk". On HPE Synergy side, if the uplink set is "Ethernet", then a network should be included in the uplink set with the same VLAN id. The following capture shows the valid configuration on both sides for VLAN 170.





If EPG encapsulation VLAN is a trunk native VLAN, on ACI side it should be configured as "802.1P" mode and on HPE Synergy side, if using Ethernet uplink set, then a network with the matching VLAN id should be included in the uplink set but "Native" option should be checked to match what's configured on ACI side. The following capture shows the same valid configuration for VLAN 170 in the trunk native mode.





For Synergy “Ethernet”/”Mapped” uplink, if you have multiple ACI EPG encapsulation VLANs facing Synergy, then you should define equivalent Synergy mapped networks and include them in the uplink set like what above examples illustrated.

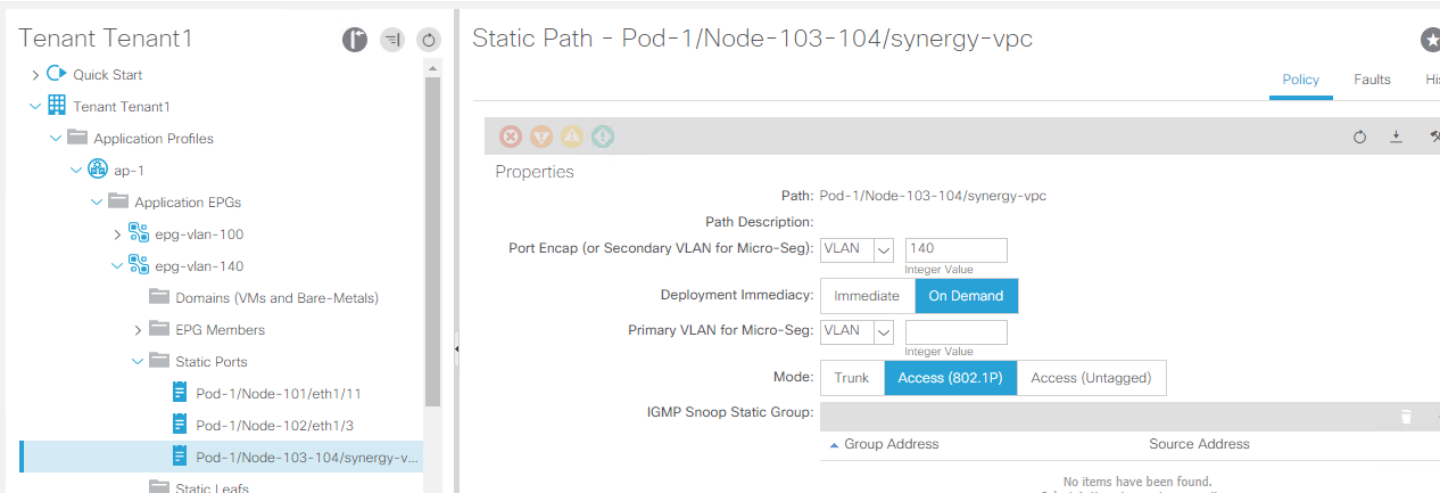
ACI EPG Encapsulation VLAN Configuration with Synergy Tunnel Networks

When Synergy uses tunnel networks to carry VLAN traffic between ACI and hypervisor, the configuration is very simple. Users only need to define a single tunnel network inside Synergy and include it in a tunnel uplink set facing ACI.

This configuration is totally transparent for ACI and Hypervisor side. On ACI and hypervisor side, admins configure encapsulation VLANs and port-groups like ACI leaf nodes are directly connected with hypervisor hosts so the right VLAN tagging matching configuration need to be ensured on both sides, just like when ACI leaf nodes are directly connected with rack mount servers.

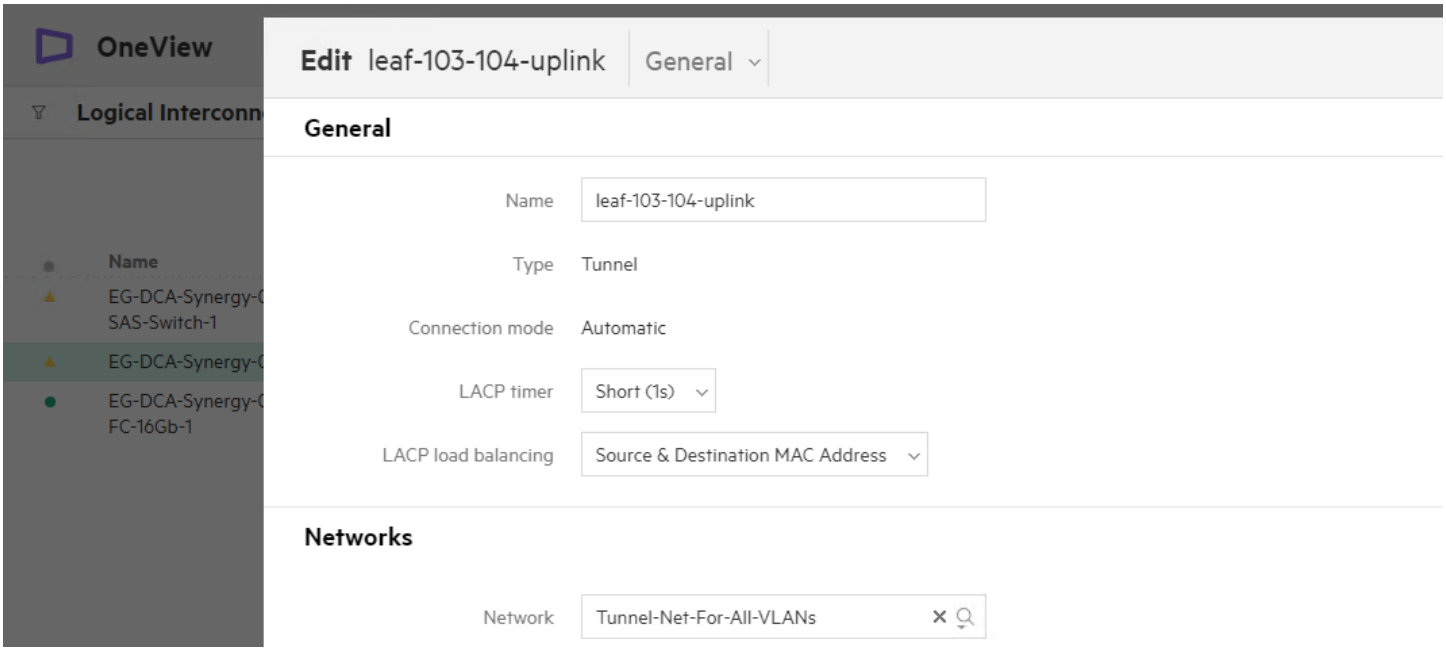
In the following capture, on ACI side, one of EPG encap VLAN is 140 and it’s configured as trunk native mode. The EPG uses static ports binding off “synergy-vpc” virtual port-channel. The vPC has multiple domains attached and the active encap VLANs include VLAN 140, other static/dynamic VLANs, and ACI infra VLAN(if users want to deploy AVE vxlan-mode in Synergy hosts).

On Synergy side, a single tunnel uplink set with a single tunnel network is defined to “tunnel” these VLANs transparently to hypervisor hosts.




```
apic3# show run template port-channel synergy-vpc
# Command: show running-config template port-channel synergy-vpc
# Time: Fri Aug  2 12:14:46 2019
  template port-channel synergy-vpc
    cdp enable
    vlan-domain member aci-ave type vmware
    vlan-domain member aci-dvs type vmware
    vlan-domain member domain-vlan1-300 type phys
    channel-mode active
    exit
```

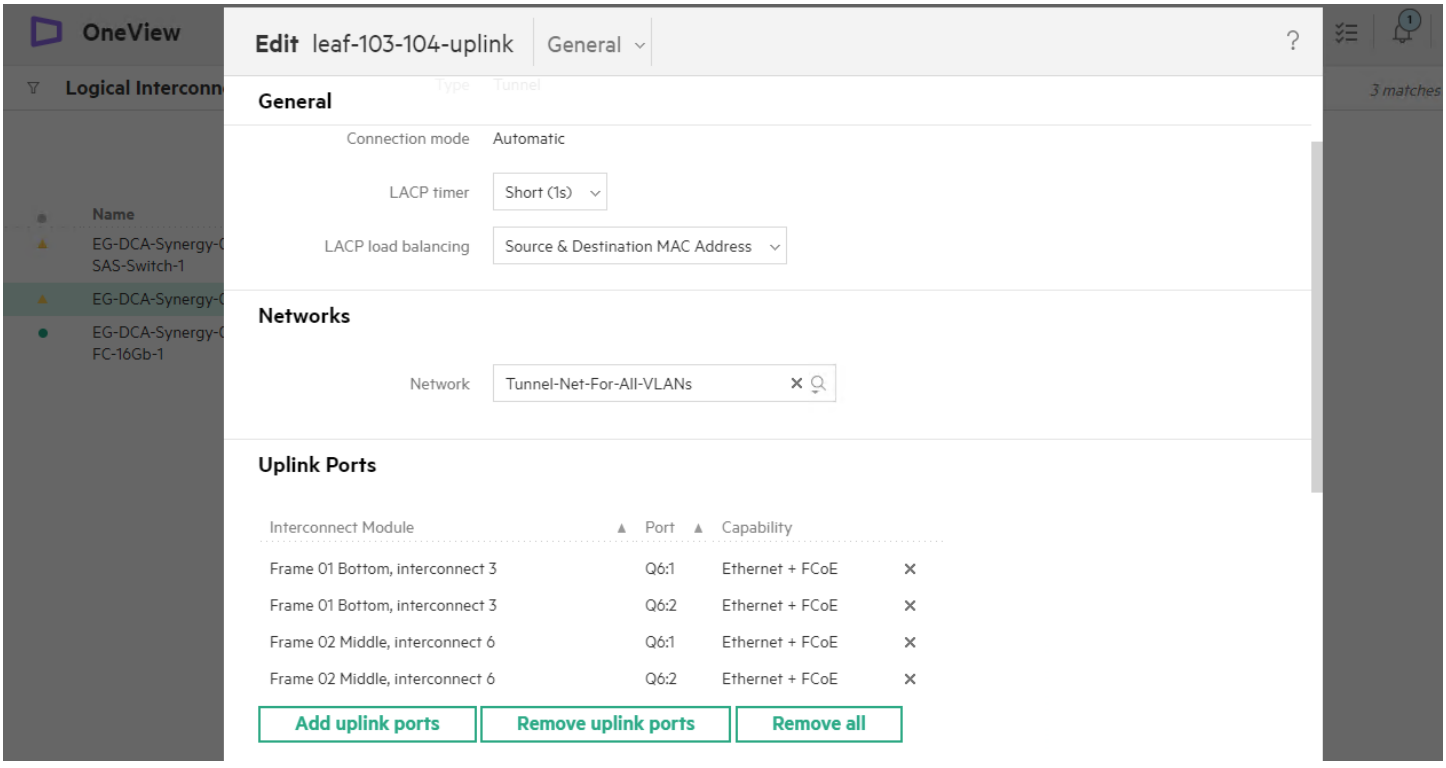
```
apic3# fabric 103 show vpc extended vpc 1
--------------------------------------------------------------------
 Node 103 (DCA-93180-01)
--------------------------------------------------------------------


vPC status
----------------------------------------------------------------------------
id   Port   Status Consistency Reason                Active vlans Bndl Grp Name
--   ----   ------ ----------- ------                ------------ ----------------
1    Po1    up     success     success               140,165-166, synergy-vpc
                                                     200-201,901,
                                                     2001,3001,40
                                                     00
```

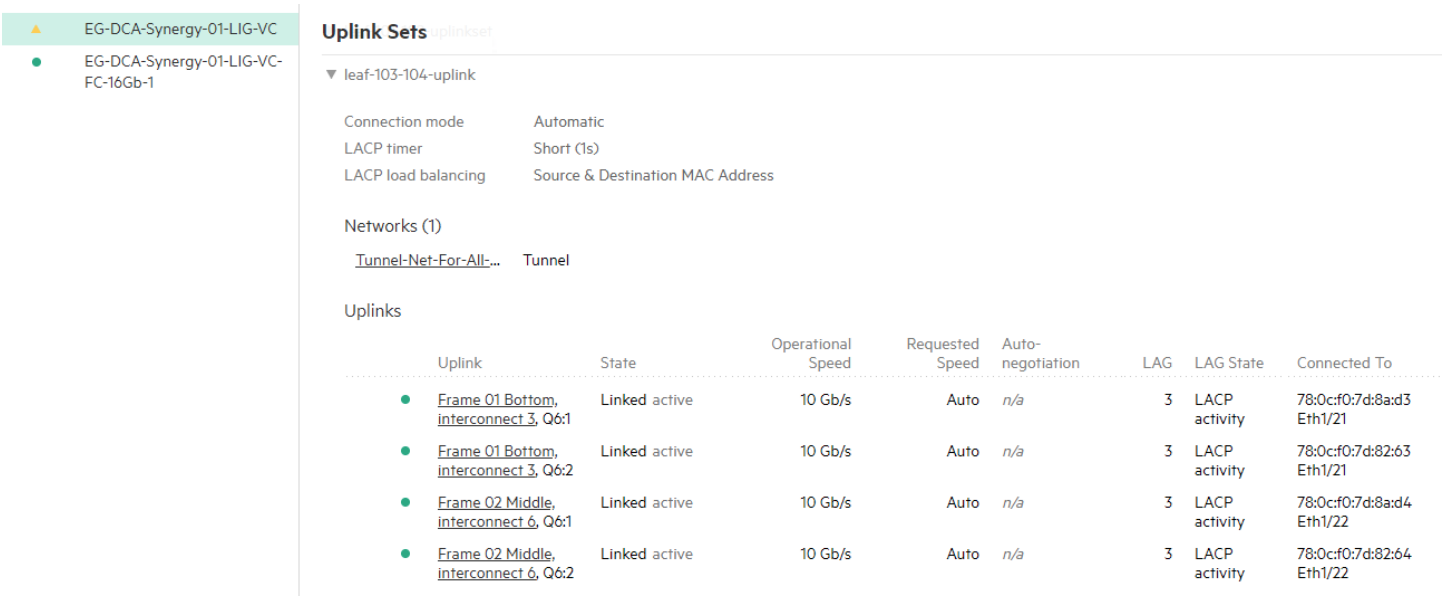


LACP Configuration

LACP configuration on ACI is under Fabric->Access Policies->Leaf Interfaces->Policy groups. Users can define a LACP Port Channel Policy under the policy group.

LACP configuration on Synergy is automatically done after users include physical ports under uplink set. An internal LAG is created for LACP negotiation.





LLDP Configuration

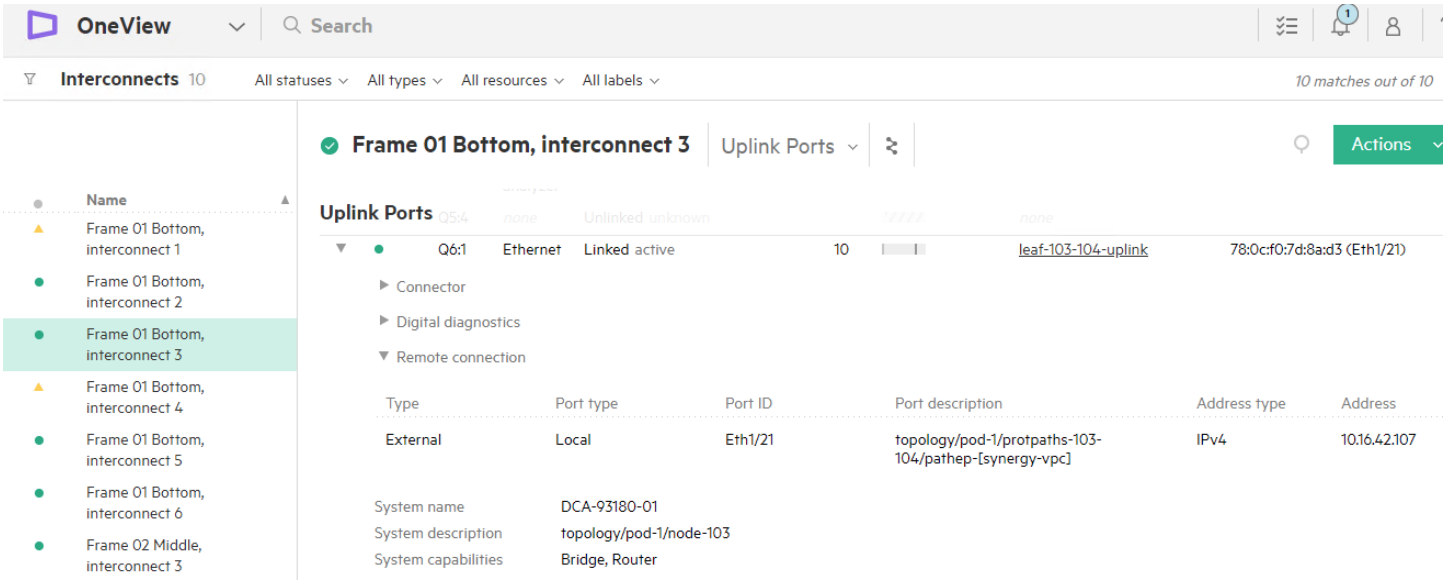
LLDP configuration on ACI is under Fabric->Access Policies->Leaf Interfaces->Policy groups. Users can define LLDP Policy under the policy group.

Synergy LLDP Uplinks Configuration

LLDP for Synergy uplinks are always enabled in Tx/Rx direction with no configuration needed. LLDP for Synergy internal downlinks to hypervisors can be optionally enabled. We'll cover it in ACI VMM integration section later.

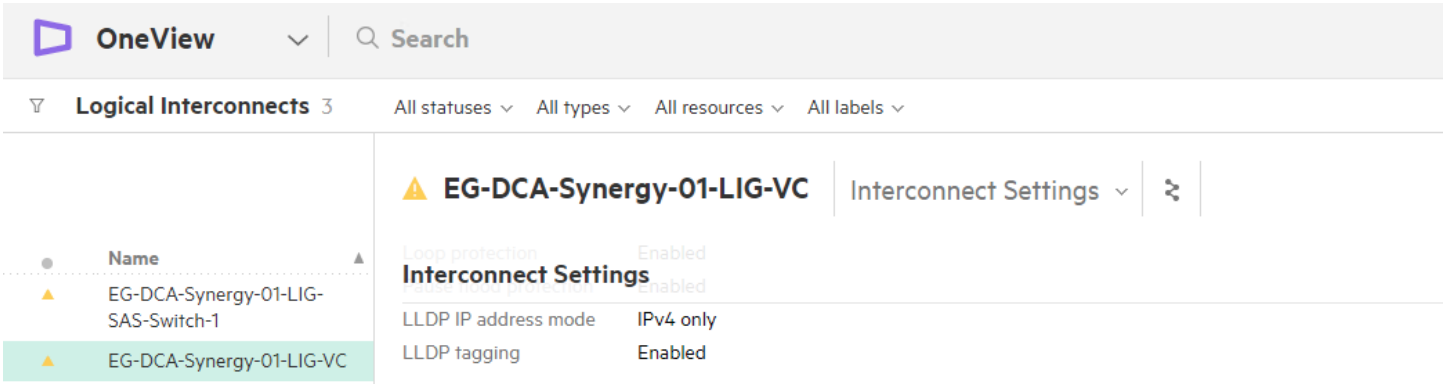
```
apic3# fabric 103 show lldp neigh int eth 1/21 detail
------------------------------------------------------------------
 Node 103 (DCA-93180-01)
------------------------------------------------------------------
Capability codes:
  (R) Router, (B) Bridge, (T) Telephone, (C) DOCSIS Cable Device
  (W) WLAN Access Point, (P) Repeater, (S) Station, (O) Other
Device ID           Local Intf      Hold-time  Capability  Port ID

Chassis id: 00fd.4551.b8a7
Port id: Ten-GigabitEthernet0/0/6:1
Local Port id: Eth1/21
Port Description: Ethernet Interface Port 123
System Name: not advertised
System Description: VC SE 40Gb F8 Module
Time remaining: 49 seconds
System Capabilities: not advertised
Enabled Capabilities: not advertised
Management Address: 10.16.43.118
Vlan ID: not advertised

Total entries displayed: 1
```

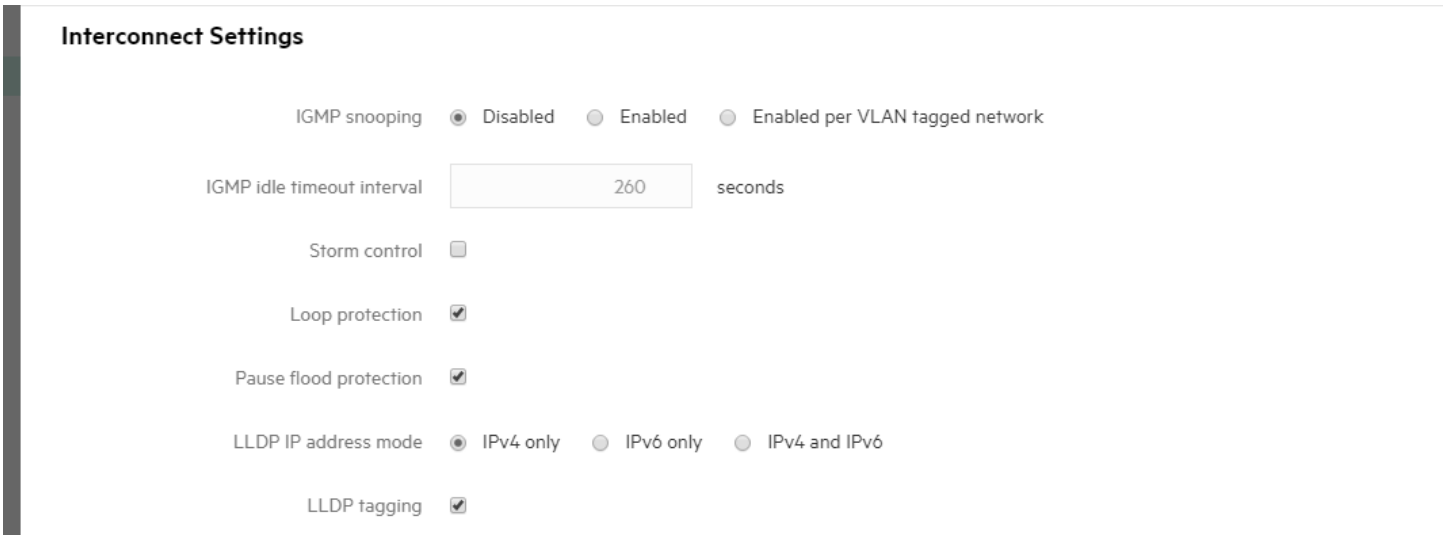
LLDP neighbor details on Synergy uplinks can be viewed under InterConnect->Uplink Ports.



Synergy LLDP Downlinks Configuration

LLDP for Synergy downlinks can be optionally enabled as “LLDP tagging” under Logical Interconnects. This is the direction facing internal compute nodes/hypervisor. For ACI doing static physical domains with servers/hypervisors, this setting doesn’t need to be enabled. However, if ACI is doing VMM integration with hypervisors with resolution immediacy set to “Immediate” or On Demand”(Other than Pre-Provision), then this setting has to be enabled so APIC can have end to end LLDP neighbor information among leaf nodes/Virtual Connect/Hypervisor hosts. This setting will be discussed later in more details in VMM integration section.





ACI Inter-VLAN Bridging scenario with Synergy Tunnel Network

ACI Inter-VLAN Bridging

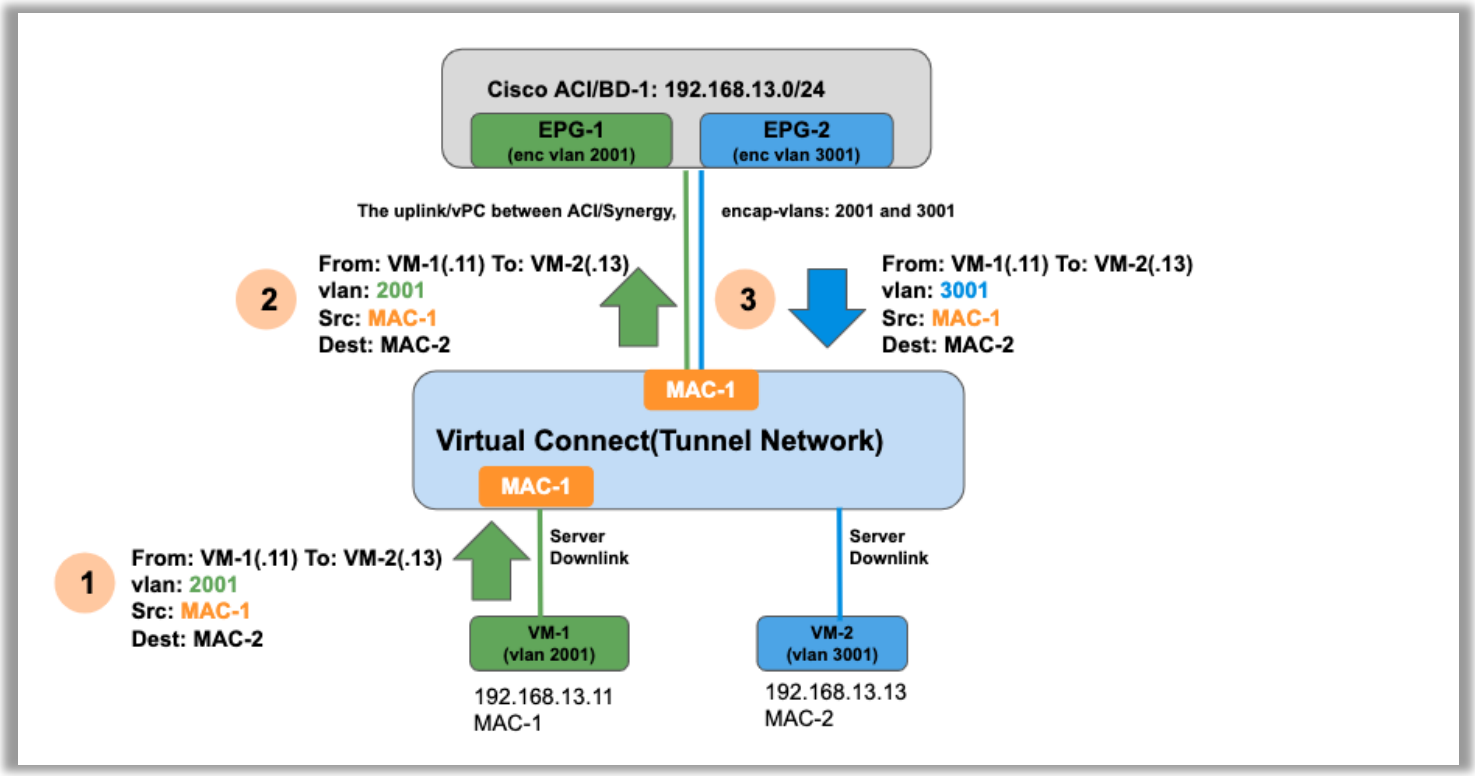
So far we have gone through ACI/Synergy interop concepts and configurations. In summary, Synergy mapped/tunnel networks can work well with ACI in network-centric and application-centric mode.

Synergy tunnel networks provides a unique and powerful feature to tunnel all ACI networks to hypervisors/servers. It greatly simplifies interop for ACI physical domain and VMM domain integration.

However, there is one ACI use case requiring user attention to have ACI/Synergy interop correctly and that is when ACI is configured as

- Multiple EPGs are under one BD
- These EPGs share the same IP subnet defined on BD

In the above case, ACI is essentially doing inter-VLAN bridging in this scenario. Let’s dig deeper to see what happens in this use case.



In above diagram, we have two EPGs(EPG-1 and EPG-2) under one BD. The two EPGs share the same IP subnet 192.168.13.0/24 and the subnet default GW is defined on ACI BD. The two EPGs use different encapsulation VLANs(VLAN 2001 and 3001) to connect Synergy and hypervisors inside it. On Synergy we have one single tunnel network defined to tunnel traffic between ACI leaf nodes and hypervisors/VMs.

In this use case, traffic across these two EPGs are neither L2 bridging nor L3 routing in a typical networking world, instead they are forwarded like inter-VLAN bridging traffic because traffic maintain their src/dst MAC address while across different VLANs.

This behavior can work well with Synergy Mapped network where Synergy has the knowledge on which user VLANs are travelling through as we already predefined all ACI encapsulation VLANs like 2001 and 3001 inside Synergy. However, in Tunnel network, it's by design not to look into user VLAN id when tunneling traffic. Synergy only learns L2 MAC address in tunnel network so it can correctly tunnel traffic to the right uplinks/downlinks and let ACI/Hypervisors handle VLAN tagging on their ends.

The above diagram simplifies some traffic forwarding/discovery details. It's assumed that users want VM-1 in EPG-1 communicate with VM-2 in EPG-2 and VMs ARP table and ACI endpoints table are populated already to show the confusion to Synergy when traffic is forwarded through.

In step 1, VM-1 192.168.13.11 sent a packet to VM-2 192.168.13.13. The traffic will come into Synergy tunnel network with VLAN id 2001 and src MAC MAC-1. Synergy will register VM-1 MAC address learned through its internal server downlink without looking into VLAN id 2001. So the only knowledge for Synergy for VM-1 MAC-1 is that MAC-1 is on internal server link.

In step 2, synergy tunneled this packet to ACI leaf nodes without changing any user packet field like a tunnel or pseudo-wire.

In Step 3, ACI forwarded the same packet from EPG-1 to EPG-2(assuming all ACI contracts allowed), the same packet will come in from ACI leaf node into the same Synergy uplink/vpc. The packet L3 IPs and L2 MACs remain the same. The only change is that ACI swapped encap VLAN from EPG-1 VLAN 2001 to EPG-2 VLAN 3001.

At this moment, Synergy tunnel network received a packet with src MAC as MAC-1, that will cause confusion for Synergy because Synergy learned the same MAC address from both internal server link and uplink. Even though the packets from internal link and uplink come in as different VLAN-id but Synergy tunnel network doesn't look into VLAN id.

When this happens, Synergy tunnel network can't successfully forward traffics for this MAC.

Solutions

There are several solutions to work with ACI inter-VLAN bridging with Synergy tunnel networks.

Option One:

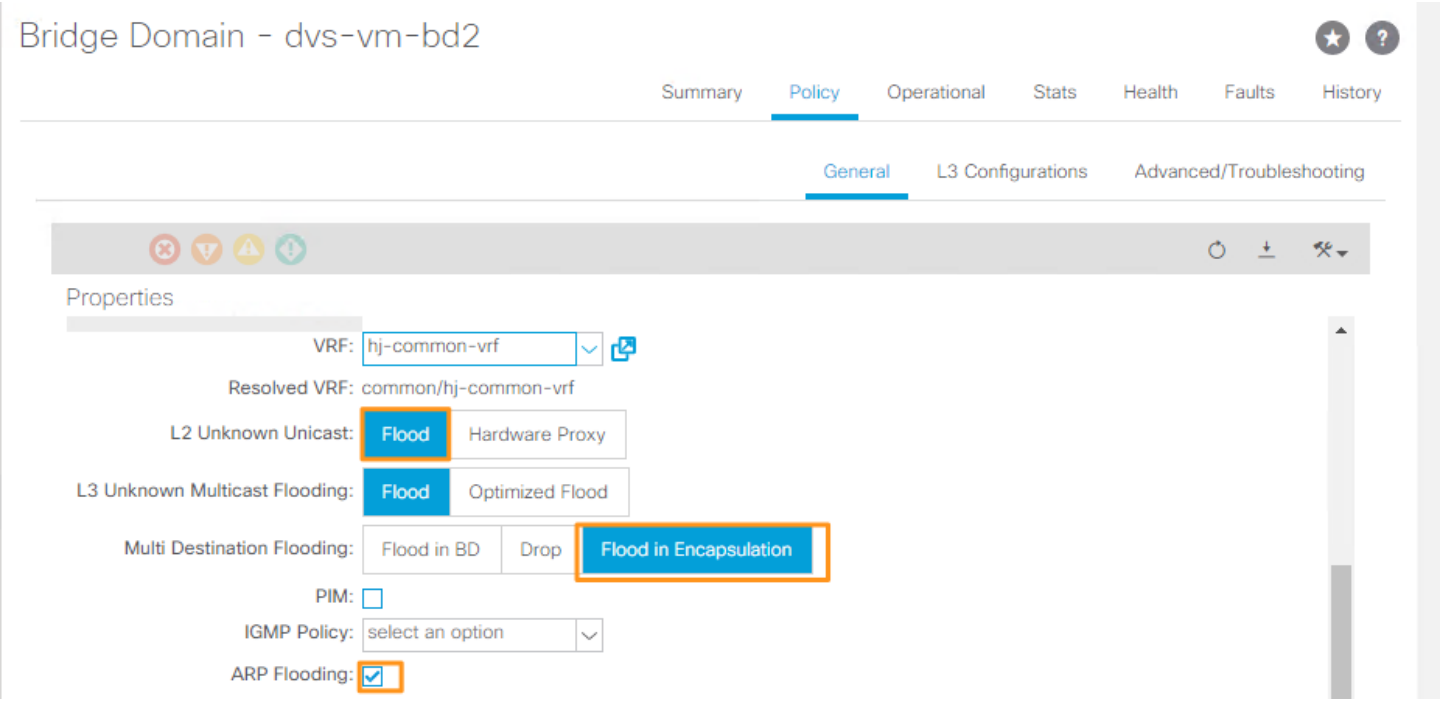
The first solution was introduced back in 2017 following the APIC 3.1(1) release.

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/3-x/release_notes/apic_rn_311.html

The feature is called "Flood in Encapsulation" under BD and it's specifically added to work with Synergy VC tunnel network.

Configuring flood in encapsulation for all protocols and proxy ARP across encapsulations	In this release, on the Cisco ACI switches with the Application Leaf Engine (ALE), all protocols are flooded in encapsulation. Multiple EPGs are now supported under one bridge domain with an external switch. When two EPGs share the same bridge domian and the **Flood in Encapsulation** option is turned on, the EPG flooding traffic does not reach the other EPG. It overcomes the challenges of using the Cisco ACI switches with the Virtual Connect (VC) tunnel network. For more information, see the *Cisco APIC Layer 2 Networking Configuration Guide*.	None.

The setting "Flood in Encapsulation" is under ACI GUI Bridge Domain section.



Note:

According to Cisco APIC Layer 2 Networking Configuration Guide, flood in encapsulation is only supported when BD is in flood mode so "L2 Unknown Unicast" will need be set to "Flood"(which will automatically enable ARP flooding)

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/2-x/L2_config/b_Cisco_APIC_Layer_2_Configuration_Guide/b_Cisco_APIC_Layer_2_Configuration_Guide_chapter_010.html#id_59068

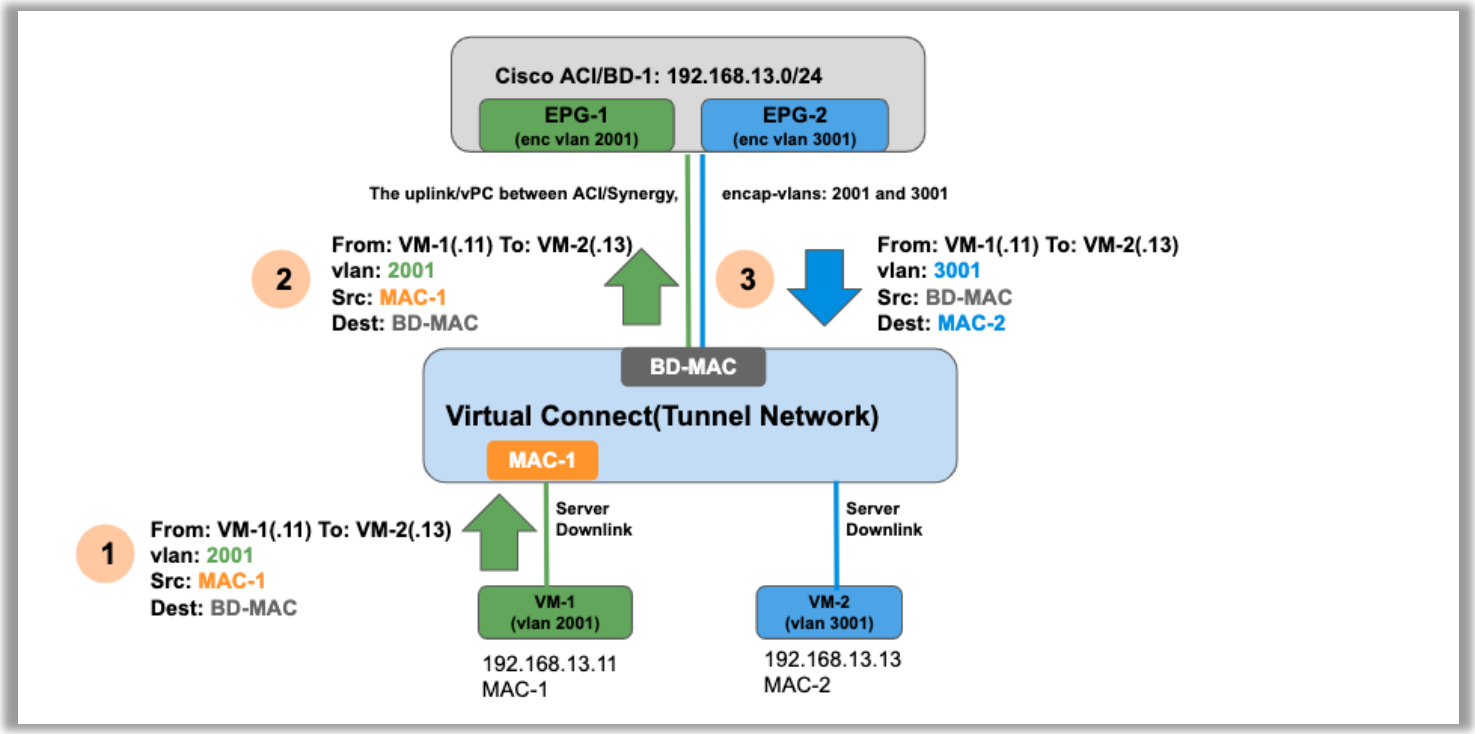
> - Flood in encapsulation is supported only in bridge domain in flood mode and ARP in flood mode. bridge domain spine proxy mode is not supported.

Behind the scene, the reason of "Flood in Encapsulation" enabling Synergy tunnel network to work in inter-VLAN bridging scenario is by implementing ACI "Proxy-ARP" feature. Proxy ARP within Cisco ACI fabric is different from traditional proxy ARP.

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_Cisco_APIC_and_ProxyARP.html
"…Proxy ARP in Cisco ACI enables endpoints within a network or subnet to communicate with other endpoints without knowing the real MAC address of the endpoints. Proxy ARP is aware of the location of the traffic destination, and offers its own MAC address as the final destination instead…"

Proxy ARP feature is also implemented in Cisco intra-EPG isolation and Micro-segmentation(uSeg) features.

In the same inter-VLAN bridging scenario discussed above, for inter-EPG traffic, Cisco ACI BD will offer its own MAC address to endpoints when it receives ARP inquiries so VM-1 will know VM-2 MAC as BD MAC and vice versa, VM-2 will know VM-1 MAC as BD MAC. ACI fabric knows the true endpoint MACs.



For the same traffic flow discussed previously:

In step 1, right now if VM-1 wants to send traffic to VM-2, it will send traffic from IP 192.168.13.11 to 192.168.13.13 with VLAN id 2001, src MAC as MAC-1 and dst MAC as BD-MAC because VM-1 learned VM-2 MAC as BD-MAC through its ARP inquiry.

In step 2, VC tunnel network will tunnel the this packet transparently to ACI leaf nodes.

In step 3, ACI will send this packet in EPG-2 with VLAN id 3001 and also swap src MAC address from MAC-1 to BD-MAC.

At this moment, VC tunnel mode will consistently learn MAC-1 from server downlink and BD-MAC from its uplink and be able to successfully tunnel packets back and forth between ACI and endpoints.

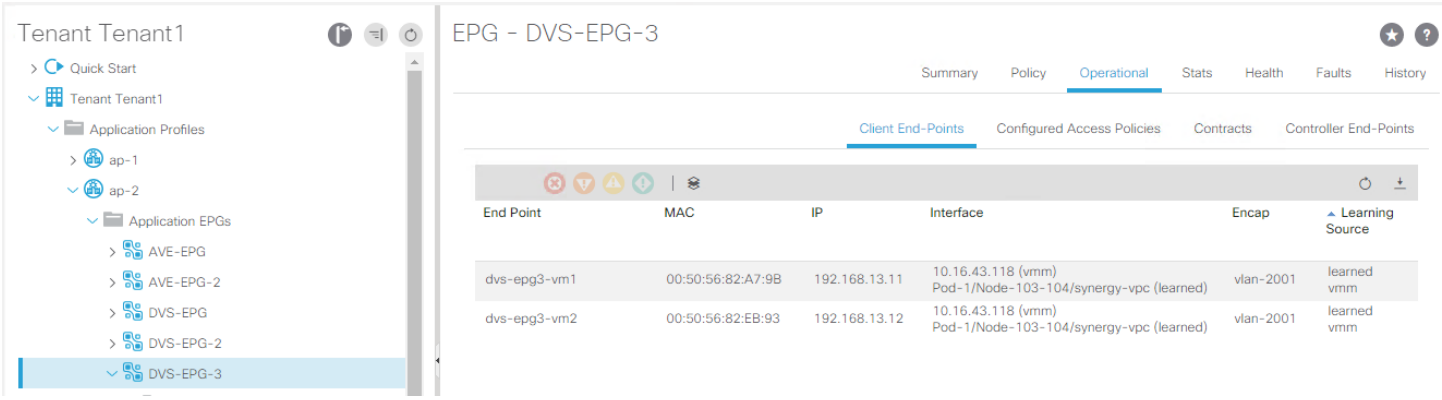
The capture below on VM-1 shows VM-1 learned VM-2 as BD-MAC 19:ff when ACI implements flood in encapsulation(proxy ARP). For the other VM in the same EPG as VM-1, the original VM MAC was learned.

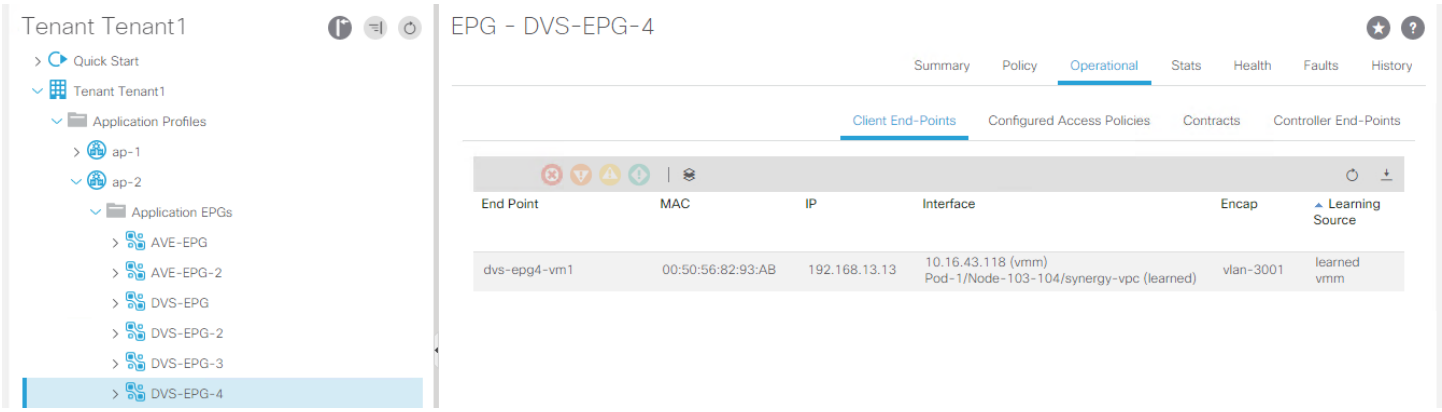
```
[root@dvs-epg3-vm1 ~]# ip addr show eno16780032 | egrep "ether|inet "
    link/ether 00:50:56:82:a7:9b brd ff:ff:ff:ff:ff:ff
    inet 192.168.13.11/24 brd 192.168.13.255 scope global eno16780032
[root@dvs-epg3-vm1 ~]# arp
Address                  HWtype  HWaddress           Flags Mask            Iface
192.168.13.1             ether   00:22:bd:f8:19:ff   C                     eno16780032
192.168.13.12            ether   00:50:56:82:eb:93   C                     eno16780032
192.168.13.13            ether   00:22:bd:f8:19:ff   C                     eno16780032
```

VM-1 arp tables shows inter-EPG VM-2(13.13) MAC as BD MAC 19:ff.

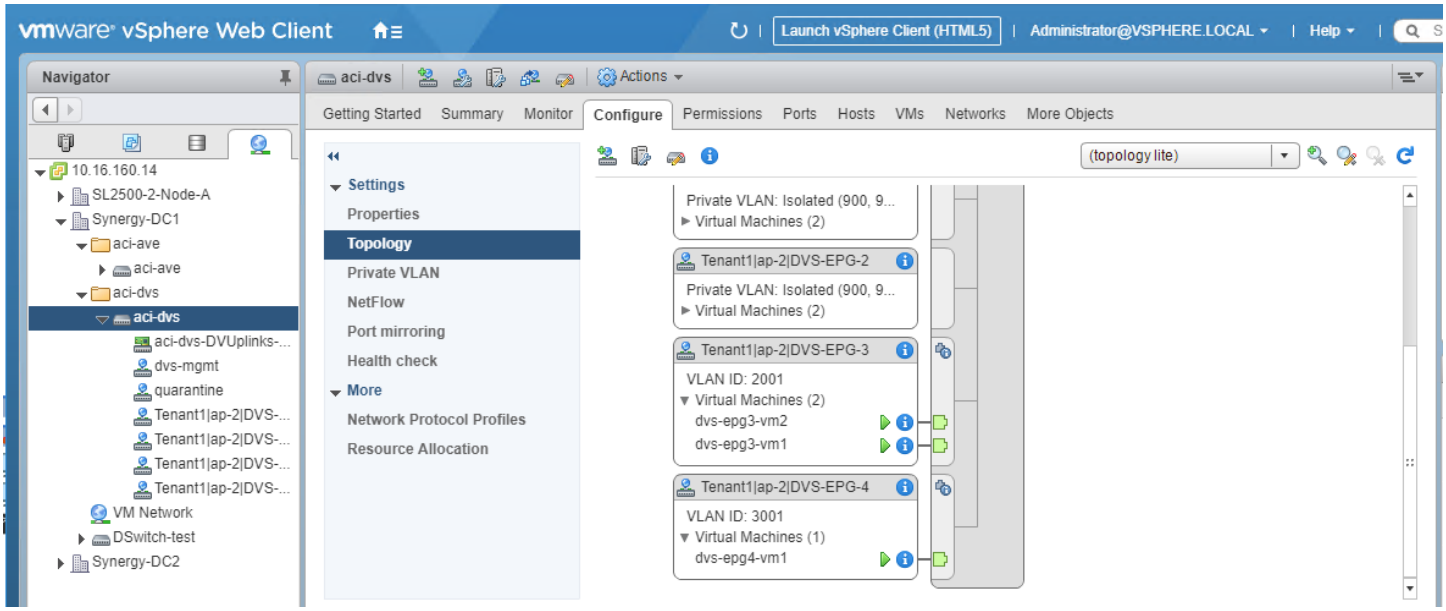
The other intra-EPG VM(13.12) has original VM MAC

ACI EPG endpoints table show true VM MACs on both EPGs





VMWare DVS shows VM-1(dvs-epg3-vm1) is in VMM port-group with VLAN id 2001 and VM-2(dvs-epg4-vm1) is in VMM port-group with VLAN id 3001.



VM-2 learned VM-1 MAC also as BD MAC 19:ff.

```
[root@dvs-epg4-vm1 ~]#  ip addr show eno16780032 | egrep "ether|inet "
    link/ether 00:50:56:82:93:ab brd ff:ff:ff:ff:ff:ff
    inet 192.168.13.13/24 brd 192.168.13.255 scope global eno16780032
[root@dvs-epg4-vm1 ~]# arp 192.168.13.11
Address                  HWtype  HWaddress           Flags Mask            Iface
192.168.13.11            ether   00:22:bd:f8:19:ff   C                     eno16780032
[root@dvs-epg4-vm1 ~]# arp 192.168.13.1
Address                  HWtype  HWaddress           Flags Mask            Iface
192.168.13.1             ether   00:22:bd:f8:19:ff   C                     eno16780032
```

VM-1 can successfully communicate inter-EPG with 13.13, intra-EPG with 13.12 and internet.

```
[root@dvs-epg3-vm1 ~]# ping 192.168.13.13
PING 192.168.13.13 (192.168.13.13) 56(84) bytes of data.
64 bytes from 192.168.13.13: icmp_seq=1 ttl=63 time=1.77 ms
64 bytes from 192.168.13.13: icmp_seq=2 ttl=63 time=0.423 ms
64 bytes from 192.168.13.13: icmp_seq=3 ttl=63 time=0.341 ms
64 bytes from 192.168.13.13: icmp_seq=4 ttl=63 time=0.387 ms
^C
--- 192.168.13.13 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 3002ms
rtt min/avg/max/mdev = 0.341/0.731/1.776/0.604 ms
[root@dvs-epg3-vm1 ~]# ping 192.168.13.12
PING 192.168.13.12 (192.168.13.12) 56(84) bytes of data.
64 bytes from 192.168.13.12: icmp_seq=1 ttl=64 time=0.335 ms
64 bytes from 192.168.13.12: icmp_seq=2 ttl=64 time=0.334 ms
64 bytes from 192.168.13.12: icmp_seq=3 ttl=64 time=0.391 ms
64 bytes from 192.168.13.12: icmp_seq=4 ttl=64 time=0.338 ms
^C
--- 192.168.13.12 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 2999ms
rtt min/avg/max/mdev = 0.334/0.349/0.391/0.030 ms
[root@dvs-epg3-vm1 ~]# ping www.google.com
PING www.google.com (172.217.9.4) 56(84) bytes of data.
64 bytes from dfw28s02-in-f4.1e100.net (172.217.9.4): icmp_seq=1 ttl=50 time=6.92 ms
64 bytes from dfw28s02-in-f4.1e100.net (172.217.9.4): icmp_seq=2 ttl=50 time=6.99 ms
64 bytes from dfw28s02-in-f4.1e100.net (172.217.9.4): icmp_seq=3 ttl=50 time=7.03 ms
64 bytes from dfw28s02-in-f4.1e100.net (172.217.9.4): icmp_seq=4 ttl=50 time=7.00 ms
64 bytes from dfw28s02-in-f4.1e100.net (172.217.9.4): icmp_seq=5 ttl=50 time=6.94 ms
^C
```

Option Two:

As previous mentioned, Proxy ARP is also enabled behind the scene in APIC features like micro-segmentation(uSeg) and intra-EPG to ensure traffics are forwarded to ACI from hosts to ensure ACI security policy. In these scenarios, users don't need to configure "flood in encapsulation" in BD.

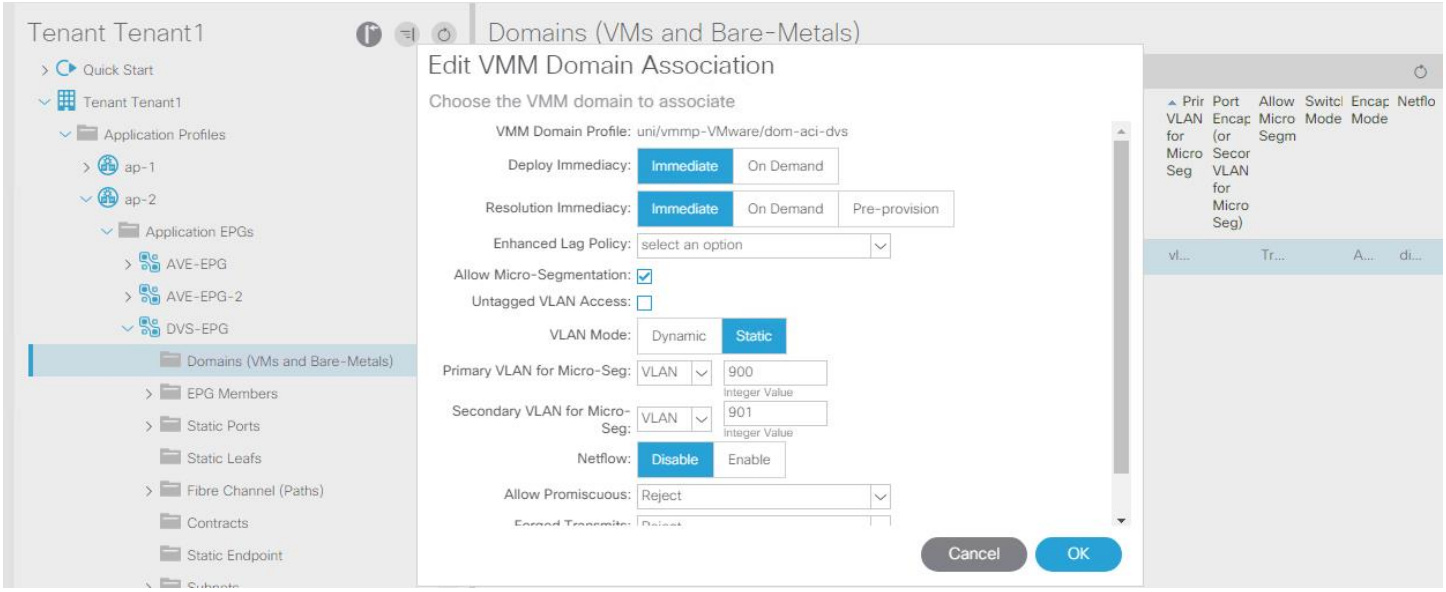
Proxy-ARP in ACI Micro-Segmentation

In the following example, two EPGs DVS-EPG and DVS-EPG-2 have micro-Segmentation enabled in their VMM association.

Note:

No uSeg EPGs were created. Only allow micro-segmentation was checked in VMM to enable ACI Proxy-ARP.

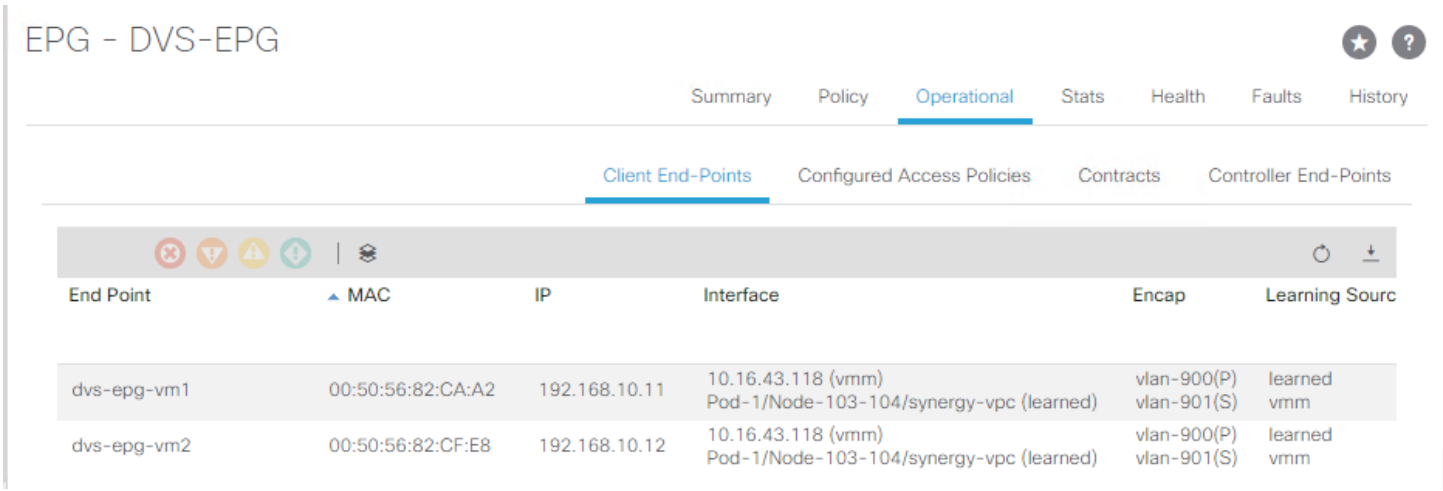
In this case, static VLAN pairs are used as PVLAN pairs. The other option is to use dynamic VLAN mode so ACI will pick dynamic VLAN pairs from its VLAN pool and provisioned on VMware DVS.



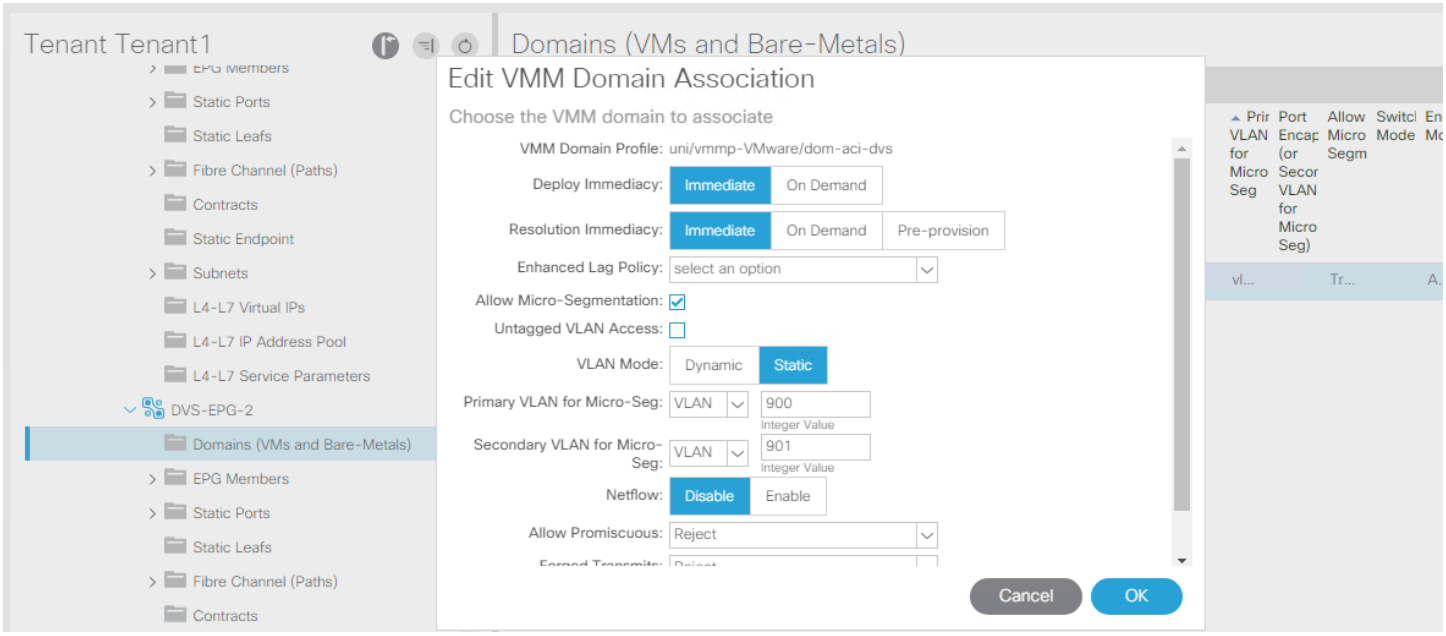
Note:

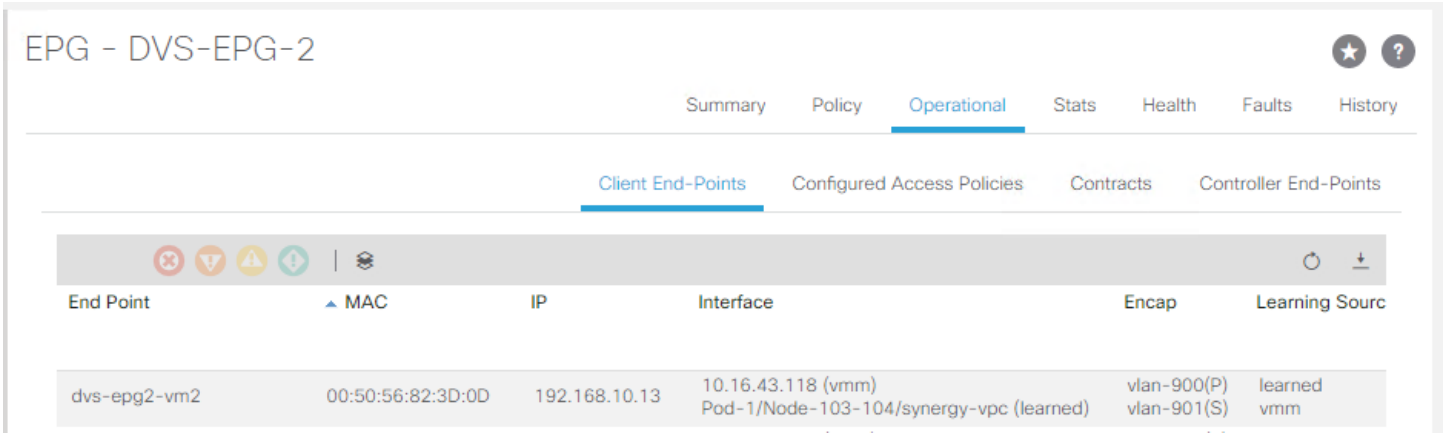
According to Cisco documentation, VMM Deploy Immediacy and Resolution Immediacy have to set as "Immediate" to enable ACI micro segmentation.

DVS-EPG is aware of its endpoints real MAC address.

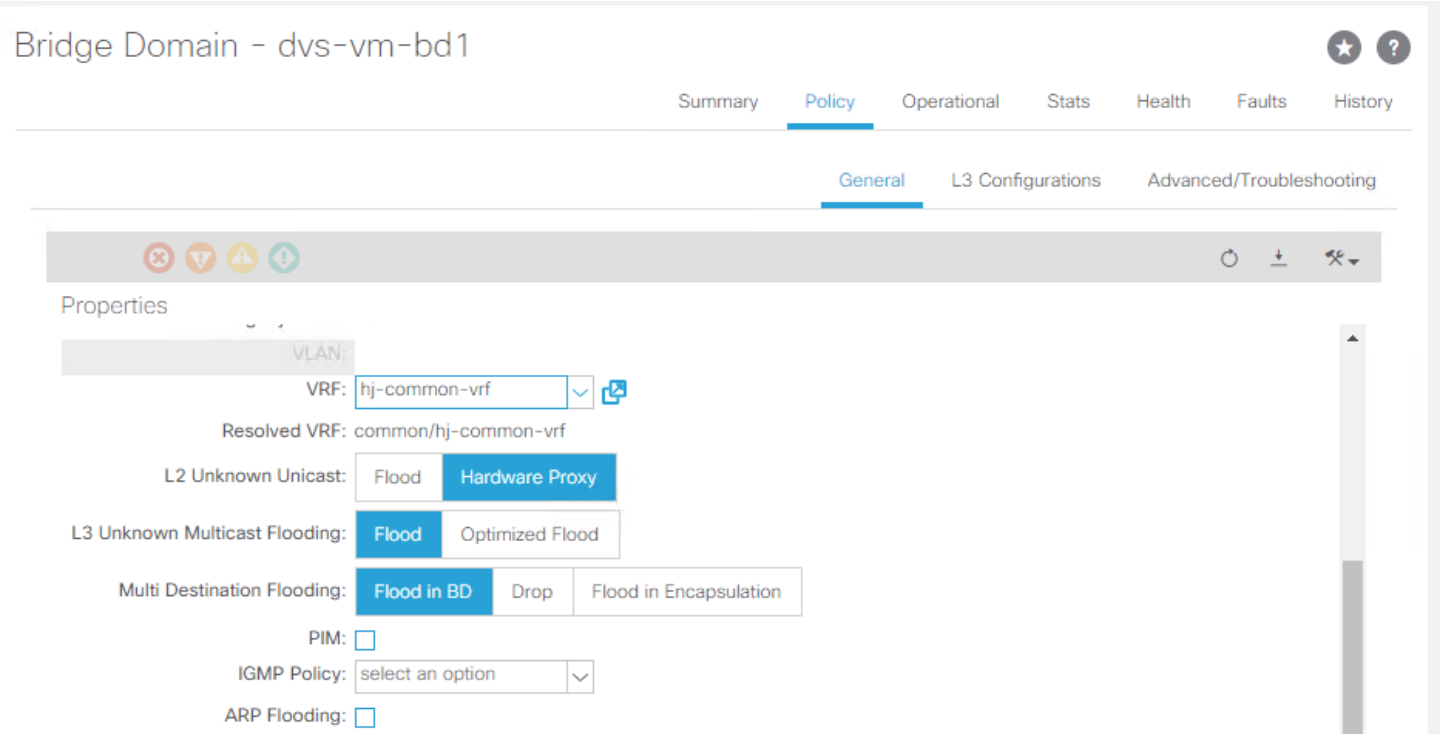


DVS-EPG-2 is also configured as micro-segmentation and aware of its endpoint real MAC address.

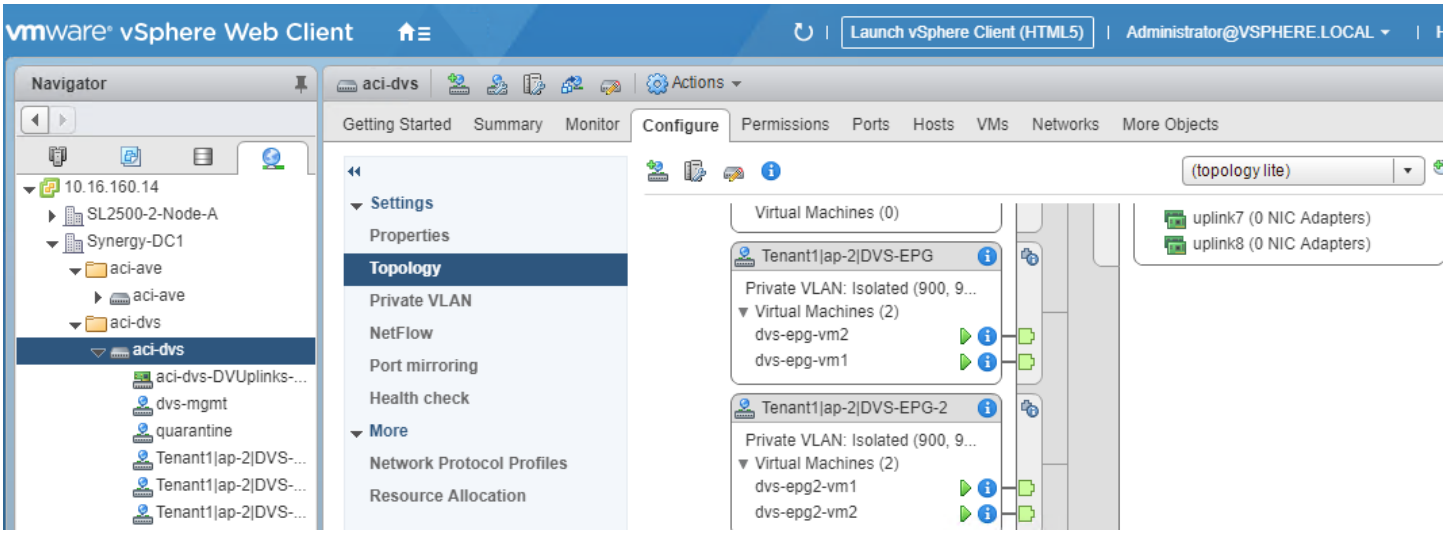


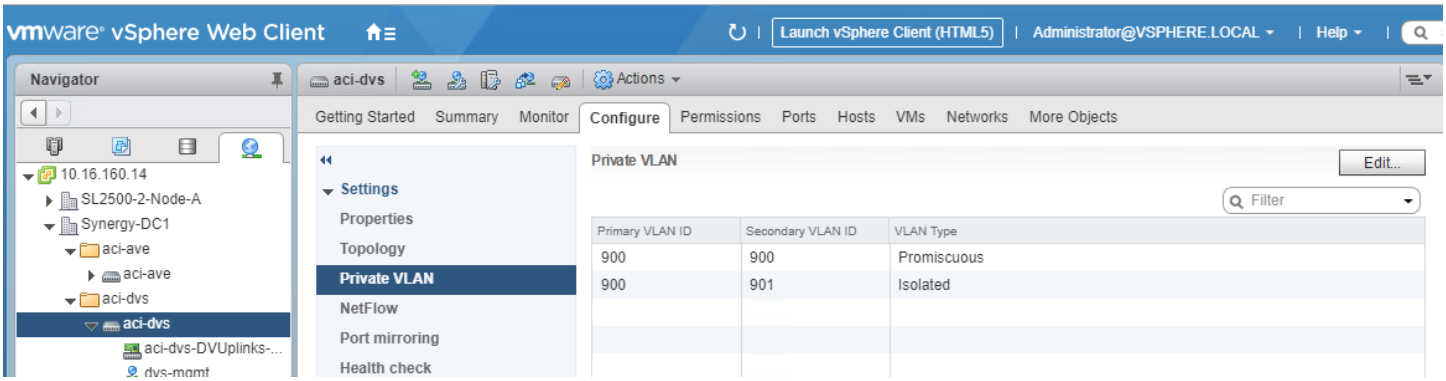


The BD two EPGs belong to is configured as default HW-proxy mode.



VMWare DVS port groups are provisioned with static VLAN pairs for port group private VLAN implementation.





EPG VM only knows other VMs by BD MAC address so all traffic to other VMs intra or inter EPGs will be sent to ACI BD to make sure centralized mirco-segmentation security policy is enforced.

```
[root@dvs-epg-vm1 ~]# arp
Address                  HWtype  HWaddress           Flags Mask            Iface
192.168.10.1             ether   00:22:bd:f8:19:ff   C                     eno16780032
192.168.10.13            ether   00:22:bd:f8:19:ff   C                     eno16780032
192.168.10.12            ether   00:22:bd:f8:19:ff   C                     eno16780032
```

```
[root@dvs-epg-vm1 ~]# ping 192.168.10.12
PING 192.168.10.12 (192.168.10.12) 56(84) bytes of data.
64 bytes from 192.168.10.12: icmp_seq=1 ttl=63 time=0.416 ms
64 bytes from 192.168.10.12: icmp_seq=2 ttl=63 time=0.308 ms
64 bytes from 192.168.10.12: icmp_seq=3 ttl=63 time=0.432 ms
^C
--- 192.168.10.12 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 1999ms
rtt min/avg/max/mdev = 0.308/0.385/0.432/0.057 ms
[root@dvs-epg-vm1 ~]# ping 192.168.10.13
PING 192.168.10.13 (192.168.10.13) 56(84) bytes of data.
64 bytes from 192.168.10.13: icmp_seq=1 ttl=63 time=0.377 ms
64 bytes from 192.168.10.13: icmp_seq=2 ttl=63 time=0.405 ms
64 bytes from 192.168.10.13: icmp_seq=3 ttl=63 time=0.399 ms
^C
--- 192.168.10.13 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 1999ms
rtt min/avg/max/mdev = 0.377/0.393/0.405/0.025 ms
[root@dvs-epg-vm1 ~]# ping 8.8.8.8
PING 8.8.8.8 (8.8.8.8) 56(84) bytes of data.
64 bytes from 8.8.8.8: icmp_seq=1 ttl=50 time=6.96 ms
64 bytes from 8.8.8.8: icmp_seq=2 ttl=50 time=6.88 ms
64 bytes from 8.8.8.8: icmp_seq=3 ttl=50 time=6.96 ms
^C
--- 8.8.8.8 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2003ms
rtt min/avg/max/mdev = 6.887/6.937/6.963/0.102 ms
```

Option Three

The other option besides leveraging ACI Proxy-ARP is to use ACI AVE hypervisor modules. The AVE will communicate with ACI leaf nodes using a single ACI infra-VLAN carrying VXLAN traffic. AVE integration will be discussed in the section later of this white paper.

ACI VMM with Synergy

ACI VMM integration with Synergy doesn't require any extra configuration on ACI and VMware side. Users can configure ACI VMM domain and associate EPGs to VMM as standard procedure.

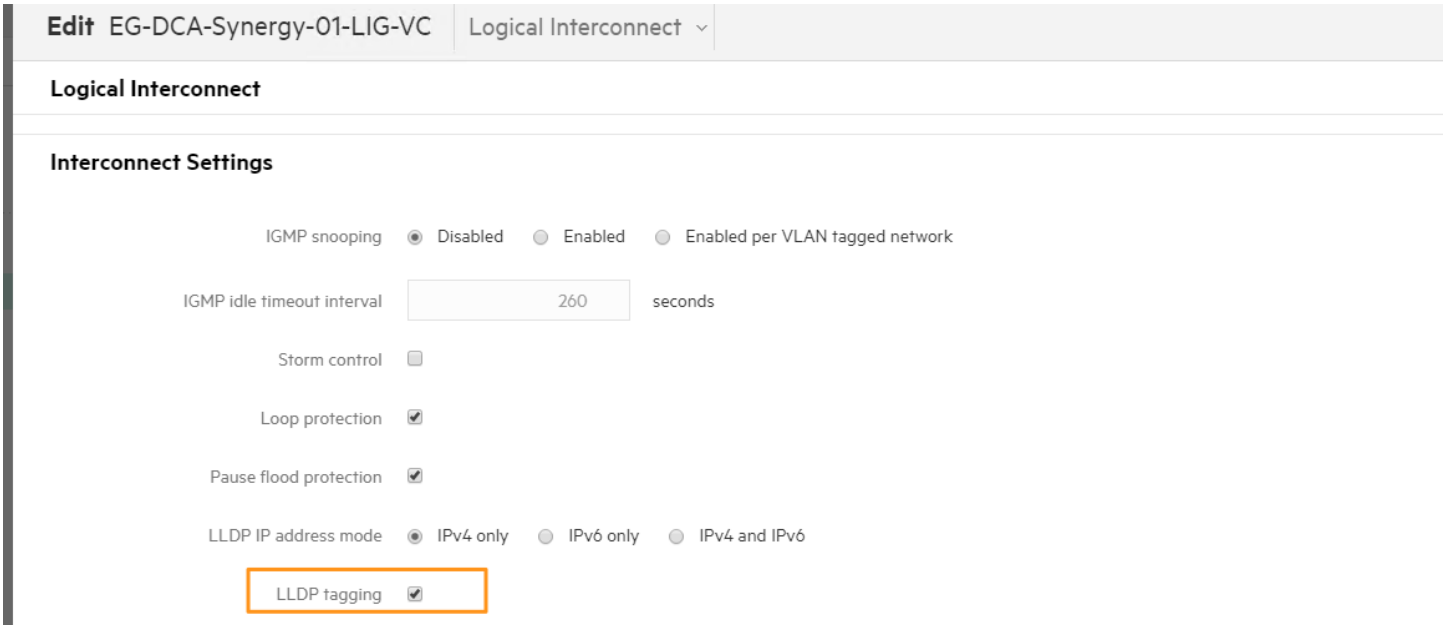
On Synergy side, If mapped networks and uplink set are used, then admins need to make sure that Synergy networks match what's being configured on ACI and VMware DVS side. If tunnel network and uplink are used, a single tunnel network can be applied to uplink and downlink to tunnel traffic through. The configuration details have been discussed in Synergy Virtual Connect Networking section.

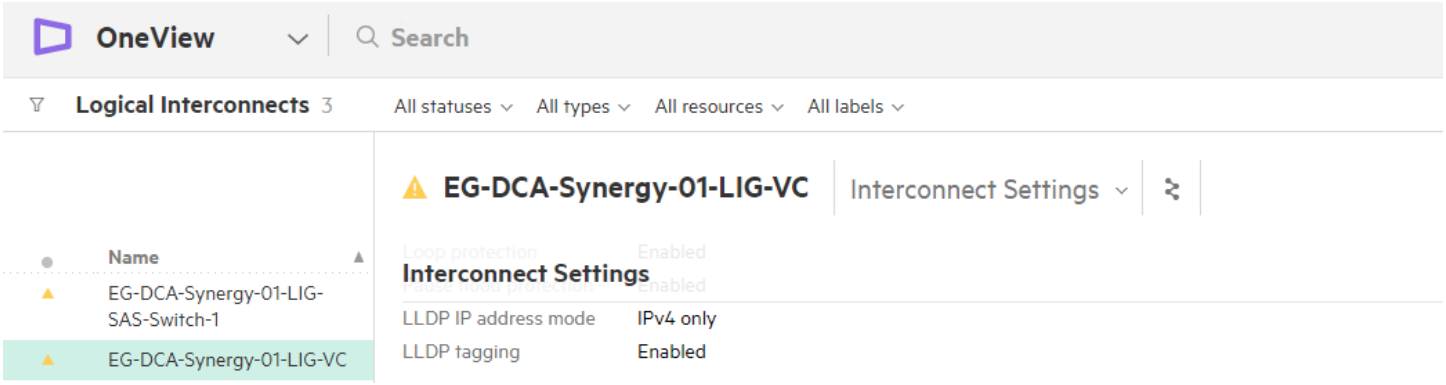
For ACI VMM integration, unless VMM "Resolution Immediacy" is set to "Pre-provision", then users need to enable Synergy downlink LLDP to ensure ACI has an end-to-end LLDP view in order to push policy from APIC to leaf nodes.

Note:

Synergy LLDP on uplink towards ACI is always enabled bidirectionally. No users configuration are needed.

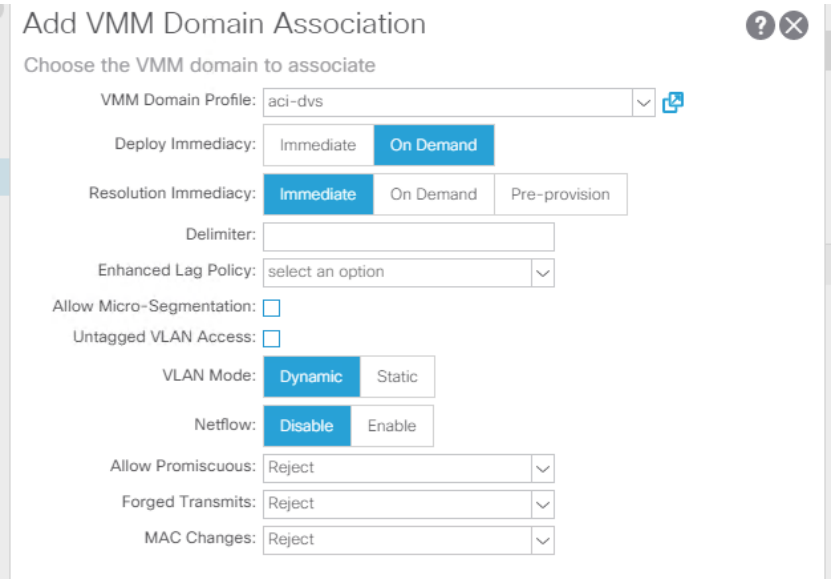
Synergy downlink LLDP feature can be enabled under Logical Interconnect section.





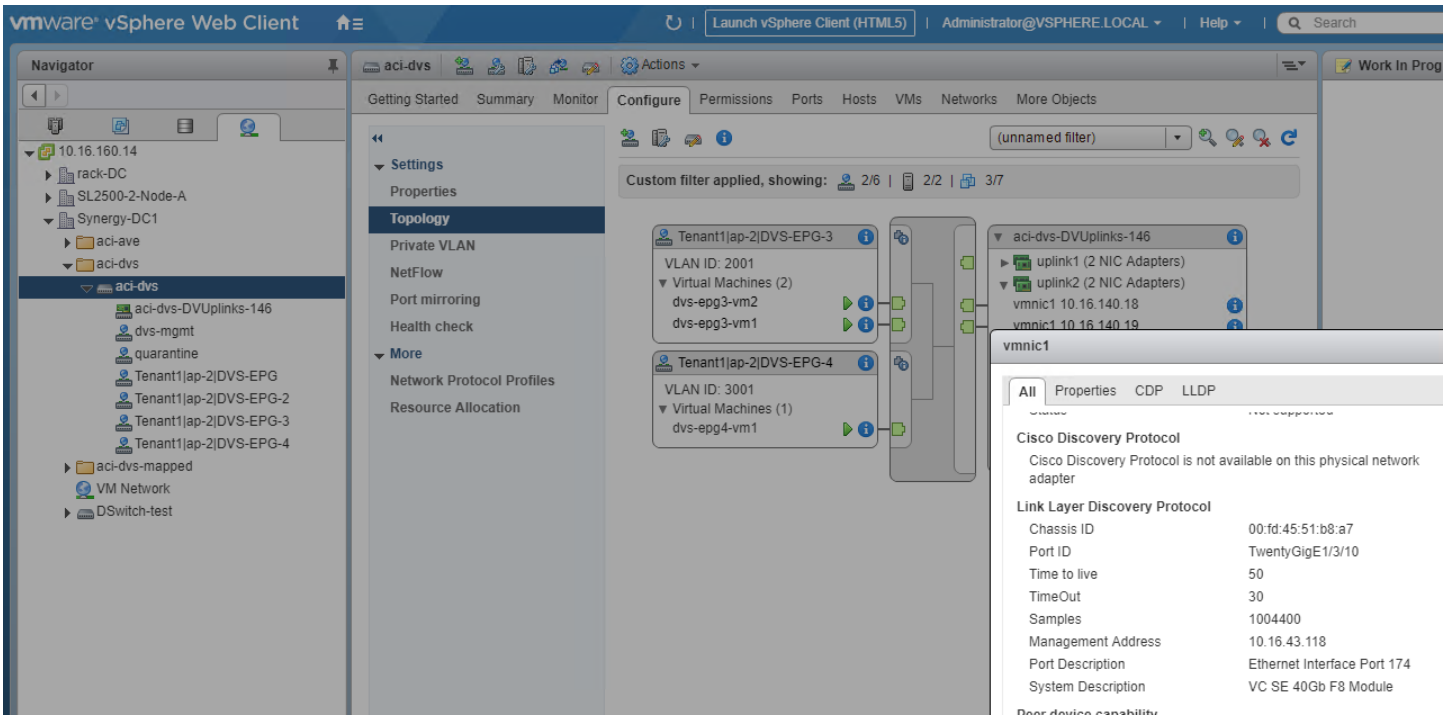
The following captures are from APIC GUI side when creating a new EPG VMM association. "Deploy Immediacy" is between APIC and ACI leaf notes and not related or affected by any Synergy setting.

"Resolution Immediacy" can be set as "Immediate", "On Demand" and "Pre-Provision". If set as former two settings, then ACI will depend on its view of LLDP neighbor to determine which leaf nodes to push APIC policy. For these two settings, Synergy LLDP downlink feature should be enabled.



Deploy Immediacy	Once policies are downloaded to the leaf software, deployment immediacy can specify when the policy is pushed into the hardware policy CAM. The options are: • **Immediate**—Specifies that the policy is programmed in the hardware policy CAM as soon as the policy is downloaded in the leaf software. • **On Demand**—Specifies that the policy is programmed in the hardware policy CAM only when the first packet is received through the data path. This process helps to optimize the hardware space. The default is **On Demand**.
Resolution Immediacy	Specifies whether policies are resolved immediately or when needed. The options are: • **Immediate**—Specifies that EPG policies (including contracts and filters) are downloaded to the associated leaf switch software upon hypervisor attachment to the VMware vSphere Distributed Switch (VDS). LLDP or OpFlex permissions are used to resolve the hypervisor to leaf node attachments. • **On Demand**—Specifies that a policy (for example, VLAN, VXLAN bindings, contracts, or filters) is pushed to the leaf node only when a hypervisor is attached to VDS and a VM is placed in the port group (EPG). • **Pre-provision**—Specifies that a policy (for example, VLAN, VXLAN binding, contracts, or filters) is downloaded to a leaf switch even before a hypervisor is attached to the VDS, thereby pre-provisioning the configuration on the switch.

The following VMware DVS capture shows ACI VMM provisioned two port-groups with VLAN 2001 and 3001. Synergy tunnel configuration will automatically include these two VLANs on synergy downlink and uplink side. Synergy mapped configuration need to make sure to provision these two VLANs on server profile and uplink set to forward traffic correctly to ACI and VMware.

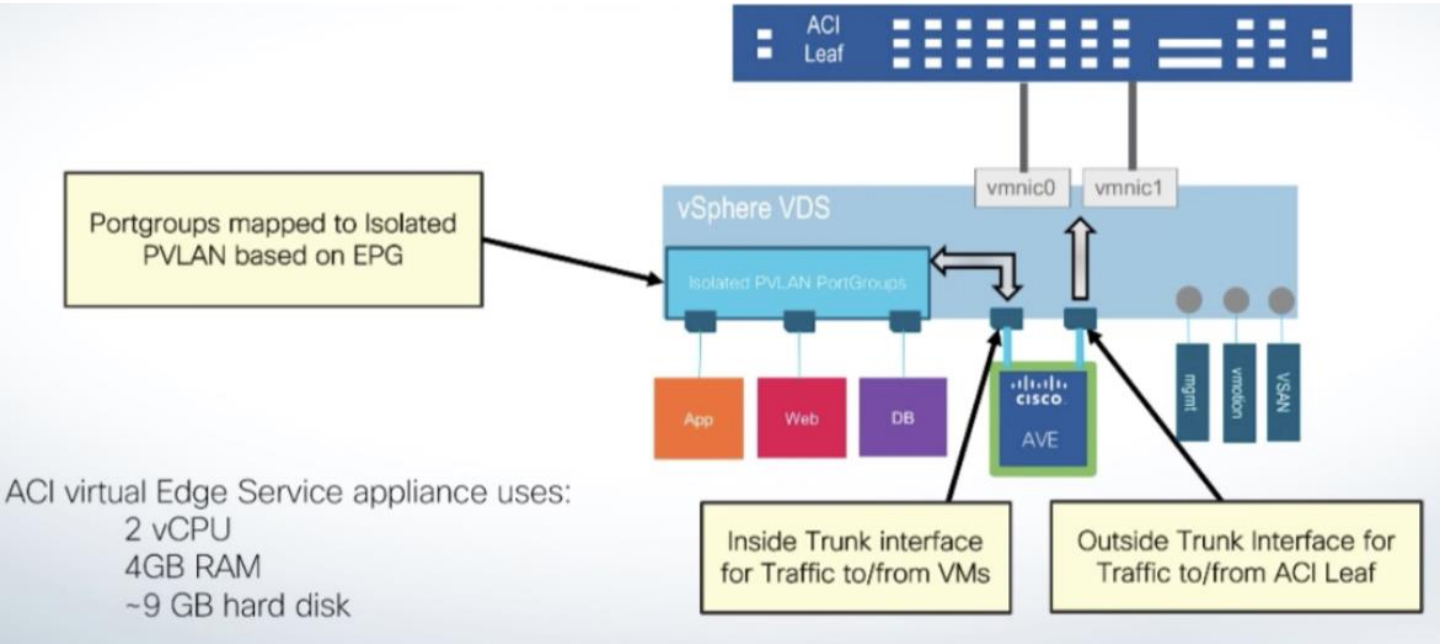


ACI AVE with Synergy

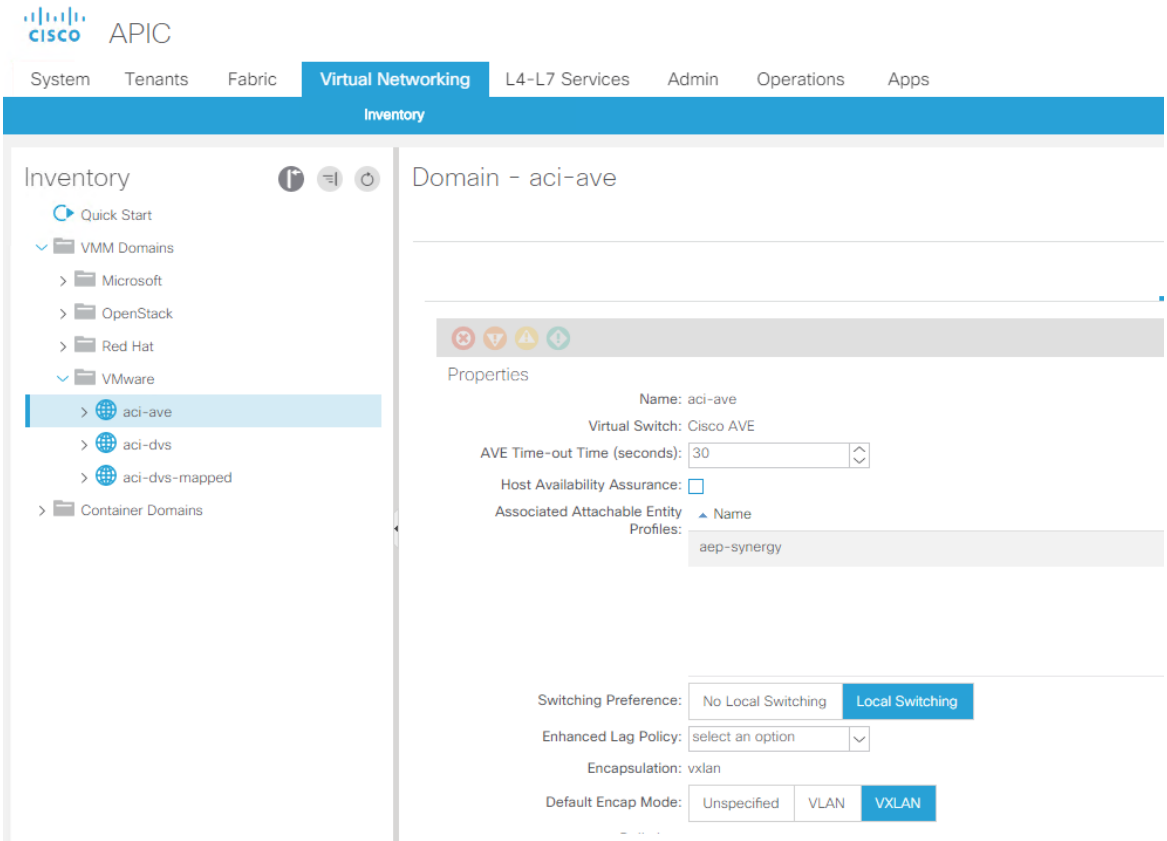
Cisco ACI Virtual Edge is a distributed virtual node that resides in the hypervisor user space of a virtualized host, providing switching and policy enforcement.

AVE deployment in Synergy requires no extra configuration from Synergy tunnel or mapped networks. A typical AVE deployment uses VXLAN to carry traffic between AVEs and ACI leaf nodes. The VXLAN traffic is encapsulated using a single ACI infra VLAN through Synergy. Therefore, users only need to config a single mapped or tunnel network to allow this infra VLAN traffic through Synergy. All underlying VXLAN operation is transparent to Synergy and that's the design principle for AVE so that its deployment is independent of any intermediate devices like Synergy Virtual Connect module.

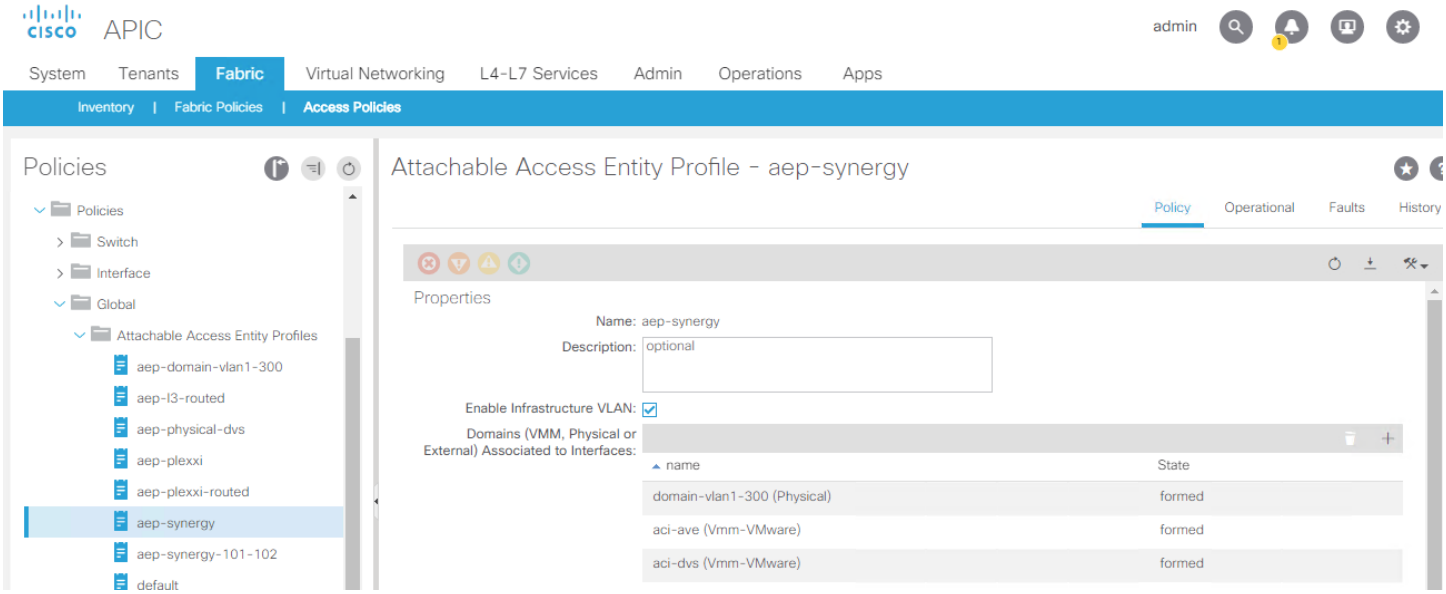
The following AVE architecture diagram is from Cisco publication. It shows EPG VMs are connected to AVE internal VMs and the AVE VM uses the DVS vmnics to pass the traffic to upstream devices like Synergy Virtual Connect modules.



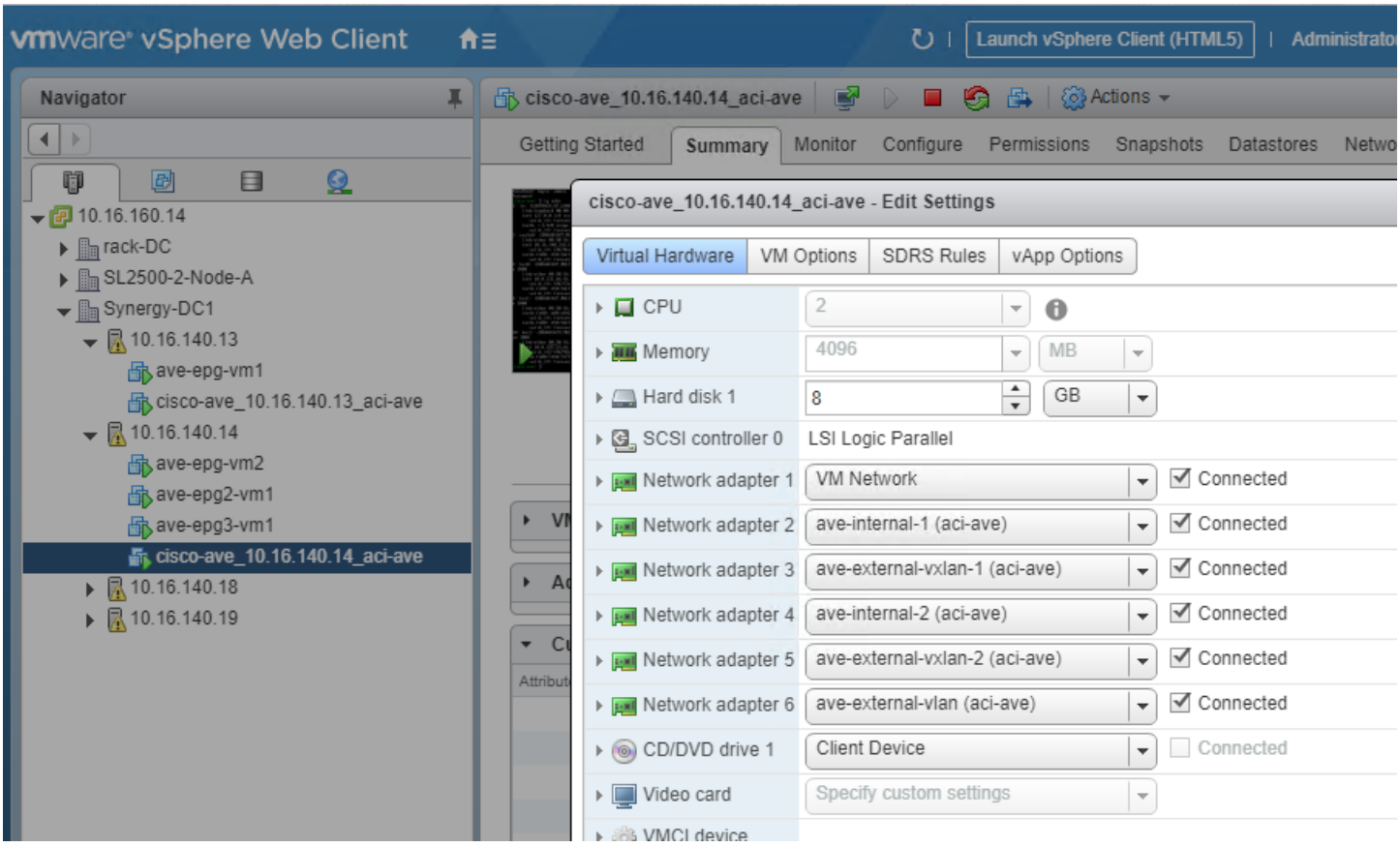
ACI AVE configuration includes creating a VMM domain from Cisco ACI GUI with defined encapsulation mode.



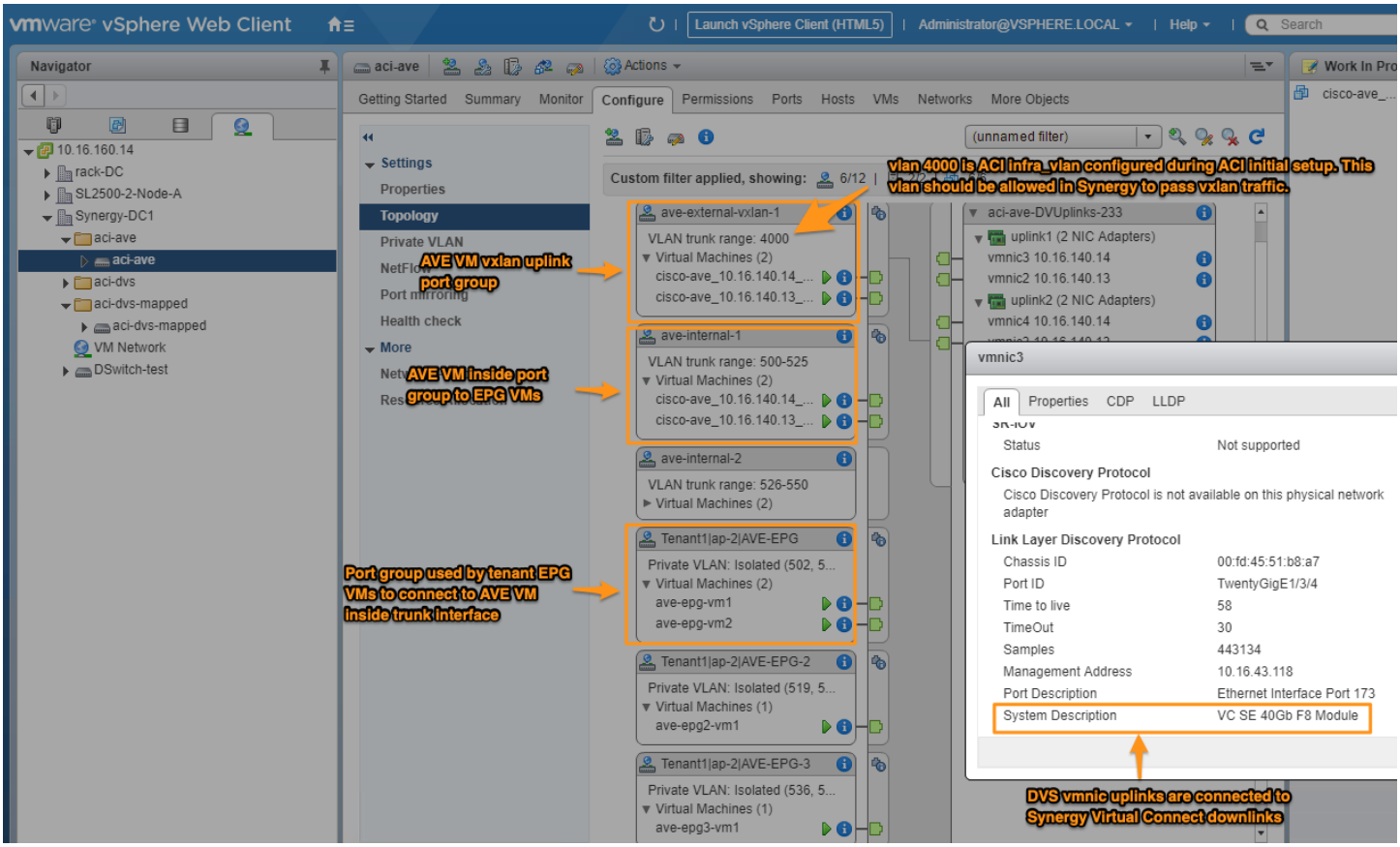
ACI Fabric Access AEP policy should allow ACI “Infrastructure VLAN” and this VLAN, which is defined in initial ACI provisioned is the only VLAN needs to be allowed on Synergy Virtual Connect by either tunnel or mapped mode.



After AVE VM installation on ESXi hosts, the VM will show multiple NICs.



The following DVS capture demonstrate how AVE VM communicates with uplinks and EPG VMs.



All EPG VMs use PVLANs to communicate with AVE VM and it will use FTEP and VTEP address as VXLAN tunnel UDP IP destination and source address to forward traffic to ACI physical leaf nodes.

```
cisco-ave:~$ vemcmd show sod
data-version 2.0
control-protocol open-flex
open-flex port 8000
open-flex ipaddr 10.0.0.30
ftep ipaddr 10.0.0.32
dvs-name comp/prov-VMware/ctrlr-[aci-ave]-hj-vcenter/sw-dvs-233
profile infra encap vlan 4000
profile infra encap pvlan 0 0
profile infra alias dvportgroup-242
profile ave-ctrl encap  vxlan 8912896
profile ave-ctrl mcast-ip 224.1.1.2
profile uplink port-channel mac-pinning
profile uplink mtu 9000
cisco-ave:~$ vemcmd show vxlan vtep

        VTEP info table
=======================================
        IP     LTL  MAC
=======================================
10.0.232.65     51   00:50:56:82:45:53
10.0.232.68     53   00:50:56:82:10:88
cisco-ave:~$ vemcmd show  port
  LTL     Port  Admin Link  State     Cause  PC-LTL  SGID  ORG  svcpath  Type              Owner  Vem Port
   19   Eth2/1     UP   UP    FWD         -       0          0        0                int-uplink  dpdk
   20   Eth2/2     UP   UP    FWD         -    1041     1    0        0                ext-uplink  dpdk
   21   Eth2/3     UP   UP    FWD         -       0          0        0              int-uplink-2  dpdk
   22   Eth2/4     UP   UP    FWD         -    1041     3    0        0              ext-uplink-2  dpdk
   23   Eth2/5     UP   UP    FWD         -       0          0        0            ext-uplink-vlan  dpdk
   51              UP   UP    FWD         -       0     1    0        0                   kni-vtep  dpdk
   52              UP   UP    FWD         -       0     1    0        0               kni-ave-ctrl  dpdk
   53              UP   UP    FWD         -       0     3    0        0                 kni-vtep-2  dpdk
   54              UP   UP    FWD         -       0     3    0        0         00:50:56:82:b6:bd  ave-epg-vm1:1
 1041      Po1     UP   UP    FWD         -       0          0        0
```

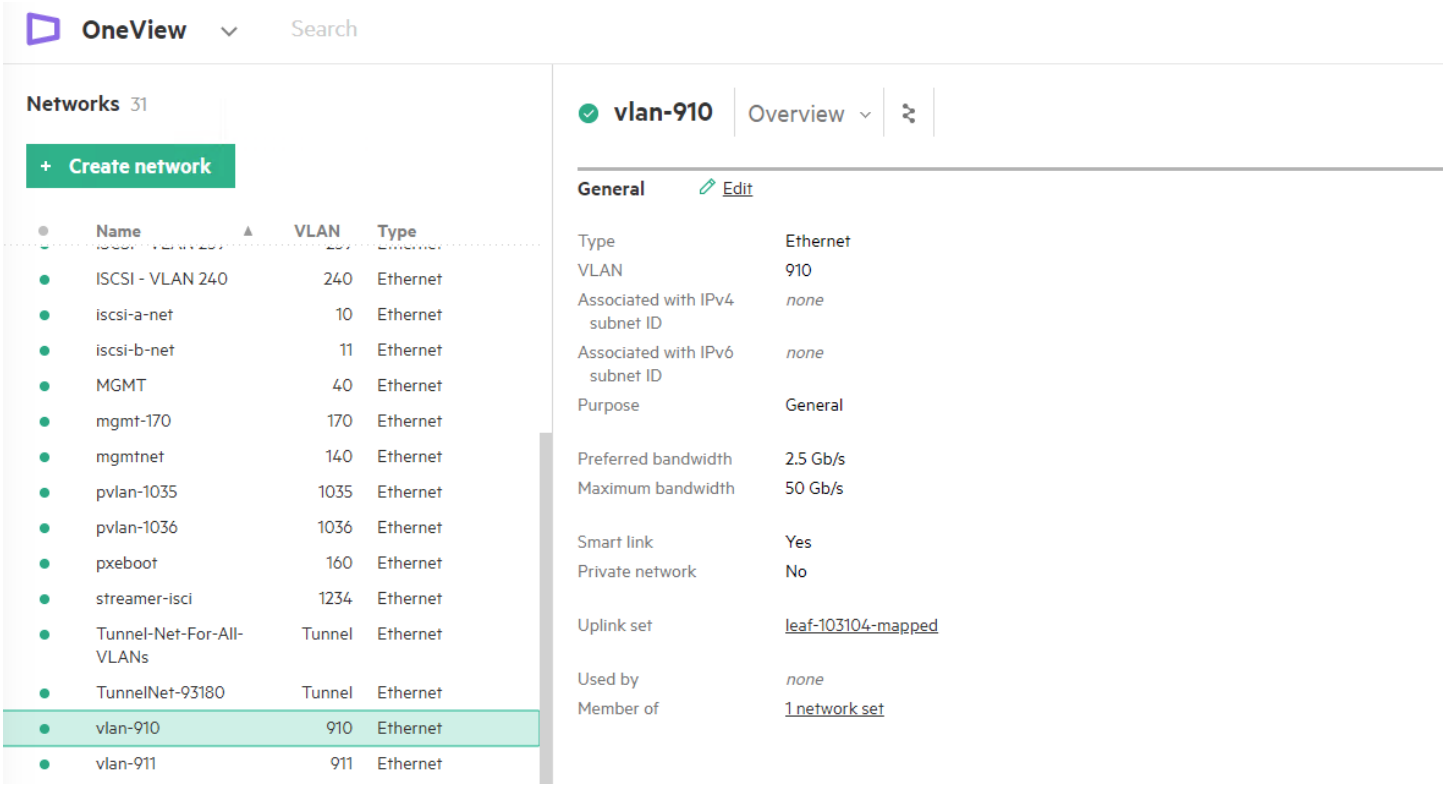
Synergy Private VLAN

OneView for Synergy 4.20 introduced Private VLAN feature. Private VLANs (PVLANs) provide Ethernet traffic isolation between all the members of the same VLAN. It diverts all the traffic between individual compute modules to flow through the TOR Leaf node unless explicitly permitted by a predefined policy. This capability enables interoperability with HPE Synergy and Cisco ACI Intra-EPG (end-point groups) isolation and micro-segmentation features.

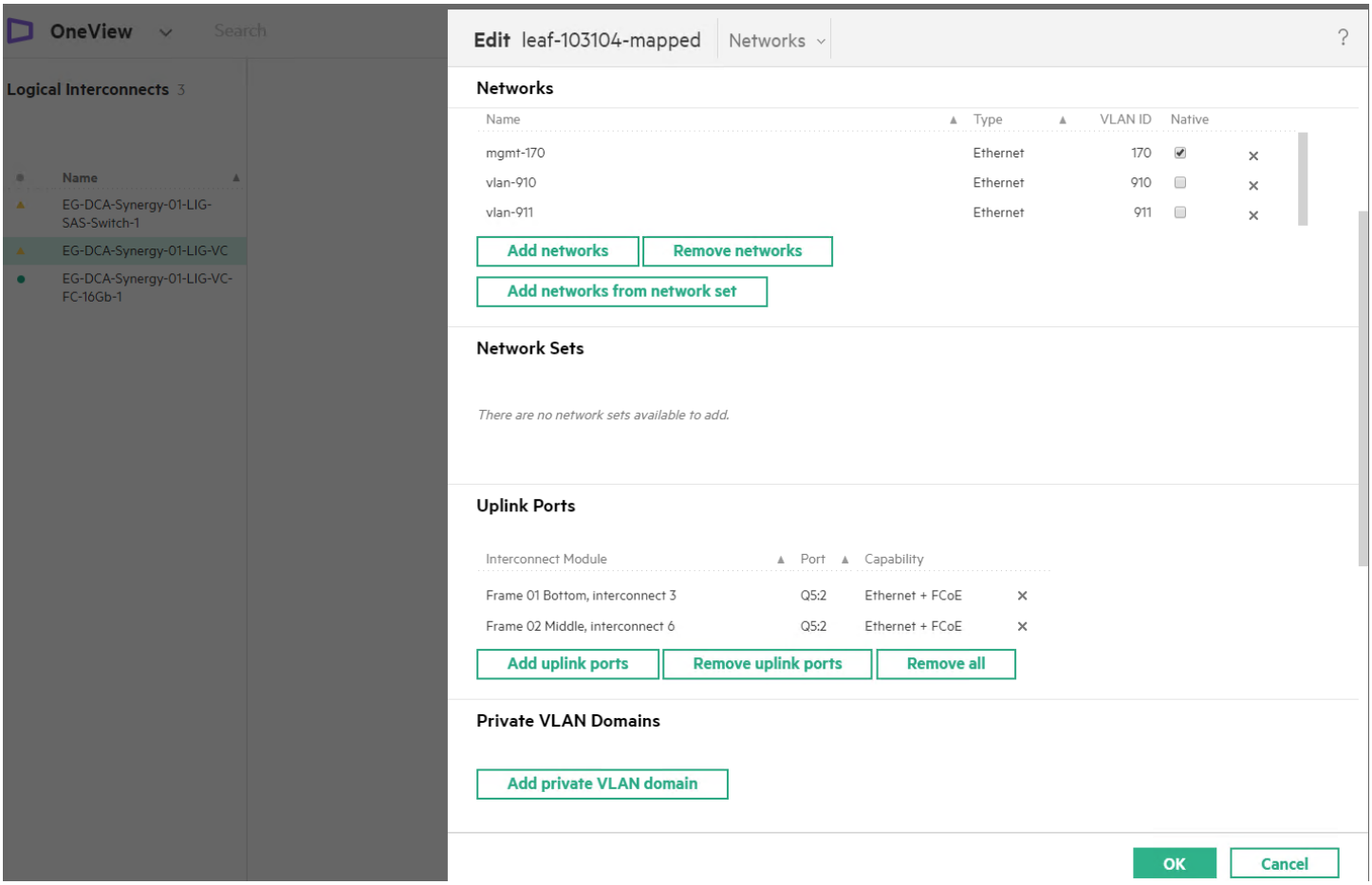
Note:

This feature is introduced for Synergy mapped networks to support PVLAN feature. Synergy tunnel networks already support PVLAN forwarding as it simply forwards traffic to ACI and VMware for PVLAN handling as discussed in the previous sections.

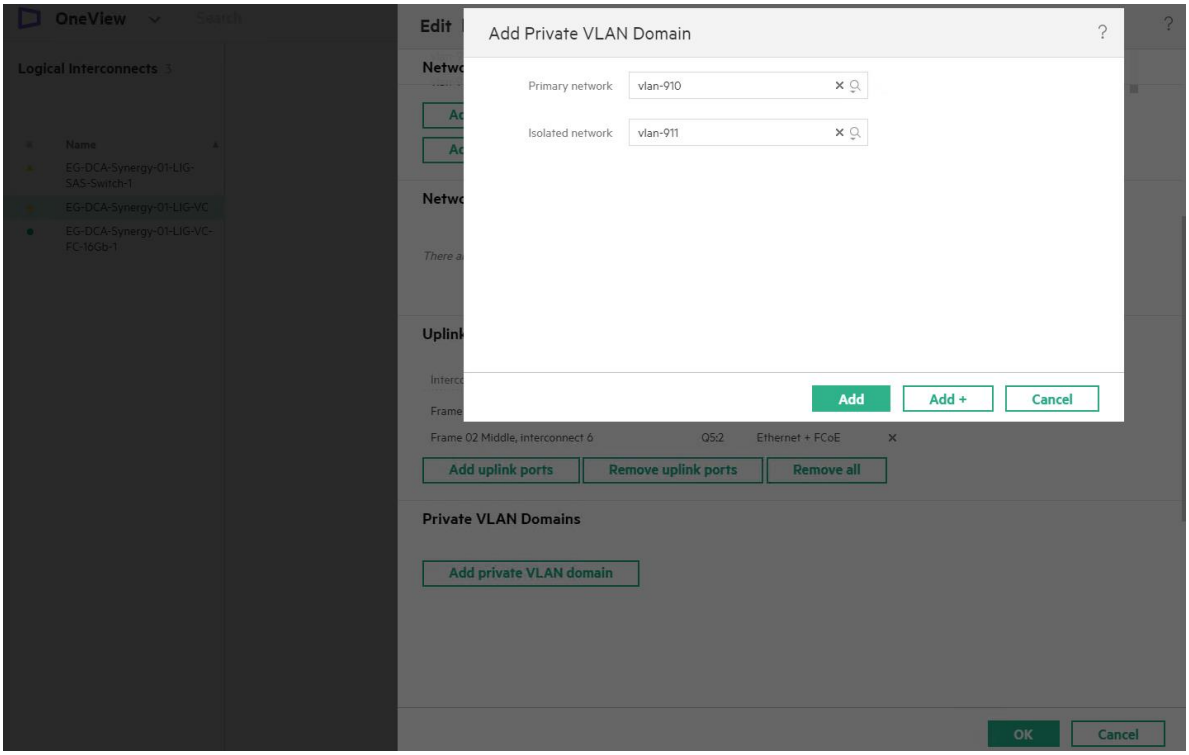
The configuration of Synergy private VLAN starts with provisioning the primary VLAN and the secondary VLAN as regular mapped network in OneView. In the example below, VLAN 910 will be the primary VLAN and VLAN 911 will be the secondary VLAN.



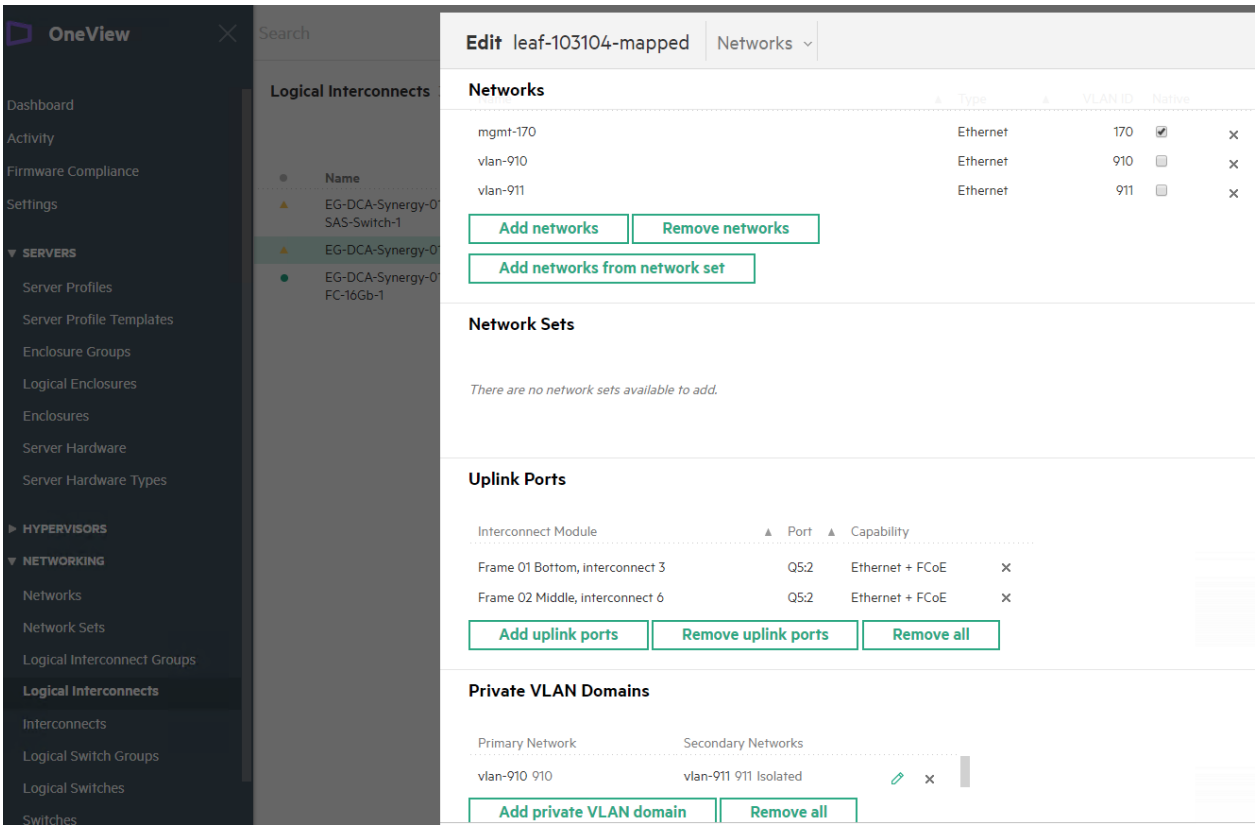
The PVLAN configuration is under Network->Logical Interconnects->Uplink Sets. In the desired uplink set configuration, after adding the two VLANs, users need to click “Add Private VLAN domain” button shown below.



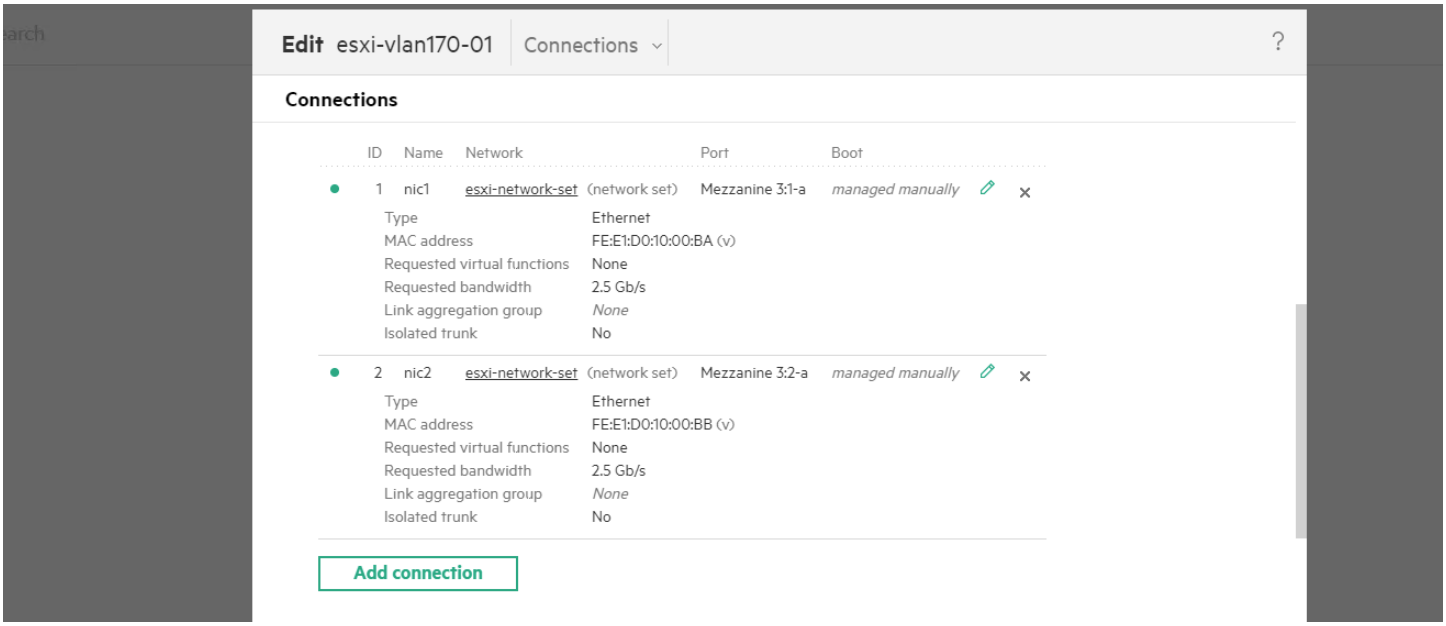
In the pop-up window, users need to choose the VLANs for the Primary and Isolated Network.



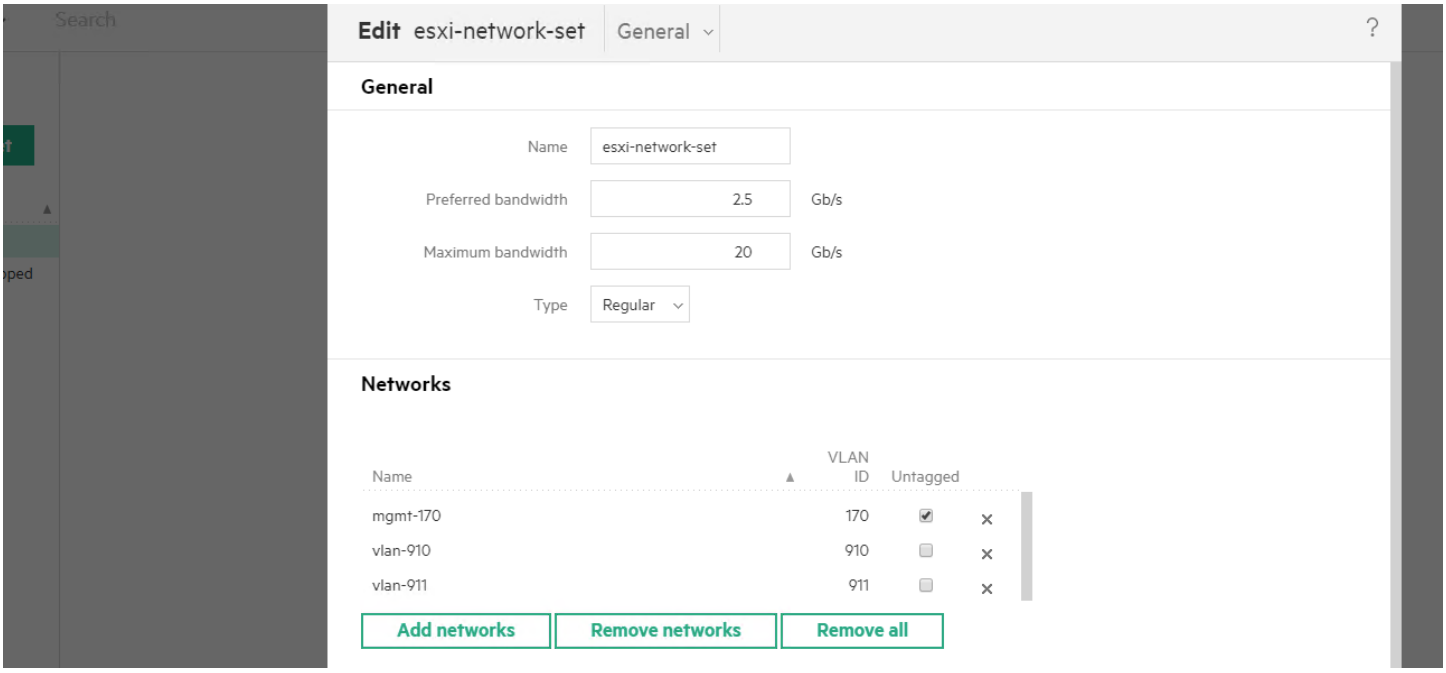
The completed uplinkset configuration with PVLANs should look like the following capture.



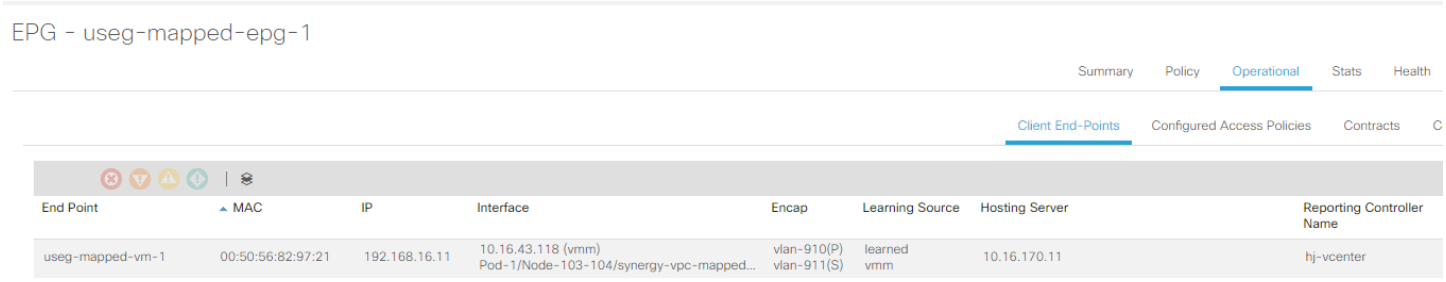
The above configuration steps should finish all of PVLAN provision for Synergy. The server profile should look like the normal network or network set for each connection.

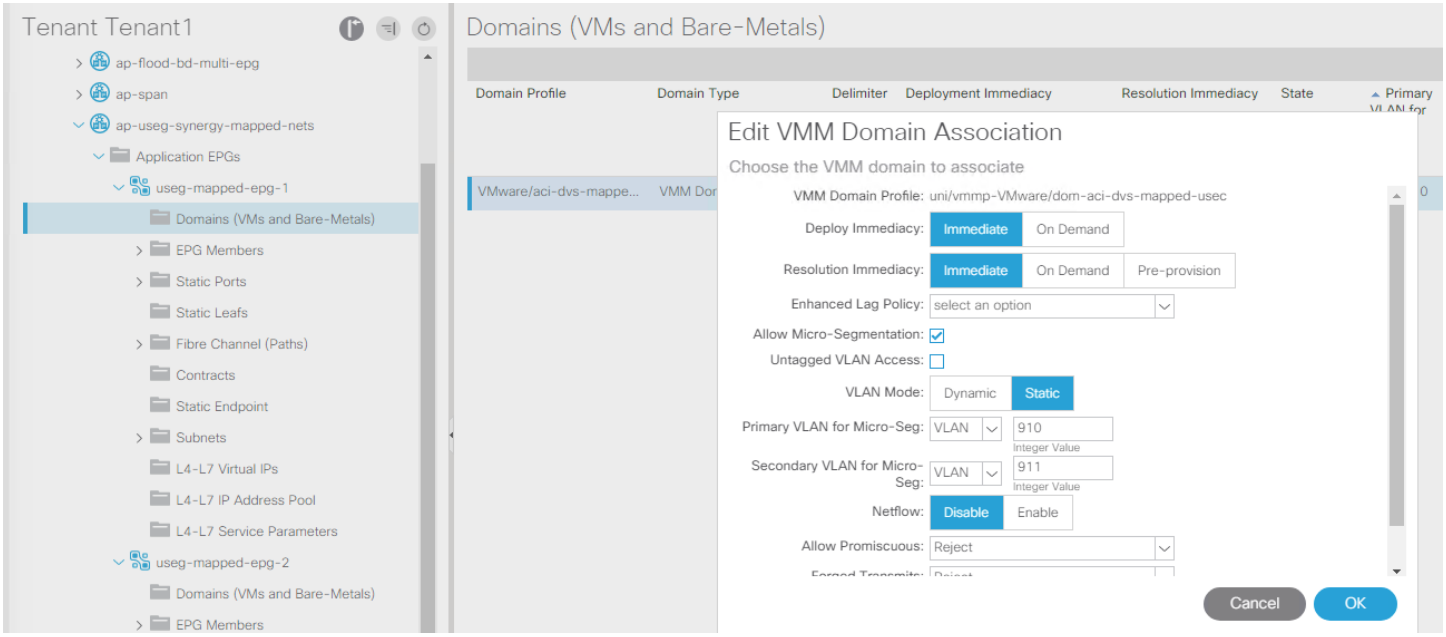


The network set definition is just regular configuration without any PVLAN reference like below.

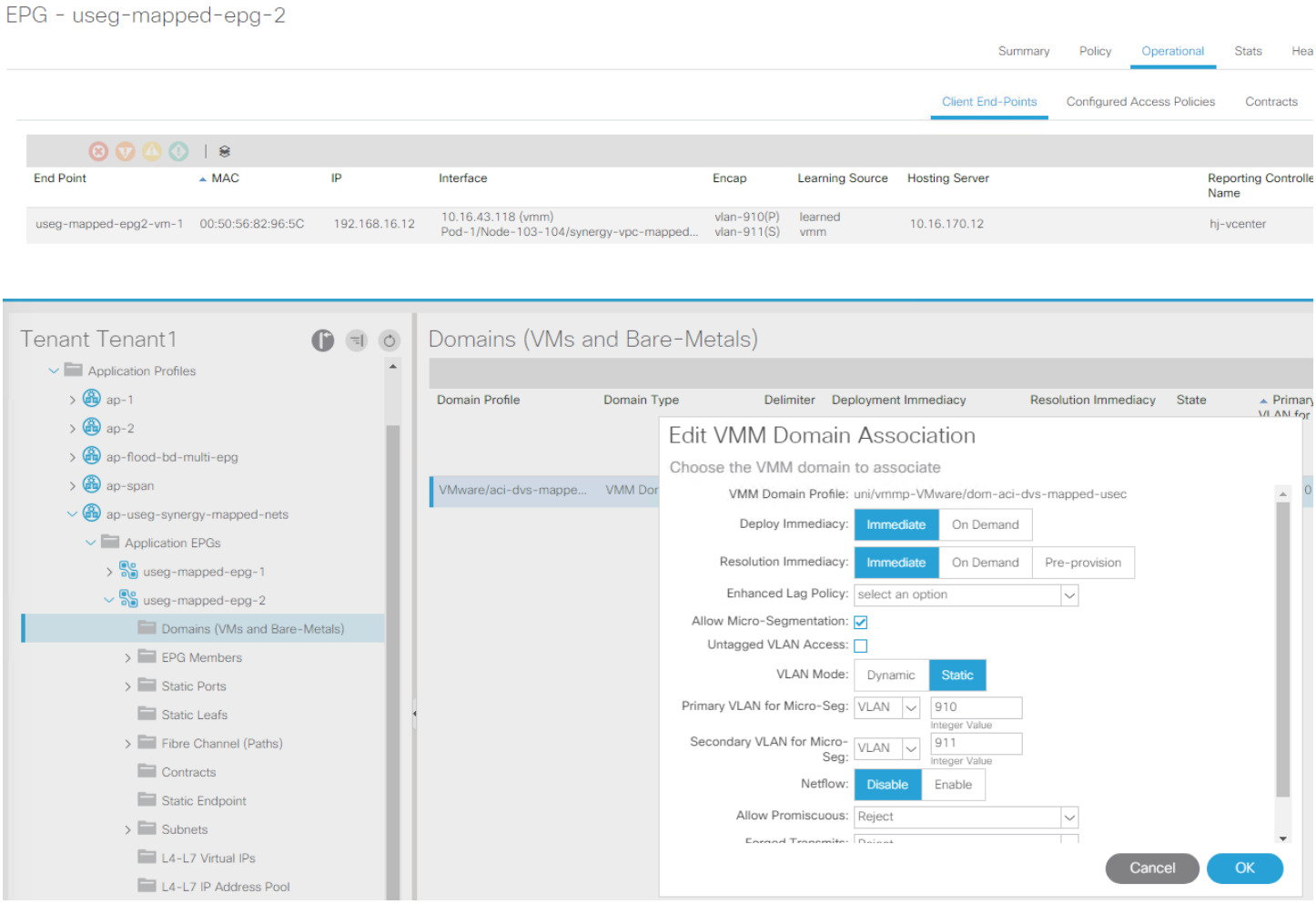


ACI is provisioned with two EPGs for micro segmentation using PVLANs with VMware DVS. The following capture shows EPG1 endpoint learning and configuration.

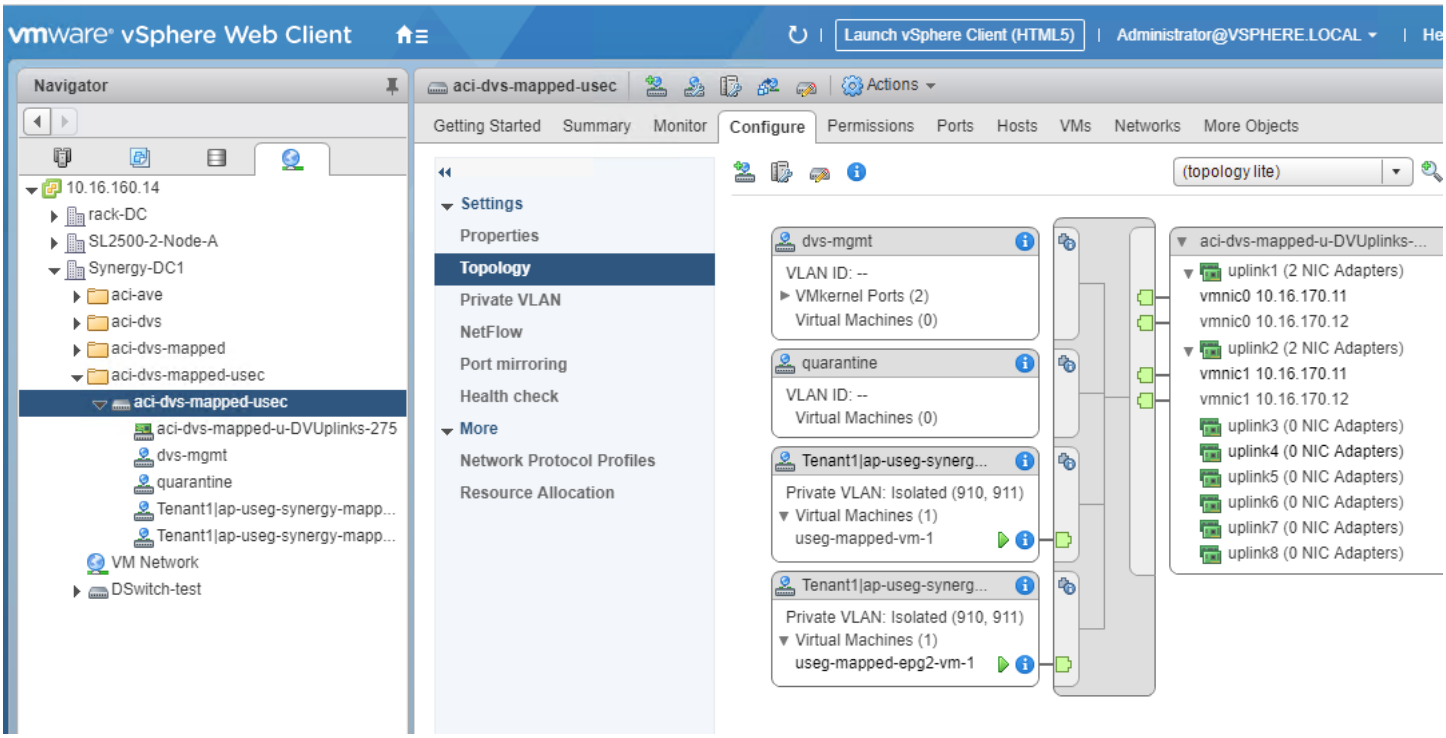




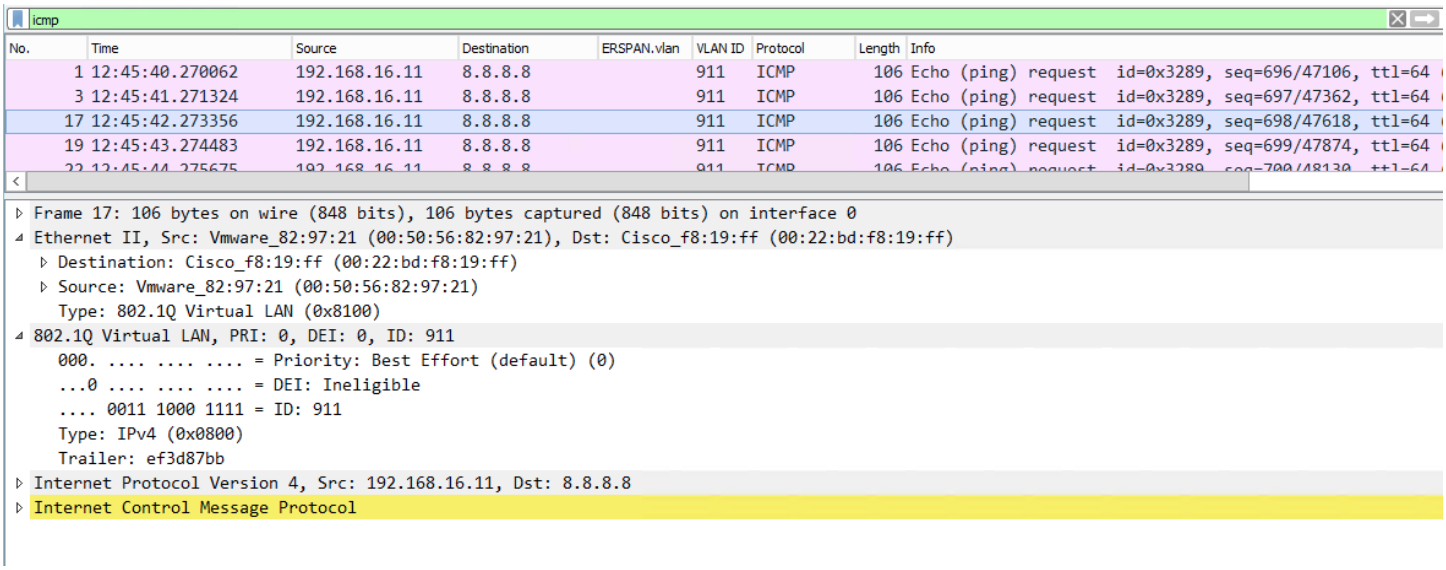
The following capture shows EPG2 endpoint learning and configuration.



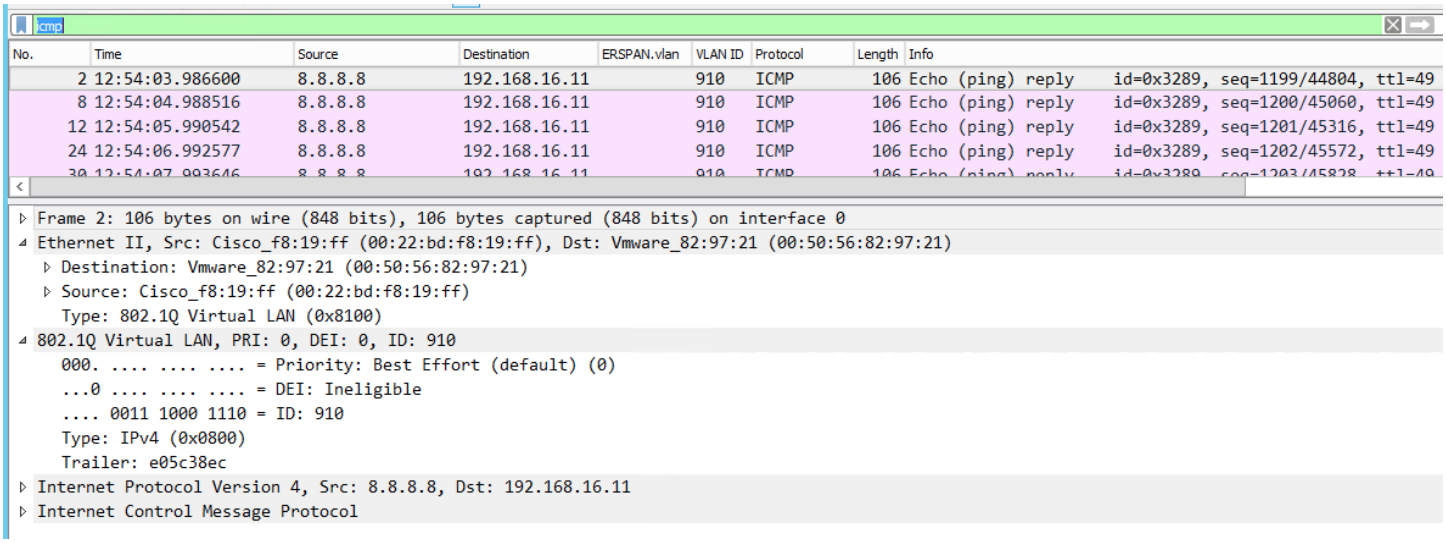
The following capture shows the PVLAN configuration on the DVS created by Cisco ACI. The primary VLAN 910 and secondary VLAN 911 match what are defined on ACI EPG VMM association configuration.



Synergy capture below verified that VMs used the secondary PVLAN 911 when it sent traffic to ACI default GW.



When ACI GW sent traffic to VMs, it used the primary VLAN 910 shown as the capture below.



No.	Time	Source	Destination	ERSPAN.vlan	VLAN ID	Protocol	Length	Info
2	12:54:03.986600	8.8.8.8	192.168.16.11		910	ICMP	106	Echo (ping) reply id=0x3289, seq=1199/44804, ttl=49
8	12:54:04.988516	8.8.8.8	192.168.16.11		910	ICMP	106	Echo (ping) reply id=0x3289, seq=1200/45060, ttl=49
12	12:54:05.990542	8.8.8.8	192.168.16.11		910	ICMP	106	Echo (ping) reply id=0x3289, seq=1201/45316, ttl=49
24	12:54:06.992577	8.8.8.8	192.168.16.11		910	ICMP	106	Echo (ping) reply id=0x3289, seq=1202/45572, ttl=49

```
▷ Frame 2: 106 bytes on wire (848 bits), 106 bytes captured (848 bits) on interface 0
◢ Ethernet II, Src: Cisco_f8:19:ff (00:22:bd:f8:19:ff), Dst: Vmware_82:97:21 (00:50:56:82:97:21)
  ▷ Destination: Vmware_82:97:21 (00:50:56:82:97:21)
  ▷ Source: Cisco_f8:19:ff (00:22:bd:f8:19:ff)
    Type: 802.1Q Virtual LAN (0x8100)
◢ 802.1Q Virtual LAN, PRI: 0, DEI: 0, ID: 910
    000. .... .... .... = Priority: Best Effort (default) (0)
    ...0 .... .... .... = DEI: Ineligible
    .... 0011 1000 1110 = ID: 910
    Type: IPv4 (0x0800)
    Trailer: e05c38ec
▷ Internet Protocol Version 4, Src: 8.8.8.8, Dst: 192.168.16.11
▷ Internet Control Message Protocol
```

Further PVLAN monitoring can be obtained by using Synergy Virtual Connect CLI. By default Synergy VC CLI is not enabled for security reason. Enabling Synergy VC CLI involves communicating with Synergy OneView using Restful API and it's beyond the scope of this white paper. Users can get help from HPE support team for further VC CLI troubleshooting.

```
OneView# show vlan isolated-user-vlan

switch default
VlanId   Type                  Super VlanId
-------  ------               --------------
910      super vlan             -
911      Isolated sub vlan      910
```

The following Synergy VC CLI showed primary VLAN 910 MAC table and it verified ACI GW MAC is learned from VC uplink "po5" and two VMs were off different compute nodes(S-Channel).

```
OneView# show mac-address-table vlan 910

Vlan    Mac Address         Type     Ports
----    -----------         ----     -----
910     00:50:56:82:96:5c   Learnt   S-Channel1/2/12:2
910     00:22:bd:f8:19:ff   Learnt   po5
910     00:50:56:82:97:21   Learnt   S-Channel1/2/11:2

Total Unicast Mac Addresses displayed: 3


Total Multicast Mac Addresses displayed: 0
```

The two VMs ARP table shows each VM knew the other VM by ACI BD MAC address. This is correct behavior when ACI is doing micro segmentation and using Proxy-ARP behind the scene.

```
[root@usec-mapped-vm-1 ~]# arp
Address                  HWtype  HWaddress           Flags Mask            Iface
192.168.16.12            ether   00:22:bd:f8:19:ff   C                     eno16780032
192.168.16.1             ether   00:22:bd:f8:19:ff   C                     eno16780032
```

```
[root@useg-mapped-epg2-vm-1 ~]# arp
Address                  HWtype  HWaddress           Flags Mask            Iface
192.168.16.11            ether   00:22:bd:f8:19:ff   C                     eno16780032
192.168.16.1             ether   00:22:bd:f8:19:ff   C                     eno16780032
[root@useg-mapped-epg2-vm-1 ~]#
```

Synergy Fabric Managers

Synergy Fabric Manager feature was introduced in Synergy OneView 4.10 version. It is a resource manager in HPE OneView that enables integration of a Cisco ACI fabric with HPE Synergy. A fabric manager aligns HPE OneView resources as defined by Cisco ACI APIC policies.

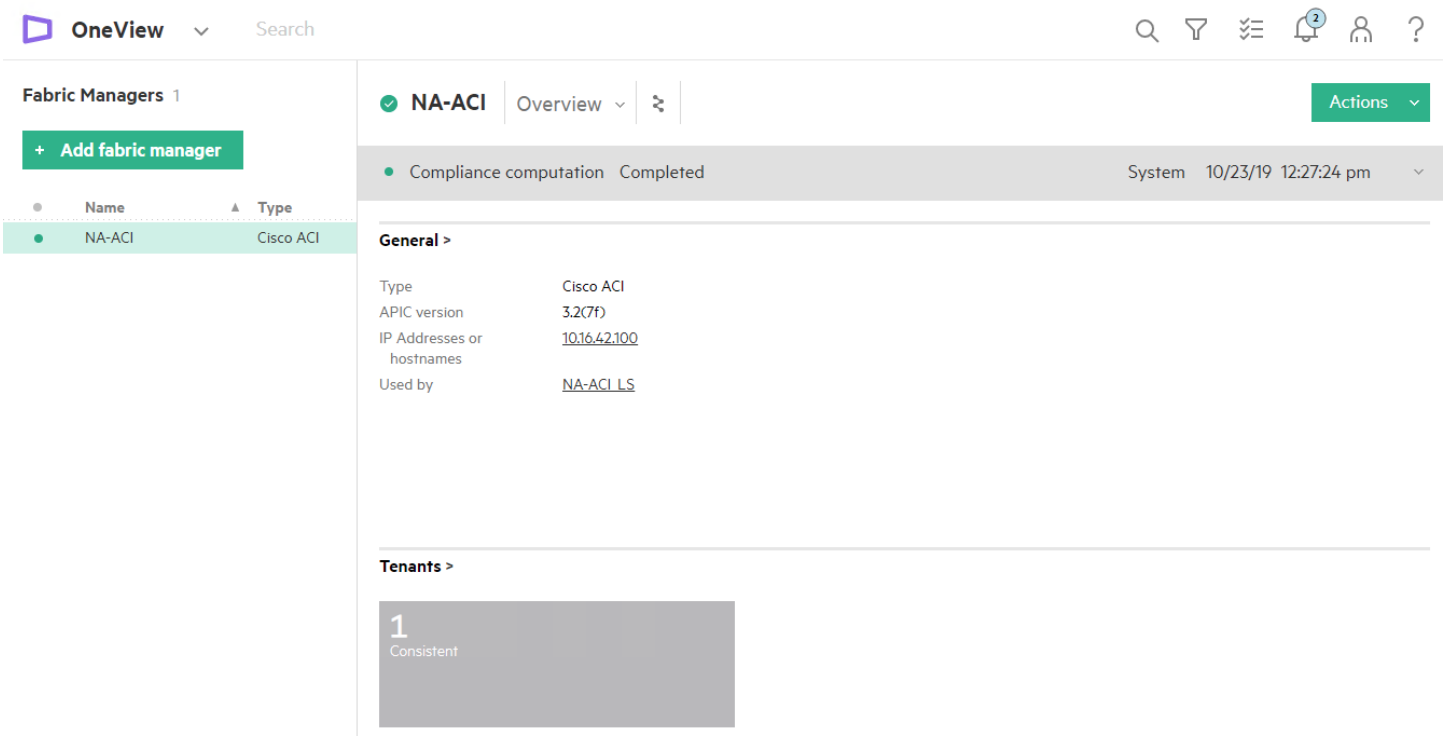
FM feature working scope is based on ACI tenant. In summary, the order of Fabric Manager operation is as the following

- Import ACI APIC controller cluster into FM. (Only ACI read access is required, FM won't modify any ACI policies).
- Select the tenant(s) in interest
- FM will check the ACI policies defined under the tenant and report any policy mismatch between ACI and Synergy, and suggest action items for any mismatch items.

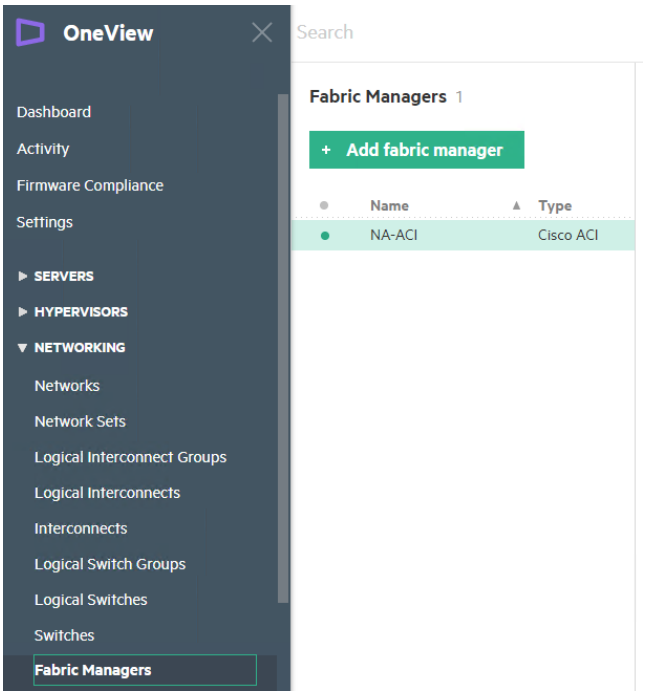
Note:

Fabric Manager feature is to help Synergy admins to match Synergy network configurations with APIC polices so network configuration mismatch can be prevented or reduced. This operation is through control pane and doesn't affect any data plane forwarding behavior discussed earlier in this white paper. Users don't have to config Fabric Manager feature in order to pass traffic successfully between ACI and Synergy.

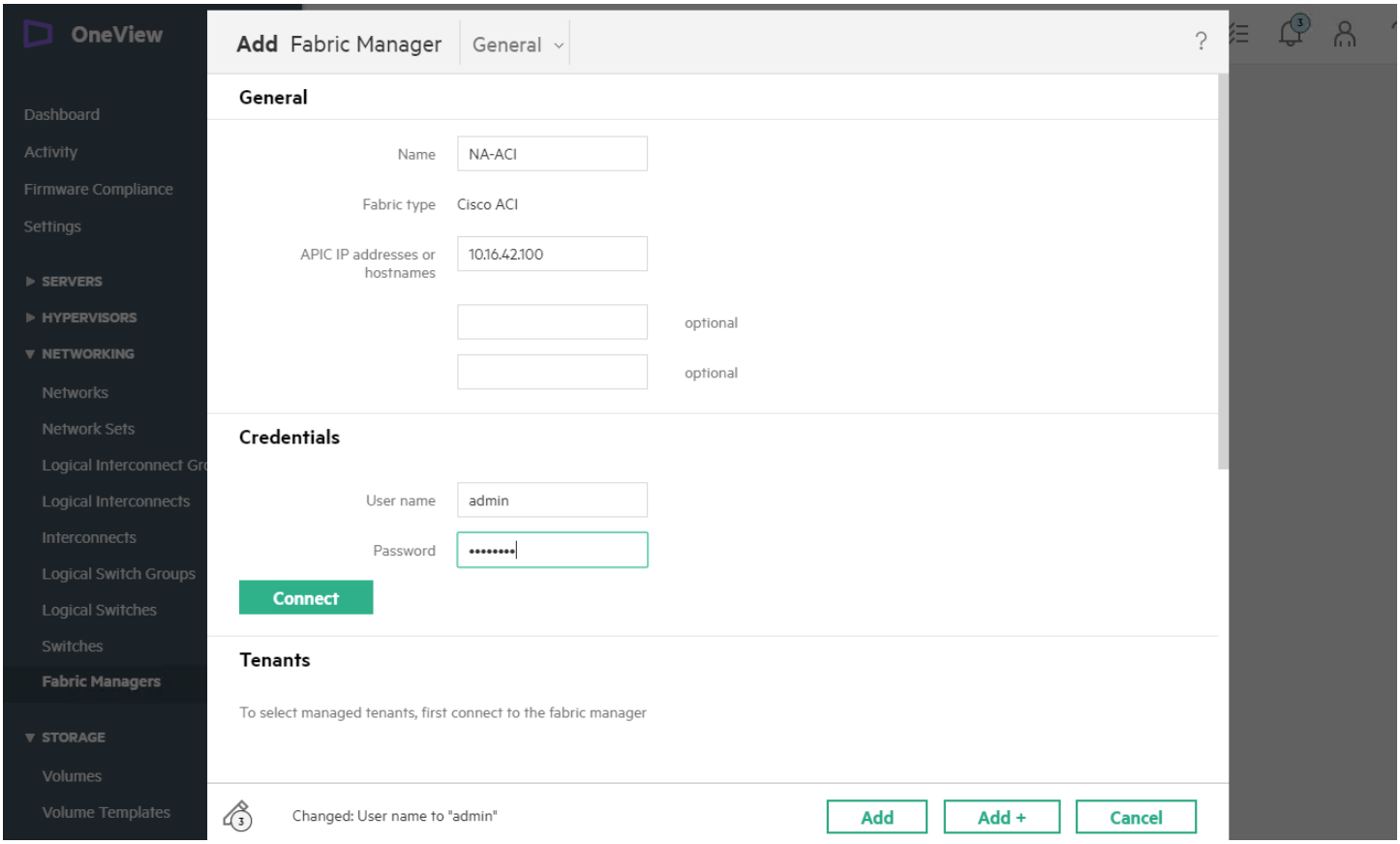
A running fabric manager for Cisco ACI is shown below.



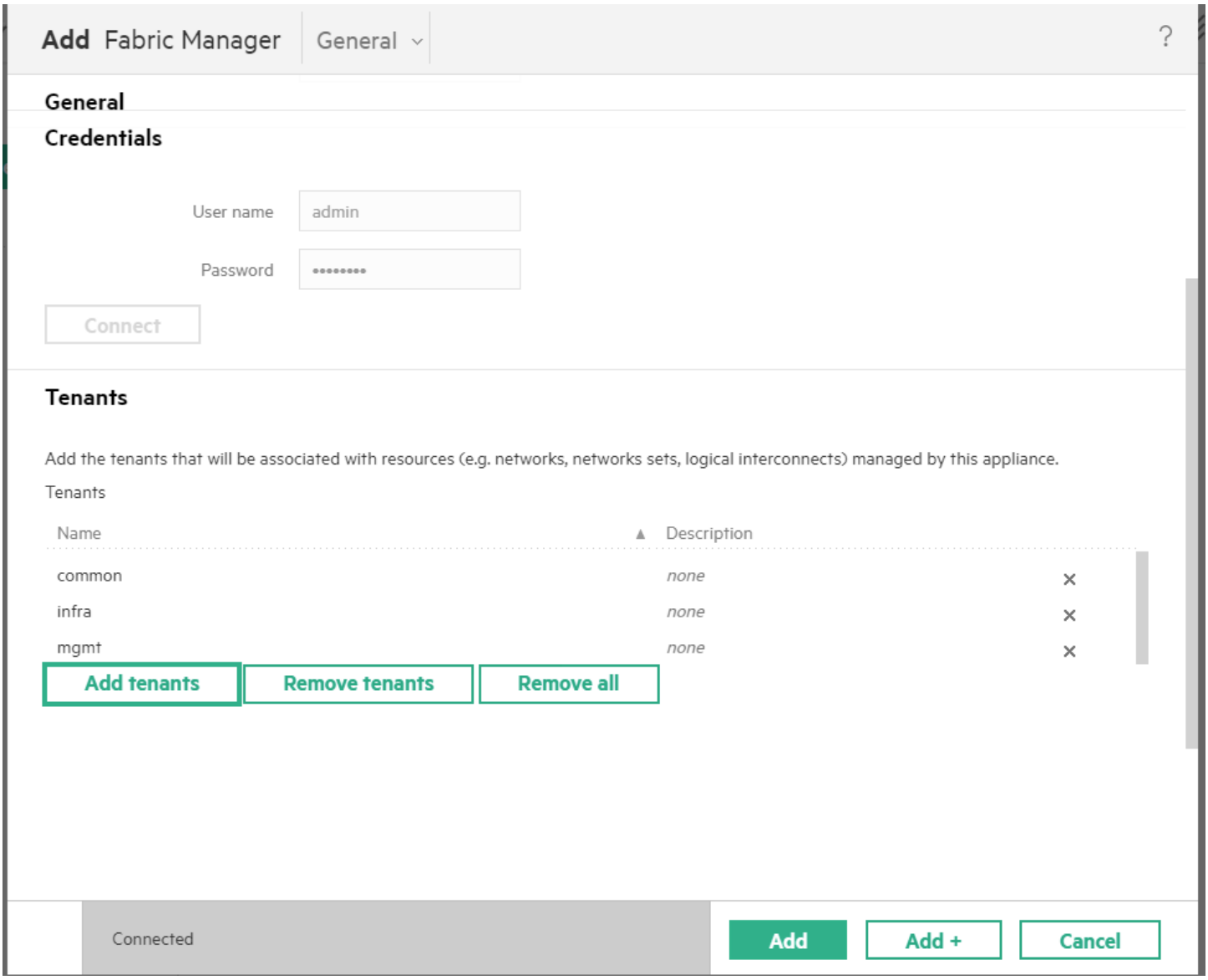
To add a new fabric manager, users can go to main menu and click "Fabric Managers" under Networking.



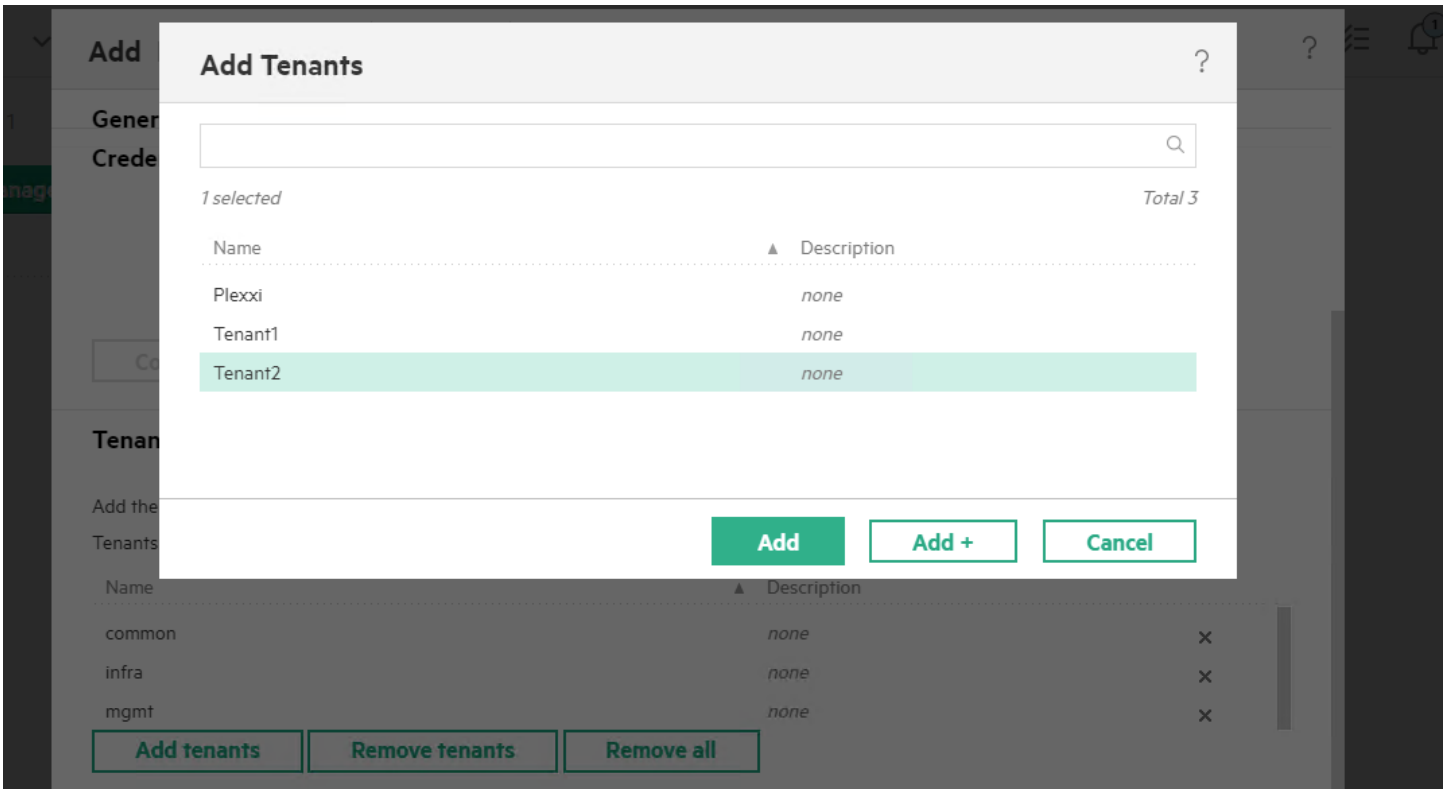
Users will then provide APIC controller IP or hostname along with login credential(only read-only privilege is required).



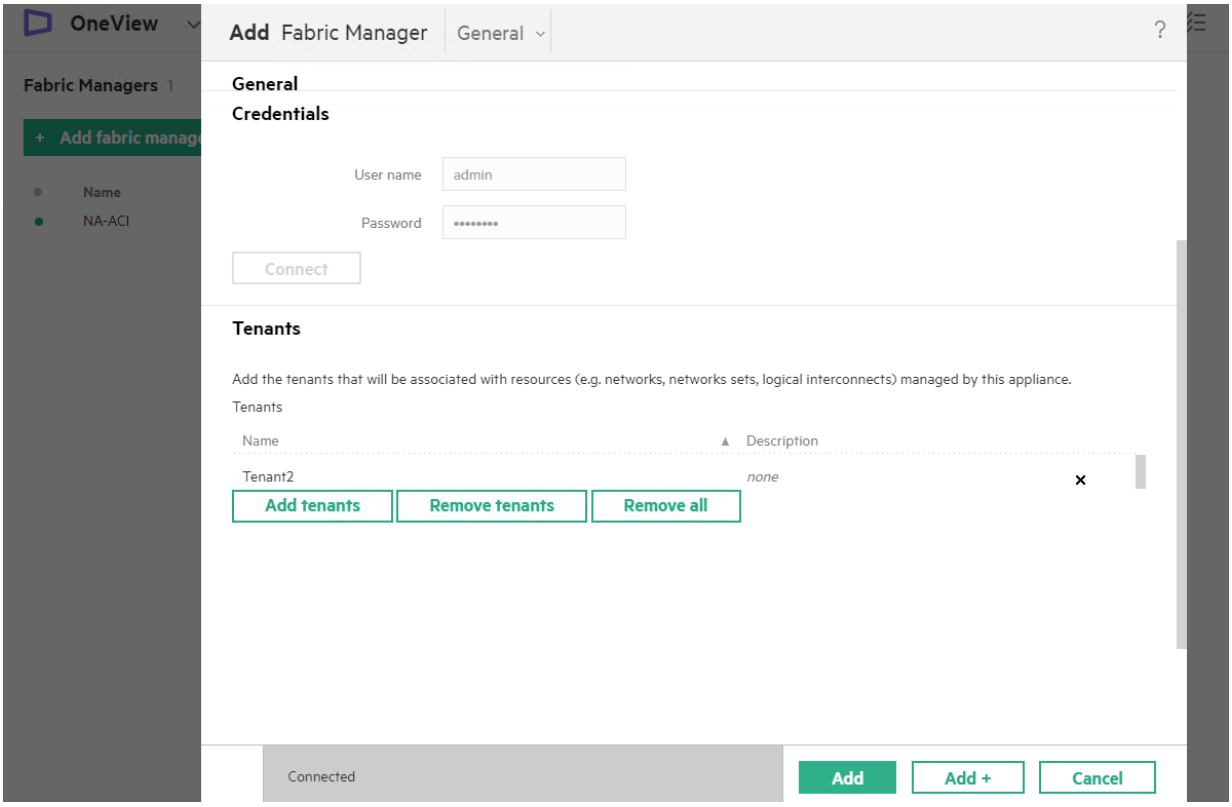
After successful login, Synergy will present default tenants defined on ACI. Users need to click “Add tenants” to add the tenant in interest.



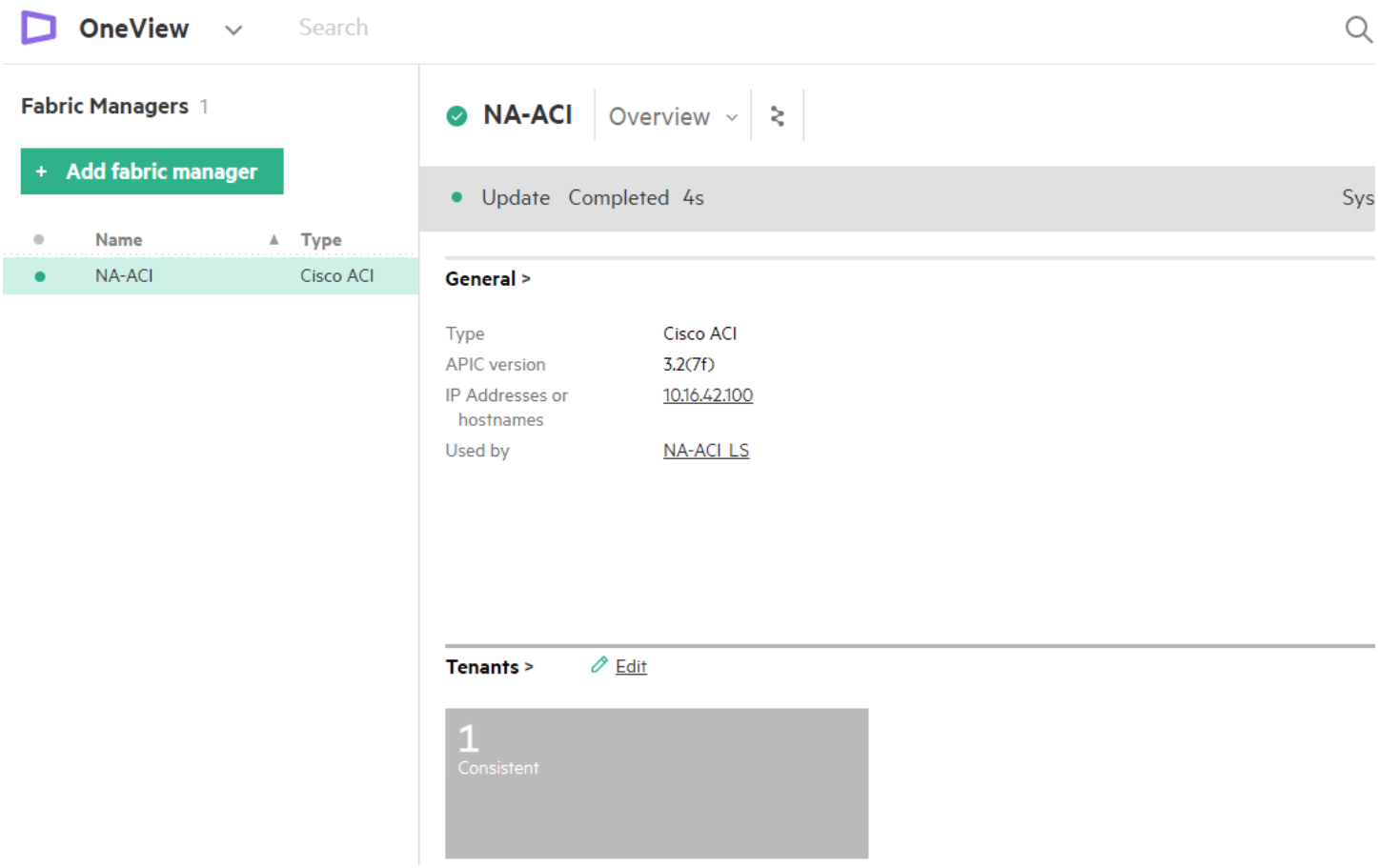
In the following capture, the Tenant2 was added.



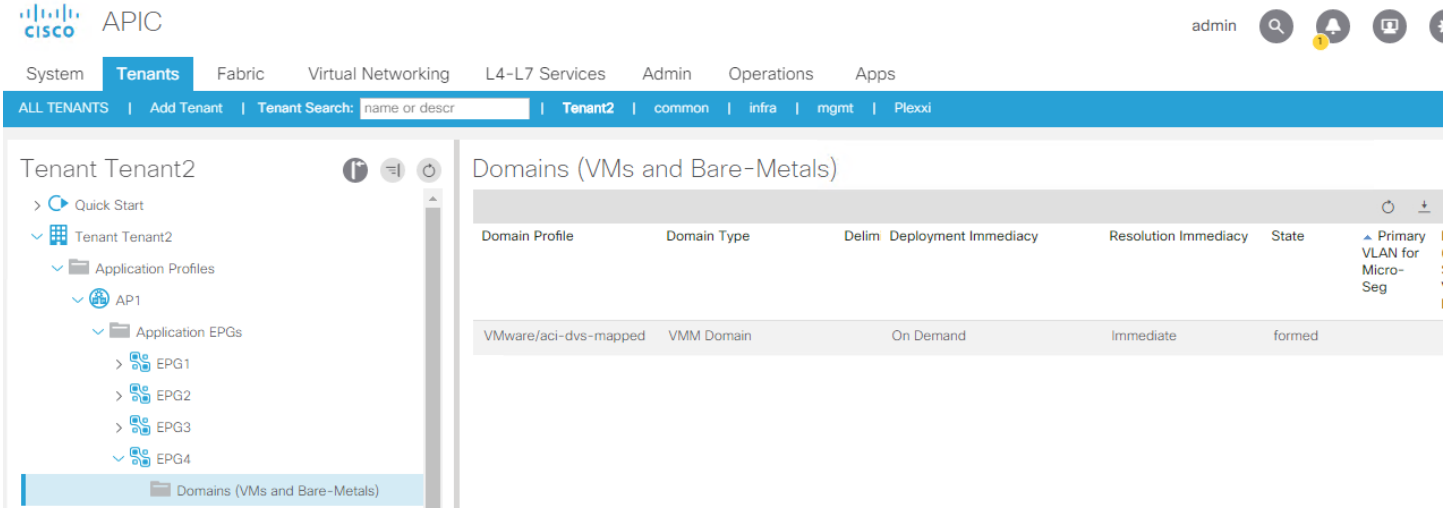
Users can remove default tenants and only leave the tenant interested in the tenant list.



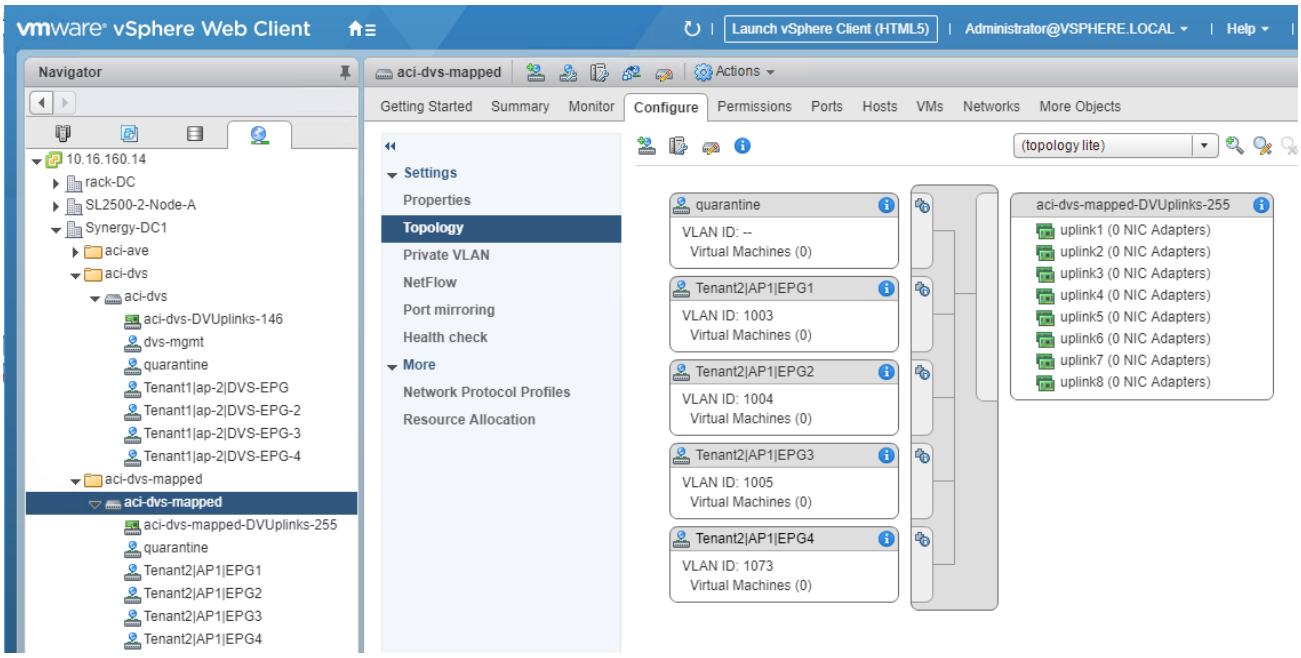
A successful fabric manager import is shown in the following screen.



Fabric Manager can dynamically detect changes for the tenant EPG encapsulation VLANs. In the following example, on ACI side, a new EPG was added associated with ACI VMM.



VMware DVS had a new port-group created with EPG name and the VLAN id 1073.



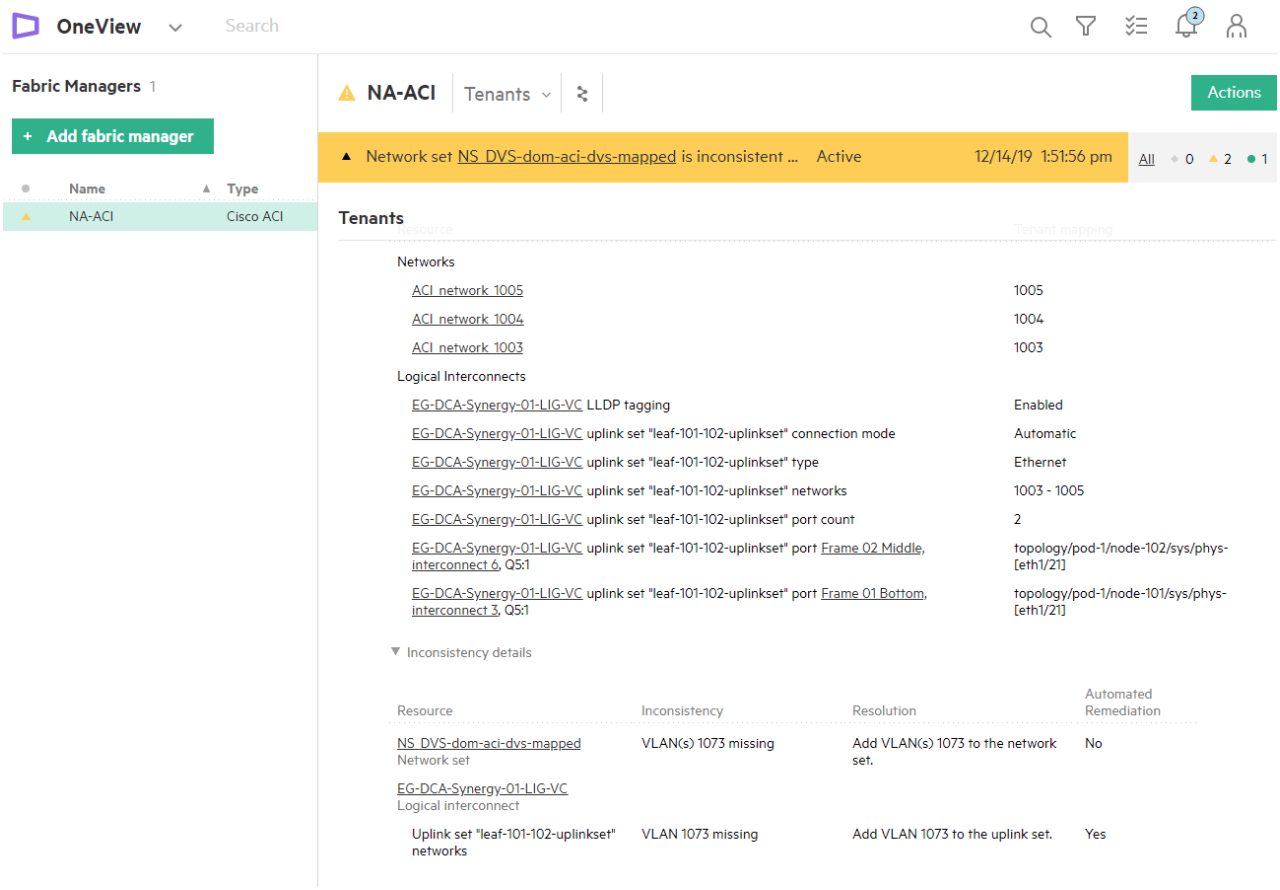
Synergy Fabric Manager dynamically refresh the status and reported the inconsistency between ACI EPG encapsulation VLAN and Synergy uplink set VLAN provision.

From the inconsistency details report, users can see Synergy uplink set is being provisioned for VLAN 1003, 1004 and 1005, which match ACI side EPG1/EPG2/EPG3 encapsulation VLAN. However, the newly created EPG4 encapsulation VLAN 1073 has not been provisioned on Synergy uplink set.

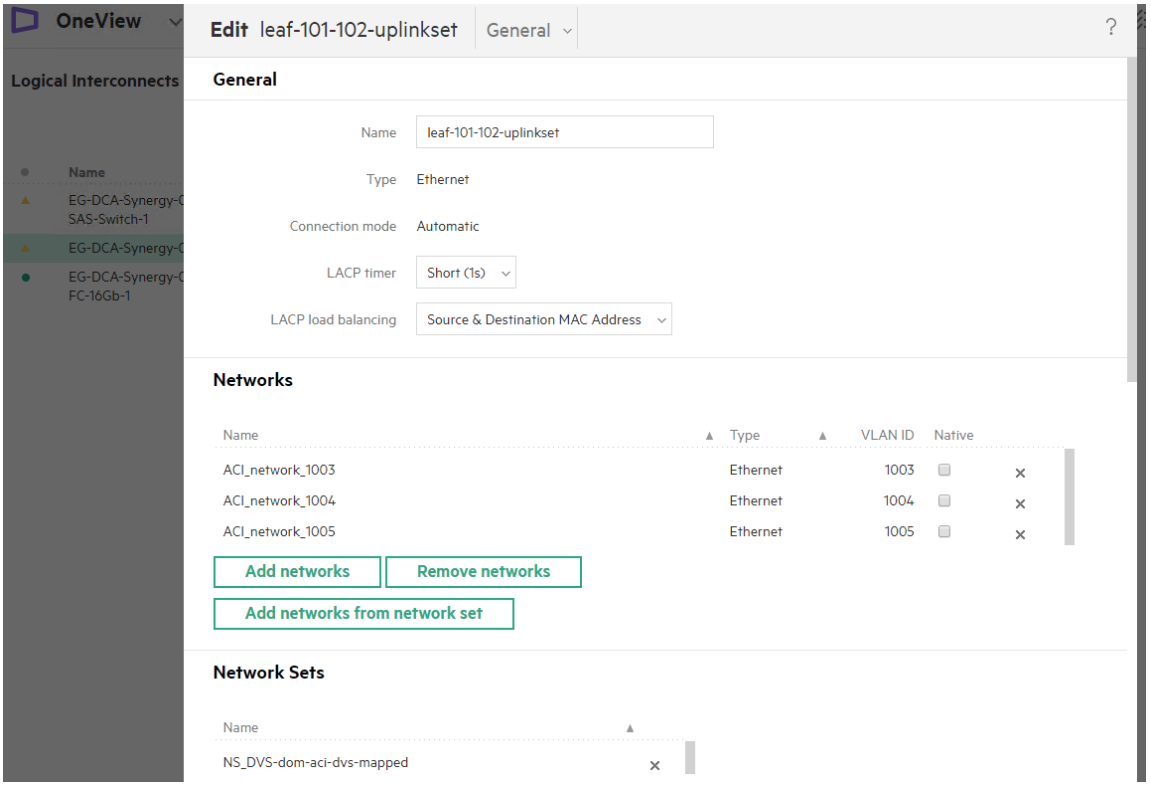
Fabric Manager also indicates that it can automatically add the VLAN 1073 automatically in the uplink set if users choose to do so. For network set applied on server profile side, Fabric Manager suggested that VLAN 1073 should be added in the network set but users have to do it manually.

Note:

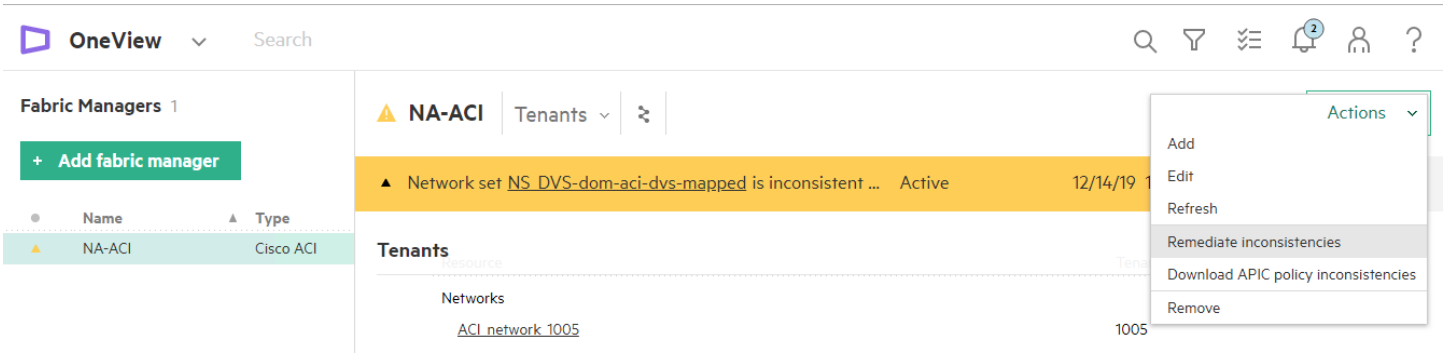
The above feature of Fabric Manager VLAN mismatch detection is meant to used with Synergy mapped mode where Synergy need to have matching VLAN configuration with ACI and hypervisor side. Synergy Tunnel VLAN doesn't really need this function as Synergy will not try to differentiate external user VLANs as we discussed in the previous sections in this paper.



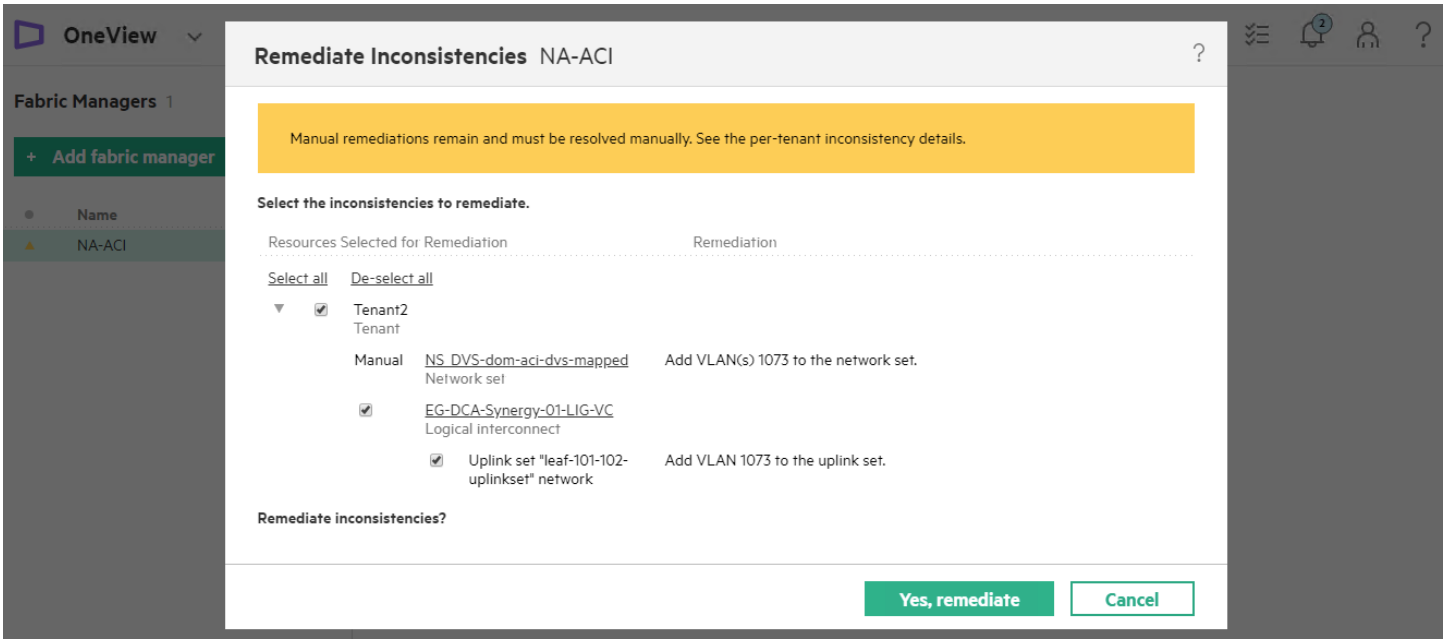
User can verify the inconsistency in uplink set configuration details, where only VLAN 1003, 1004 and 1005 are configured.



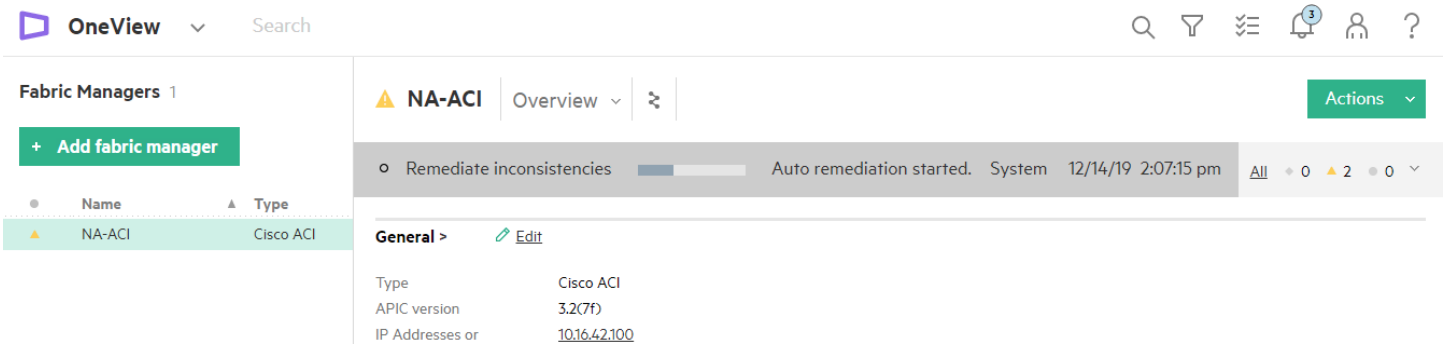
Users can click “Remediate inconsistencies” menu if they want to let FM add the VLAN automatically.



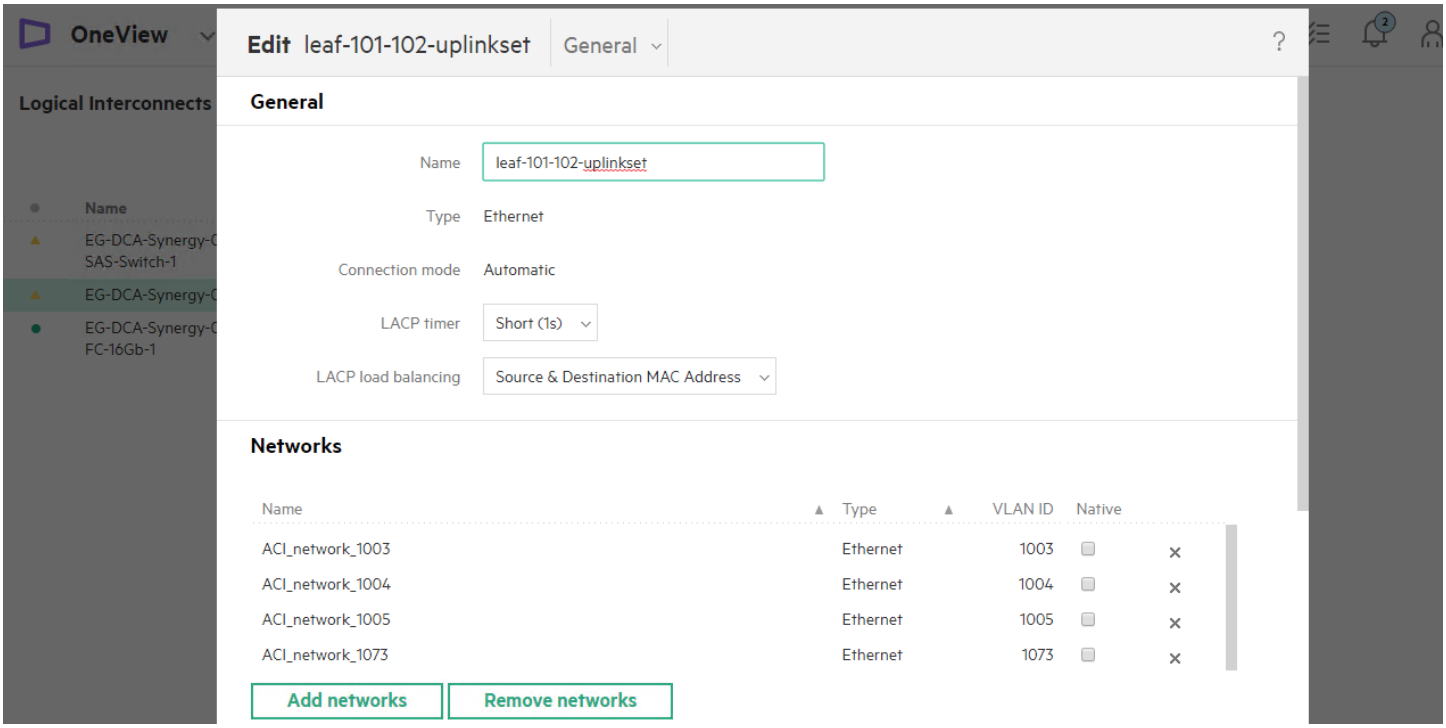
FM will report the action items it could take.



Upon users selecting the items and click “remediate”, FM will start the remediation process.



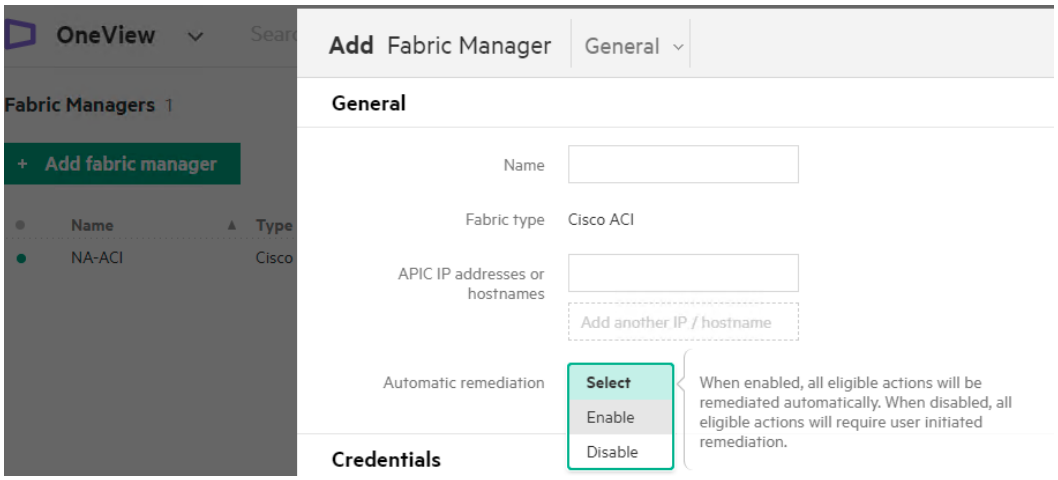
After remediation finishes, users should see VLAN 1073 added in the uplink set.



Synergy OneView 5.20 introduces more in-depth integration capabilities including:

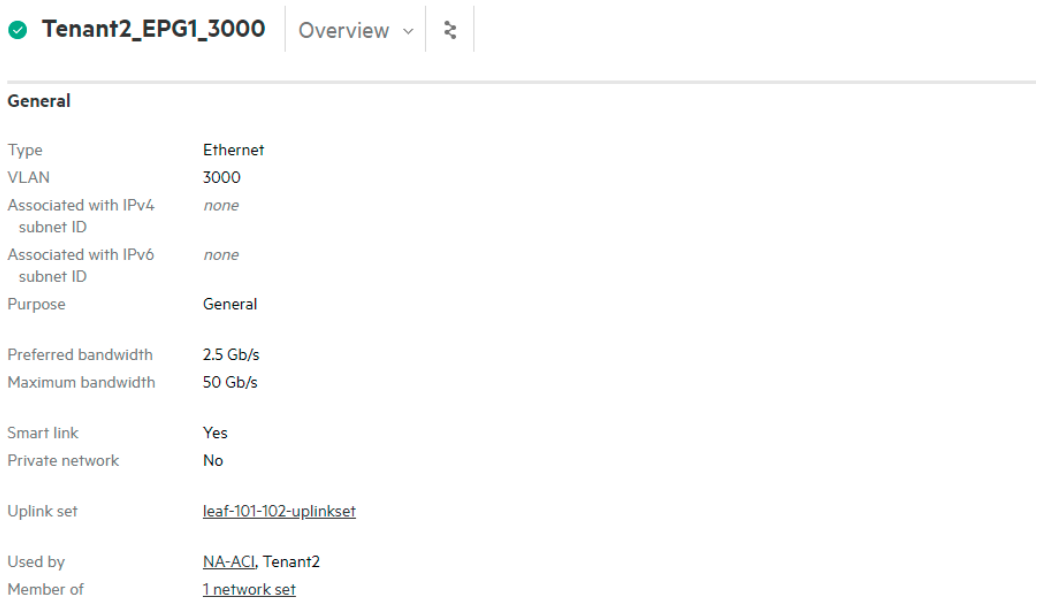
- Auto-remediation of configuration changes between the HPE OneView Fabric Manager and Cisco ACI without requiring manual user intervention.

  In the following screen capture, users can choose if they want to enable the feature of “Automatic Remediation” when importing the APIC controller. When enabled, for tenants being monitored by Fabric Manager, ACI EPG VLAN configuration changes will trigger Fabric Manager to automatically modify OneView uplink set VLAN config to match the new EPG VLAN configuration.



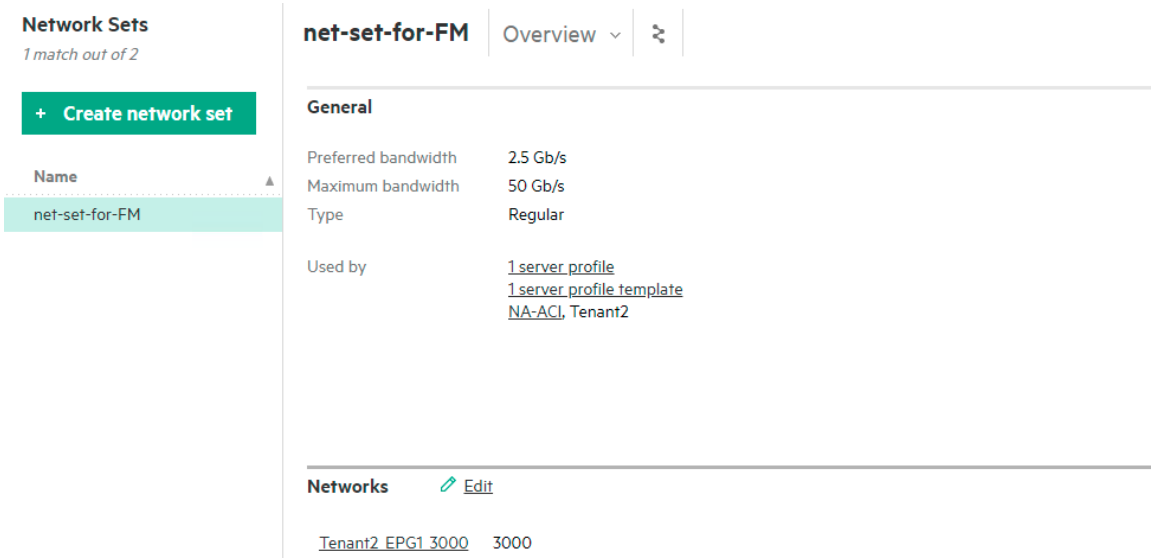
- Adds custom naming to the auto-created network names to include both the tenant and end-point group naming.

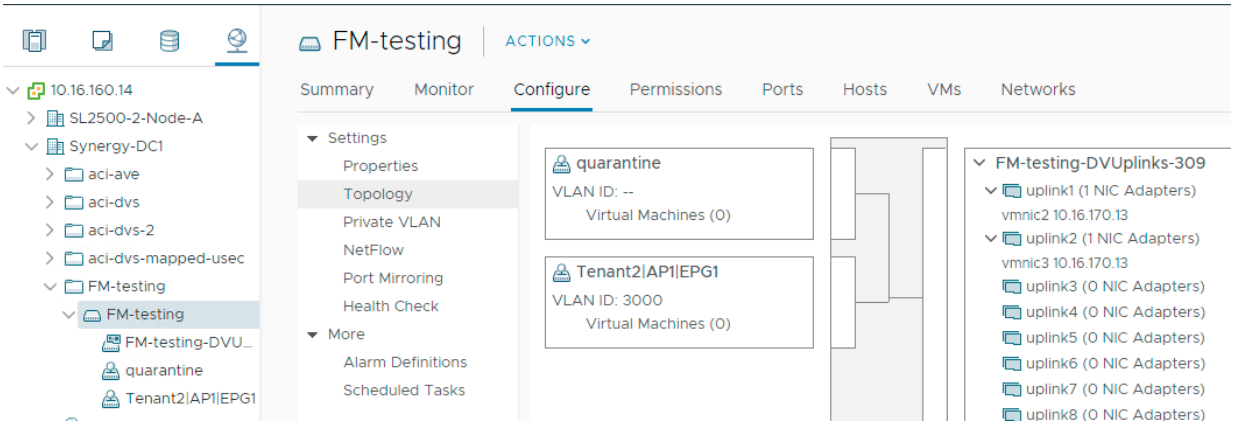
In the following screen capture, the new naming convention used by Fabric Manager indicates the VLAN is created according to what's configured under ACI Tenant 2, EPG1 with VLAN tag 3000.



- HPE OneView Fabric Manager provides the ability to auto-remediate inconsistencies by adding networks directly to already deployed network sets

In the following capture, it's shown that FM manager added the newly created VLAN in the network set used by the server profile to make sure that the VLAN configuration matches what's being configured on ACI-VMM-created DVS port-group.




FM-testing
ACTIONS
10.16.160.14
SL2500-2-Node-A
Synergy-DC1
aci-ave
aci-dvs
aci-dvs-2
aci-dvs-mapped-usec
FM-testing
FM-testing
FM-testing-DVU...
quarantine
Tenant2|AP1|EPG1
Summary
Monitor
Configure
Permissions
Ports
Hosts
VMs
Networks
Settings
Properties
Topology
Private VLAN
NetFlow
Port Mirroring
Health Check
More
Alarm Definitions
Scheduled Tasks
quarantine
VLAN ID: --
Virtual Machines (0)
Tenant2|AP1|EPG1
VLAN ID: 3000
Virtual Machines (0)
FM-testing-DVUplinks-309
uplink1 (1 NIC Adapters)
vmnic2 10.16.170.13
uplink2 (1 NIC Adapters)
vmnic3 10.16.170.13
uplink3 (0 NIC Adapters)
uplink4 (0 NIC Adapters)
uplink5 (0 NIC Adapters)
uplink6 (0 NIC Adapters)
uplink7 (0 NIC Adapters)
uplink8 (0 NIC Adapters)

Resources and additional links

For in-depth Virtual Connect traffic flow analysis, please check

HPE Virtual Connect traffic flow with HPE Synergy

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

Sign up for updates

★ Rate this document

4AA4-xxxxENW, Month 20XX, Rev. #