ACI Troubleshooting Documentation

Release 1.0

Edi Wibowo

Aug 31, 2020

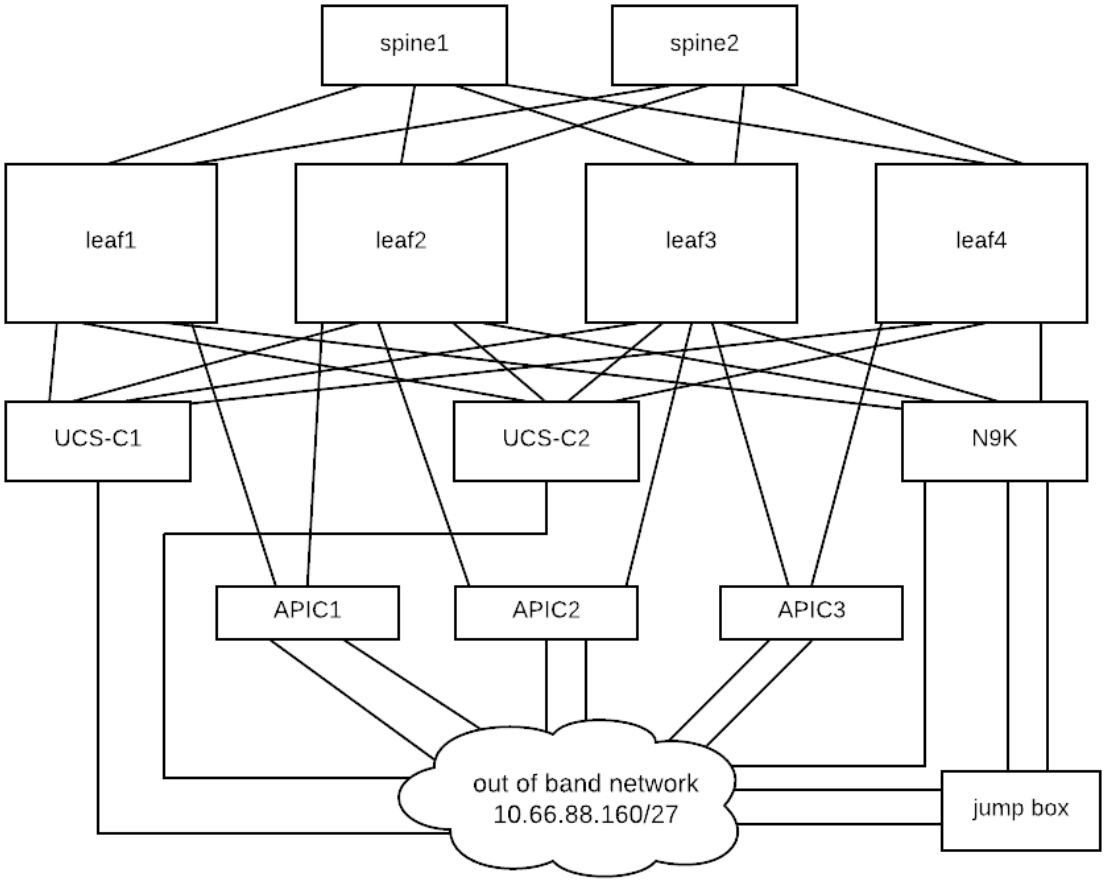
Contents

Welcome to ACI Troubleshoot Lab documentation

CHAPTER 1

Physical Topology

This lab documentation uses the following physical connectivity.



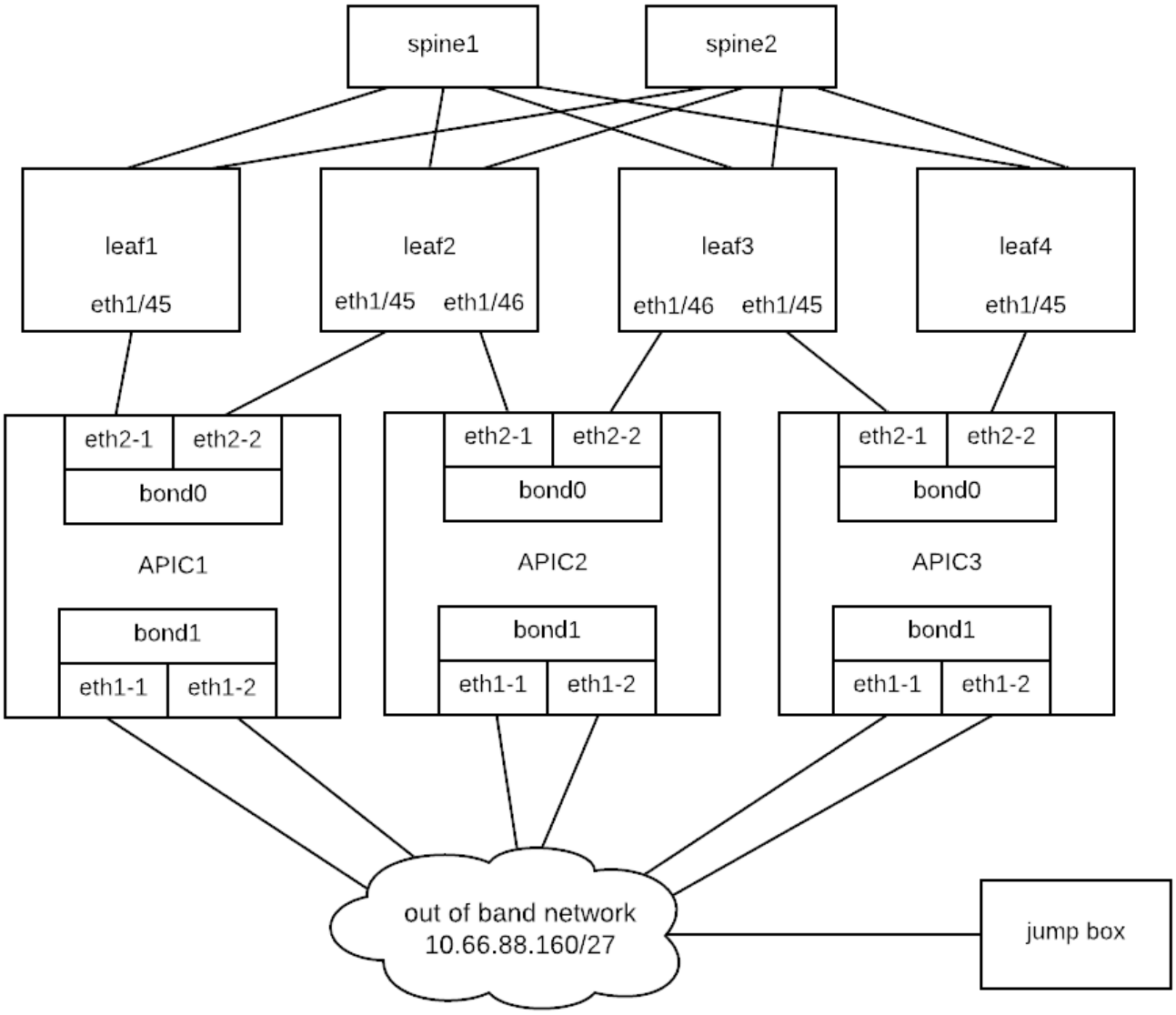
CHAPTER 2

Table of Contents

2.1 Fabric Discovery

The ACI fabric is brought up in a cascading manner, starting with the leaf nodes that are directly attached to the APIC. LLDP and control-plane IS-IS convergence occurs in parallel to this boot process. The ACI fabric uses LLDP- and DHCP-based fabric discovery to automatically discover the fabric switch nodes, assign the infrastructure VXLAN tunnel endpoint (VTEP) addresses.

2.1.1 APIC Cluster Connectivity



2.1.2 Erase Configuration

APIC Config Erase

To erase configuration of APIC so that we can re-setup APIC:

Sometimes KVM cannot launch because of Java issues. If you encounter such a problem, you can use Serial Over LAN as follows.

SSH to CIMC of the APIC:

```
ssh admin@<cimc IP addr>
```

Enable the Serial Over LAN (SoL):

```
cimc#
cimc# scope sol
cimc /sol # set enabled yes
cimc /sol *# set baud-rate 115200
cimc /sol *# commit
cimc /sol # connect host
CISCO Serial Over LAN:
Press Ctrl+x to Exit the session

Application Policy Infrastructure Controller
apic1 login: admin
Password:
Last login: Thu Mar 15 00:31:36 on tty1
apic# acidiag touch setup
apic# acidiag reboot
```

Switch Config Erase

To erase configuration of leaf/spine switch so that they can automatically retrieve configuration from APIC:

```
switch# acidiag touch clean
switch# reload
```

2.1.3 Fabric Initial Setup

Once the APIC is rebooted, it will start in the initial config wizard:

```
Starting Setup Utility


This setup utility will guide you through the basic configuration of
the system. Setup configures only enough connectivity for management
of the system.

*Note: setup is mainly used for configuring the system initially,
when no configuration is present. So setup always assumes system
defaults and not the current system configuration values.

Press Enter at anytime to assume the default values. Use ctrl-c
at anytime to restart from the beginning.


Cluster configuration ...
  Enter the fabric name [ACI Fabric1]: ACI Training
  Enter the fabric ID (1-128) [1]:
  Enter the number of controllers in the fabric (1-9) [3]:
  Enter the POD ID (1-9) [1]:
  Enter the controller ID (1-3) [1]:
  Enter the controller name [apic1]:
  Enter address pool for TEP addresses [10.0.0.0/16]:
```

(continues on next page)

(continued from previous page)

```
  Note: The infra VLAN ID should not be used elsewhere in your environment
        and should not overlap with any other reserved VLANs on other platforms.
  Enter the VLAN ID for infra network (2-4094): 4094
  Enter address pool for BD multicast addresses (GIPO) [225.0.0.0/15]:

Out-of-band management configuration ...
  Enable IPv6 for Out of Band Mgmt Interface? [N]:
  Enter the IPv4 address [192.168.10.1/24]: 10.66.88.181/27
  Enter the IPv4 address of the default gateway [None]: 10.66.88.161
  Enter the interface speed/duplex mode [auto]:

admin user configuration ...
  Enable strong passwords? [Y]: N
  Enter the password for admin:

  Reenter the password for admin:

Cluster configuration ...
  Fabric name: ACI Fabric1
  Fabric ID: 1
  Number of controllers: 3
  Controller name: apic1
  POD ID: 1
  Controller ID: 1
  TEP address pool: 10.0.0.0/16
  Infra VLAN ID: 4094
  Multicast address pool: 225.0.0.0/15

Out-of-band management configuration ...
  Management IP address: 10.66.88.181/27
  Default gateway: 10.66.88.161
  Interface speed/duplex mode: auto

admin user configuration ...
  Strong Passwords: N
  User name: admin
  Password: ********

The above configuration will be applied ...

Warning: TEP address pool, Infra VLAN ID and Multicast address pool
         cannot be changed later, these are permanent until the
         fabric is wiped.

Would you like to edit the configuration? (y/n) [n]:n
```

2.1.4 Configuration Verification

Ensure the bond interace is up

Check which active interface is connected to the leaf:

```
apic1# cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v3.7.1 (April 27, 2011)
```

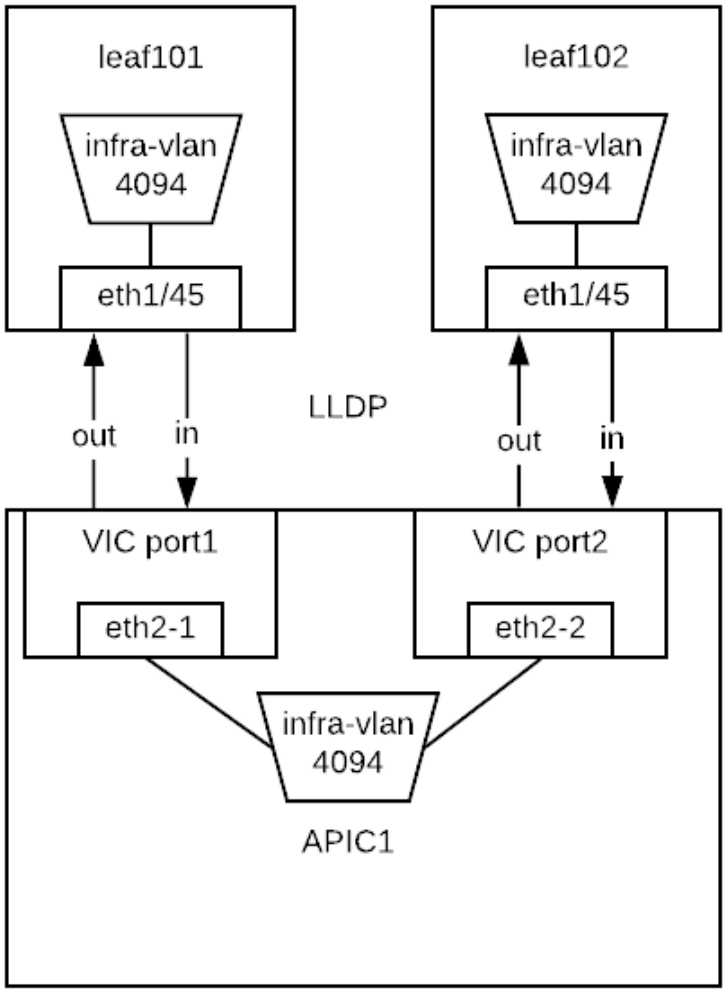
(continues on next page)

(continued from previous page)

```
Bonding Mode: fault-tolerance (active-backup)
Primary Slave: None
Currently Active Slave: eth2-1 <<< Check the active interface
MII Status: up
MII Polling Interval (ms): 60
Up Delay (ms): 0
Down Delay (ms): 0

Slave Interface: eth2-1
MII Status: up <<< Ensure the bond member interface is up
Speed: 10000 Mbps
Duplex: full
Link Failure Count: 0
Permanent HW addr: d8:b1:90:61:30:74
Slave queue ID: 0

Slave Interface: eth2-2
MII Status: up <<< Ensure the bond member interface is up
Speed: 10000 Mbps
Duplex: full
Link Failure Count: 0
Permanent HW addr: d8:b1:90:61:30:75
Slave queue ID: 0
```

Ensure the lldp information is correct

Leaf switch discovers the attached APIC via LLDP and requests a TEP address via DHCP from the APIC.



Check the incoming lldp information that APIC receives from Leaf switch:

```
apic1# acidiag run lldptool in eth2-1 | grep topo
  topology/pod-1/paths-101/pathep-[eth1/45]
  topology/pod-1/node-101

apic1# acidiag run lldptool in eth2-2 | grep topo
  topology/pod-1/paths-102/pathep-[eth1/45]
  topology/pod-1/node-102

apic1# acidiag run lldptool in eth2-1 | grep -A 1 -i vlan
Cisco Infra VLAN TLV
  4094

apic1# acidiag run lldptool in eth2-2 | grep -A 1 -i vlan
Cisco Infra VLAN TLV
  4094
```

Check the outgoing lldp information that APIC sends to Leaf switch:

```
apic1# acidiag run lldptool out eth2-1 | grep topo
  topology/pod-1/node-1

apic1# acidiag run lldptool out eth2-2 | grep topo
  topology/pod-1/node-1

apic1# acidiag run lldptool out eth2-1 | grep -A 1 -i vlan
Cisco Infra VLAN TLV
  4094

apic1# acidiag run lldptool out eth2-2 | grep -A 1 -i vlan
Cisco Infra VLAN TLV
  4094
```

Check the lldp neighbours on connected Leaf:

```
leaf101# show lldp neighbors
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
apic1                Eth1/45         120                    eth2-1 <<< apic1 is a␣
↪LLDP neighbor
spine201             Eth1/53         120        BR          Eth1/29
spine202             Eth1/54         120        BR          Eth1/29
Total entries displayed: 3
```

Ensure that the infra VLANs on APIC and Leaf match. If they do not match, please run the following to reset switch to manufacture config (bug CSCvd67346). Use prepare-mfg.sh on all switches in the environment and reload at the same time. For example:

```
leaf101# dir bootflash/
aci-n9000-dk9.12.1.2e.bin

leaf101# prepare-mfg.sh aci-n9000-dk9.12.1.2e.bin
```

If the incoming LLDP is empty (shown below), that means the VIC port has consumed the LLDP and the APIC port does not receive it. The reason is that the LLDP is enabled on VIC card. We need to disable the LLDP on the VIC card so that the LLDP information is passed to the APIC port (eth2-1).

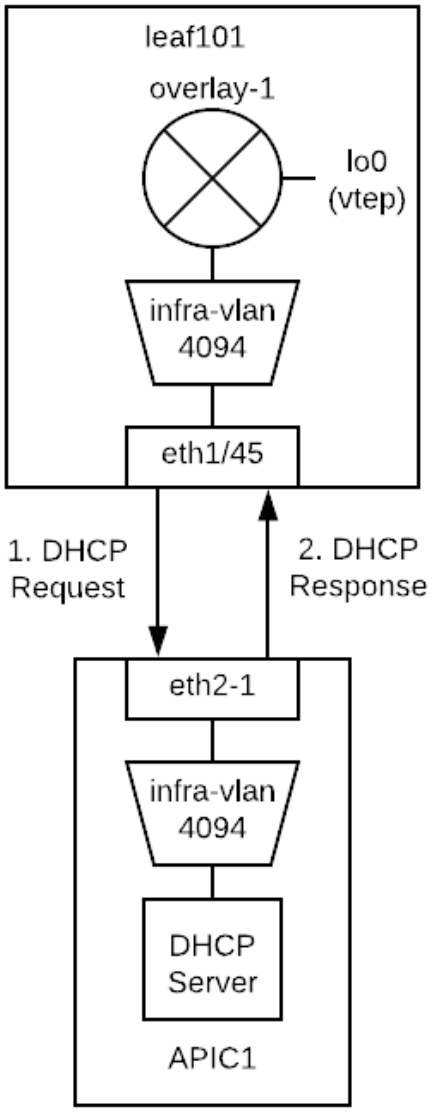
```
 apic1# acidiag run lldptool in eth2-1

 apic1#

leaf101# show lldp neighbors
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
d8b1.9061.3071       Eth1/45         120                    d8b1.9061.3075  <<< The␣
↪device is shown as mac address instead of APIC hostname.
spine201             Eth1/53         120        BR          Eth1/29
spine202             Eth1/54         120        BR          Eth1/29
Total entries displayed: 3
```

To disable LLDP on VIC, SSH as user admin to CIMC of the APIC:

```
CIMC# scope chassis
CIMC /chassis # show adapter
PCI Slot Product Name Serial Number Product ID Vendor
-------- -------------- -------------- -------------- --------------------
1 UCS VIC 1225 FCHxxxxxxxx UCSC-PCIE-C... Cisco Systems Inc
CIMC /chassis # scope adapter 1
CIMC /chassis/adapter # show detail | grep LLDP
LLDP: Enabled
CIMC /chassis/adapter # set lldp disabled
CIMC /chassis/adapter *# commit
New VNIC adapter settings will take effect upon the next server reset
CIMC /chassis/adapter # exit
CIMC /chassis # power cycle
```

Ensure that the VTEP is assigned to the leaf switch

When leaf is registered, it will request VTEP address for loopback0 interface via DHCP.



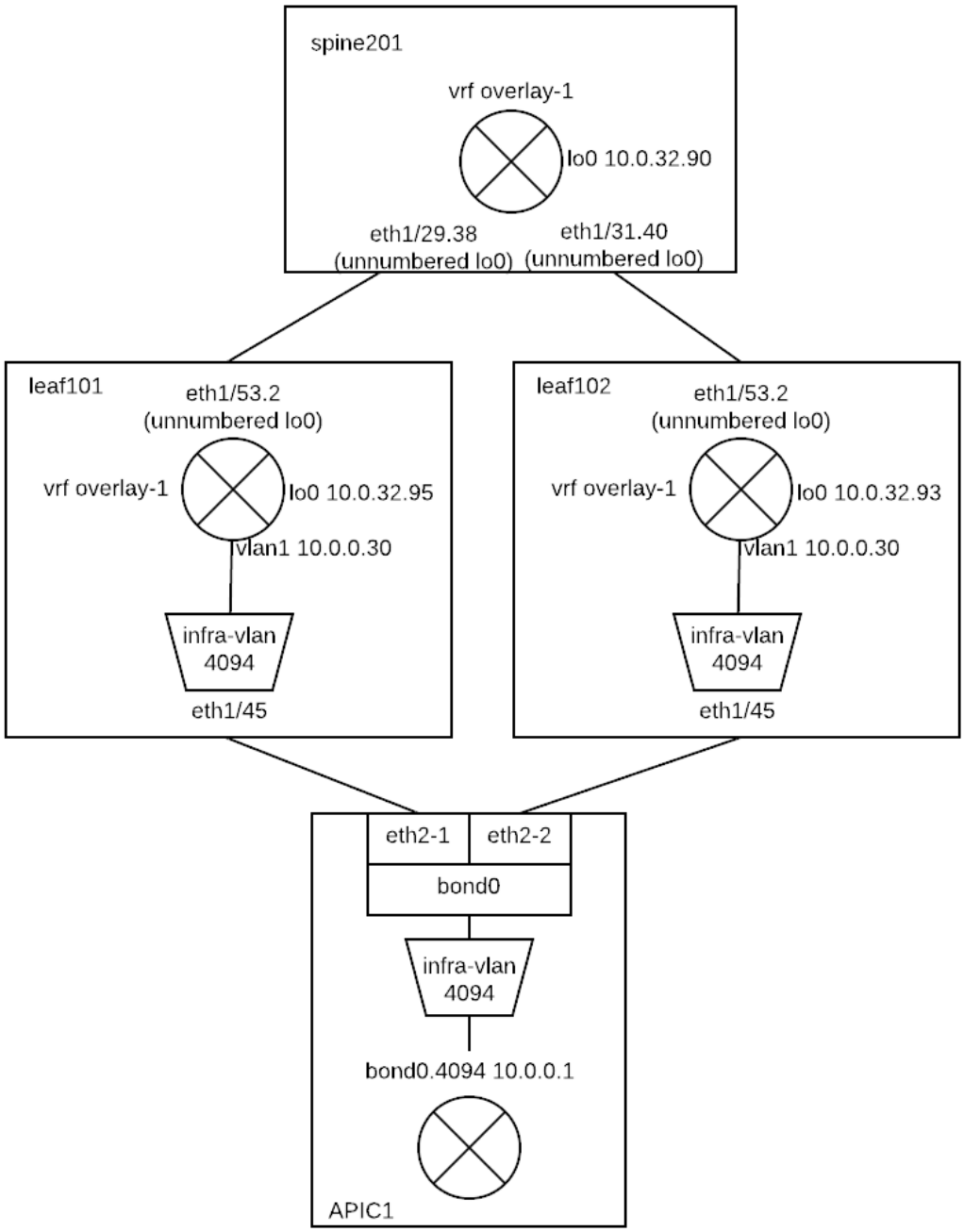

```
leaf101# show ip interface brief vrf overlay-1
IP Interface Status for VRF "overlay-1"(4)
Interface            Address              Interface Status
eth1/49              unassigned           protocol-down/link-down/admin-up
eth1/50              unassigned           protocol-down/link-down/admin-up
eth1/51              unassigned           protocol-down/link-down/admin-up
eth1/52              unassigned           protocol-down/link-down/admin-up
eth1/53              unassigned           protocol-up/link-up/admin-up
eth1/53.2            unnumbered           protocol-up/link-up/admin-up
                     (lo0)
eth1/54              unassigned           protocol-up/link-up/admin-up
eth1/54.3            unnumbered           protocol-up/link-up/admin-up
                     (lo0)
vlan1                10.0.0.30/27         protocol-up/link-up/admin-up
lo0                  10.0.32.95/32        protocol-up/link-up/admin-up <<< VTEP
```

(continues on next page)

(continued from previous page)

```
lo1023                  10.0.0.32/32          protocol-up/link-up/admin-up
```

Once all switches are registered, we can see their VTEPs (loopback lo0 interfaces):




```
leaf101# acidiag fnvread
      ID   Pod ID                 Name    Serial Number         IP Address    Role␣
↪     State   LastUpdMsgId
```

(continues on next page)

(continued from previous page)

```
--------------------------------------------------------------------------------
↪-------------------------
     101        1              leaf101      FDO20231J7L      10.0.32.95/32    leaf
↪      active   0
     102        1              leaf102      SAL1946SWK8      10.0.32.93/32    leaf
↪      active   0
     103        1              leaf103      SAL1946SWNT      10.0.32.92/32    leaf
↪      active   0
     104        1              leaf104      SAL1946SWNU      10.0.32.91/32    leaf
↪      active   0
     201        1             spine201                       10.0.32.90/32   spine
↪      active   0
     202        1             spine202      SAL18391DXP      10.0.32.94/32   spine
↪      active   0

Total 6 nodes
```

Also we can see the Dynamic Tunnel End Points are created in IS-IS:

```
leaf101# show isis dteps vrf overlay-1

IS-IS Dynamic Tunnel End Point (DTEP) database:
DTEP-Address       Role    Encapsulation   Type
10.0.64.64         SPINE   N/A             PHYSICAL,PROXY-ACAST-V4
10.0.64.65         SPINE   N/A             PHYSICAL,PROXY-ACAST-MAC
10.0.64.66         SPINE   N/A             PHYSICAL,PROXY-ACAST-V6
10.0.32.93         LEAF    N/A             PHYSICAL
10.0.32.92         LEAF    N/A             PHYSICAL
10.0.32.91         LEAF    N/A             PHYSICAL
10.0.32.90         SPINE   N/A             PHYSICAL
10.0.32.94         SPINE   N/A             PHYSICAL
```

The gateway of the APIC to reach other VTEPs is 10.0.0.30.

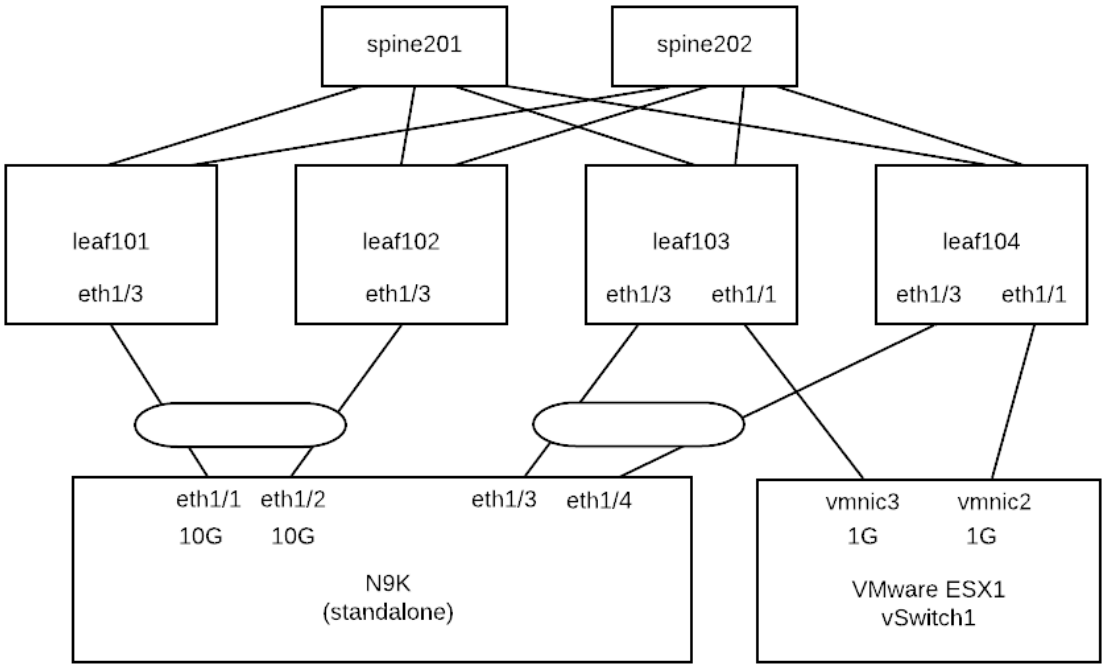
```
apic1# netstat -rn
Kernel IP routing table
Destination     Gateway         Genmask         Flags   MSS Window  irtt Iface
0.0.0.0         10.66.88.161    0.0.0.0         UG        0 0          0 oobmgmt
10.0.0.0        10.0.0.30       255.255.0.0     UG        0 0          0 bond0.4094
10.0.0.30       0.0.0.0         255.255.255.255 UH        0 0          0 bond0.4094
10.0.64.64      10.0.0.30       255.255.255.255 UGH       0 0          0 bond0.4094
10.0.64.65      10.0.0.30       255.255.255.255 UGH       0 0          0 bond0.4094
10.66.88.160    0.0.0.0         255.255.255.224 U         0 0          0 oobmgmt
169.254.1.0     0.0.0.0         255.255.255.0   U         0 0          0 teplo-1
169.254.254.0   0.0.0.0         255.255.255.0   U         0 0          0 lxcbr0
apic1#
```

2.1.5 Reference

1. Disable LLDP on VIC https://supportforums.cisco.com/legacyfs/online/attachments/document/files/apic-vic-lldp-fn.pdf
2. CNA Data Center DCICT 200-155 Official Cert Guide by Ahmed Afrose et. al.

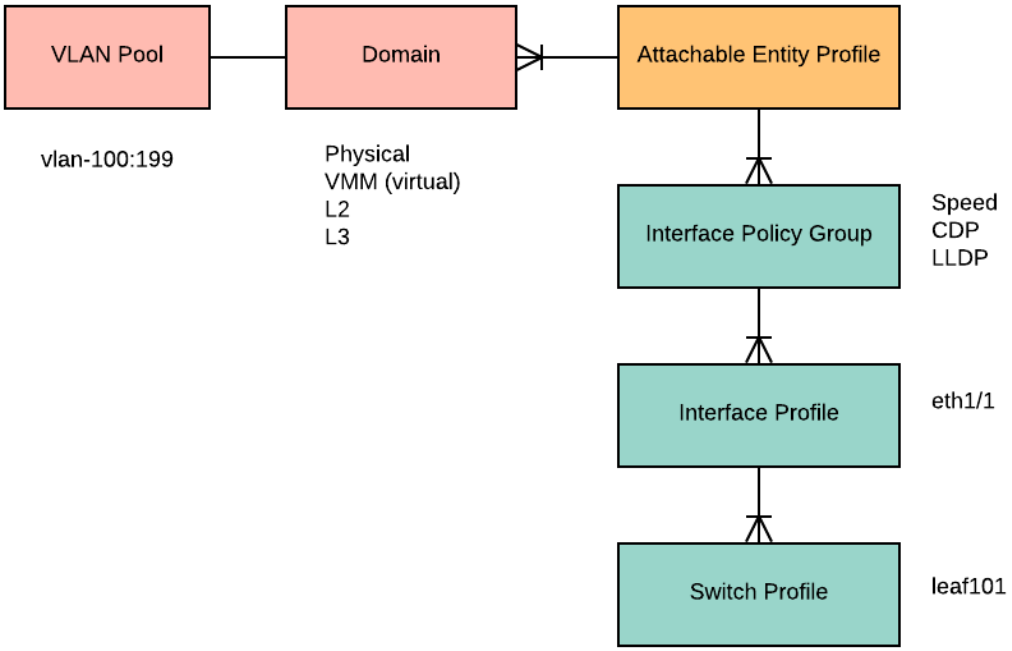
2.2 Access Policies

2.2.1 Lab Topology



2.2.2 Access Policies for leaf front panel ports

Access policies define the connectivity from external devices to ACI leaf switches such as interfaces, VLANs, CDP, LLDP, etc.



Attachable Entiry Profile is used for linking many to many relationships between Domains and Interface Policy Groups.

A domain determines the type of bridge domain that is deployed to the leaf port.

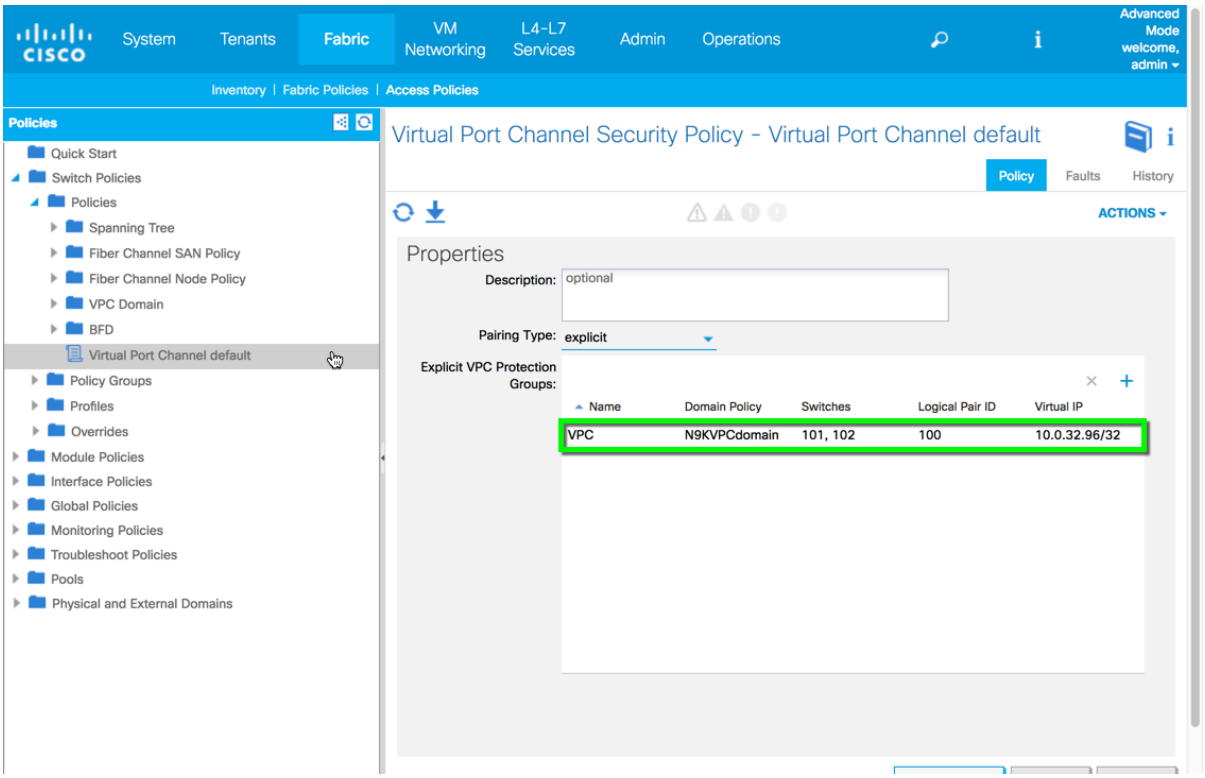
Interface Policy Group

There are several types of interface policy groups:

- Physical Ports
- Port Channels
- Virtual Port Channels

VPC

First of all, we will need to create a VPC domain for a pair of leaf switches:



Note: See the below reference for a VPC config guide

To find out which Interface Policy Group is used for a VPC:

```
leaf101# show vpc extended
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 100
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : Disabled
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 1
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Enabled (timeout = 240 seconds)
Operational Layer3 Peer           : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1           up     -

vPC status
----------------------------------------------------------------------------------
id   Port   Status Consistency Reason                Active vlans Bndl Grp Name
```

(continues on next page)

(continued from previous page)

```
--   ----   ------ ----------- ------                --------------- -----------------
1    Po1    up     success     success               -               UplinkForN9KVPC1 <<<␣
↪Interface Policy Group

leaf101# show port-channel extended
Flags:  D - Down        P - Up in port-channel (members)
        I - Individual  H - Hot-standby (LACP only)
        s - Suspended   r - Module-removed
        S - Switched    R - Routed
        U - Up (port-channel)
        M - Not in use. Min-links not met
        F - Configuration failed
------------------------------------------------------------------------------
Group Port-       BundleGrp                Protocol  Member Ports
      Channel
------------------------------------------------------------------------------
1     Po1(SU)     UplinkForN9KVPC1         NONE      Eth1/3(P)
```

To check LACP messages:

```
leaf101# show lacp int e1/3 | grep -i pdu
  PDUs sent: 10
  PDUs rcvd: 0
```

The leaf101 does not receive and LACP PDUs.

2.2.3 Common Problems

- Speed mismatch
- MCP - Duplicate VLAN
- A VPC policy group represent 1 virtual port-channel.
- LACP Mismatch

2.2.4 Reference

- VPC config guide https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/Operating_ACI/guide/b_Cisco_Operating_ACI/b_Cisco_Operating_ACI_chapter_0110.html

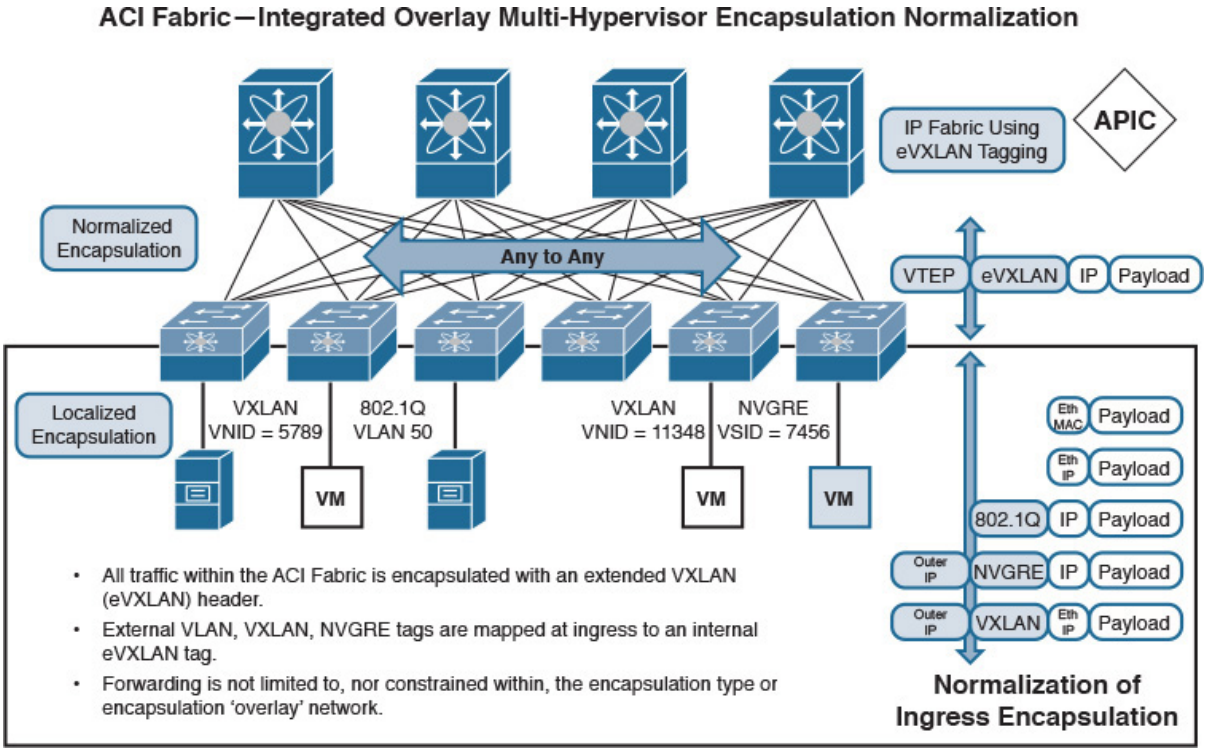
2.3 End Point Group

So far all the fabric nodes have been discovered (VTEPs are assigned) and access policies have been created (port speed, cdp, lldp and other leaf port properties). Now, we are ready to assign the ownership of leaf ports to EPGs.

EPG classification can be based on:

- Access (untagged) = Access VLAN
- Source IP address
- Trunk = Trunk
- Access (802.1p) = Native VLAN

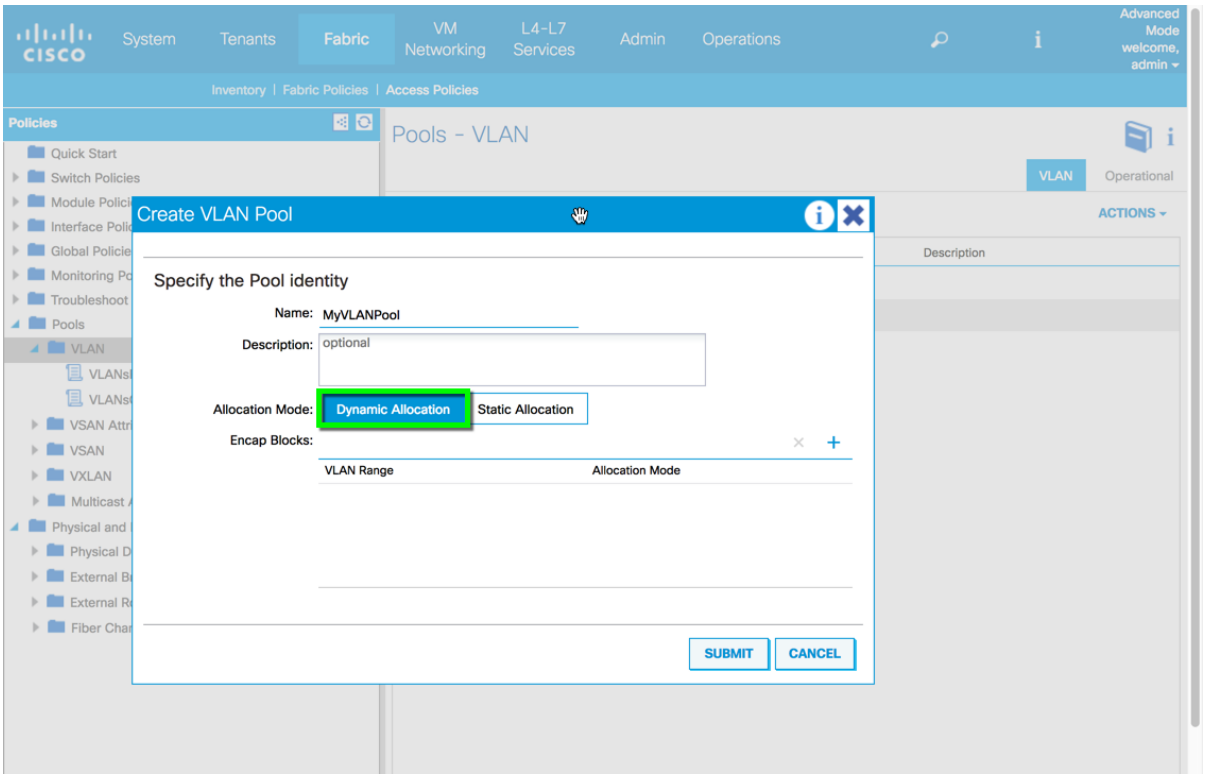
- NVGRE
- VXLAN



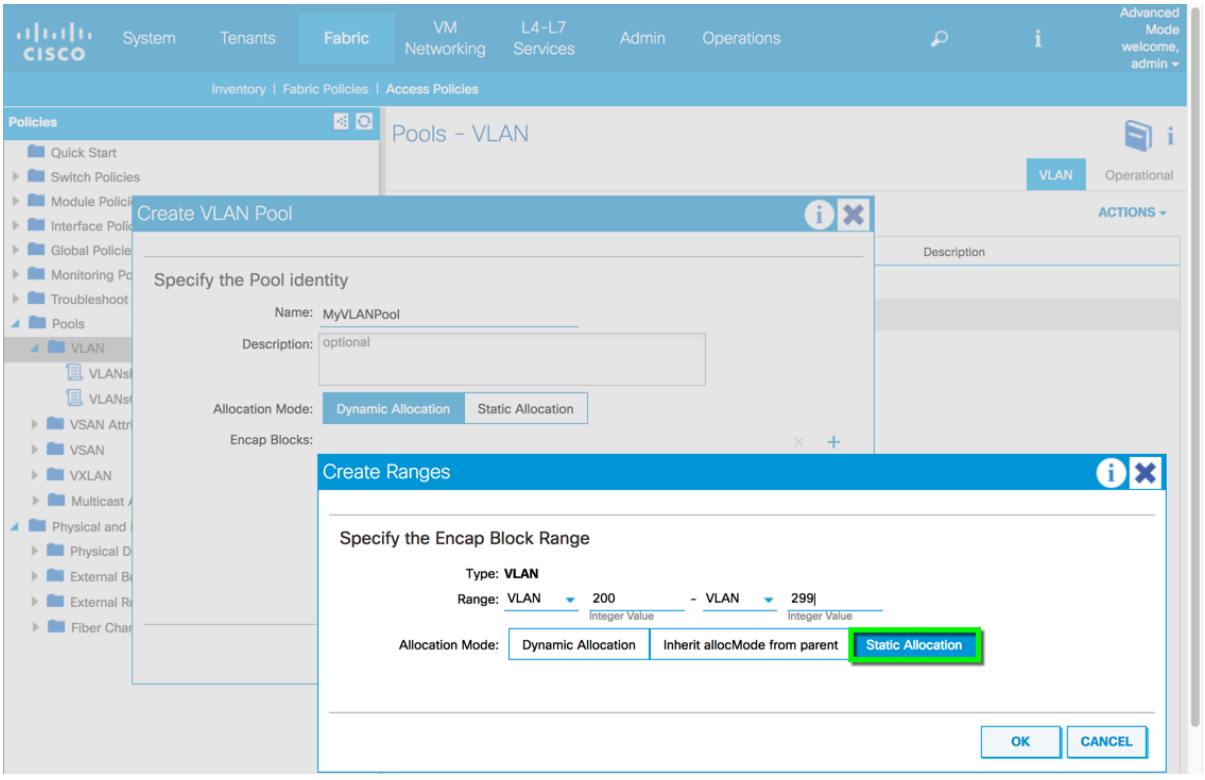
In this lab, we will use VLAN as an EPG classifier. Therefore, we will need to create a VLAN pool

2.3.1 VLAN Pool Creation

When you create a VLAN pool, it is a good practice to set allocation mode to dynamic.



Then when you add an encap block, you can choose either static or dynamic. In that way, you will have flexibility to add both dynamic and static encap blocks. Dynamic encap blocks are used for Virtual Machine Manager (VMM) domain.

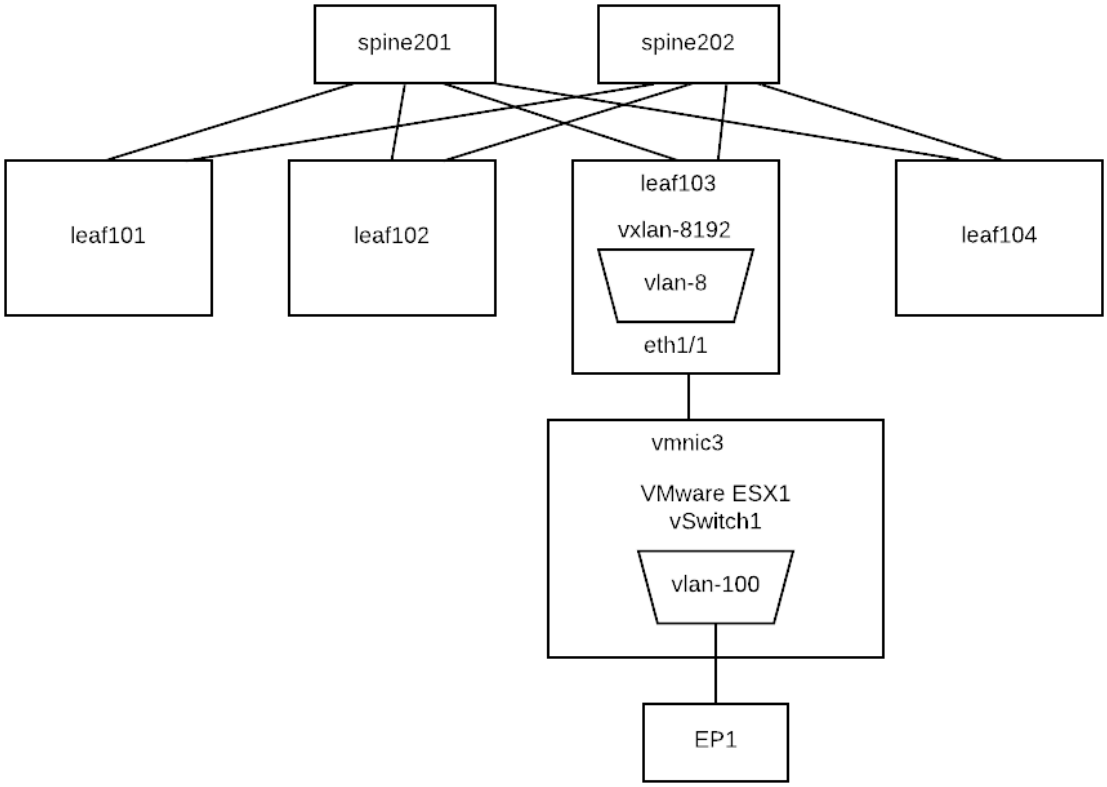


Make sure VLAN pools do not have overlapping vlans. The reason is that ACI floods STP Bridge Protocol Data Units

(BPDUs) to the VXLAN network identifier (VNID) assigned to the FD VLAN. VNID is assigned through the VLAN pool so encapsulation has to be part of same VLAN pool to be in part of same STP domain. Otherwise STP BPDU can be dropped by ACI.

2.3.2 Static binding

We can statically classify EPG by assigning an encap vlan on leaf ports. In below example, port eth1/1 on leaf103 is statically bound with encap vlan 100. That means any incoming traffic with vlan tag 100 is classified as EPG "tshoot-epg".




```
leaf103# show endpoint
Legend:
 s - arp              O - peer-attached    a - local-aged       S - static
 V - vpc-attached     p - peer-aged        M - span             L - local
 B - bounce           H - vtep
+-----------------------------------+---------------+-----------------+--------------
↪+-------------+
      VLAN/                           Encap           MAC Address       MAC Info/
↪  Interface
      Domain                          VLAN            IP Address        IP Info
+-----------------------------------+---------------+-----------------+--------------
↪+-------------+
8                                       vlan-100     0050.5696.609a L
↪       eth1/1
tshoot:tshoot-vrf                       vlan-100      192.168.1.101 L
↪       eth1/1
```

(continues on next page)

(continued from previous page)

```
overlay-1                                                    10.0.32.92 L
↪          lo0
3/overlay-1                          vxlan-16777209    d8b1.9061.1e65 L
↪      eth1/46


leaf103# show vlan id 8

 VLAN Name                                 Status    Ports
 ---- -------------------------------- --------- -------------------------------
 8    tshoot:tshoot-ap:tshoot-epg      active    Eth1/1

 VLAN Type  Vlan-mode
 ---- ----- ----------
 8    enet  CE

 leaf103# show system internal epm vlan 8


+----------+---------+-----------------+----------+------+----------+-----------
   VLAN ID     Type      Access Encap     Fabric    H/W id  BD VLAN    Endpoint
                         (Type Value)     Encap                         Count
+----------+---------+-----------------+----------+------+----------+-----------
 8            FD vlan 802.1Q        100 8192       7      7          1
```

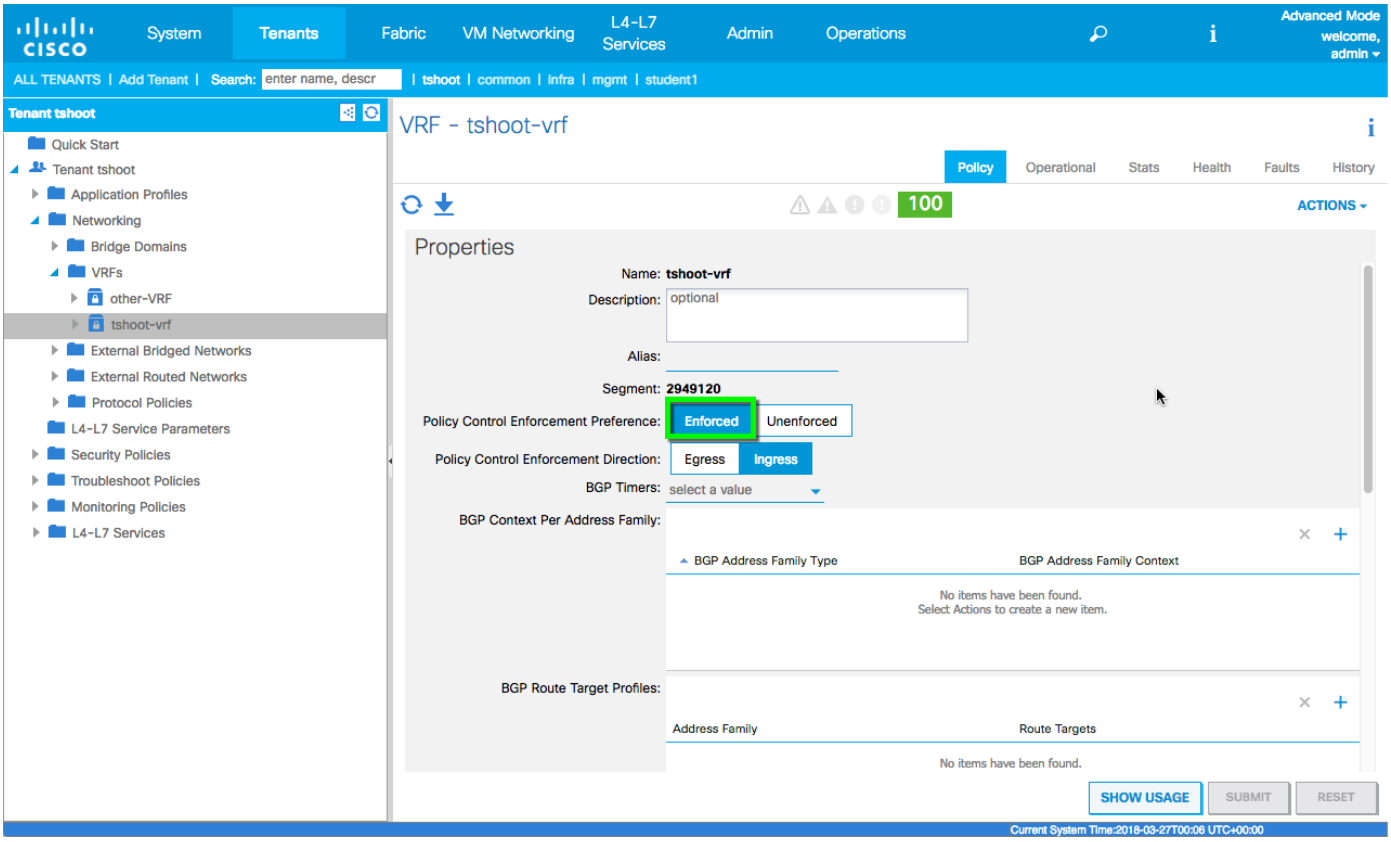
In above example, encap vlan-100 has been mapped to ACI platform independent (PI) vlan 8 which is mapped to vxlan-8192.

2.3.3 Reference

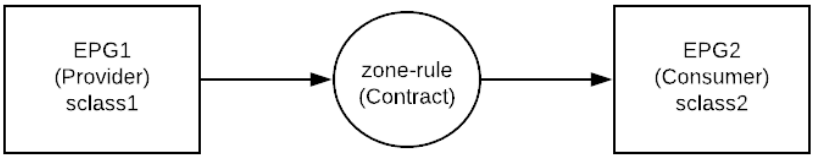
1. CNA Data Center DCICT 200-155 Official Cert Guide by Ahmed Afrose et. al.
2. ACI Operation with L2 Switches and Spanning Tree Link Types https://www.cisco.com/c/en/us/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/211236-ACI-operation-with-L2-switches-and-Spann.html

2.4 Contract

The requirement for contracts to be applied in the zoning-rule, the VRF must be inthe enforced mode.



In order for different EPGs to be able to communicate, they must have a contract. Server provides the contract and Client consumes a contract.



Each EPG has a source class. To find out the source class of an EPG:

For example, the vlan encap for an EPG is vlan-100.

```
leaf103# show vlan extended | grep vlan-100
 8     enet  CE         vlan-100

        leaf103# show vlan id 8

 VLAN Name                             Status    Ports
 ---- -------------------------------- --------- -------------------------------
 8    tshoot:tshoot-ap:A-epg           active    Eth1/3, Po1

 VLAN Type  Vlan-mode
 ---- ----- ----------
 8    enet  CE
```

In above example, encap vlan-100 has been mapped to ACI platform independent (PI) vlan 8.

Let us check the sclass ID that will be used in zone-rule (contract):

```
    leaf103# vsh_lc
    module-1# show system internal eltmc info vlan 8


          vlan_id:              8   :::      hw_vlan_id:             39
        vlan_type:        FD_VLAN   :::         bd_vlan:              7
access_encap_type:         802.1q   :::    access_encap:            100
         isolated:              0   :::   primary_encap:              0
fabric_encap_type:           VXLAN   :::    fabric_encap:           8192
           sclass:          49156   :::           scope:              4
          bd_vnid:           8192   :::        untagged:              0
  acess_encap_hex:           0x64   :::  fabric_enc_hex:         0x2000
  pd_vlan_ft_mask:           0x4f
 fd_learn_disable:              0
    bcm_class_id:             16   :::  bcm_qos_pap_id:           1024
      qq_met_ptr:             18   :::       seg_label:              0
   ns_qos_map_idx:              0   :::  ns_qos_map_pri:              1
  ns_qos_map_dscp:              0   :::   ns_qos_map_tc:              0
    vlan_ft_mask:         0x7830

   NorthStar Info:
       qq_tbl_id:           1441   :::         qq_ocam:              0
  seg_stat_tbl_id:              0   :::        seg_ocam:              0
  ::::
```

We can see that the source class (sclass) is 49156 for EPG tshoot:tshoot-ap:A-epg.

To know the sclass of an external EPG of L3out:

```
leaf103# vsh_lc
module-1# show system internal aclqos prefix | grep 2949120
Vrf Vni Addr           Mask     Scope Class  Shared Remote

======= ============== ======== ===== ====== ====== ======
2949120 0::/0 0::/0 4     15     FALSE FALSE
2949120 0.0.0.0        ffffffff 4     15     FALSE FALSE
2949120 9.9.9.9        0        4     16388  FALSE FALSE
```

From above, we can see that ip address 9.9.9.9/32 has sclass 16388.

Contracts exist in VRF. To know the VRF ID, you can run the following command:

```
leaf103# show system internal epm vrf all


+--------------------------------+--------+----------+----------+------+--------
               VRF                  Type    VRF vnid  Context ID Status Endpoint
                                                                      Count
+--------------------------------+--------+----------+----------+------+--------
 black-hole                       Tenant   16777200   3          Up     0
 tshoot:tshoot-vrf                Tenant   2949120    6          Up     1
 overlay-1                        Infra    16777199   4          Up     2
```

To check the zoning rule of a contract that is applied:

```
leaf103# show zoning-rule scope 2949120 | grep 49156
4186           16387           49156           6               enabled
↪2949120         permit                                  fully_qual(6)
```

(continues on next page)

(continued from previous page)

```
leaf103# show zoning-filter filter 6
FilterId  Name          EtherT      ArpOpc      Prot        MatchOnlyFrag Stateful␣
↪SFromPort   SToPort     DFromPort   DToPort     Prio        Icmpv4T     Icmpv6T     ␣
↪TcpRules
========  ==========   ======      =========   =======     ======        =======  ␣
↪=======     ====        ====        ====        =========   =======     ========   ␣
↪========
6         6_0           ip          unspecified tcp         no            no       ␣
↪unspecified unspecified http       http        dport       unspecified unspecified
```

To get the hit statistics of a particular filter:

```
leaf103# show system internal policy-mgr stats | grep 4186
Rule (4186) DN (sys/actrl/scope-2949120/rule-2949120-s-16387-d-49156-f-6) Ingress: 0,␣
↪Egress: 0, Pkts: 0  RevPkts: 0
```

To check whether policy enforcement process denies:

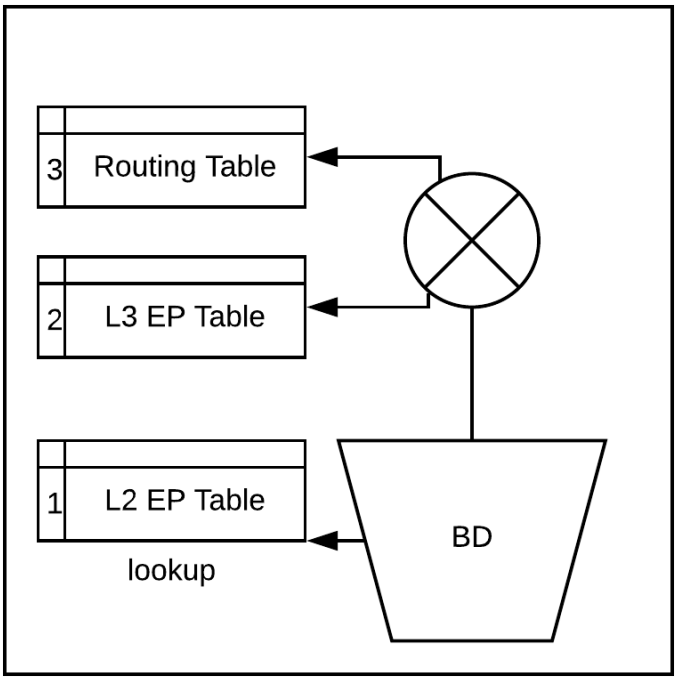
```
leaf103# show logging ip access-list internal packet-log deny
[ Wed Mar 21 00:10:53 2018 434710 usecs]: CName: tshoot:tshoot-vrf(VXLAN: 2949120),␣
↪VlanType: FD_VLAN, Vlan-Id: 5, SMac: 0x641225750331, DMac:0x0022bdf819ff, SIP: 9.9.
↪9.9, DIP: 192.168.200.254, SPort: 0, DPort: 0, Src Intf: port-channel1, Proto: 1,␣
↪PktLen: 98
```

In above example, a packet with source IP 9.9.9.9 is denied to access destination IP 192.168.200.254

2.4.1 Reference

1. Verify Contracts and Rules in the ACI Fabric https://www.cisco.com/c/en/us/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/119023-technote-apic-00.pdf
2. Working with Contracts https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/Operating_ACI/guide/b_Cisco_Operating_ACI/b_Cisco_Operating_ACI_chapter_01000.pdf

2.5 End Point Learning



In ACI, the lookup is done in this order:

1. L2 EP table
2. If it is a routed traffic, L3 EP table
3. If not found in the L3 EP table, routing table

Let us look at this below example.

First, find out the platform independent vlan that is mapped for vlan-199.

```
leaf103# show vlan extended | grep vlan-199
 5    enet  CE         vlan-199
```

We can see that vlan-199 is mapped to vlan 5.

Let us confirm that vlan 5 corresponds to the correct EPG:

```
leaf103# show vlan id 5

 VLAN Name                              Status    Ports
 ---- --------------------------------- --------- -------------------------------

 5    tshoot:tshoot-ap:B-epg            active    Eth1/3, Po1

 VLAN Type  Vlan-mode
 ---- ----- ----------
 5    enet  CE
```

To check the content of end point table for the EPG:

```
leaf103# show endpoint vlan 5
Legend:
 s - arp              O - peer-attached    a - local-aged       S - static
 V - vpc-attached     p - peer-aged        M - span             L - local
 B - bounce           H - vtep
+------------------------------------+---------------+-----------------+--------------
↪+-------------+
      VLAN/                           Encap           MAC Address       MAC Info/
↪  Interface
      Domain                          VLAN            IP Address        IP Info
+------------------------------------+---------------+-----------------+--------------
↪+-------------+
5/tshoot:tshoot-vrf                        vlan-199    6412.2575.0334 LpV
↪          po1
5                                          vlan-199    6412.2575.0331 LpV
↪          po1
tshoot:tshoot-vrf                          vlan-199     192.168.199.2 LV
↪          po1


+------------------------------------------------------------------------------+
                            Endpoint Summary
+------------------------------------------------------------------------------+
Total number of Local Endpoints     : 2
Total number of Remote Endpoints    : 0
Total number of Peer Endpoints      : 0
Total number of vPC Endpoints       : 2
Total number of non-vPC Endpoints   : 0
Total number of MACs                : 2
Total number of VTEPs               : 0
Total number of Local IPs           : 1
Total number of Remote IPs          : 0
Total number All EPs                : 2
```

To show more details about a local end point:

```
leaf103# show system internal epm endpoint mac 6412.2575.0331

MAC : 6412.2575.0331 ::: Num IPs : 1
IP# 0 : 192.168.199.2 ::: IP# 0 flags :
Vlan id : 5 ::: Vlan vnid : 8291 ::: VRF name : tshoot:tshoot-vrf
BD vnid : 16285610 ::: VRF vnid : 2949120
Phy If : 0x16000000 ::: Tunnel If : 0
Interface : port-channel1
Flags : 0x80005c25 ::: sclass : 16387 ::: Ref count : 5
EP Create Timestamp : 03/20/2018 21:27:35.632579
EP Update Timestamp : 03/20/2018 21:54:44.324243
EP Flags : local|vPC|peer-aged|IP|MAC|host-tracked|sclass|timer|
::::
```

To show the routing table:

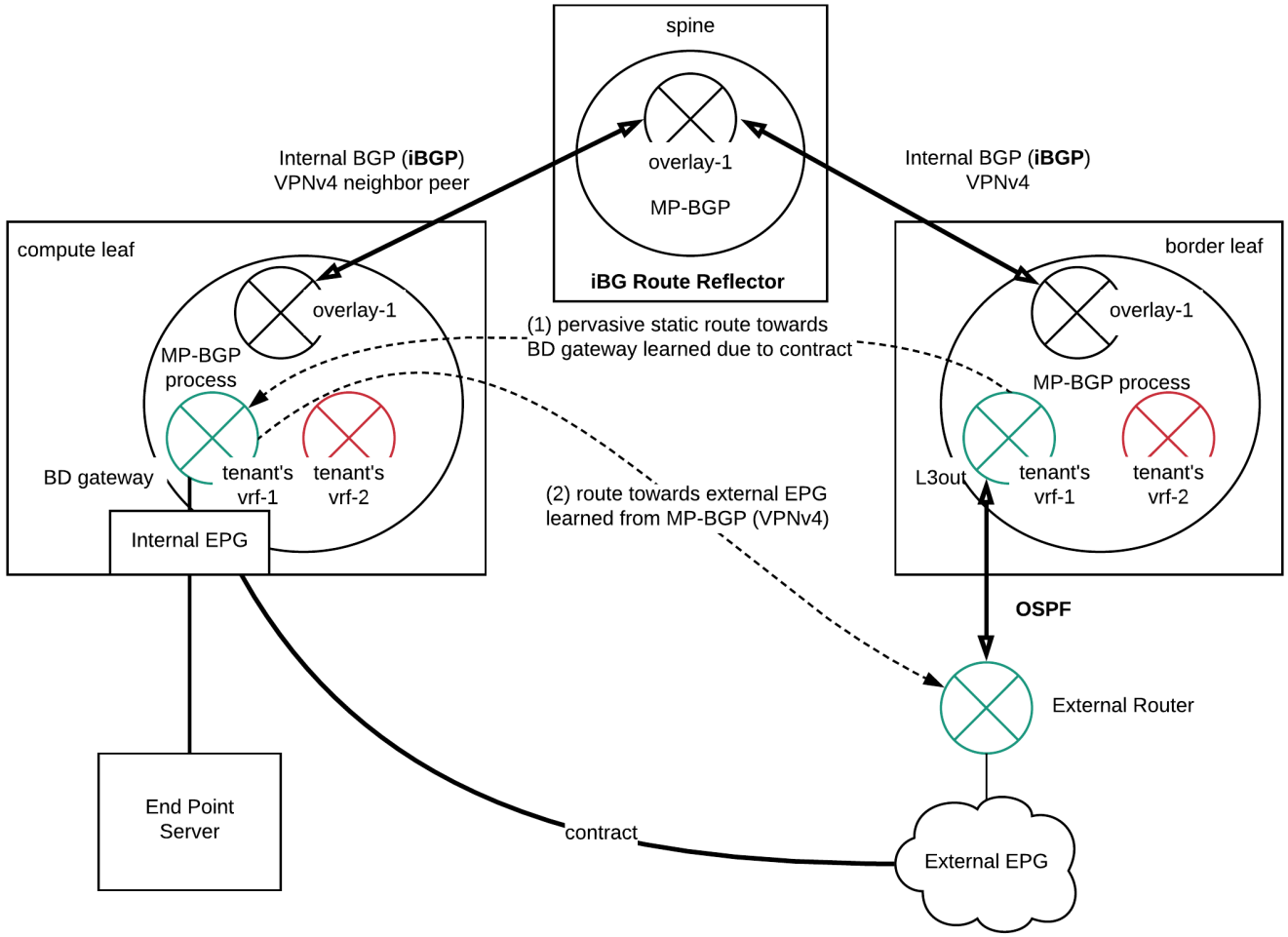
```
leaf103# show ip route vrf tshoot:tshoot-vrf
IP Route Table for VRF "tshoot:tshoot-vrf"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
```

(continues on next page)

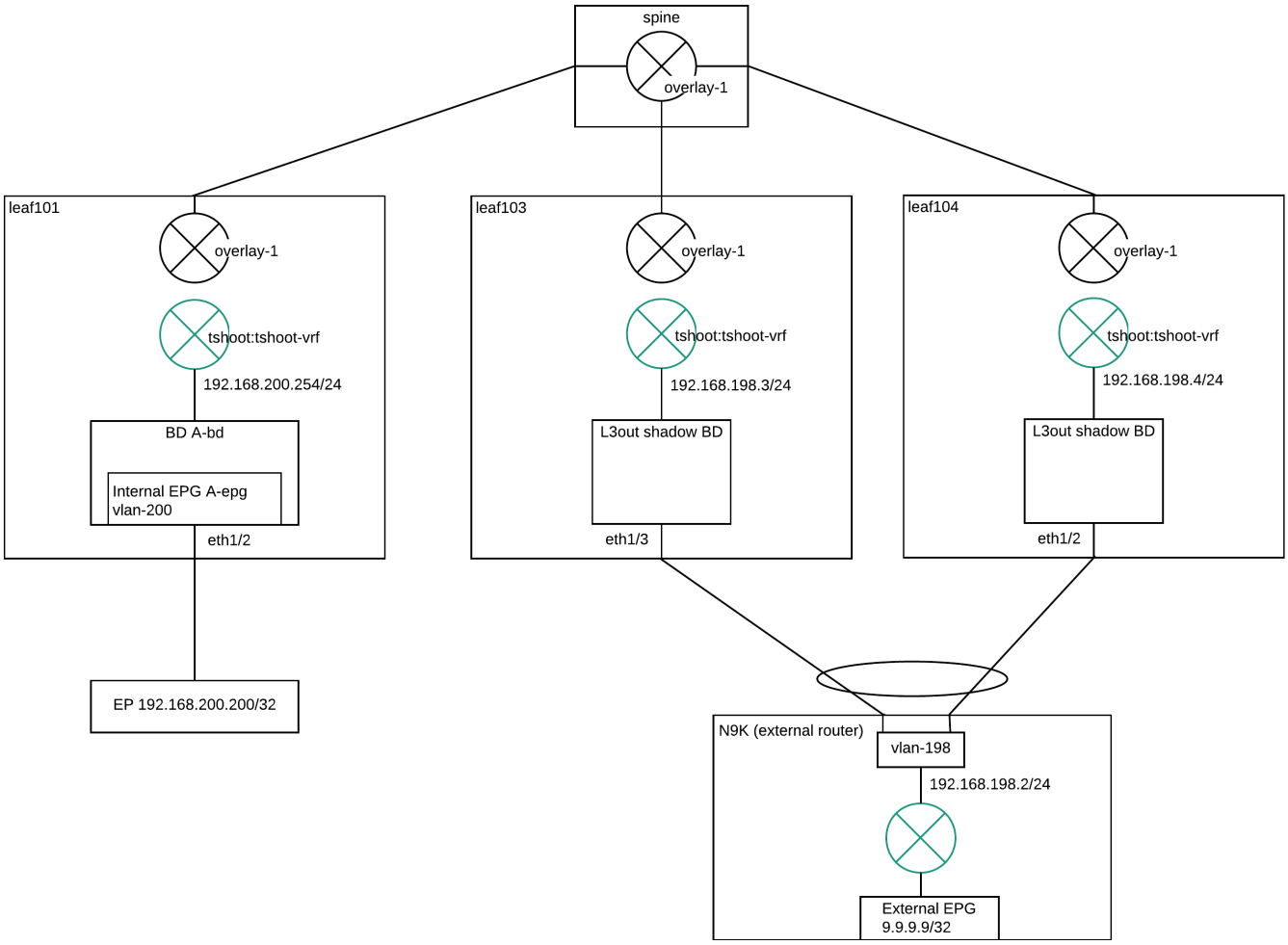
(continued from previous page)

```
'%<string>' in via output denotes VRF <string>

192.168.199.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.64.64%overlay-1, [1/0], 00:35:37, static, tag 4294967295
192.168.199.254/32, ubest/mbest: 1/0, attached, pervasive
    *via 192.168.199.254, vlan4, [1/0], 00:35:37, local, local
192.168.200.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.64.64%overlay-1, [1/0], 00:00:06, static, tag 4294967295
```

2.6 L3out



2.6.1 Lab Topology



2.6.2 Troubleshooting

L3out configuration checklist:

- L3out - External EPG
- Contract between Internal EPG and External EPG
- BD subnet advertised externally and L3out association
- Fabric Policy – BGP Route Reflector

L3out - External EPG

The indication that the external EPG has been correctly configured is the L3out shadow BD is deployed to the border leaf switches.

```
leaf103# show vlan extended | grep vlan-198
VLAN Name                             Status    Ports
```

(continues on next page)

(continued from previous page)

```
---- -------------------------------- --------- --------------------------------
9    enet  CE         vxlan-14974940, vlan-198
```

VLAN 198 has been mapped to platform independent VLAN 9 on leaf103.

```
leaf103# show vlan id 9

 VLAN Name                             Status    Ports
 ---- -------------------------------- --------- --------------------------------
 9    tshoot:tshoot-vrf:l3out-N9K-     active    Eth1/3, Po1
      OSPF:vlan-198

 VLAN Type  Vlan-mode
 ---- ----- ----------
 9    enet  CE

leaf103# show system internal epm vlan all | grep 9
+----------+---------+-----------------+----------+------+----------+-----------+
   VLAN ID    Type      Access Encap     Fabric    H/W id  BD VLAN    Endpoint
                        (Type Value)     Encap                         Count
+----------+---------+-----------------+----------+------+----------+-----------+
 9            Ext. BD 802.1Q        198 14974940   19     9          1
```

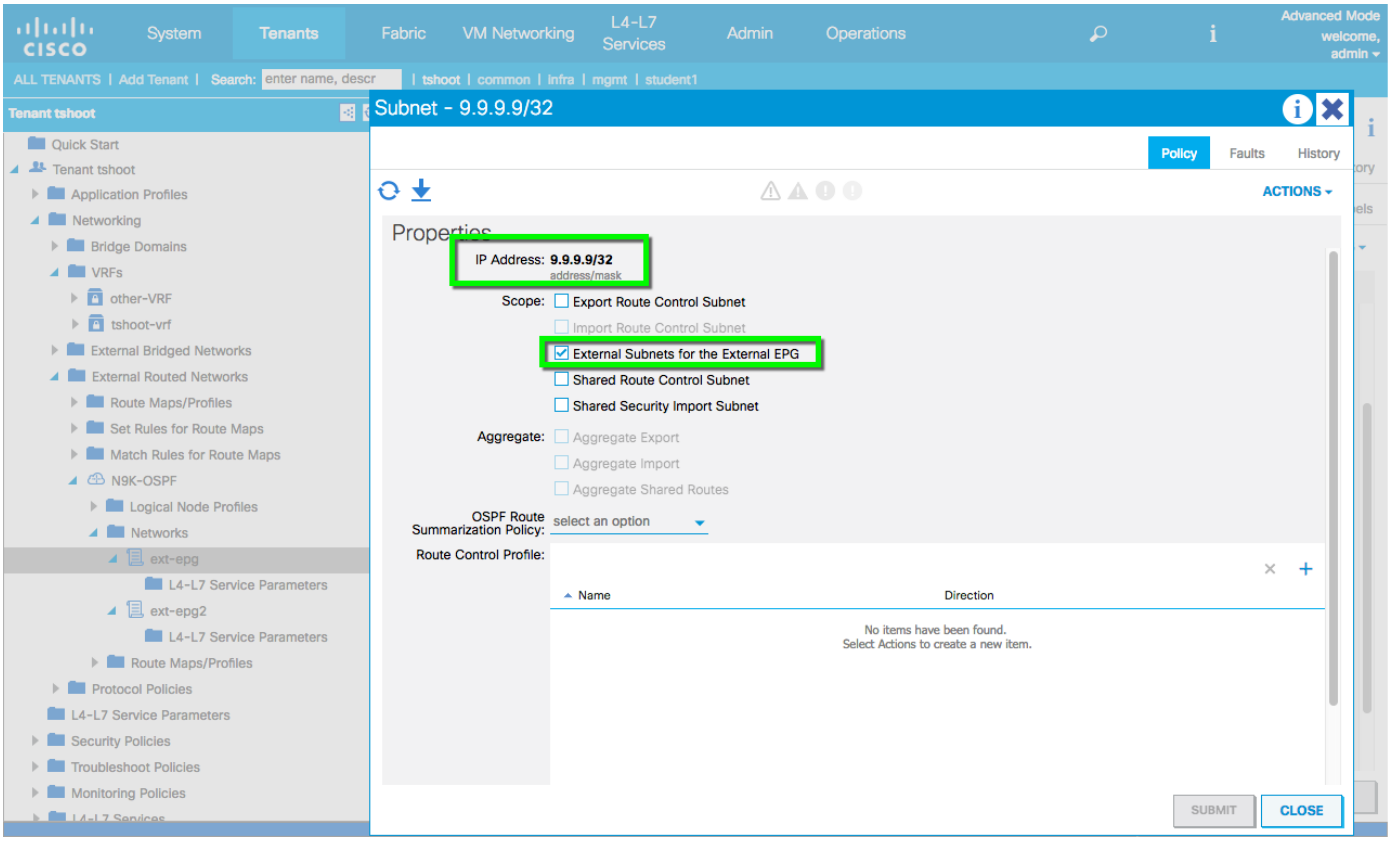
We can see that the L3out shadow BD has been deployed with Access Encap Vlan 198 and Fabric Encap (VxLAN ID) 14974940.

Contract between Internal EPG and External EPG

Applying a contract to the internal EPG and the external EPG will create zoning-rules and pervasive static route:

```
leaf103# show ip route vrf tshoot:tshoot-vrf
192.168.200.0/24, ubest/mbest: 1/0, attached, direct, pervasive
    *via 10.0.64.64%overlay-1, [1/0], 00:38:14, static

leaf103# show isis dteps vrf overlay-1
IS-IS Dynamic Tunnel End Point (DTEP) database:
DTEP-Address       Role    Encapsulation   Type
10.0.64.64         SPINE   N/A             PHYSICAL,PROXY-ACAST-V4
```

To check the EPG source class (sclass) ID for external EPGs which are classified based on source IP addresses:



We need to know the VRF VXLAN Network ID (VNI):

```
leaf103# show system internal epm vrf all
+--------------------------------+--------+----------+----------+------+--------
               VRF                   Type    VRF vnid  Context ID Status Endpoint
                                                                        Count
+--------------------------------+--------+----------+----------+------+--------
 tshoot:tshoot-vrf                Tenant   2949120    6          Up      6
```

```
leaf103# vsh_lc
module-1# show system internal aclqos prefix

Vrf Vni Addr            Mask     Scope Class  Shared Remote

======= ============== ======== ===== ====== ====== ======
2719745 0::/0 0::/0 3      15      FALSE FALSE
2719745 0.0.0.0         ffffffff 3     15     FALSE FALSE
2949120 0::/0 0::/0 4      15      FALSE FALSE
2949120 0.0.0.0         ffffffff 4     15     FALSE FALSE
2949120 9.9.9.9         0        4     16388  FALSE FALSE <<< External EPG

Shared Addr     Mask      Scope Class  RefCnt

============== ======== ===== ====== ======
module-1#
```

The sclass of external EPG 9.9.9.9/32 is 16388.

To check the zoning rule (contract), we need to check on the compute leaf:

```
leaf101# show zoning-rule scope 2949120 | grep 16388
4221            49156           16388           default         enabled
↪2949120          permit                                src_dst_any(9)
4222            16388           49156           default         enabled
↪2949120          permit                                src_dst_any(9)
```

BD subnet advertised externally and L3out association

To check whether the BD subnet is externally advertised and associated to the L3out:

```
leaf103# show ip ospf vrf  tshoot:tshoot-vrf
...
 Redistributing External Routes from
   static route-map exp-ctx-st-2949120

leaf103# show route-map exp-ctx-st-2949120
route-map exp-ctx-st-2949120, deny, sequence 1
  Match clauses:
    tag: 4294967295
  Set clauses:
route-map exp-ctx-st-2949120, permit, sequence 15801
  Match clauses:
    ip address prefix-lists: IPv4-st16388-2949120-exc-int-inferred-export-dst
    ipv6 address prefix-lists: IPv6-deny-all
  Set clauses:

leaf103# show ip prefix-list IPv4-st16388-2949120-exc-int-inferred-export-dst
ip prefix-list IPv4-st16388-2949120-exc-int-inferred-export-dst: 1 entries
   seq 1 permit 192.168.200.254/24
```

We can see that subnet 192.168.200.254/24 is permited to be redistibuted from static to OSPF.

Fabric Policy – BGP Route Reflector

To check whether BGP route reflector has been configured, we can check the BGP VPNV4 neigborship in vrf overlay-1. 10.0.32.90 is the spine which is configured as a BGP route reflector.

```
leaf103# show bgp vpnv4 unicast summary vrf overlay-1
BGP summary information for VRF overlay-1, address family VPNv4 Unicast
BGP router identifier 10.0.32.92, local AS number 6500
BGP table version is 47, VPNv4 Unicast config peers 1, capable peers 1
6 network entries and 8 paths using 1032 bytes of memory
BGP attribute entries [2/288], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [1/4]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.32.90      4  6500     400     407       47    0    0 06:22:35 2
```

To check whether the external routes from external EPGs has been learnt in BGP VPNV4:

```
leaf103# show bgp vpnv4 unicast vrf overlay-1
BGP routing table information for VRF overlay-1, address family VPNv4 Unicast
BGP table version is 47, local router ID is 10.0.32.92
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
```

(continues on next page)

(continued from previous page)

```
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath

   Network            Next Hop            Metric     LocPrf     Weight Path

Route Distinguisher: 10.0.32.92:2     (VRF tshoot:tshoot-vrf)
*>r9.9.9.9/32         0.0.0.0                  5        100      32768  ?
* i                   10.0.32.91               5        100          0  ?
```

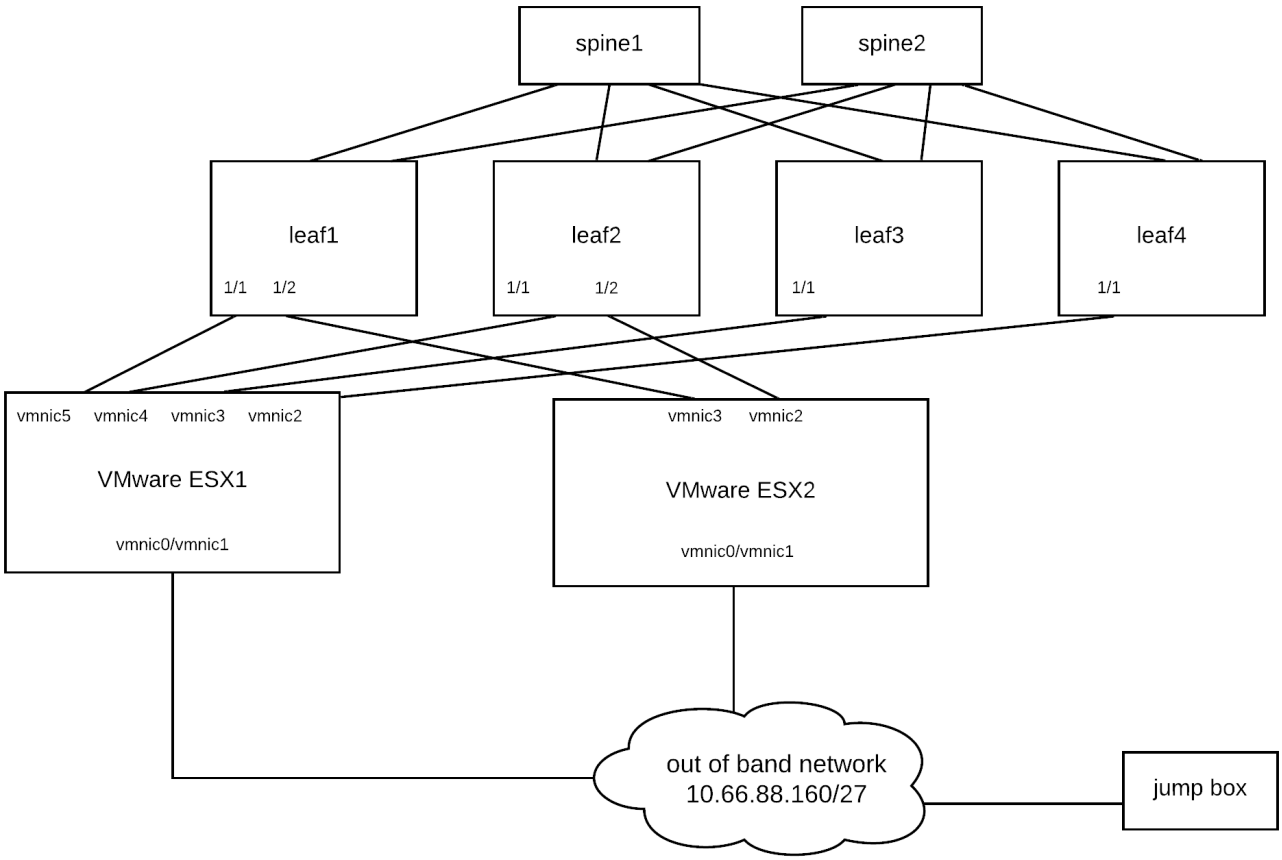
BD Subnets

To check existing BD subnets:

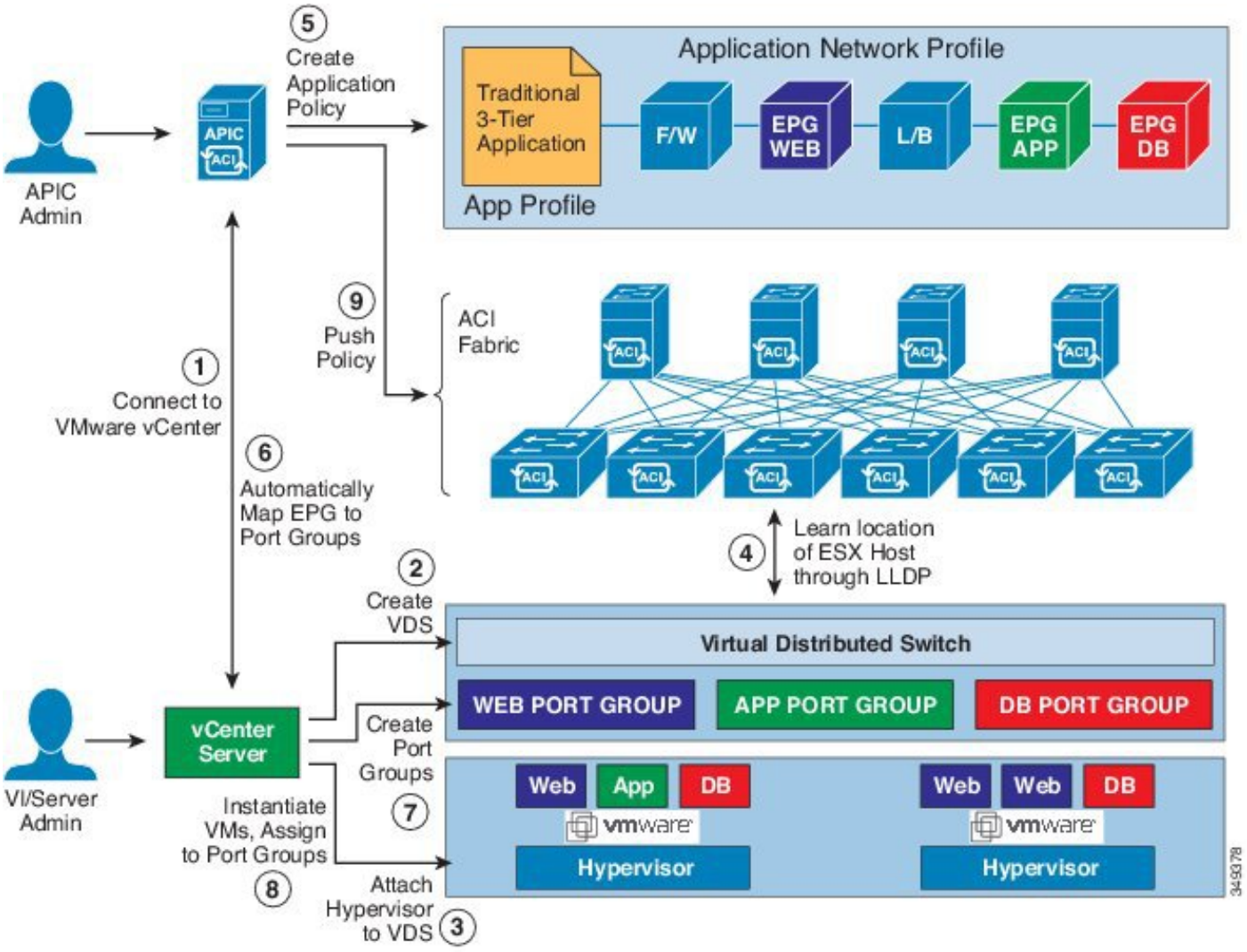
```
apic# moquery -c fvSubnet
```

2.7 Virtual Machine Manager Domain

2.7.1 Lab Topology



2.7.2 Resolution Immediacy



- Pre-provision—Specifies that a policy (for example, VLAN, VXLAN binding, contracts, or filters) is downloaded to a leaf switch even before a VM controller is attached to the virtual switch (for example, VMware VDS). This pre-provisions the configuration on the switch. This helps the situation where management traffic for hypervisors/VM controllers are also using the virtual switch associated to APIC VMM domain (VMM switch). Deploying a VMM policy such as VLAN on ACI leaf switch requires APIC to collect CDP/LLDP information from both hypervisors via VM controller and ACI leaf switch. However if VM Controller is supposed to use the same VMM policy (VMM switch) to communicate with its hypervisors or even APIC, the CDP/LLDP information for hypervisors can never be collected because the policy required for VM controller/hypervisor management traffic is not deployed yet. When using pre-provision immediacy, policy is downloaded to ACI leaf switch regardless of CDP/LLDP neighborship. Even without a hypervisor host connected to the VMM switch.
- Immediate—Specifies that EPG policies (including contracts and filters) are downloaded to the associated leaf switch software upon ESXi host attachment to a DVS. LLDP or OpFlex permissions are used to resolve the VM controller to leaf node attachments. The policy will be downloaded to leaf when you add host to the VMM switch. CDP/LLDP neighborship from host to leaf is required.
- On Demand—Specifies that a policy (for example, VLAN, VXLAN bindings, contracts, or filters) is pushed to the leaf node only when an ESXi host is attached to a DVS and a VM is placed in the port group (EPG). The policy will be downloaded to leaf when host is added to VMM switch and virtual machine needs to be placed

into port group (EPG). CDP/LLDP neighborship from host to leaf is required. With both immediate and on demand, if host and leaf lose LLDP/CDP neighborship the policies are removed.

2.7.3 Deployment Immediacy

- Once the policies are downloaded to the leaf software, deployment immediacy can specify when the policy is pushed into the hardware policy content-addressable memory (CAM).
- 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.

2.7.4 Reference

1. VMM Domain https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/aci-fundamentals/b_ACI-Fundamentals/b_ACI-Fundamentals_chapter_01011.html

2.8 REST API

This is an example of using REST API to interact with APIC. This script is written in Python. However, you can also use Postman to send the http request.

```
import requests
requests.packages.urllib3.disable_warnings()

if __name__ == '__main__':
# variables
    apic_ip = '192.168.1.1' # OOB mgmt
    apic_user = 'admin'
    apic_pw = 'xyz'
    apic_apic_url = 'https://' + apic_ip + '/api/'

    # login data
    login_data = '''<?xml version="1.0" encoding="UTF-8"?>
                        <imdata totalCount="1">
                            <aaaUser name="''' + apic_user + '''" pwd="''' + apic_pw␣
↪+ '''"/>
                    </imdata>'''

    # create requests session
    session = requests.session()

# login to apic (store cookie in requests session)
    result = session.post(apic_apic_url + 'aaaLogin.xml', data=login_data,␣
↪verify=False)
```

2.9 Firmware Upgrade

During upgrade you can monitor the leaf log:

```
tail -f /mnt/pss/installer.log
```

CHAPTER 3

Indices and tables

- genindex
- modindex
- search

CHAPTER 4

Attachments

Here is a link to a lab excerside `aci-tshoot-lab.docx`

CHAPTER 5

Author

Edi Wibowo Email: ewibowo@live.com LinkedIn: https://www.linkedin.com/in/ediwibowo