

..!!..!!..
cisco



VXLAN and EVPN

Complete Technology Guide for TAC CSE

Kyuhwan Kim

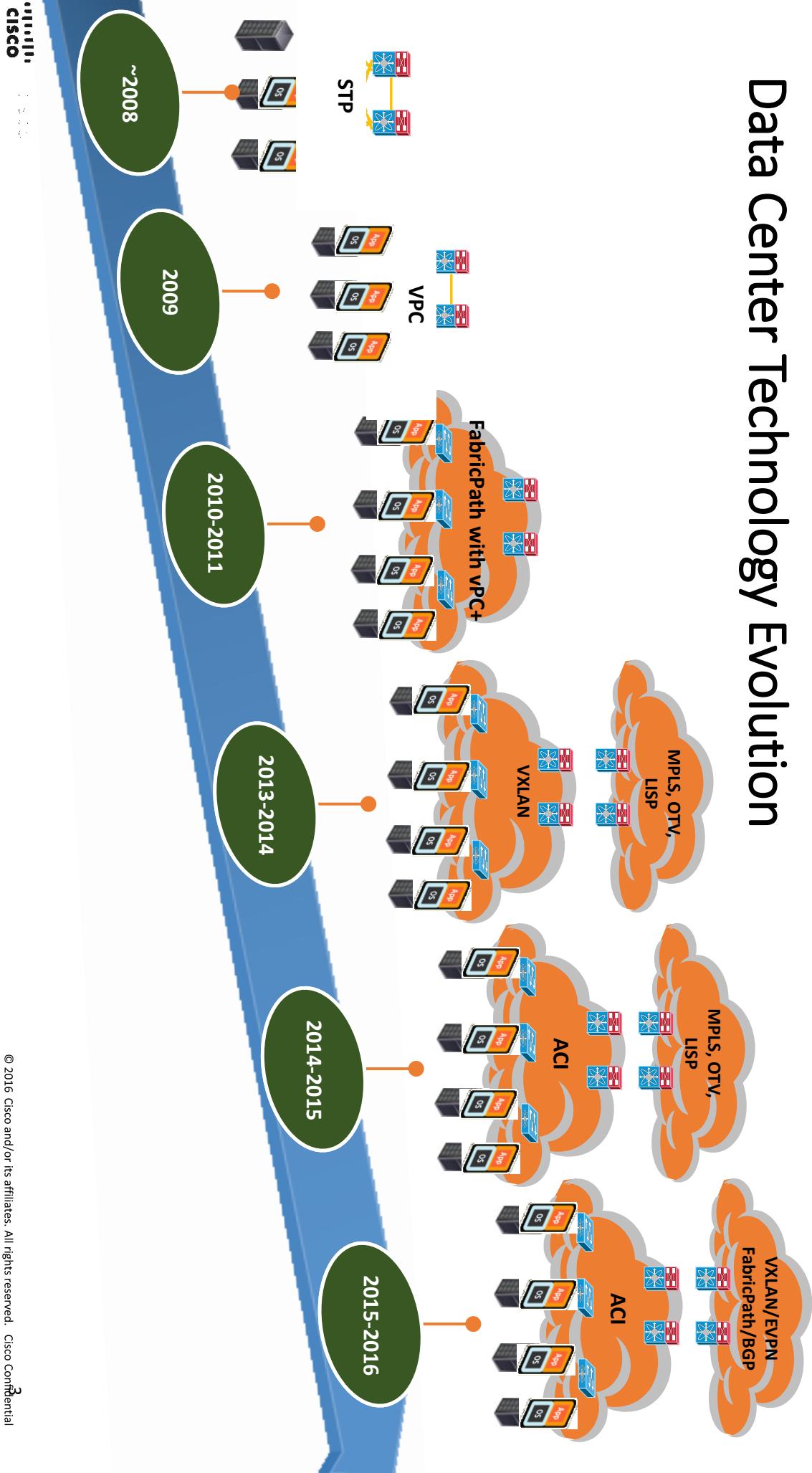
TAC R&S Leadership

Version 3: Jan 21, 2016

Agenda

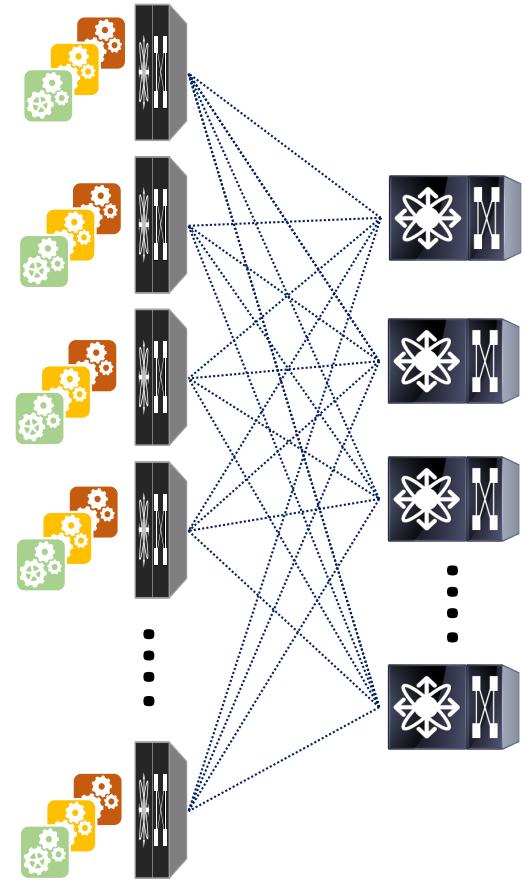
- Introduction
- Data Center Fabric
- VXLAN Overview
- EVPN Overview
- EVPN Automation Solutions
- Conclusion

Data Center Technology Evolution



Trend: Flexible Data Center Fabrics

■ Desirable Attributes:

- Mobility
 - Segmentation
 - Scale
 - Automated & Programmable
 - Abstracted consumption models
 - Full Cross Sectional Bandwidth
 - Layer-2 + Layer-3 Connectivity
 - Physical + Virtual
- 

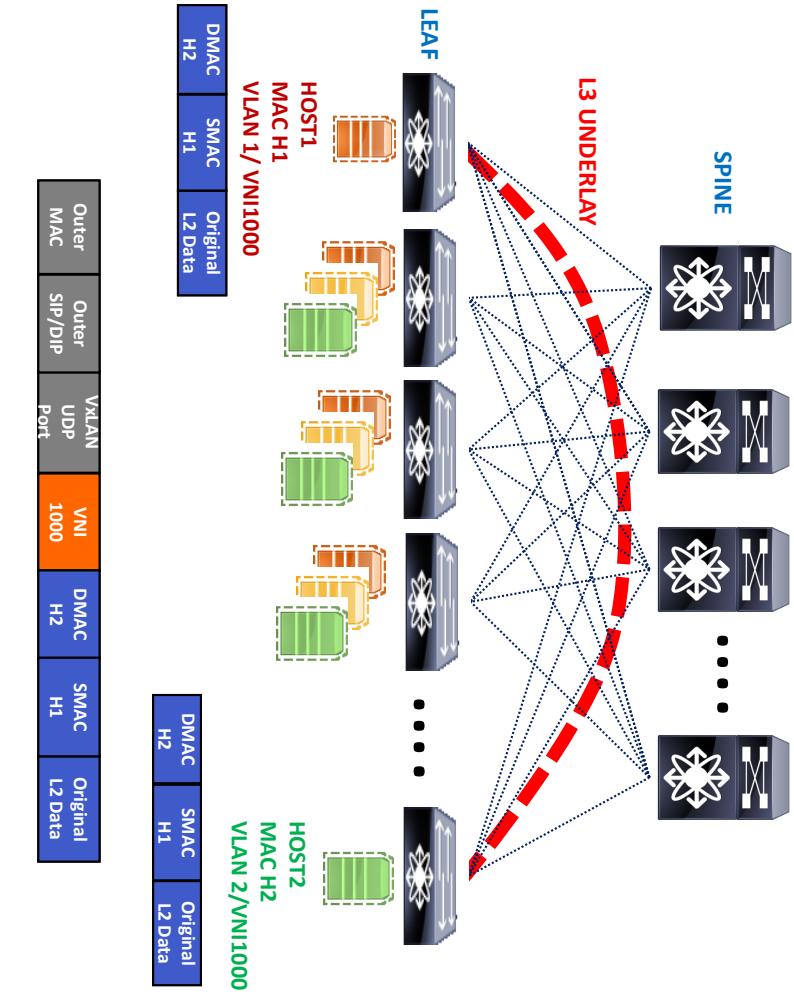
VXLAN

Why VXLAN?

VXLAN provides a Network with Segmentation, IP Mobility, and Scale

- “Standards” based Overlay (RFC 7348)
- Leverages Layer-3 ECMP – all links forwarding
- Increased Name-Space to 16M identifier
- Integration of Physical and Virtual
- Base for SDN

VXLAN Overview



Layer 2 overlay on top of your Layer 3 underlay

- Each VXLAN Segment is identified by a unique 24-bit segment ID called a **VXLAN Network Identifier (VNI)**
- Only hosts on the same VNI are allowed to communicate with each other
- Original L2 packet is encapsulated with VXLAN header in a [UDP->IP->Ethernet](#)

Overcome 4094 VLAN Scale Limitation

- VLANs use a 12-bit VLAN ID
- Multi-tenancy and vApps driving the need for more than 4K VLANs

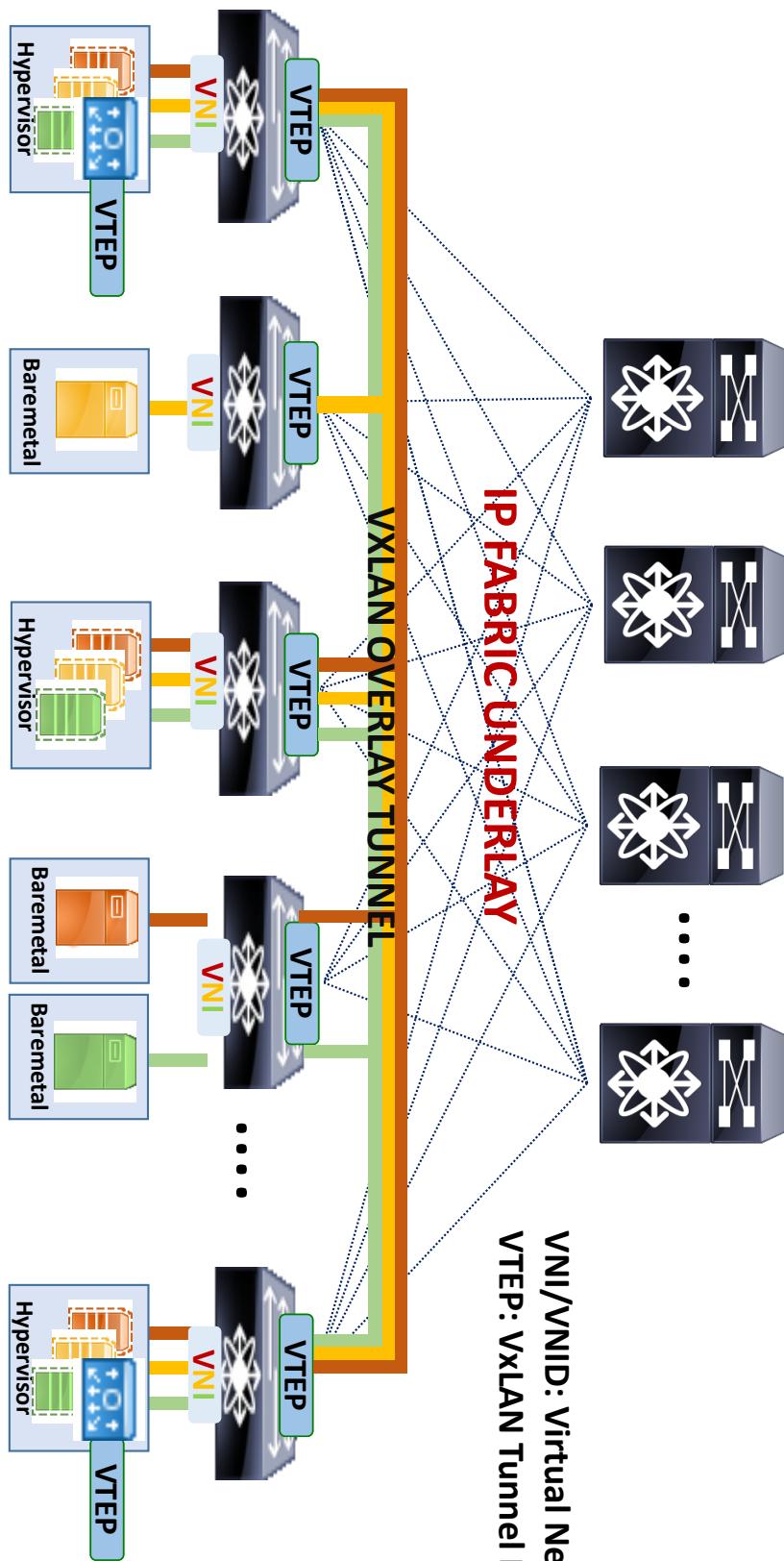
Better utilization of available network paths

- Static VLAN trunk provisioning doesn't work well for Cloud Computing and VM mobility
- No need of Spanning Tree (blocks paths)
 - Utilize L3 underlay network (ECMP, Link Agg,...)
- **Multi-Tenant with virtualization**
 - Isolation of network traffic by a tenant and need to use same VLANs on different tenants

VxLAN Terminology

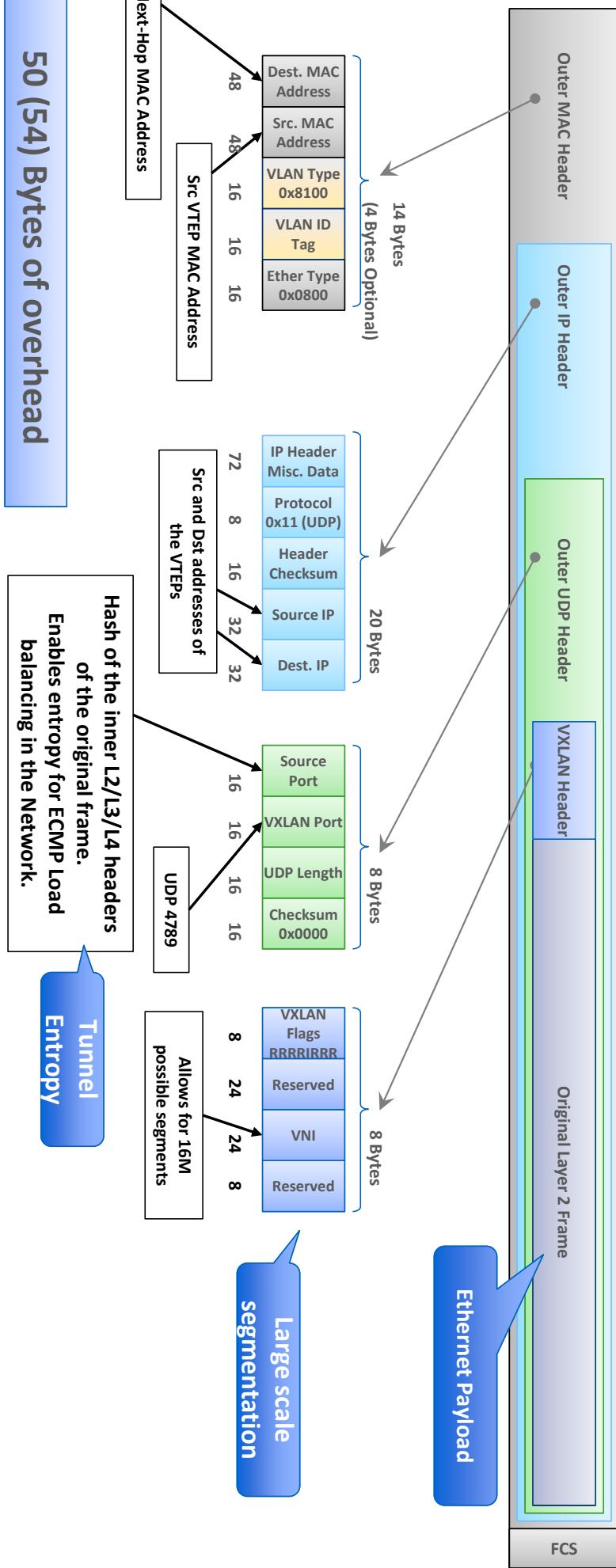
IP FABRIC UNDERLAY

VNI/VNID: Virtual Network Identifier
VTEP: VxLAN Tunnel End Point



VXLAN Packet Structure

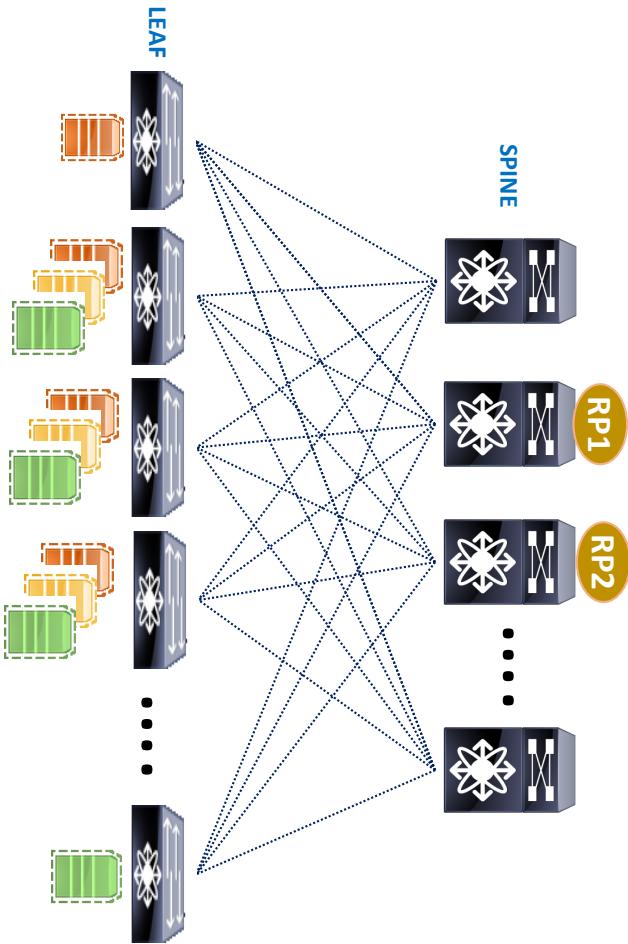
Ethernet in IP with a shim for scalable segmentation



50 (54) Bytes of overhead

VXLAN Underlay Consideration

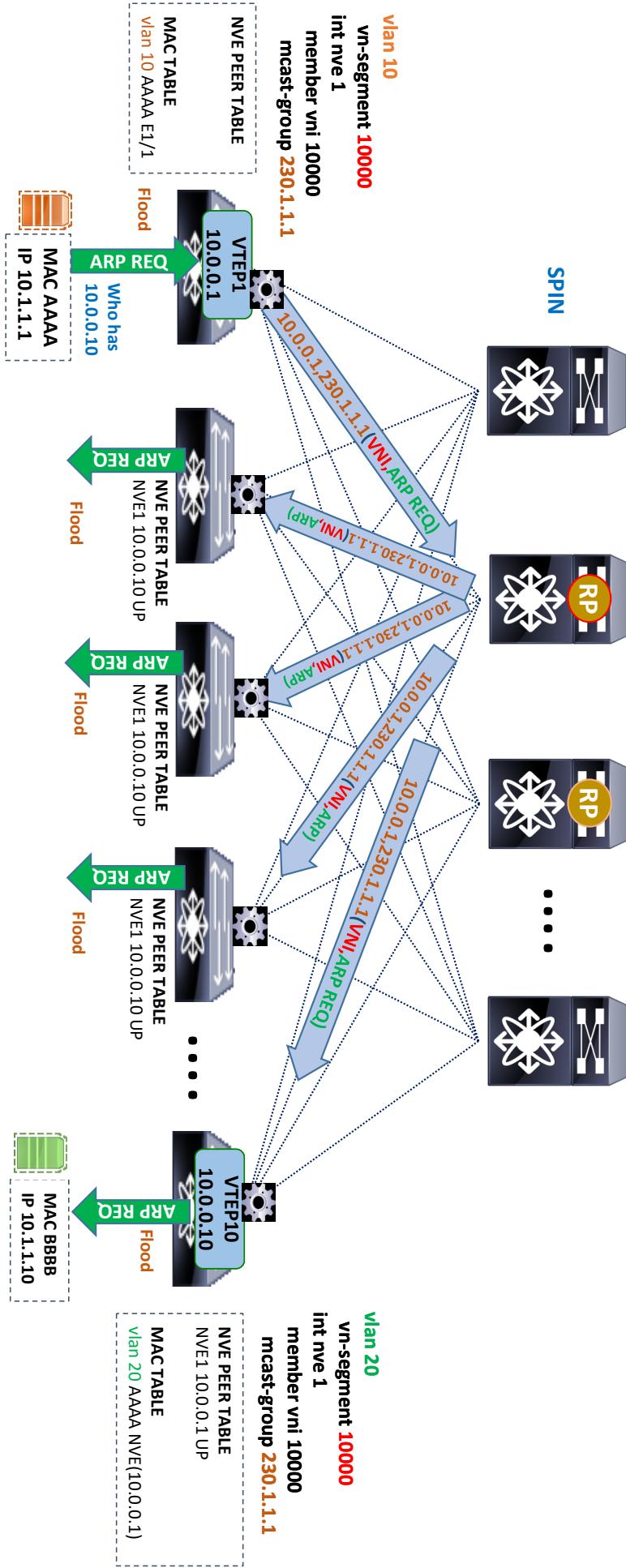
- Separate VTEP Source Loopback from Routing Protocol Loopback
- Spine to Leaf: use P2P /30 subnet
 - Use OSPF P-t-P network type
- Avoid fragmentation
 - VXLAN doesn't support fragmentation
 - VXLAN adds 50 or 54 bytes
 - Use MTU 9216 (9192 for N5K/N6K)
- Use PIM Anycast RP
 - With MSDP or with NXOS Anycast PIM RP
- IPv4 only currently in underlay



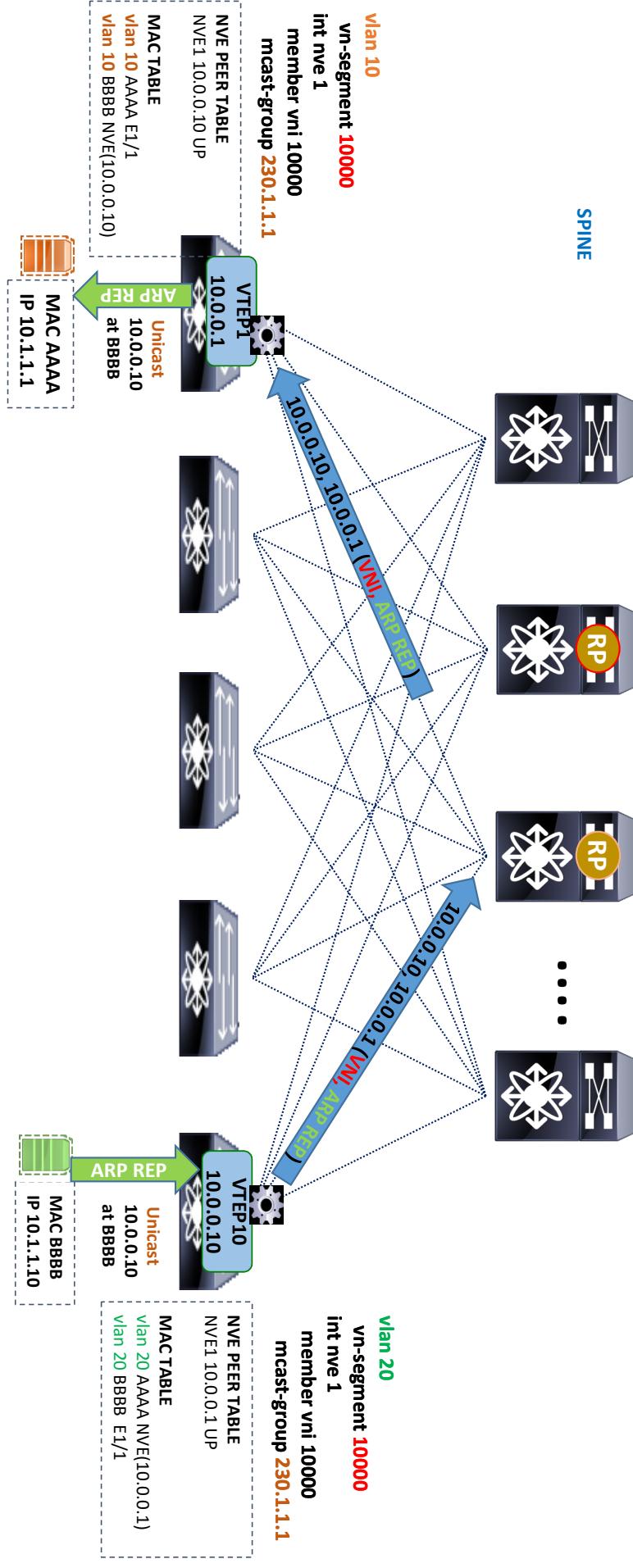
VXLAN Flooding & Learn Establishment

- Flooding of Broadcast/Unknown-unicast/Multicast (BUM) packets across underlying core network:
 - Option 1: Use **multicast IP core**
 - Each VNI is mapped to a multicast group.
 - For BUM packets coming on access side interfaces, they are encapsulated with VNI group address as DIP and sent out along the multicast tree to core
 - Option 2: Use **Ingress Replication (IR)**
 - Some customers would want to avoid using multicast in their core
 - Remote VTEPs are statically configured per VNI
 - BUM packets ingress on access side are replicated to all remote static VTEPs in the VNI, each is encapsulated with one VTEP IP as unicast DIP
 - Support multiple VTEPs per VNI and a VTEP in multiple VNIs
 - Up to 16 static IR VTEPs is recommended
- Multicast and IR configuration can co-exist on the same switch, but on different VNIs.
- Multicast-core learned VTEP vs. IR static VTEP:
 - Learned VTEP tunnel will be removed when all dynamically learned MACs that are associated with the VTEP are aged out.
 - IR static VTEP tunnel is kept alive as long as the route to the VTEP is available.

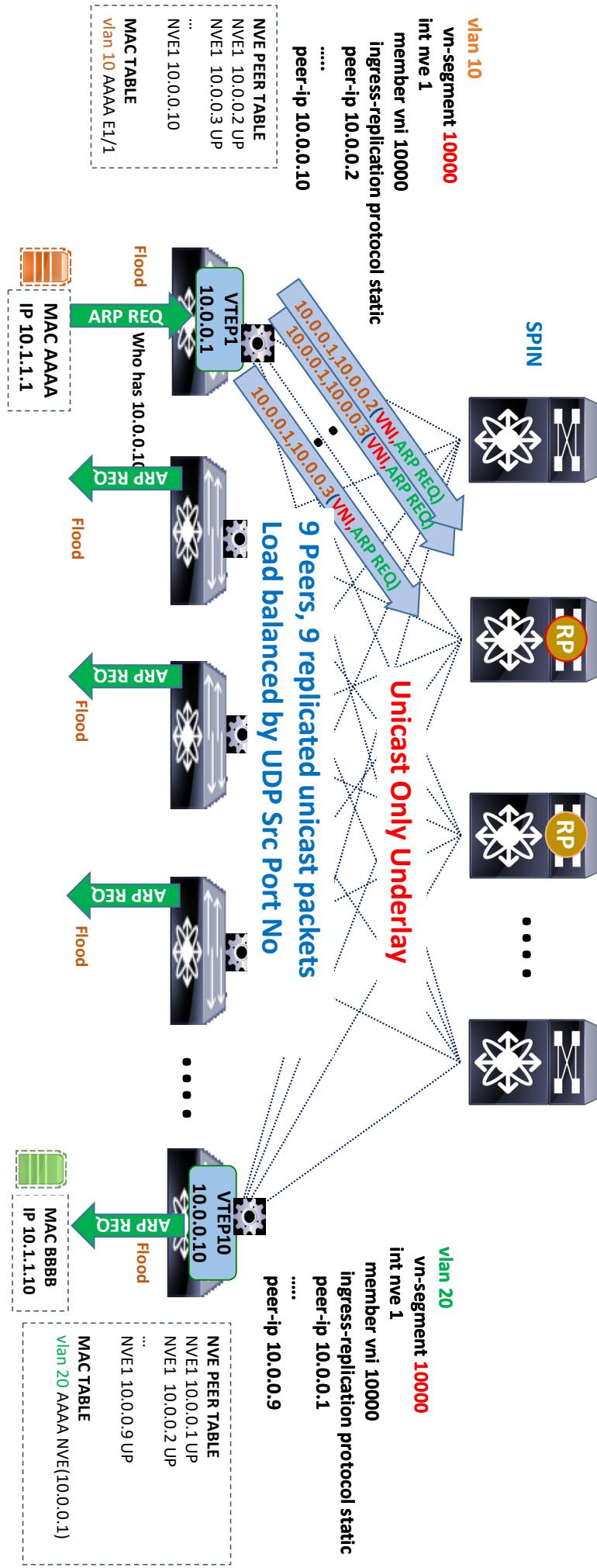
VXLAN Flood & Learn: Multicast IP CORE



VXLAN Flood and Learn: Multicast IP CORE



VxLAN Flood and Learn: Ingress Replication



Multicast Enabled Underlay

- May use PIM-ASM or PIM-BiDir (Different hardware has different capabilities)

	N1KV	Nexus 7K with F3 LC	Nexus 3K	Nexus 5K/6K	Nexus 9K Standalone	CSR 1000V ASR1K	ASR9K
Mcast mode	IGMP v2/v3	PIM-ASM & Bidir-PIM	PIM-ASM & Bidir (no VPC)	Bidir-PIM	PIM-ASM & Bidir (no VPC)	Bidir-PIM (ASM –Future)	PIM-ASM & Bidir-PIM

- Spine and Aggregation switches make good RP locations in clos and traditional topologies respectively
- Reserve a range of multicast channels to service the overlay and optimize for diverse VNIs
- In clos topologies, using multiple RPs across the multiple spines and mapping different VNIs to different RPs will provide a simple load balancing measure
- Design a multicast underlay for a network overlay, host VTEPs will simply leverage this network

VXLAN Gateway Types

- **VXLAN to VLAN Bridging**

- **(Layer-2 Gateway)**

Ingress VXLAN packet
on RED segment

Egress packet is
DOT1Q tagged and
bridged to VLAN

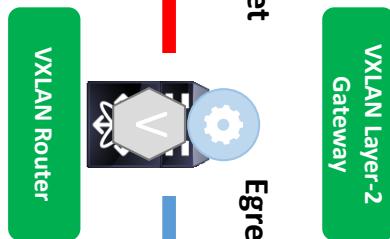


- **VXLAN-to-VXLAN Routing**

- **(Layer-3 Gateway)**

Ingress VXLAN packet
on RED segment

Egress VXLAN packet is ROUTED
to new VXLAN segment

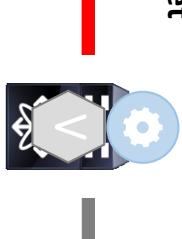


- **VXLAN-to-VLAN Routing**

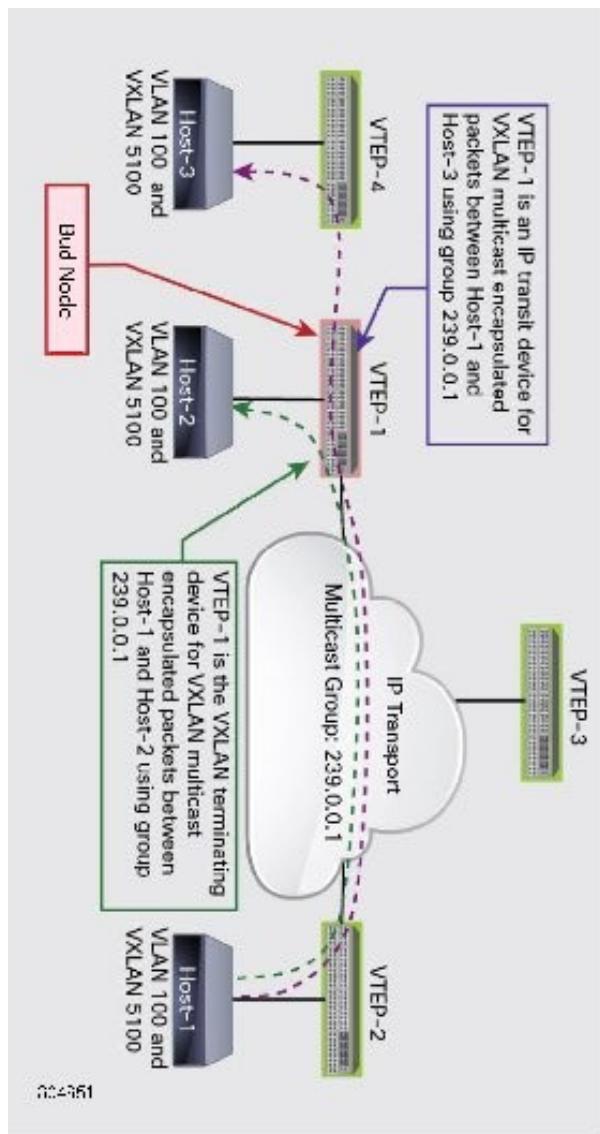
- **(Layer-3 Gateway)**

Ingress VXLAN packet
on RED segment

Egress packet is DOT1Q tagged
and ROUTED to new VLAN



VXLAN Feature: BUD NODE or HOST VTEP Support



■ **VTEP act as VTEP + IP Transport Device**

- VTEP-1 joined underlay VXLAN Delivery Multicast Group, 239.0.0.1
- VTEP-4 joined same group, 239.0.0.1 and VTEP-1 is transport device for VTEP-4
- VTEP-1 needs to do
 - Decap 239.0.0.1 traffic and forward traffic to proper local VXLAN attached devices
 - Don't decap 239.0.0.1 traffic and just multicast forward to VTEP-4
- Usually requires recirculation

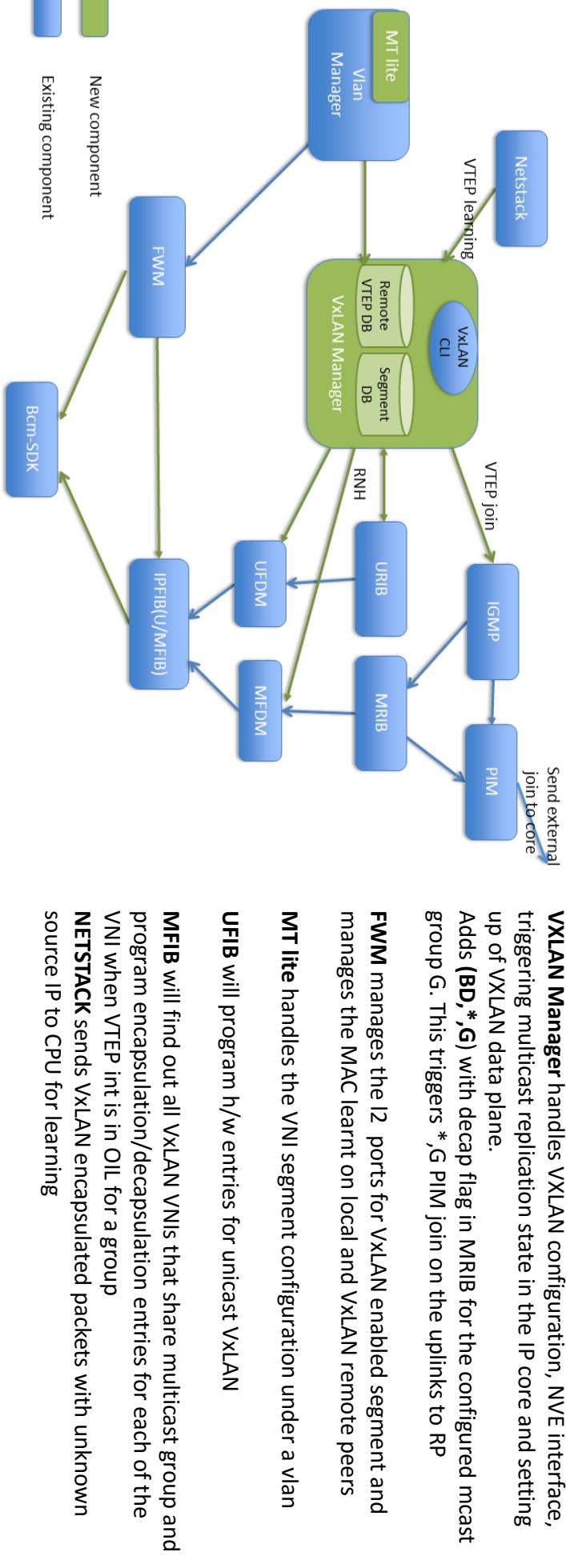
VXLAN Features & Platforms

	VXLAN L2 GW w/F&L	VXLAN L3 GW w/F&L	VXLAN Ingress Replication	VXLAN w/EVPN	Bud Node (Host VTEP)	BGP EVPN/ PBB-EVPN as RR
NEXUS1K	VXGW on N1100	Use CSR1K as GW		Q2 2016		
NEXUS31XX	Goa	No, HW LIMIT	Goa, only one peer	No	Goa	
NEXUS56XX /6000	Iluka 7.1(0)N1)	7.1(0)N1	N/A, Q1 CY16	7.3.0.N1.1	7.1(0)N1	
N7000/7700 w/F3	Gibraltar 7.2	Gibraltar 7.2	Gibraltar 7.2	Gibraltar 7.2	Gibraltar 7.2	
NEXUS9000	Bronte 7.0.3.I1	Camden 7.0.3.I2(2Q2015)	Bronte 7.0.3.I1	Bronte 7.0.3.I1	7.0.3.I1	
CSR1000	3.11S	Polaris 16.3 3Q 2016	Polaris 16.3 3Q 2016	Polaris 16.3 3Q 2016		3.11S
ASR1000	3.13S			Polaris 16.3 3Q 2016		3.12S
ASR9000	XR 5.3.1			XR 5.3.2		XR 4.3.2

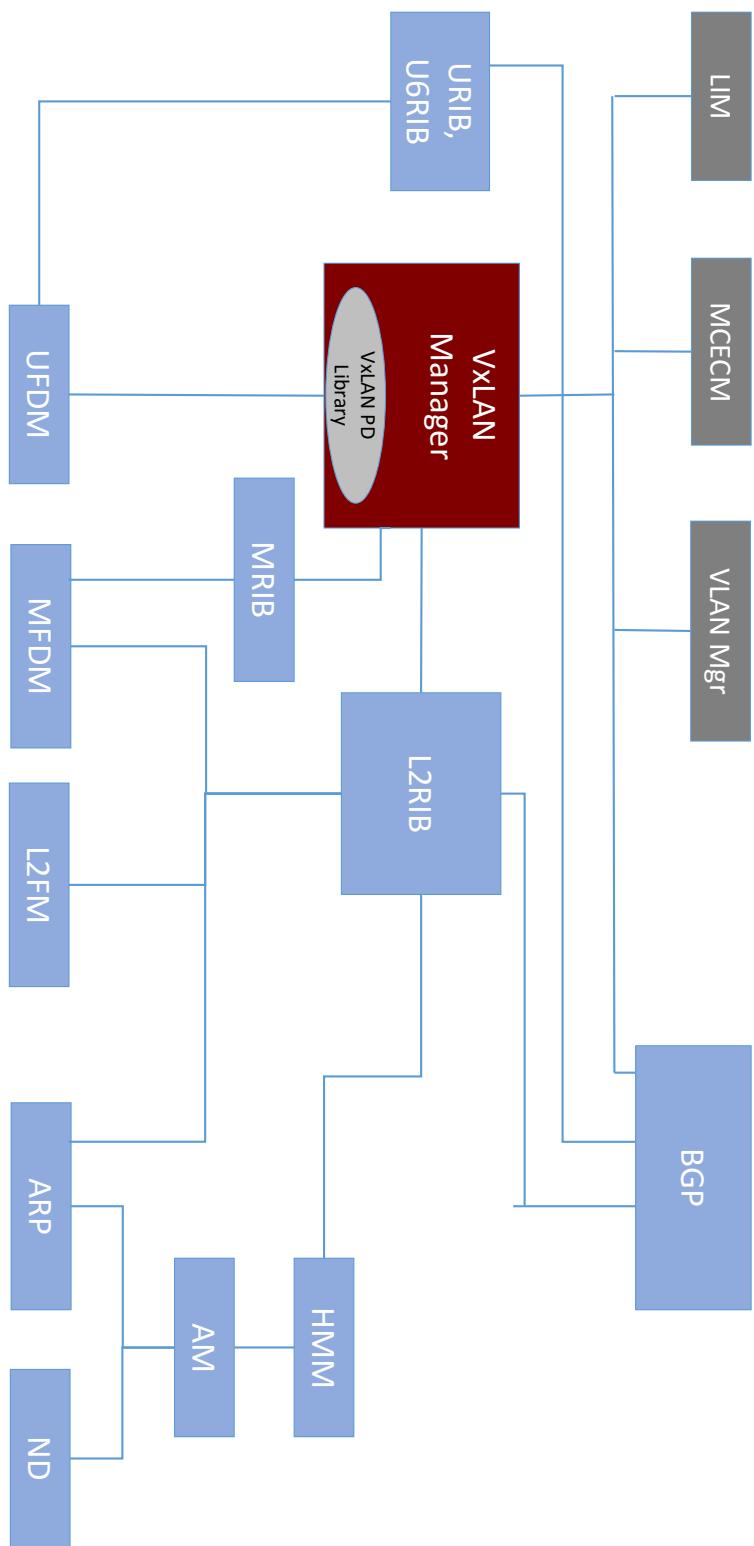
VXLAN: Flood-&-Learn vs EVPN Control Plane

	Flood-&-Learn	EVPN Control Plane
Overlay Services	L2+L3	L2+L3
Underlay Network	IP network with ECMP	IP network with ECMP
Encapsulation	MAC in UDP	MAC in UDP
Peer Discovery	Data-driven flood-&learn	MP-BGP
Peer Authentication	Not available	MP-BGP
Host Route Learning	Local hosts: Data-driven flood-&learn Remote hosts: Data-driven flood-&learn	Local Host: Data-driven Remote host: MP-BGP
Host Route Distribution	No route distribution.	MP-BGP
L2/L3 Unicast Forwarding	Unicast encapsulation	Unicast encapsulation
BUM Traffic forwarding	Multicast replication Unicast/Ingress replication	Multicast replication Unicast/Ingress replication

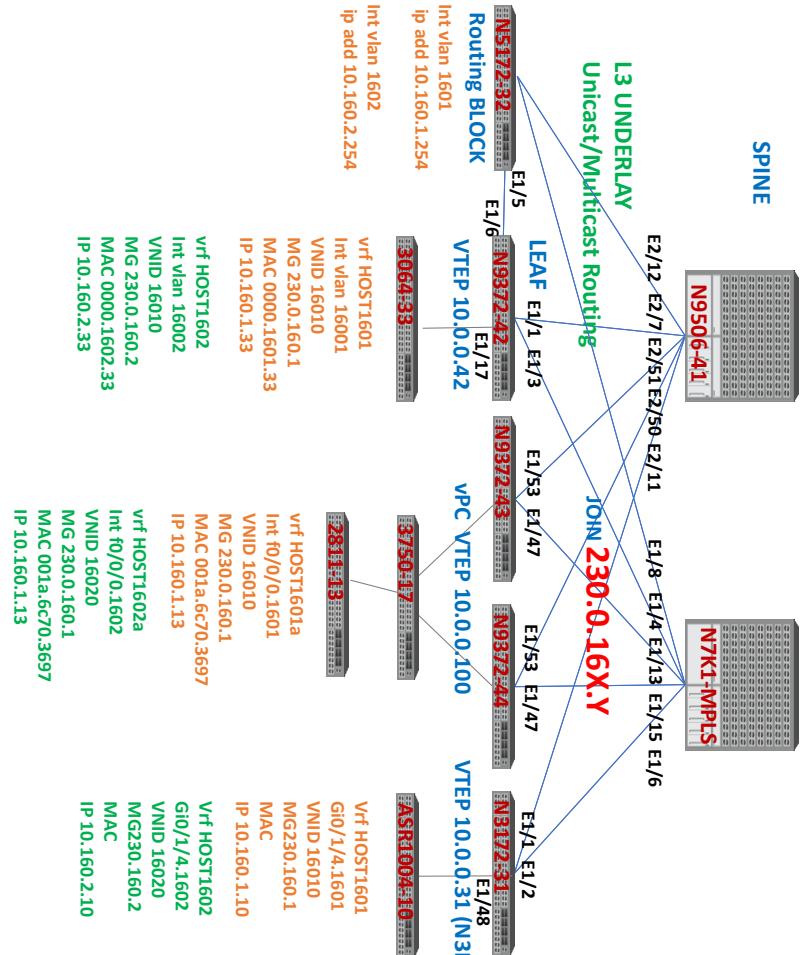
N3K VXLAN Components



N9K VXLAN Component



SJ RP VXLAN Static LAB – Flood & Learn



- Underlay L3 Setup
- Two Spines and 5 Leafs
 - Run OSPF as main routing protocol
 - Two Spines are BSR RP with PIM Anycast RP (not MSDP)
- N9K vPC
 - N9372-43 and 44 are vPC peers
 - vPC uses virtual VTEP IP
 - Lo0 has 10.0.0.100 as 2nd IP
 - N3K don't use 2nd IP but both vPC peers has same ip address
- N31XX supports VXLAN with T2
- N5672 supports VXLAN with only **PIM Bidir**
 - N9K doesn't support PIM BiDIR in VPC so N9K and N6K can't work together on VXLAN
 - N5672 is working as separated routing block for VXLAN by configure SVI for vlan 16XY
 - N9K can't function as SVI and VXLAN GW at the same time (Plan to support this feature)
- Hosts in vlan 1601 can ping hosts in vlan 1601 or vlan 1602
 - Mapped vlan 16XY to vnid 16XY0
 - Vlan 1601 → VNID 16010, 10.160.1.0/24, 230.0.160.1
 - Vlan 1602 → VNID 16020, 10.160.2.0/24, 230.0.160.2

SJ RP VXLAN Static LAB

■ Underlay PIM Anycast RP with BSR

SPINE as RP

interface loopback99

description anycast-rp

ip address 10.0.0.99/32

ip router ospf 1 area 0.0.0.0

ip pim sparse-mode

ip pim bsr-candidate loopback99

ip pim rp-candidate loopback99 group-list 230.0.0.0/16 priority 100

ip pim anycast-rp 10.0.0.99 10.0.0.3

ip pim anycast-rp 10.0.0.99 10.0.0.41

ip pim bsr listen forward

Leaf - how to check which RP is being used for VXLAN delivery MG.

N9372-42# sh ip pim rp 230.0.0.0

Vrf HOST1601

610/1/4.1601

VNID 16010

MG 230.0.160.1

MAC 0000.1601.33

IP 10.160.1.33

MG 230.0.160.1

MAC 001a.6c70.3697

Vrf HOST1602

610/1/4.1602

VNID 16020

MG 230.0.160.2

MAC 0000.1602.33

IP 10.160.2.33

Vrf HOST1602a

610/1/4.1602

VNID 16020

MG 230.0.160.1

MAC 001a.6c70.3697

IP 10.160.1.13

N3172-31# sh ip pim rp 230.0.0.0

PIM RP Status Information for VRF "default"

PIM RP Information for group 230.0.0.0 in VRF "default"

RP: 10.0.0.99, (0), uptime: 22:52:45, expires: 00:02:20,

priority: 100, RP-source: 10.0.0.99 (B), group ranges:

230.0.0.0/16

RP: 10.0.0.99, (0), uptime: 22:52:45, expires: 00:02:20,

priority: 100, RP-source: 10.0.0.99 (B), group ranges:

230.0.0.0/16

VXLAN Flood & Learn – Multicast

1. Enable features
feature vn-segment-vlan-based
feature nv overlay



2. Configure VLAN and Map to VN-Segment

```
vlan 1601
vn-segment 16010
vlan 1602
vn-segment 16020
```

2. Configure NVL and Map to VN-Segment

```
interface nve1
no shutdown
source-interface loopback0
member-vni 16010
mcast-group 230.0.160.1
member-vni 16020
mcast-group 230.0.160.2
```

NVE came up after you no shut int NVL

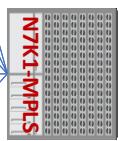
```
N9372-42# sh nve int
Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VIP-Only [not-notified]
Local Router MAC: 881d.fcca.3f4f
Host Learning Mode: Data-Plane
Source-Interface: loopback0 (primary: 10.0.0.42, secondary: 0.0.0.0)
N9372-42# sh int nve 1
nve1 is up
admin state is up, Hardware: NVE
MTU 9216 bytes
Encapsulation VXLAN
Auto-mdix is turned off
RX
    ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
    These counters are helpful to
    troubleshoot
TX
    ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
```

VXLAN Flood & Learn – Multicast

CHECK basic NVE Parameters
N9372-42# sh nve vni 16010
VNI: 16010

Per VNI counter

SPINE



E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

E2/12 E2/7 E2/5 E2/9 E2/11
E1/8 E1/4 E1/13 E1/15 E1/6

TX
715 unicast packets 107007 unicast bytes
56834 multicast packets 4100720 multicast bytes

RX
295 unicast packets 34190 unicast bytes
5893 multicast packets 733527 multicast bytes

N9372-42# sh nve vxlan-params
VxLAN Dest. UDP Port: 4789

N9372-42# sh nve internal platform interface nve 1 detail
Printing Interface ifindex 0x49000001 detail

Intf	State	PriIP	SecIP	VnIs	Peers
nve1	UP	10.0.0.42	0.0.0.0	2	

SW BD/VNIs of interface nve1:

If this is VPC, you
will see 2nd IP

Sw BD	Vni	State	Intf	Type	Vrf-ID
MG 230.0.160.1	MG 230.0.160.1				
MAC 0000.1601.33	Int fo/0/0.1601				
IP 1.160.1.33	VNI 16010				
MG 230.0.160.1	MG 230.0.160.1				
MAC 001a.6e70.3697	Int fo/0/0.1602				
IP 1.160.1.33	VNI 16010				
MG 230.0.160.2	MG 230.0.160.2				
MAC 0000.1602.33	Int fo/0/0.1602				
VNI 16020	MG 230.0.160.1				
IP 10.160.2.10	MAC 001a.6c03.3697				
MG 230.0.160.1	Int fo/0/0.1602				
IP 10.160.1.13	VNI 16020				
MG 230.0.160.2	MG 230.0.160.2				
MAC 0000.1602.33	Int fo/0/0.1602				
VNI 16020	MG 230.0.160.1				
IP 10.160.2.10	MAC 001a.6c03.3697				
MG 230.0.160.1	Int fo/0/0.1602				
IP 10.160.1.13	VNI 16020				

Peers of interface nve1:

N3172-31# sh mac address-table aging-time
Vlan Aging Time

1602 1800
1601 1800

N9372-42# sh mac address-table aging-time
Aging Time

1800

NVE Peering is based on MAC learning and Aging with flood & learn
When MAC table is aged out, NVE Peering is teared down

VXLAN Flood & Learn – Multicast

N9K shows IGMP join is generated from lo0
But N9K shows nothing. Different implementation

NVE came up

```
N9372-42# sh nve vni
Codes: CP - Control Plane          DP - Data Plane
       UC - Unconfigured           SA - Suppress ARP
Interface VNI      Multicast-group State Mode Type [BD/VRF]   Flags
-----+-----+-----+-----+-----+-----+
nve1    16010   230.0.160.1        Up   DP   L2 [1601] 
nve1    16020   230.0.160.2        Up   DP   L2 [1602]
```

N9372-42# sh nve vni 16010 detail

```
VNI: 16010
NVE-Interface : nve1
Mcast-Addr: : 230.0.160.1
VNI State : Up
Mode : data-plane
VNI Type : L2 [1601]
VNI Flags : add-complete
Provision State : add-complete
Vlan-BD : 1601
SVI State : n/a
```

N9372-42# sh nve internal vni 16010

```
Ready-State : Ready [L2-vni-flood-learn-ready]
```

N9372-42# sh nve peers

No NVE Peer yet

N9372-42# sh mac add v1 1601

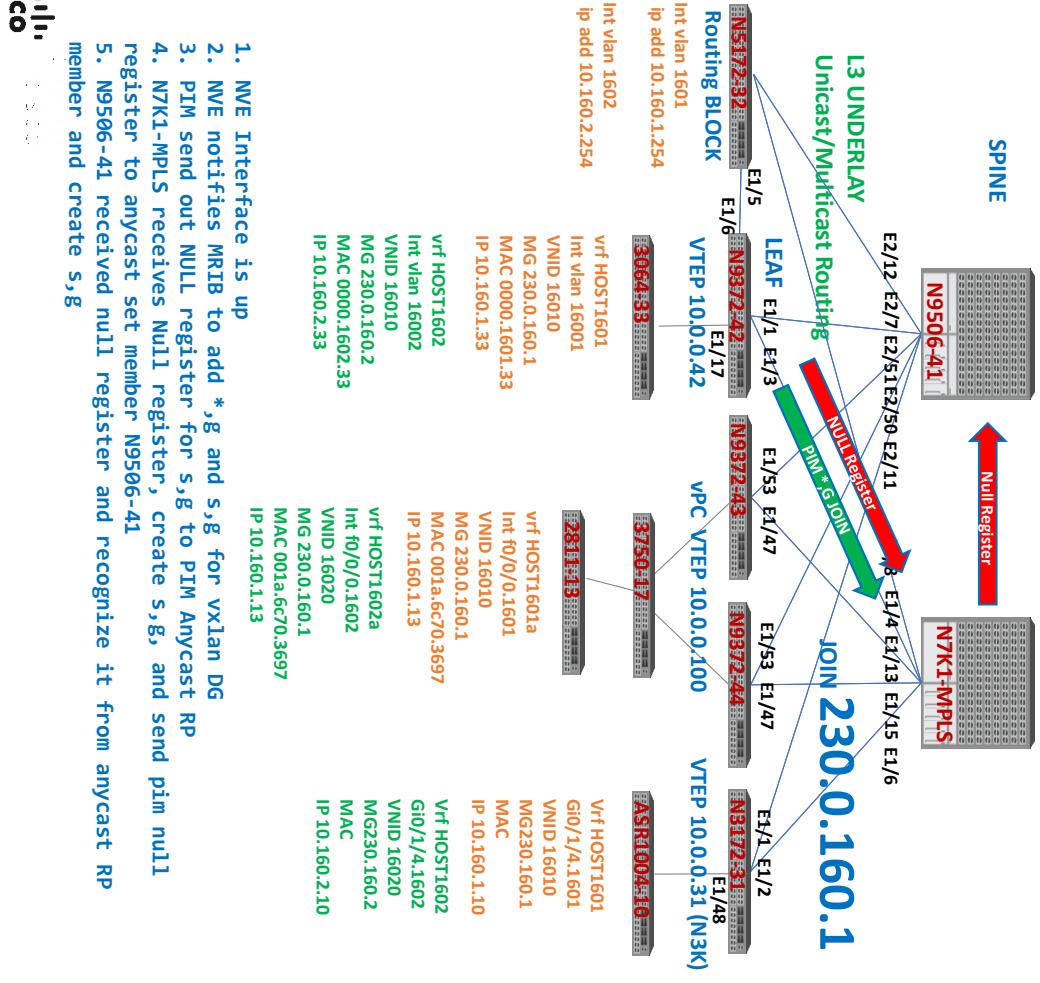
Legend: * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC

age - seconds since last seen, + - primary entry using VPC Peer-Link,
(T) - True, (F) - False
VLAN MAC Address Type age Secure NTFY Ports
-----+-----+-----+-----+-----+-----+

N9372-42#

Mac table is empty as I
shut down CE facing and
Core Facing interfaces

VXLAN Flood & Learn – Multicast



```
NVE triggered *,G JOIN for vxlan DG 230.0.160.1 and sent to N7K1-MPLS
N9372-42# sh ip pim event-history join-prune | b Send
(*, 230.0.160.1/32), uptime: 00:35:49, pim ip
[117] : Send Join-prune on Ethernet1/3, length: 42 in context 1
9) Event:E_DEBUG, length:69, at 305087 usecs after Mon Jun 8 02:05:29 2015
[117] : Put (*, 230.0.160.1/32), WRS in join-list for nbr 10.41.42.41
```

```
N7K1-MPLS has *,G as it received *,G Join from N9372-42
MPLS# sh ip mr
(*, 230.0.160.1/32), uptime: 00:35:49, pim ip
Incoming interface: loopback99, RPF nbr: 10.0.0.99
Outgoing interface list: (count: 1)
Ethernet1/4, uptime: 00:35:49, pim
(10.0.0.42/32, 230.0.160.1/32), uptime: 1d22h, pim mrib ip
Incoming interface: Ethernet1/4, RPF nbr: 10.3.42.42, internal
Outgoing interface list: (count: 0)
```

```
N9504-41# sh ip mr
(10.0.0.42/32, 230.0.160.1/32), uptime: 01:12:25, pim mrib ip
```

```
N9372-42# sh ip pim event-history null-register
null-register events for PIM process
1) Event:E_DEBUG, length:80, at 816279 usecs after Sat Jun 6 04:03:39 2015
[118] : Received Register-Stop from 10.0.0.99 for (10.0.0.42/32, 230.0.160.2/32)
4) Event:E_DEBUG, length:90, at 814630 usecs after Sat Jun 6 04:03:39 2015
[119] : Sending Null Register for vxlan-group (10.0.0.42/32, 230.0.160.2/32) Register sent
```

S,G is from null register
No OIL as there is no other receiver

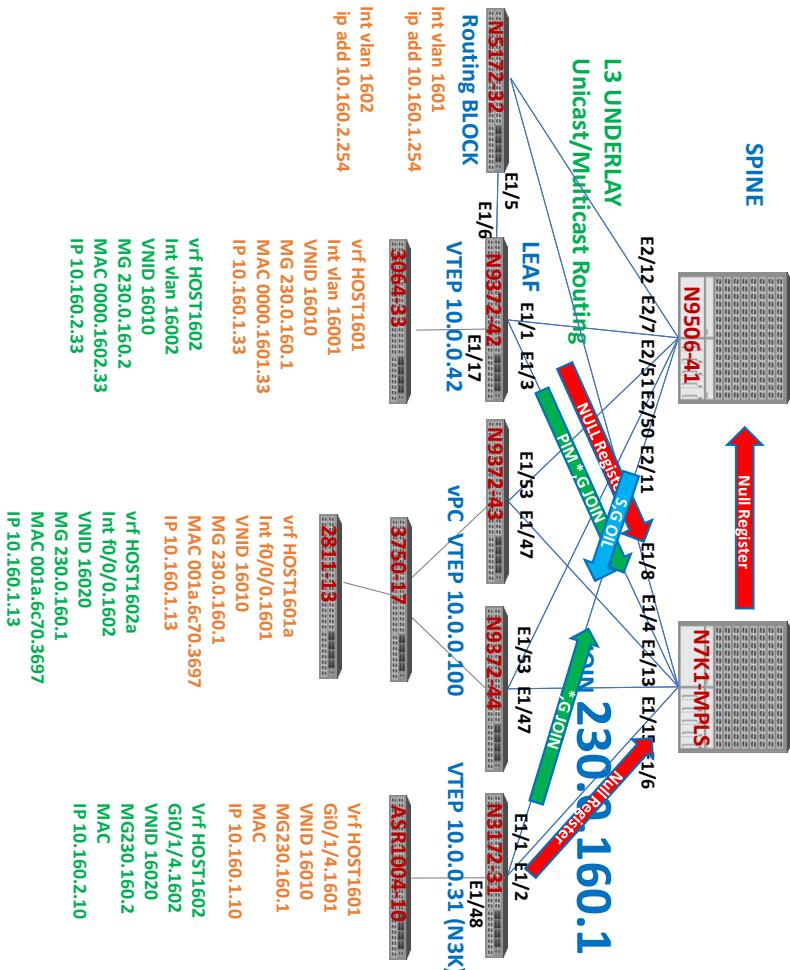
S,G is from null register
fwded from N7K1-MPLS by pim anycast

N9372-42 sent null-register
although there is no data traffic

```
MPLS# sh ip pim event-history null-register
1) Event:E_DEBUG, length:74, at 173973 usecs after Sun Jun 7 16:12:25 2015
[118] : Send Register-Stop to 10.0.0.42 for (10.0.0.42/32, 230.0.160.2/32)
2) Event:E_DEBUG, length:55, at 173910 usecs after Sun Jun 7 16:12:25 2015
[118] : Forward Register to Anycast-RP member 10.0.0.41
3) Event:E_DEBUG, length:37, at 173714 usecs after Sun Jun 7 16:12:25 2015
[118] : RP 10.0.0.99 is an Anycast-RP
4) Event:E_DEBUG, length:75, at 173701 usecs after Sun Jun 7 16:12:25 2015
[119] : Received Register from 10.0.0.42 for (10.0.0.42/32, 230.0.160.2/32)
```

```
N9504-41# sh ip pim event-history null-register
1) Event:E_DEBUG, length:57, at 372705 usecs after Mon Jun 8 07:04:41 2015
[118] : Register received from Anycast-RP member 10.0.0.3
2) Event:E_DEBUG, length:37, at 372703 usecs after Mon Jun 8 07:04:41 2015
[118] : RP 10.0.0.99 is an Anycast-RP
3) Event:E_DEBUG, length:80, at 372655 usecs after Mon Jun 8 07:04:41 2015
[119] : Received NULL Register from 10.0.0.3 for (10.0.0.42/32, 230.0.160.2/32)
```

VXLAN Flood & Learn – Multicast



Let's bring N3172-31's PIM interfaces.

3172-31(config)# int e1/1-2

N3172-31(config-if-range)# no sh

IT creates *,G and S,G immediately

N3172-31# sh ip mri

(* , 230.0.160.1/32), uptime: 01:46:51, static ip pim

Incoming interface: Ethernet1/1, RPF nbr: 10.31.41.41, uptime: 00:12:09

Outgoing interface list: (count: 1)

Loopback0, uptime: 01:46:51, static

(10.0.0.31.32, 230.0.160.1/32), uptime: 01:46:33, ip mrib pim

Incoming interface: Loopback0, RPF nbr: 10.0.0.31, uptime: 01:46:33

Outgoing interface list: (count: 2)

Ethernet1/2, uptime: 00:11:53, pim

Loopback0, uptime: 01:46:33, mrib, (RPF)

MPLS# sh ip pim event-history null-register

29) Event:E_DEBUG, Length:74, at 766991 usecs after Sun Jun 7 19:46:21 2015

[118] : Send Register-Stop to 10.0.0.31 for (10.0.0.31/32, 230.0.160.1/32)

30) Event:E_DEBUG, Length:55, at 766999 usecs after Sun Jun 7 19:46:21 2015

[118] : Forward Register to Anycast-RP member 10.0.0.41

31) Event:E_DEBUG, Length:37, at 766671 usecs after Sun Jun 7 19:46:21 2015

[118] : RP 10.0.0.99 is an Anycast-RP

32) Event:E_DEBUG, Length:75, at 766638 usecs after Sun Jun 7 19:46:21 2015

[119] : Received Register from 10.0.0.31 for (10.0.0.31/32, 230.0.160.1/32)

N9504-41 receive *,G join from N3172-31, so it copies OIL from *,G to S,G when N9504-41

receives Null Register flooded from N9K1-MPLS.

We build S,G tree for VXLAN DG although there is no traffic flows

N9504-41# sh ip mri

(* , 230.0.160.1/32), uptime: 00:20:19, pim ip

Incoming interface: Loopback99, RPF nbr: 10.0.0.99

Outgoing interface list: (count: 1)

Ethernet2/11, uptime: 00:20:19, pim

(10.0.0.31/32, 230.0.160.1/32), uptime: 00:20:02, pim mrib ip

Incoming interface: Ethernet2/11, RPF nbr: 10.31.41.31, internal

Outgoing interface list: (count: 0)

(10.0.0.42/32, 230.0.160.1/32), uptime: 01:54:04, pim mrib ip

Incoming interface: Ethernet2/7, RPF nbr: 10.41.42.42, internal

Outgoing interface list: (count: 1)

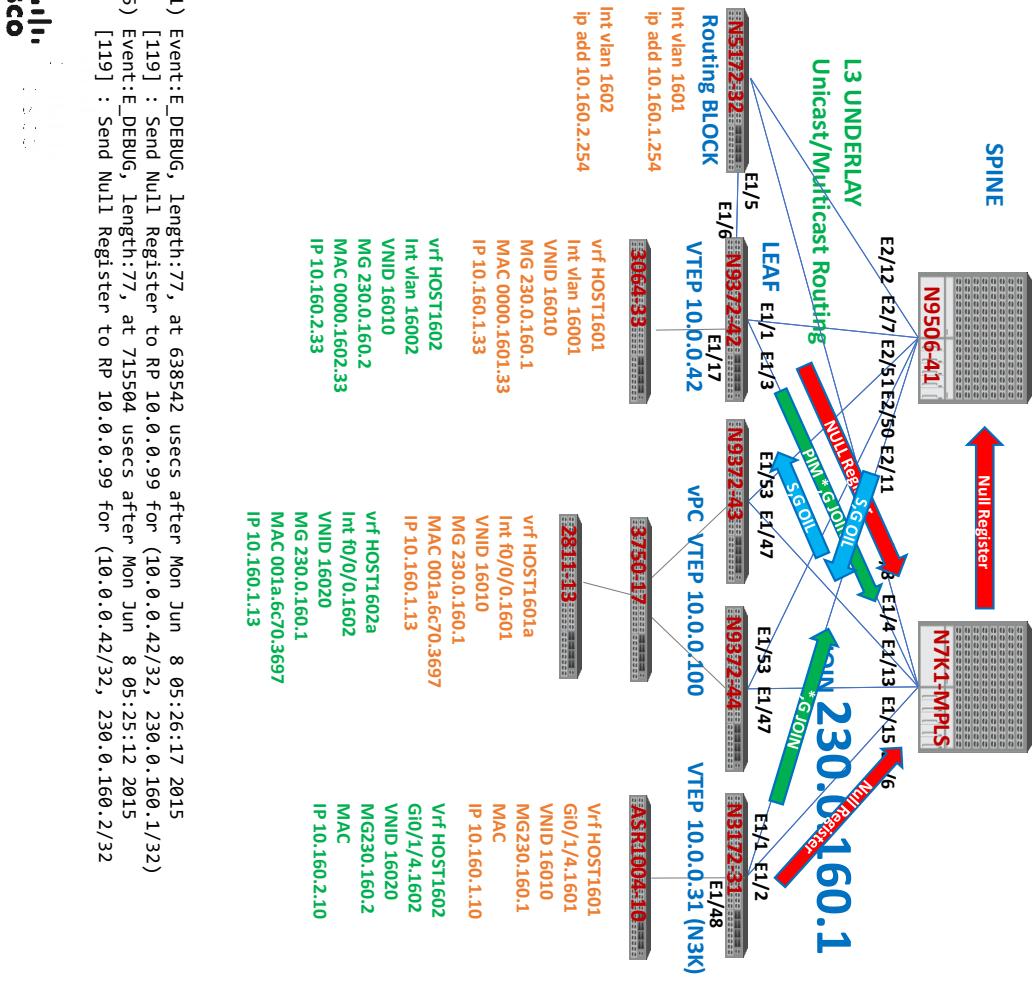
Ethernet2/11, uptime: 00:20:19, pim

VTEP-42's S,G has OIL pointing to VTEP-31

N9504-41# sh ip pim event-history join-prune

6) Event:E_DEBUG, Length:78, at 409558 usecs after Mon Jun 8 10:50:30 2015
[117] : Put (10.0.0.42/32, 230.0.160.1/32), S in join-list for nbr 10.41.42.42

VXLAN Flood & Learn – Multicast



Same for the other ANYCAST RP, N7K1-MPLS, copies OIL from *,G to S,G for 10.0.0.31

MPLS# sh ip mrr

(*, 230.0.160.1/32), uptime: 01:19:21, pim ip

Incoming interface: Loopback99, RPF nbr: 10.0.0.99

Outgoing interface list: (count: 1)

Ethernet1/4,

uptime: 01:19:21, pim

(10.0.0.31/32, 230.0.160.1/32), uptime: 00:21:27, pim mrib ip

Incoming interface: Ethernet1/6, RPF nbr: 10.3.31.31, internal

Outgoing interface list: (count: 1)

Ethernet1/4,

uptime: 00:21:27, pim

(10.0.0.42/32, 230.0.160.1/32), uptime: 1d23h, pim mrib ip

Incoming interface: Ethernet1/4, RPF nbr: 10.3.42.42, internal

Outgoing interface list: (count: 0)

N9372-42# sh ip mrr

(*, 230.0.160.1/32), uptime: 01:24:04, nve ip pim

Incoming interface: Ethernet1/3, RPF nbr: 10.3.42.3

Outgoing interface list: (count: 1)

nve1, uptime: 01:24:04, nve

(10.0.0.31/32, 230.0.160.1/32), uptime: 00:24:38, ip mrib pim nve

Incoming interface: Ethernet1/3, RPF nbr: 10.3.42.3

Outgoing interface list: (count: 1)

nve1, uptime: 00:24:38, nve

(10.0.0.42/32, 230.0.160.1/32), uptime: 01:24:04, nve ip mrib pim

Incoming interface: Loopback0, RPF nbr: 10.0.0.42

Outgoing interface list: (Count: 1)

Ethernet1/1, uptime: 00:24:54, pim

(10.0.0.31/32, 230.0.160.1/32), uptime: 00:24:54, pim

Incoming interface: Loopback0, RPF nbr: 10.0.0.42

Outgoing interface list: (Count: 1)

Ethernet1/1, uptime: 00:24:54, pim

(10.0.0.31/32, 230.0.160.1/32), uptime: 00:24:54, pim

Incoming interface: Loopback0, RPF nbr: 10.0.0.31, uptime: 00:24:54, pim

Outgoing interface list: (Count: 1)

Ethernet1/2, uptime: 00:24:54, pim

(10.0.0.31/32, 230.0.160.1/32), uptime: 00:24:54, pim

Incoming interface: Loopback0, RPF nbr: 10.0.0.31, uptime: 00:24:54, pim

Outgoing interface list: (Count: 1)

Ethernet1/2, uptime: 00:24:54, pim

(10.0.0.31/32, 230.0.160.1/32), uptime: 00:24:54, pim

Incoming interface: Loopback0, RPF nbr: 10.0.0.31, uptime: 00:24:54, pim

Outgoing interface list: (Count: 1)

Ethernet1/2, uptime: 00:24:54, pim

WHY N9372-42 sees N3172-31 although there is no Flooding traffic?

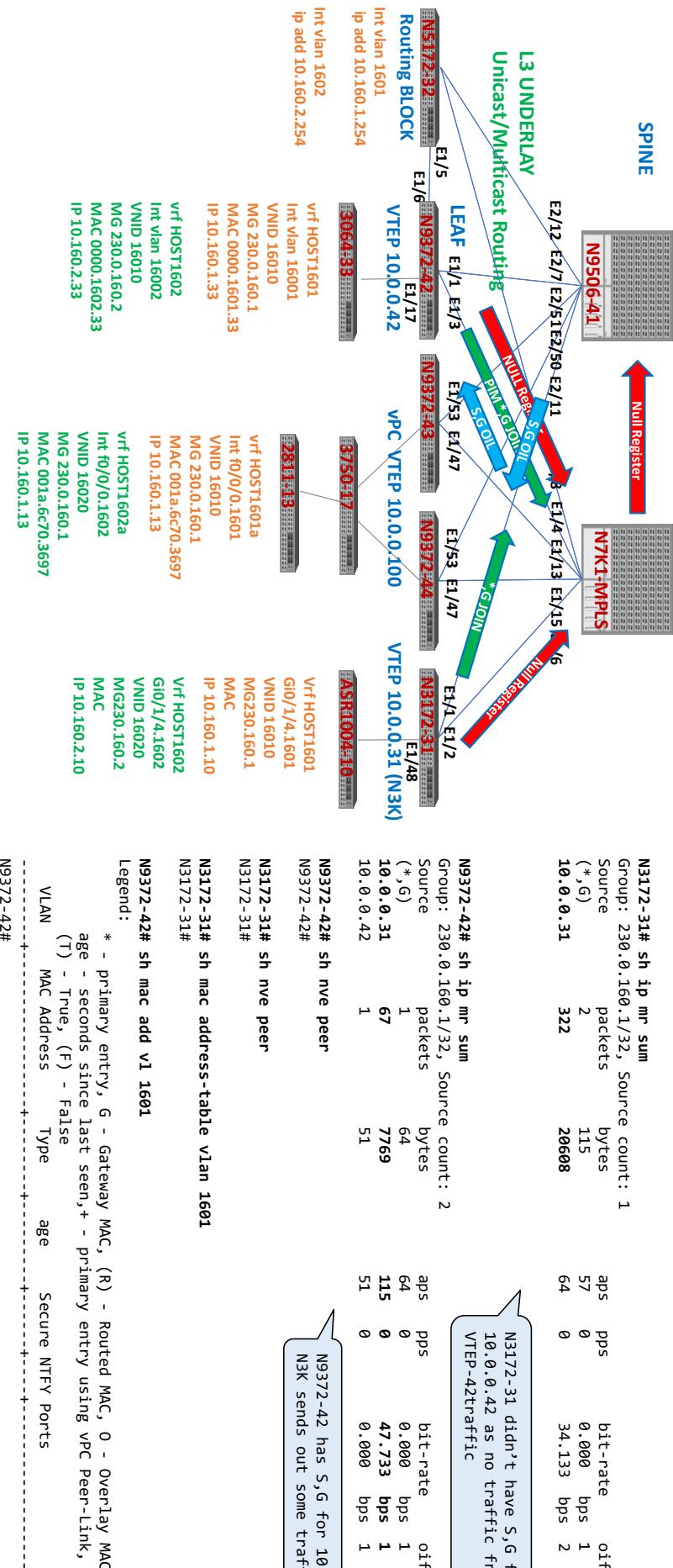
N93K seems like actually sending some bogus pkt to create S,G path

N9372-42# sh ip mrr sum

Group:	230.0.160.1/32, Source count:	2				
Source packets	bytes	aps	pps	bit-rate	oifs	
(*,6)	64	64	0	0.000 bps	1	
10.0.0.31	67	7769	115	0	47.733 bps	1
10.0.0.42	1	51	51	0	0.000 bps	1

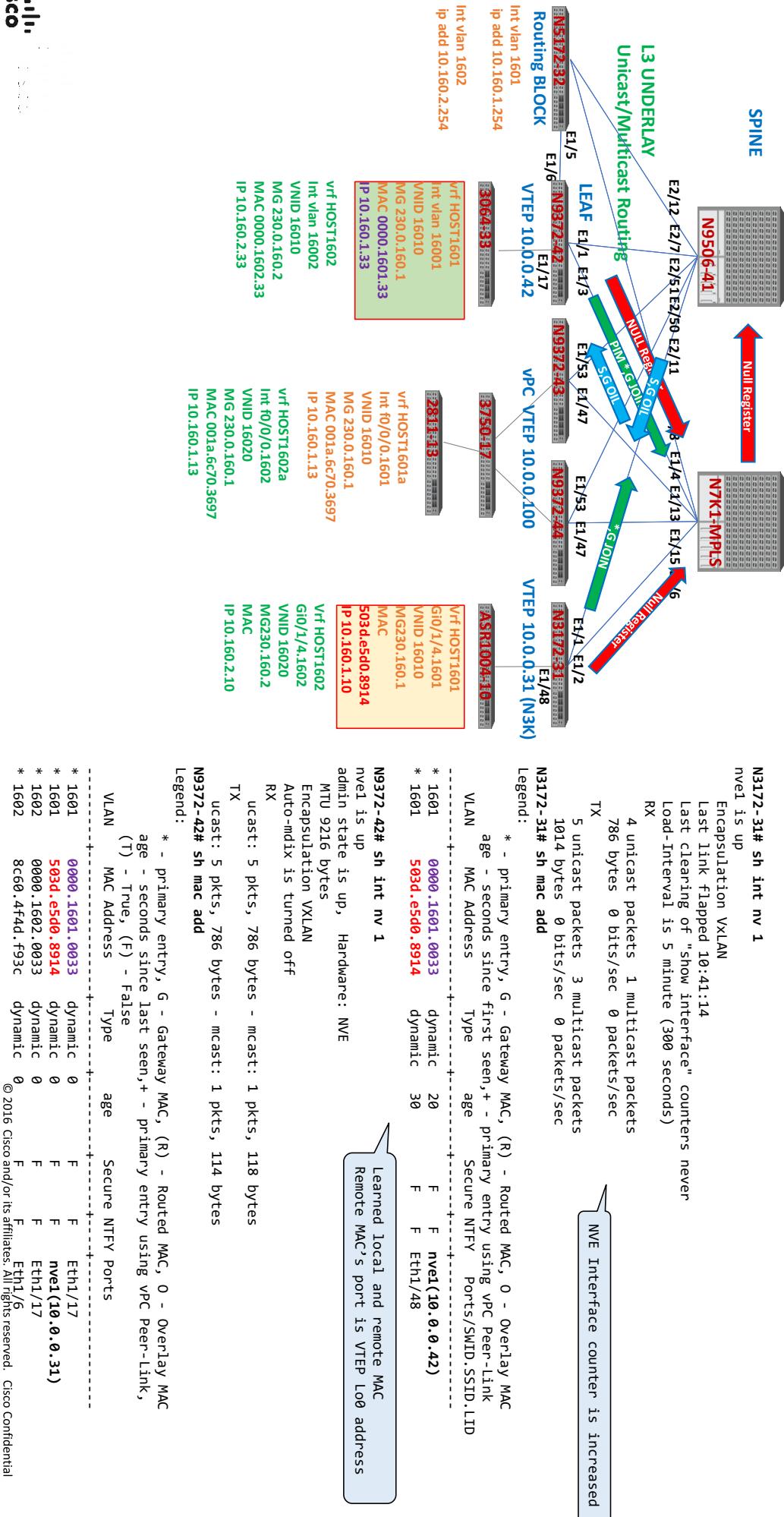
VXLAN Flood & Learn – Multicast

So, we can see N3K actually keep sending out packets over S,G although there is no actual traffic flows and no mac address is learned.



VXLAN Flood & Learn – Multicast

Let's ping from HOST ASR10 behind N3172-31 to host N3064-33 behind N9372-42
ASR1004-1# ping vrf host1601 10.160.1.33 repeat 6
 Sending 6, 100-byte ICMP Echos to 10.160.1.33, timeout is 2 seconds:
 ...!!!



VXLAN Flood & Learn – Multicast



```
N3172-31# sh nve peer detail
Peer: 10.0.0.42 Peer-IP: 10.0.0.42 Peer-State: Up
Interface: nve1
Peer learnt VNI : 16010
Configured VNI's : 16010,16020
Provision State : add-complete
Route Update : Yes
Uptime : 00:04:55
```

HM is ready

Now we can see NVE Peer is learned



```
N9372-42# sh nve peer
Interface Peer-IP: 10.0.0.31 State: Up LearnType: DP Uptime: 00:05:40 Router-Mac: n/a
```



```
N9372-42# sh nve internal platform interface nve 1 detail
Printing Interface ifindex 0x49000001 detail
=====|=====|=====|=====|=====|=====|=====|=====
Vrf HOST1601 | Intf | State | Prt-IP | Sec-IP | Vlans | Peers |
G10/1/4,1601 | ===== | ===== | ===== | ===== | ===== | ===== |
VNUID 16010 | ===== | ===== | ===== | ===== | ===== | ===== |
MG 230.0.160.1 | ===== | ===== | ===== | ===== | ===== | ===== |
MAC 0000.1601.33 | ===== | ===== | ===== | ===== | ===== | ===== |
IP 10.160.1.33 | ===== | ===== | ===== | ===== | ===== | ===== |
=====|=====|=====|=====|=====|=====|=====|=====
MAC 001a.6c70.3697 | ===== | ===== | ===== | ===== | ===== | ===== |
IP 10.160.1.13 | ===== | ===== | ===== | ===== | ===== | ===== |
=====|=====|=====|=====|=====|=====|=====|=====
Vrf HOST1602 | Intf | State | Prt-IP | Sec-IP | Vlans | Peers |
G10/1/4,1602 | ===== | ===== | ===== | ===== | ===== | ===== |
VNUID 16020 | ===== | ===== | ===== | ===== | ===== | ===== |
MG 230.0.160.2 | ===== | ===== | ===== | ===== | ===== | ===== |
MAC 0000.1602.33 | ===== | ===== | ===== | ===== | ===== | ===== |
IP 10.160.2.10 | ===== | ===== | ===== | ===== | ===== | ===== |
=====|=====|=====|=====|=====|=====|=====|=====
Peers of interface nve1:
=====|=====|=====|=====|=====|=====|=====|=====

```

```
peer_ip: 10.0.0.31, peer_id: 1, state: UP MAC-learning: Enabled
active_swbd:
add_pending_swbd:
rem_pending_swbd:
```

VXLAN Flood & Learn – Multicast

SPINE

Null Register

N3172-31# sh ip mr
(*, 230.0.160.1/32), uptime: 02:47:13, static ip pim
Incoming interface: Ethernet1/1, RPF nbr: 10.31.41.41, uptime: 01:12:32
Outgoing interface list: (count: 1)
loopback0, uptime: 02:47:13, static
(10.0.0.31/32, 230.0.160.1/32), uptime: 02:46:56, ip mrib pim
Incoming interface: loopback0, RPF nbr: 10.0.0.31, uptime: 02:46:56
Outgoing interface list: (count: 2)
Ethernet1/2, uptime: 01:12:15, pim
loopback0, uptime: 02:46:56, mrib, (RPF)

Only 1st packet is leaked to CPU for learning NVE Peer

N3172-31# sho nve internal event-history packets
NVE - Packets:
1) Event:E_DEBUG, length:104, at 804045 usecs after Mon Jun 8 03:41:39 2015
[108] [3812]: VXLAN Packet - UDP Port: 4789 VNI: 16010 VTEP: 0x22280001 Source: 10.0.0.42 Dest: 10.0.0.31

Check NVE Peer Discovery Event

N3172-31# sh nve internal event-history events
7) Event:E_DEBUG, length:54, at 805330 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: URIB RNH notification added for 10.0.0.42
8) Event:E_DEBUG, length:80, at 80586 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Received peer add done from PD for peer 10.0.0.42 vni 16010 on nve1
9) Event:E_DEBUG, length:95, at 804971 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Sending Peer 10.0.0.42 [State: add-pending] add request on vni 16010 to PD on nve1
10) Event:E_DEBUG, length:79, at 804875 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Created Peer 10.0.0.42 on VNI 16010 on interface nve1 - State Init
11) Event:E_DEBUG, length:79, at 804870 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: XOS_EVT: Added Peer 10.0.0.42 on VNI 16010 on interface 0x22280001
12) Event:E_DEBUG, length:88, at 804666 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: XOS_EVT: Request to add Peer 10.0.0.42 on VNI 16010 on interface 0x22280001
13) Event:E_DEBUG, length:81, at 804660 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Peer 10.0.0.42 discovery notification on VNI 16010 on interface nve1

Still, there is no s,g for 10.0.0.42 on VTEP-31 as N9372-42 never sent any BUM traffic to delivery multicast group yet.

1. VTEP31 sent ARP in delivery multicast group, 230.0.160.1
2. VTEP42 received multicast and decap and fwded to host10
3. Host10 ARP reply is **unicast** from VTEP42 to VTEP31 as VTEP42 learned HOST33's MAC from ARP and knows VTEP31's IP 0.0.0.42.

```

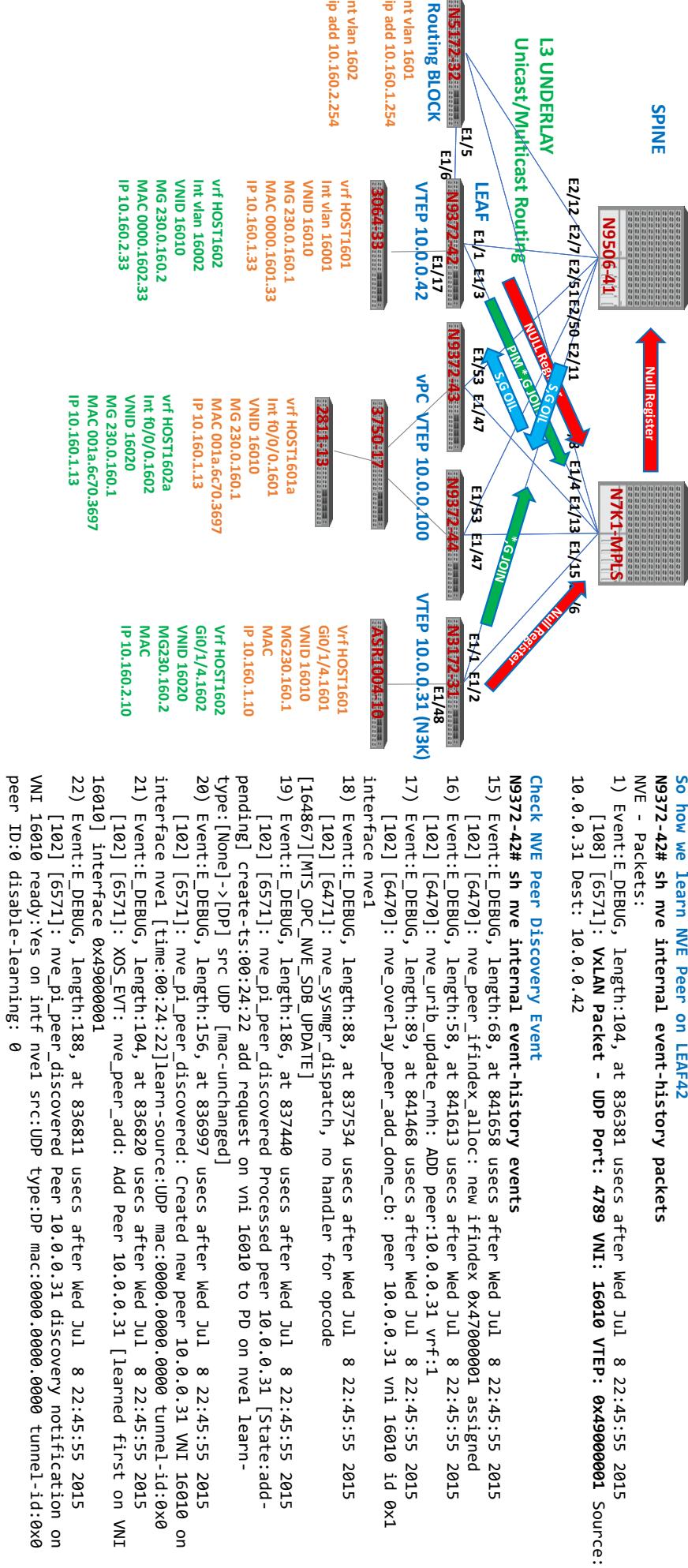
N3172-31# sh ip mr
(*, 230.0.160.1/32), uptime: 02:47:13, static ip pim
Incoming interface: Ethernet1/1, RPF nbr: 10.31.41.41, uptime: 01:12:32
Outgoing interface list: (count: 1)
loopback0, uptime: 02:47:13, static
(10.0.0.31/32, 230.0.160.1/32), uptime: 02:46:56, ip mrib pim
Incoming interface: loopback0, RPF nbr: 10.0.0.31, uptime: 02:46:56
Outgoing interface list: (count: 2)
Ethernet1/2, uptime: 01:12:15, pim
loopback0, uptime: 02:46:56, mrib, (RPF)

So how we Learn NVE Peer
N3172-31# sho nve internal event-history packets
NVE - Packets:
1) Event:E_DEBUG, length:104, at 804045 usecs after Mon Jun 8 03:41:39 2015
[108] [3812]: VXLAN Packet - UDP Port: 4789 VNI: 16010 VTEP: 0x22280001 Source: 10.0.0.42 Dest: 10.0.0.31

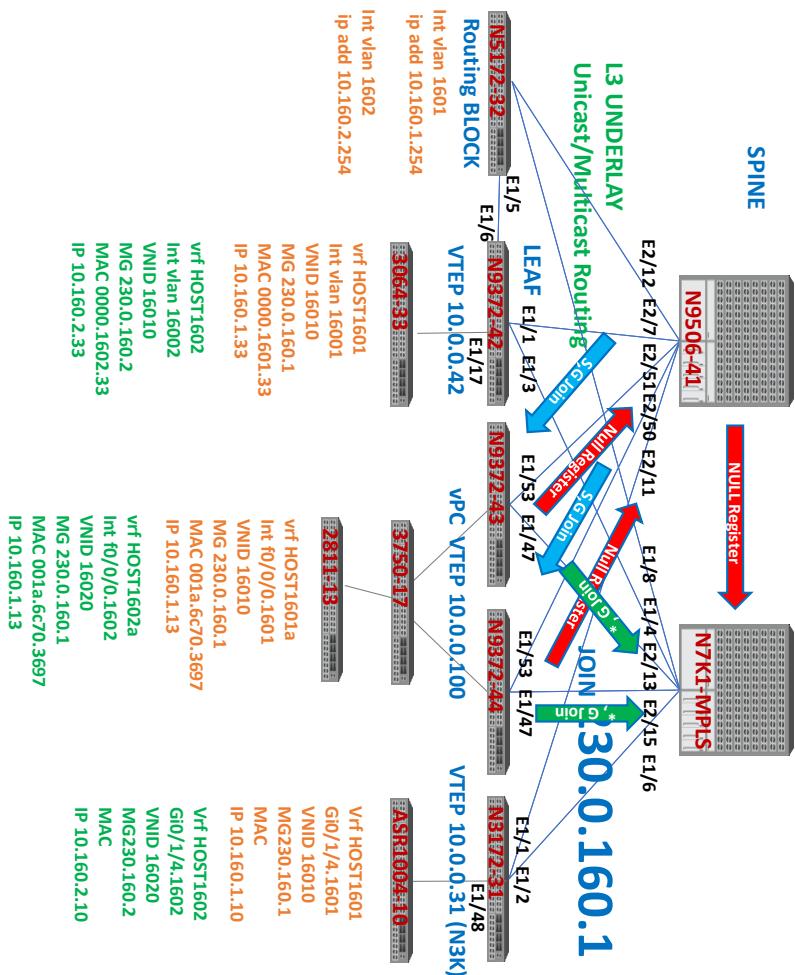
Check NVE Peer Discovery Event
N3172-31# sh nve internal event-history events
7) Event:E_DEBUG, length:54, at 805330 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: URIB RNH notification added for 10.0.0.42
8) Event:E_DEBUG, length:80, at 80586 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Received peer add done from PD for peer 10.0.0.42 vni 16010 on nve1
9) Event:E_DEBUG, length:95, at 804971 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Sending Peer 10.0.0.42 [State: add-pending] add request on vni 16010 to PD on nve1
10) Event:E_DEBUG, length:79, at 804875 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Created Peer 10.0.0.42 on VNI 16010 on interface nve1 - State Init
11) Event:E_DEBUG, length:79, at 804870 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: XOS_EVT: Added Peer 10.0.0.42 on VNI 16010 on interface 0x22280001
12) Event:E_DEBUG, length:88, at 804666 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: XOS_EVT: Request to add Peer 10.0.0.42 on VNI 16010 on interface 0x22280001
13) Event:E_DEBUG, length:81, at 804660 usecs after Mon Jun 8 03:41:39 2015
[102] [3812]: Peer 10.0.0.42 discovery notification on VNI 16010 on interface nve1

```

VXLAN Flood & Learn – Multicast



VXLAN Flood & Learn – VPC



Multicast on VPC

SRC outside of VPC (Encap)

- a. Only one LEAF became DECAPPER closer to RP
- b. Non-DECAP drop pkt

- SRC inside of VPC (Decap)
 - a. Both vPC Peers are ENCAPPER and FWD pkt if they get S,G join
 - b. PKT received on one PEER copied over to other peer by vlan 4041

vpc-domain 434
peer-switch
peer-keepalive destination 10.1.1.2 source 10.1.1.1 vrf vpck

interface Vlan80
description MCT SVI
ip address 10.1.80.43/24
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

vrf HOST1601
G10/14.1601
VNID 16010
MG 230.0.160.1
MAC 0000.1601.33
IP 10.1.60.1.33

N9372TX-43# sh run int 10 0
interface loopback0
ip address 10.0.0.43/32

ip address 10.0.0.100/32 secondary
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

N9372TX-43# sh run int 10 0
interface loopback0
ip address 10.0.0.44/32

ip address 10.0.0.100/32 secondary
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

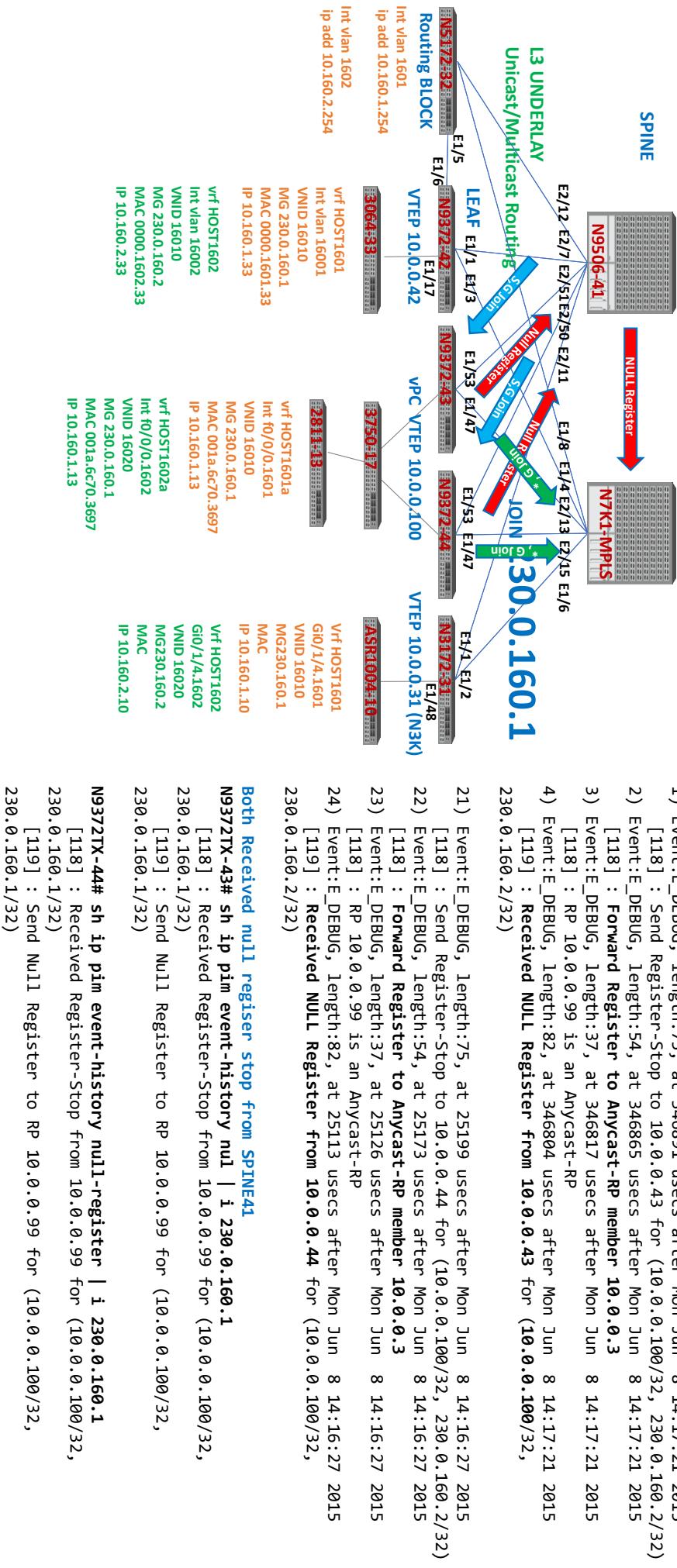
N9372TX-44# sh run int 10 0
interface loopback0
ip address 10.0.0.44/32

ip address 10.0.0.100/32 secondary
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

N9372TX-43# sh nve interface

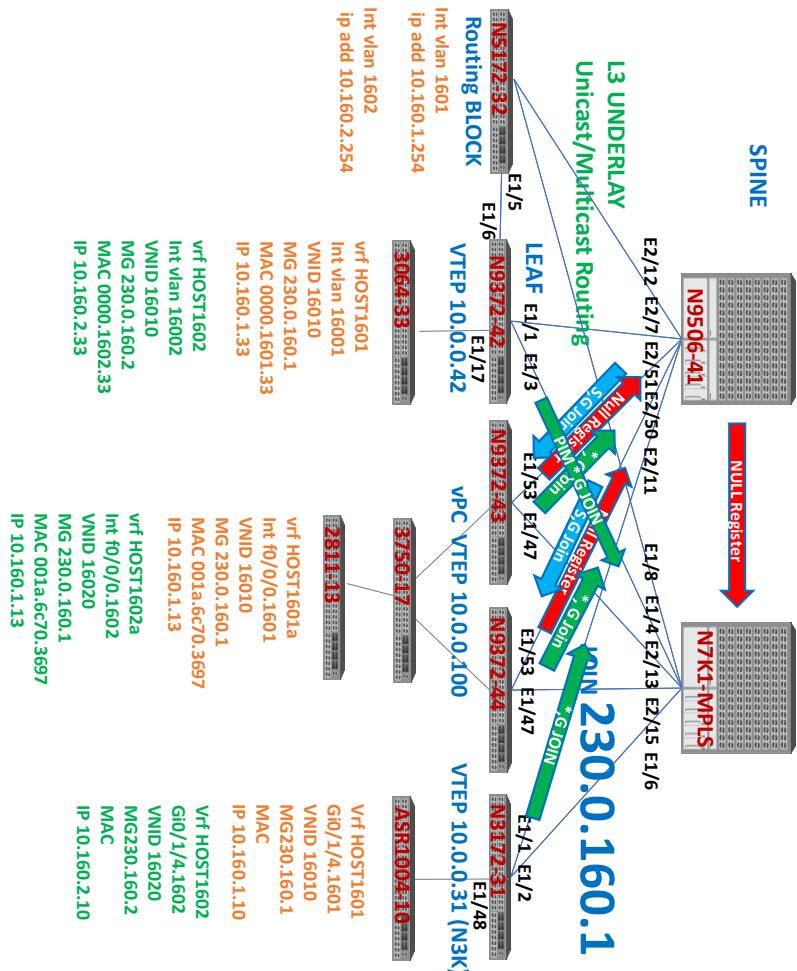
```
Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VIP-Only [notified]
Local Router MAC: 88f0.3187.8c4b
Host Learning Mode: Data-Plane
Source-Interface: loopback0 (primary: 10.0.0.43, secondary: 10.0.0.100)
```

VXLAN Flood & Learn – VPC



VXLAN Flood & Learn – VPC

Both VPC Peers sent *,G Join to N7K1-MPLS
There is no traffic as I shut down all ACCESS traffic



VXLAN Flood & Learn – VPC

All other VTEP is “non-vPC source” so vPC elects one forwarder as expected.

We call it as “decaper” as it decap packets

N9372TX-43# sh ip pim internal vpc rpf



Source: 10.0.0.31
Pref/Metric: 110/3
Source role: primary
Forwarding state: Tie (forwarding)
MRIB Forwarding state: forwarding

Source: 10.0.0.42
Pref/Metric: 110/3
Source role: secondary
Forwarding state: Tie (not forwarding)
MRIB Forwarding state: not forwarding

Source: 10.0.0.31
Pref/Metric: 110/3
Source role: primary
Forwarding state: Tie (forwarding)
MRIB Forwarding state: forwarding

Source: 10.0.0.42
Pref/Metric: 110/3
Source role: secondary
Forwarding state: Tie (not forwarding)
MRIB Forwarding state: not forwarding



Source: 10.0.0.100
Pref/Metric: 0/0
Source role: primary
Forwarding state: Win-force (forwarding)
MRIB Forwarding state: forwarding

Source: 10.0.0.100
Pref/Metric: 0/0
Source role: secondary
Forwarding state: Win-force (forwarding)
MRIB Forwarding state: forwarding

Source: 10.0.0.100
Pref/Metric: 0/0
Source role: primary
Forwarding state: Win-force (forwarding)
MRIB Forwarding state: forwarding

Source: 10.0.0.100
Pref/Metric: 0/0
Source role: secondary
Forwarding state: Win-force (forwarding)
MRIB Forwarding state: forwarding



So, only DECAPER will decap packets and non-Decaper will drop packets
N9372TX-43# sh ip mri det
(10.0.0.31/32, 230.0.160.1/32), uptime: 03:12:02, ip(0) mrib(0) nve(1) pim(0)
Data Created: Yes

VXLAN Flags
VTEP43 is decapper for 10.0.0.31,G and
10.0.0.42,G as it won FWD election



VPC Flags
RPF-Source Forwarder
Stats: 247/28469 [Packets/Bytes], 13.600 bps
Stats: Active Flow
Incoming interface: Ethernet1/47, RPF nbr: 10.3.43.3

Outgoing interface list: (count: 1)
nve1, uptime: 03:12:02, nve
(10.0.0.42/32, 230.0.160.1/32), uptime: 03:11:21, pim(1) mrib(0) ip(0) nve(1)
Data Created: Yes



VXLAN Decap
VPC Flags
RPF-Source Forwarder
Stats: 1/51 [Packets/Bytes], 0.000 bps
Stats: Inactive Flow
Incoming interface: Ethernet1/47, RPF nbr: 10.3.43.3

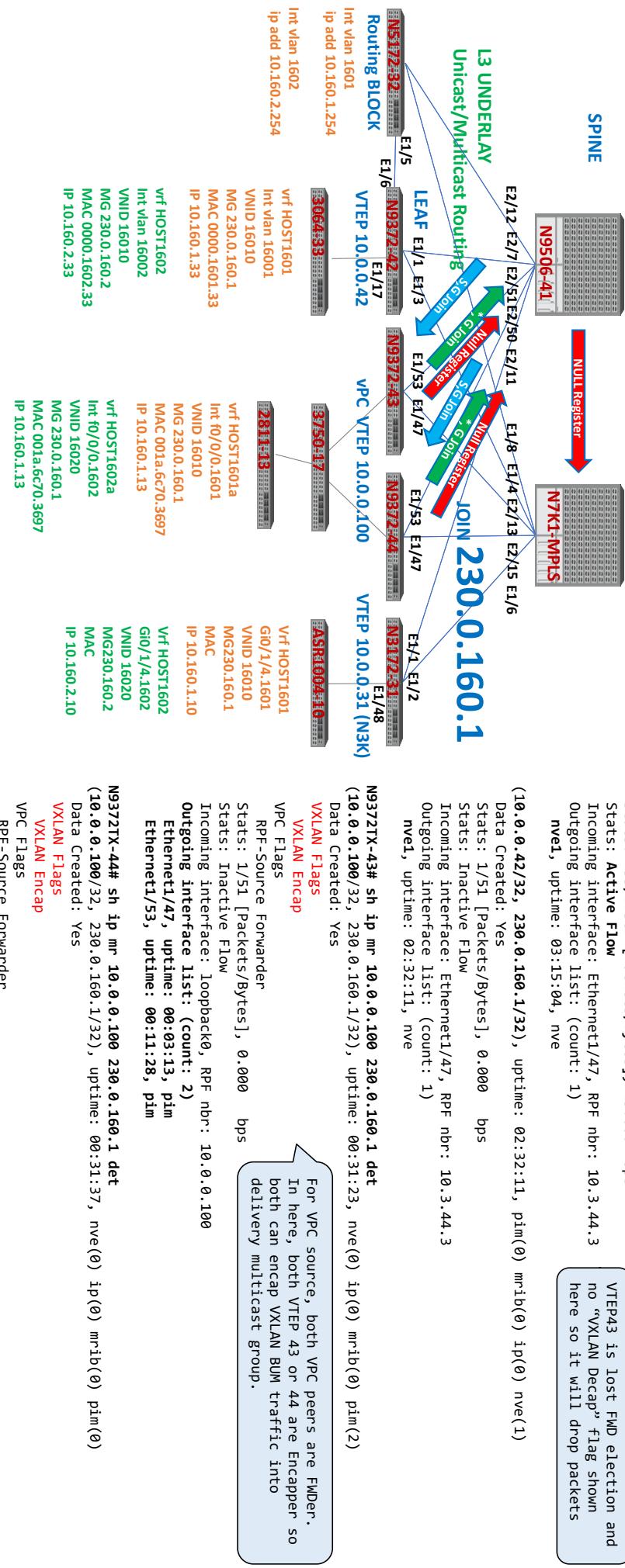
Outgoing interface list: (count: 2)
Ethernet1/53, uptime: 01:20:44, pim(1) mrib(0) ip(0) nve(1), uptime: 03:11:21, nve
nve1, uptime: 03:11:21, nve



VXLAN Decap
VPC Flags
RPF-Source Forwarder
Stats: 1/51 [Packets/Bytes], 0.000 bps
Stats: Inactive Flow
Incoming interface: Ethernet1/47, RPF nbr: 10.3.43.3

Outgoing interface list: (count: 2)
Ethernet1/53, uptime: 01:20:44, pim(1) mrib(0) ip(0) nve(1), uptime: 03:11:21, nve
nve1, uptime: 03:11:21, nve

VXLAN Flood & Learn – VPC



cisco

VXLAN Flood & Learn – Routing

Before 7.0(3)I2.2, N9K can't do both SVI routing and VXLAN GW on same box with flood & learn.
With EVPN, N9K does Anycast GW with SVI for L3 routing with VRF.

So, I created separate routing block on N5K here.

N5K is just normal L2/L3 switch without VXLAN enabled so N9372-42 is working as VXLAN L2 GW and

bridging VXLAN and VLAN to N5172-32.

All hosts points def GW as 10.160.1.254 or 10.160.2.254

And, you can test L2 GW function and L3 routing by telnetting to R13.

R13 can ping all other hosts in same vlan and different vlan as below

```
2800-13#ping vrf host1601a 10.160.2.10
Sending 5, 100-byte ICMP Echoes to 10.160.2.10, timeout is 2 seconds: .!!!
```

```
N5672-32# sh int vl 1601
Vlan1601 is up, line protocol is up
Hardware is EtherSVI, address is 8c60.4f4d.f93c
```

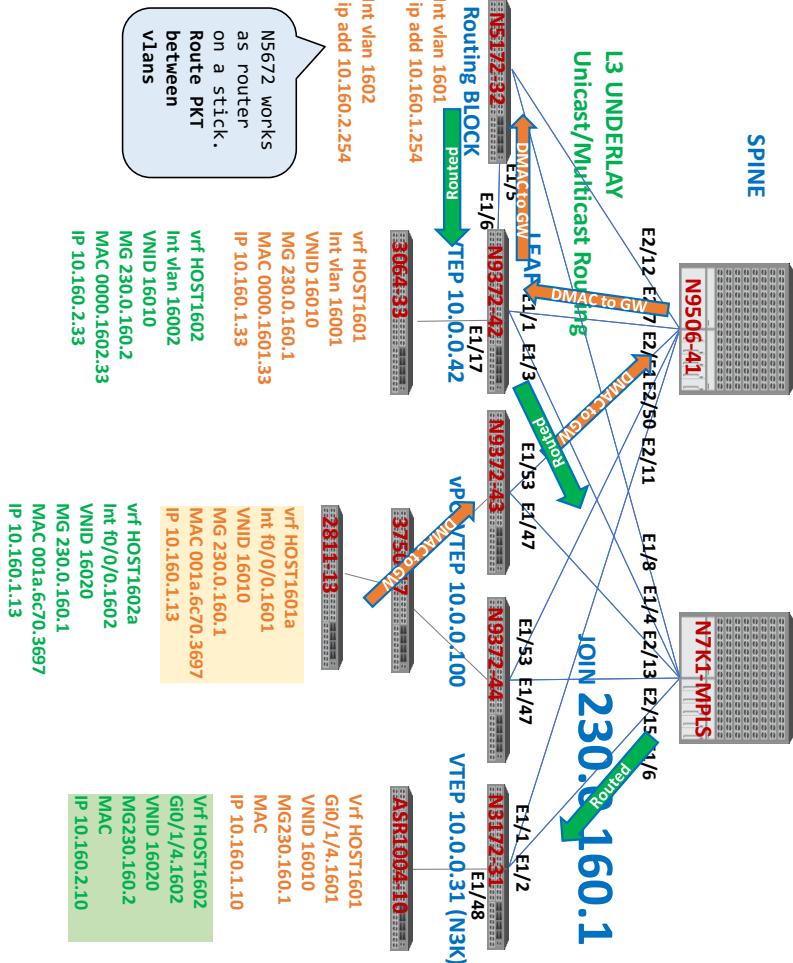
H05743 send ping with D.MAC of
def GW, 8c60.4f4d.f93c
VTEP43 found dmac is behind VTEP
10.0.0.42

```
N9372TX-43# sh mac add
VLAN MAC Address Type age Secure NTFY Ports
-----+-----+-----+-----+-----+
* 1601 001a.6c70.3697 dynamic 0 F F F nve1(10.0.0.42)
* 1601 8c60.4f4d.f93c dynamic 0 F F F nve1(10.0.0.42)
```

H05743 send ping with D.MAC of
def GW, 8c60.4f4d.f93c
VTEP43 found dmac is behind VTEP
10.0.0.42

```
N9372-42# sh mac add vl 1601
VLAN MAC Address Type age Secure NTFY Ports
-----+-----+-----+-----+-----+
* 1601 001a.6c70.3697 dynamic 0 F F F nve1(10.0.0.100)
* 1601 8c60.4f4d.f93c dynamic 0 F F F Eth1/6
```

H05743 send ping with D.MAC of
def GW, 8c60.4f4d.f93c
VTEP42 received the pkt and look
at local mac table and fwd to
N5172. VXLAN to VLAN Bridge

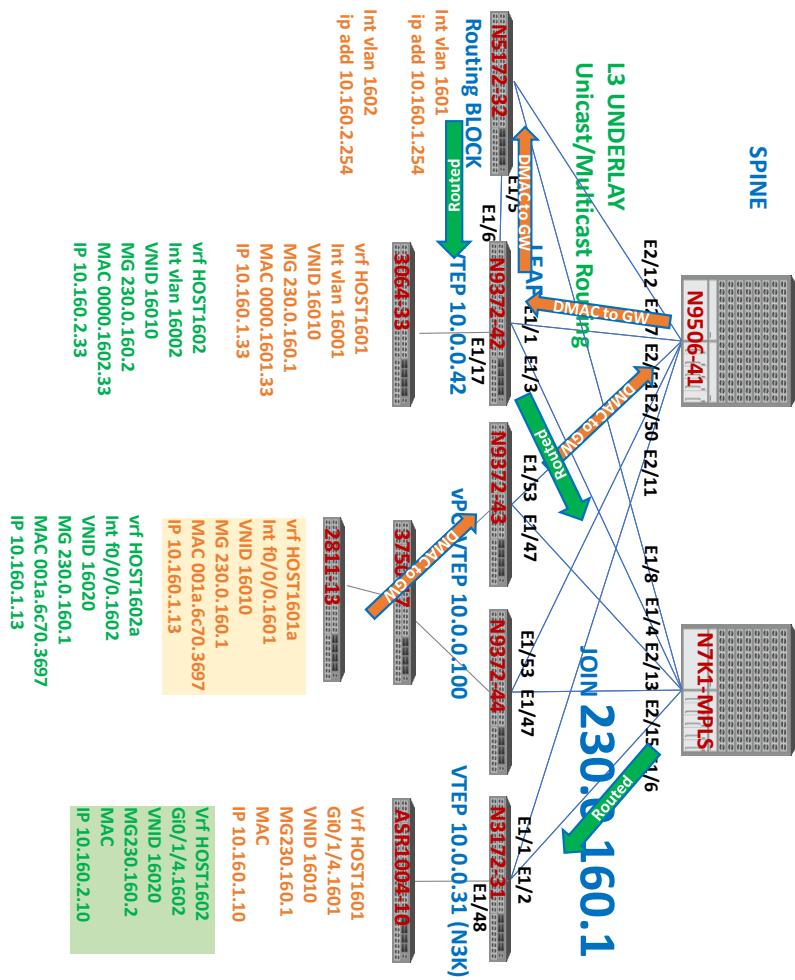


```
N9372-42# sh mac add vl 1602
VLAN MAC Address Type age Secure NTFY Ports
-----+-----+-----+-----+-----+
* 1602 503d.e5d0.8914 dynamic 320 F F Eth1/5
```

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

cisco

VXLAN Flood & Learn – Summary



1. Check VLAN to VNI Mapping
2. Check NVNE is up and configured properly
3. Check L3 Underlay configuration
 - Unicast routing
 - Multicast PIM Anycast RP
4. VXLAN Manager create *,G and S,G then send out Null Register to PIM RP
5. PIM Anycast RP will forward Null Register to all Anycast Set
6. RP that received *,G will create proper OIL and copy OIL to S,G addresses
7. Check local MAC address table to see if VTEP learns local mac addresses
8. If remote MAC is not learned, then troubleshoot multicast Core and make sure proper *,G and S,G tree is built

VXLAN Flood & Learn – Ingress Replication

Like VPLS static peering, you can manually configure all VTEPs loopbacks under

NVE interface and VTEP replicates BUM traffic to all peers

```
interface nve1
  ingress-replication protocol static
    peer-ip 10.0.0.31
    peer-ip 10.0.0.42
  member vni 16110
  member vni 16110
  ingress-replication protocol static
    peer-ip 10.0.0.31
    peer-ip 10.0.0.42
  member vni 16120
  member vni 16120
  ingress-replication 10.0.0.100
  ingress-replication 10.0.0.100
```



```
N9372-42# sh nve vni ingress-replication
```

Interface	VNI	Replication List	Source	Up Time
vne1	16110	10.0.0.31	CLI	00:08:19
vne1	16120	10.0.0.100	CLI	00:08:19
vne1	16120	10.0.0.31	CLI	00:08:19
vne1	16120	10.0.0.100	CLI	00:08:17

```
N9372-42# sh 12route evpn f1 all
```

Topology ID	Peer-id	Flood List	Service Node
16111	2	10.0.0.31	no
16111	1	10.0.0.100	no
16112	2	10.0.0.31	no
16112	1	10.0.0.100	no

```
N9372-42# sh forwarding distribution peer-id
```

UDFM Peer-id	allocations:
App: VXLAN	Vlan: 1 Id: 0x100a00001f Peer-id: 0x2
App: VXLAN	Vlan: 1 Id: 0x10a000064 Peer-id: 0x1
App: VXLAN	Vlan: 4095 Id: 0xa00001f Peer-id: 0x9db
App: VXLAN	Vlan: 4095 Id: 0xa000064 Peer-id: 0x9da

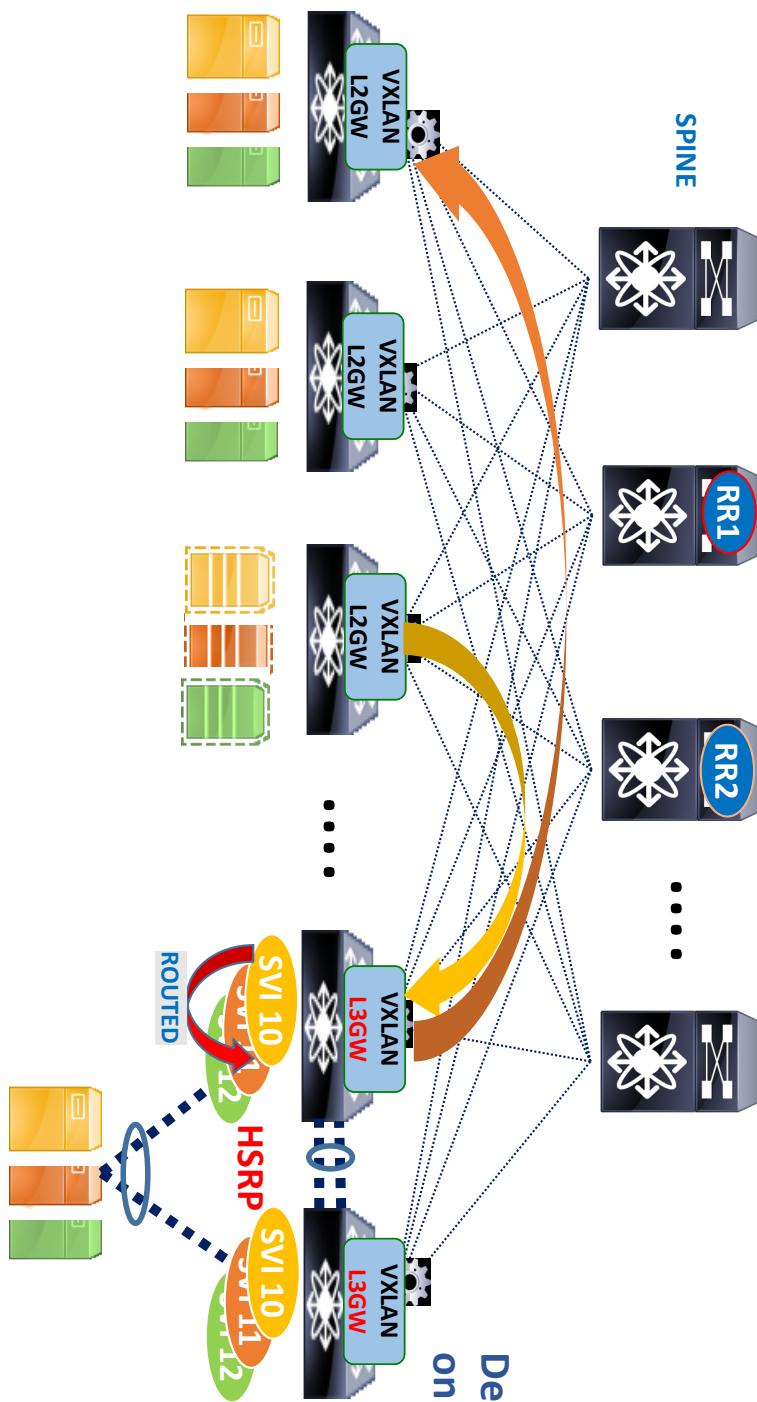
```
N9372-42# sh forwarding distribution internal ingress-replication-peer
```

```
Total Ingress-replication Peers: 2
Peer-0x9da vlan list:
  VL-1611 VL-1612
Peer-0x9db vlan list:
  VL-1611 VL-1612
Peer-0x9dc vlan list:
  VL-1611 VL-1612

2800-13#ping vrf host1611a 10.161.1.33
Sending 5, 100-byte ICMP Echoes to 10.161.1.33, timeout is 2 seconds:
.!!!
```

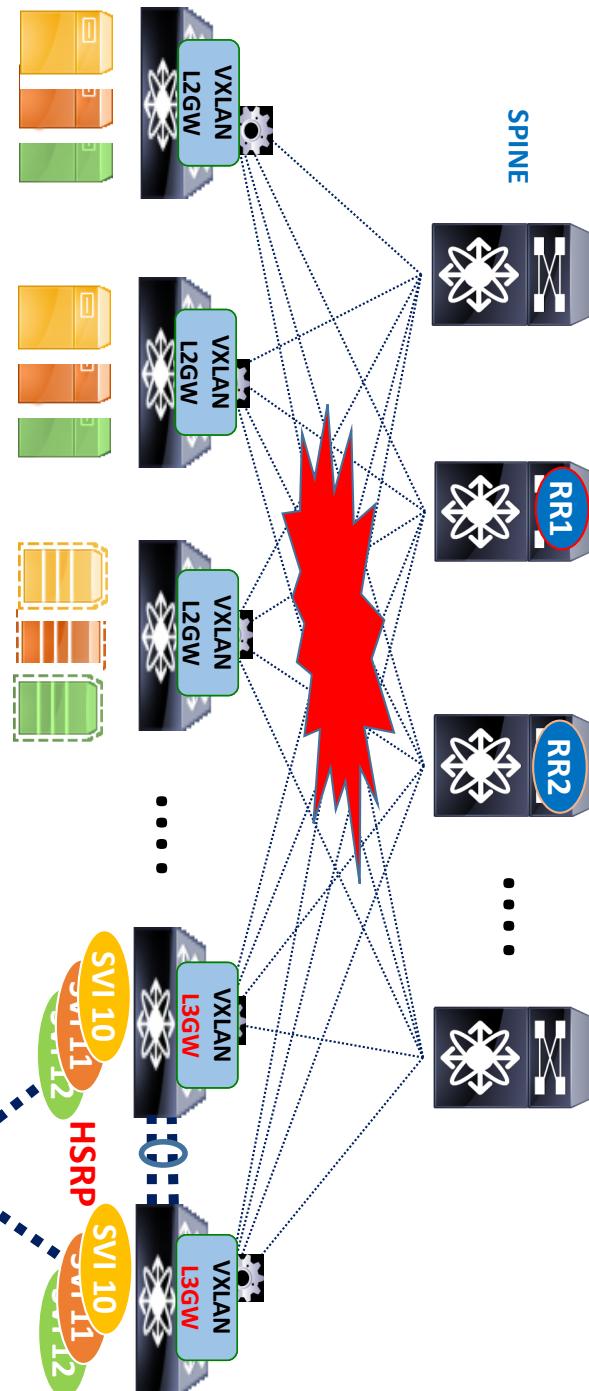
VXLAN Flood & Learn: Centralized GW

Def GW is configured only
on a VPC Pair with HSRP



L3 GW on N9K VTEP requires “**hardware access-list tcam region arp-ether 256**” and Reload

VXLAN Flood & Learn: Limitation



Rely on Flooding

- Simple multicast problem in CORE can cause whole network down
- Sub-Optimal Routing path
- Must travel to Central GW and back



VXLAN with MP-BGP EVPN

Present

VXLAN Evolution with EVPN

- Truly Scalable with protocol-driven control plane
 - Multi-Protocol BGP based Control-Plane using EVPN NLRI (**Network Layer Reachability Info**)
 - Advertise/Learn MAC, MAC-IP, IP Prefix with MP-BGP
 - No more dependency on flooding for learning MAC and Peers
- Seamless routing and bridging on any VTEP
 - Make Forwarding decisions at any VTEPs for Layer-2 (MAC) and Layer-3 (IP);
 - Symmetric Integrated Route/Bridge (IRB)
 - Distributed Anycast GW on all VTEPs
- Minimize flooding through ARP Suppression
 - Without actual data traffic, VTEPs learn all /32 Host routes and MAC in FABIRC when a HOST is coming online and sending G-ARP
- Built-in L2/L3 Multi-Tenancy Support with VRF/VNI
- Security through VTEP Peer-Authentication with BGP

Multi-Tenancy Support

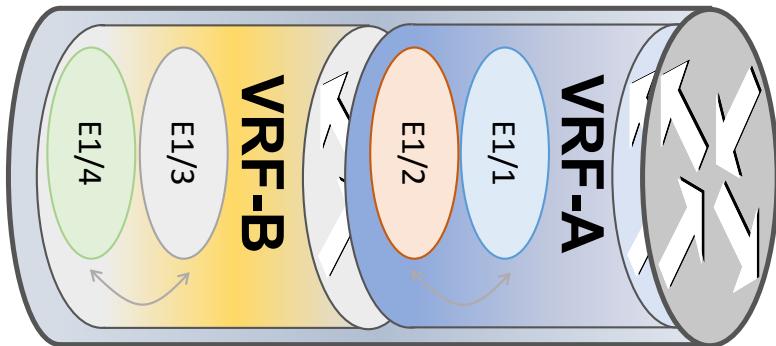
- Process of creating an environment where resources are split and combined, based on consumption, supply and policies
- Multi-Tenancy at Layer-2
 - Prevents hosts in a given Layer-2 segment from observing traffic of hosts in a different segment
 - Separation of Broadcast/Flood domains into bridge domains/segments
 - Containment of the Fault domain to a given Layer-2 bridge domain
- Multi-Tenancy at Layer 3
 - Prevents hosts in a given Layer-3 segment from communicating to hosts in a different L3 segment
 - Separation of Routing Table
 - Containment of the Fault Routing to a given VRF

Routing Domain - VRF

Virtual Routing and Forwarding (VRF)

- Allows multiple instances of a routing table to co-exist within the same router
- Independent IPv4 and IPv6 address spaces
- Full unicast and multicast routing protocols support
- All IP-based features in NX-OS are VRF aware
- Non-default VRFs are locally-significant on a router

Data traffic is not routed across VRFs with the default configuration

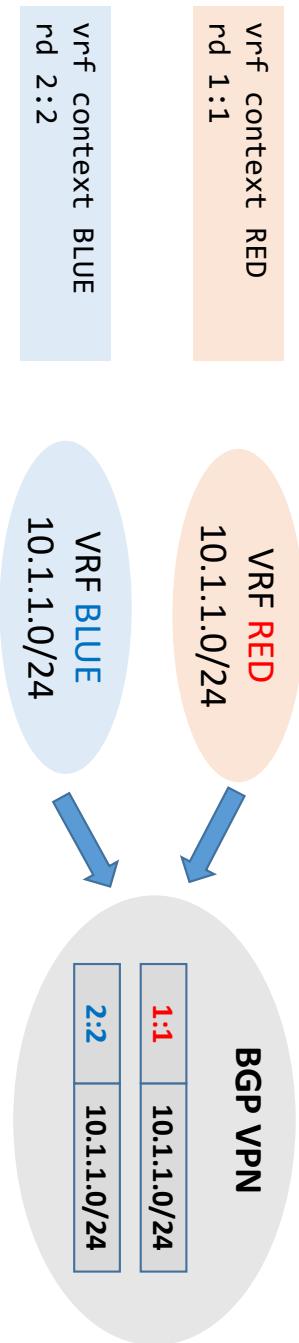


Multiprotocol BGP (MP-BGP)

- Multiprotocol BGP (MP-BGP)
- Extension to Border Gateway Protocol (BGP) - RFC 4760
- Allows BGP for carrying any “routable” protocol’s reachability information in BGP
- VPN Address-Family:
 - Allows different types of address families
 - VPNv4, VPNv6, L2VPN EVPN, MVPN
 - Information transported across single BGP peering

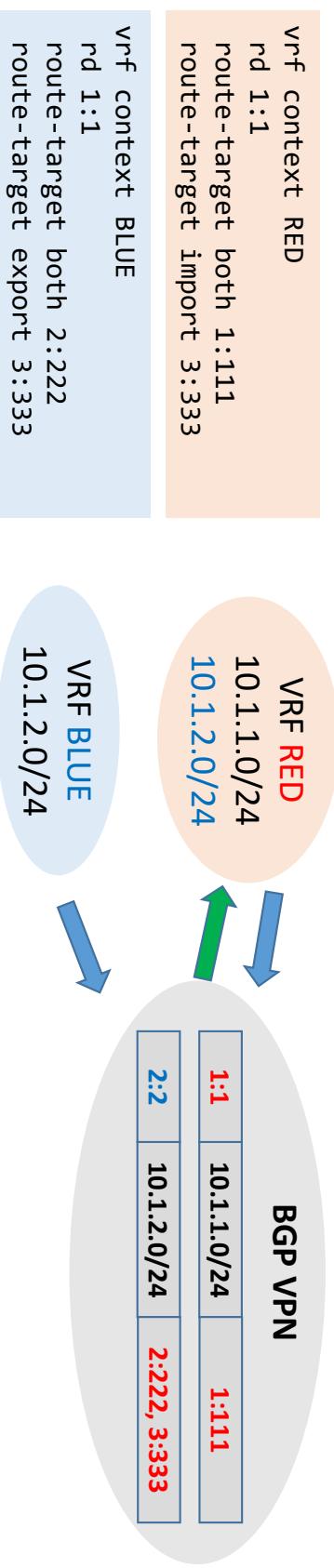
MP-BGP VPN: Route Distinguisher

- VRF allows overlapping IP addresses with isolated routing domains
- Need to track which route is belong to which VRF
- RD is a NLRI field
- Assign RD to each VRF and add RD to IP address to maintain uniqueness among identical routes in different VRFs
 - VRF "RED" has RD 1:1 and VRF "BLUE" has RD 2:2
 - 10.1.1.0/24 in VRF "RED" is exported into BGP VPN AF as "**1:1:10.1.1.0**"
 - 10.1.1.0/24 in VRF "Blue" is exported into BGP VPN AF as "**2:2:10.1.1.0**"

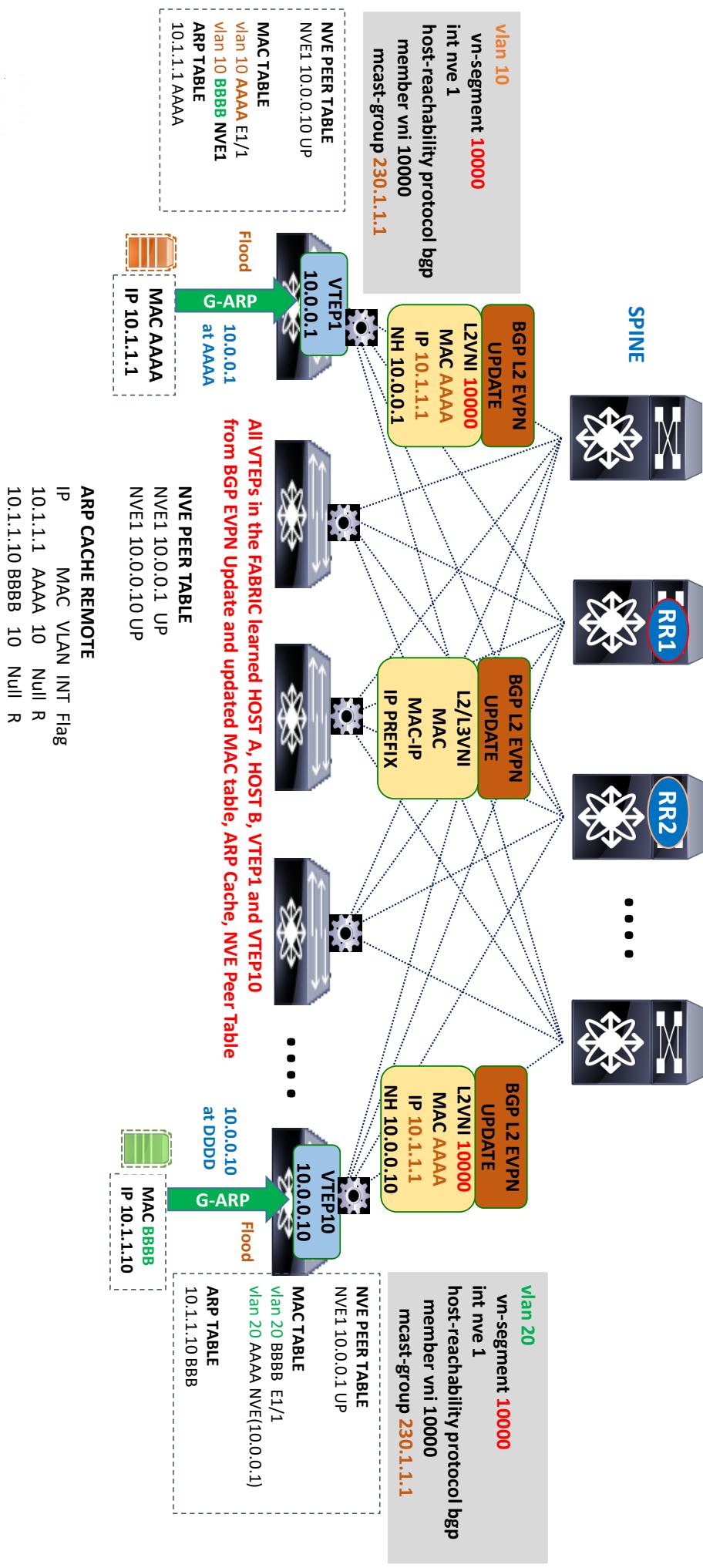


MP-BGP VPN: Route Target

- RT is Extended BGP Community
- VRF routes are exported from BGP VRF into VPN Address Family or imported from VPN Address Family into BGP VRF
- RT is attached to routes when exported by “route-target export x:y”
 - Multiple RT can be attached to a route
- Import routes based “route-target import x:y”
 - Leak routes from VRF BLUE to RED
 - “route-target export 3:333” on vrf RED
 - “route-target import 3:333 on vrf BLUE

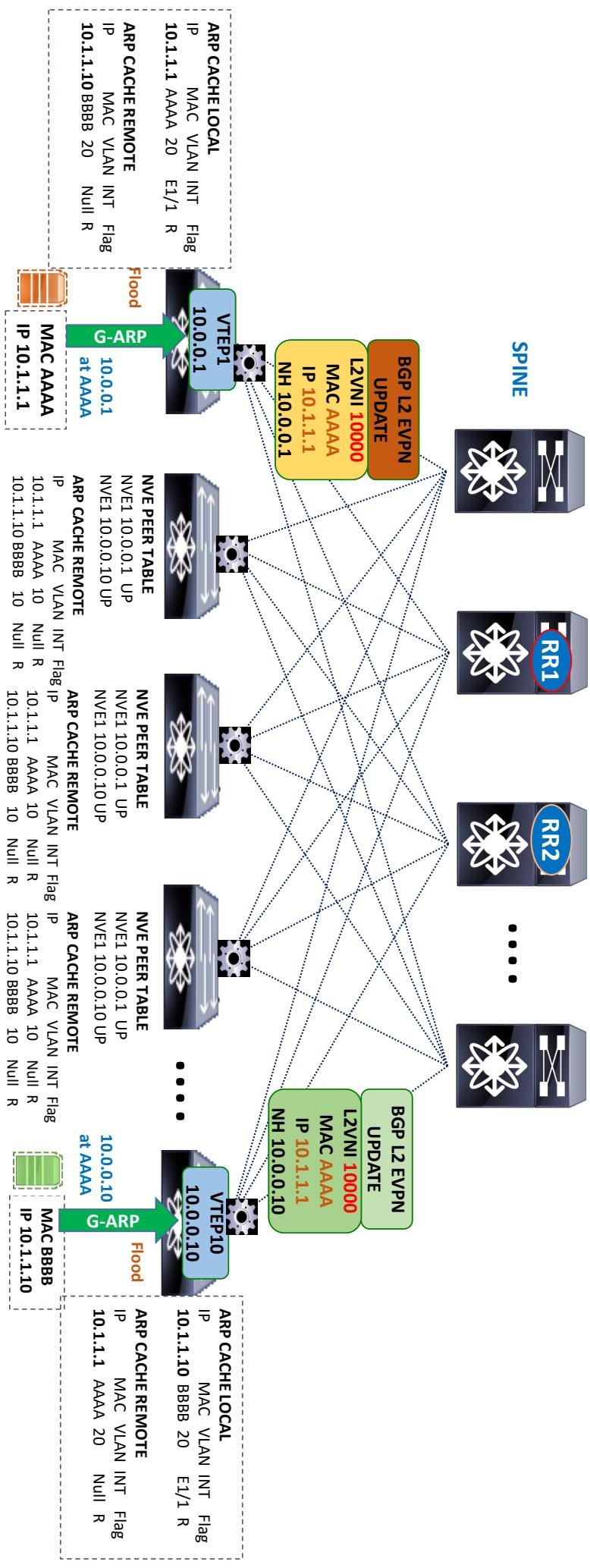


VXLAN BGP EVPN: Learning

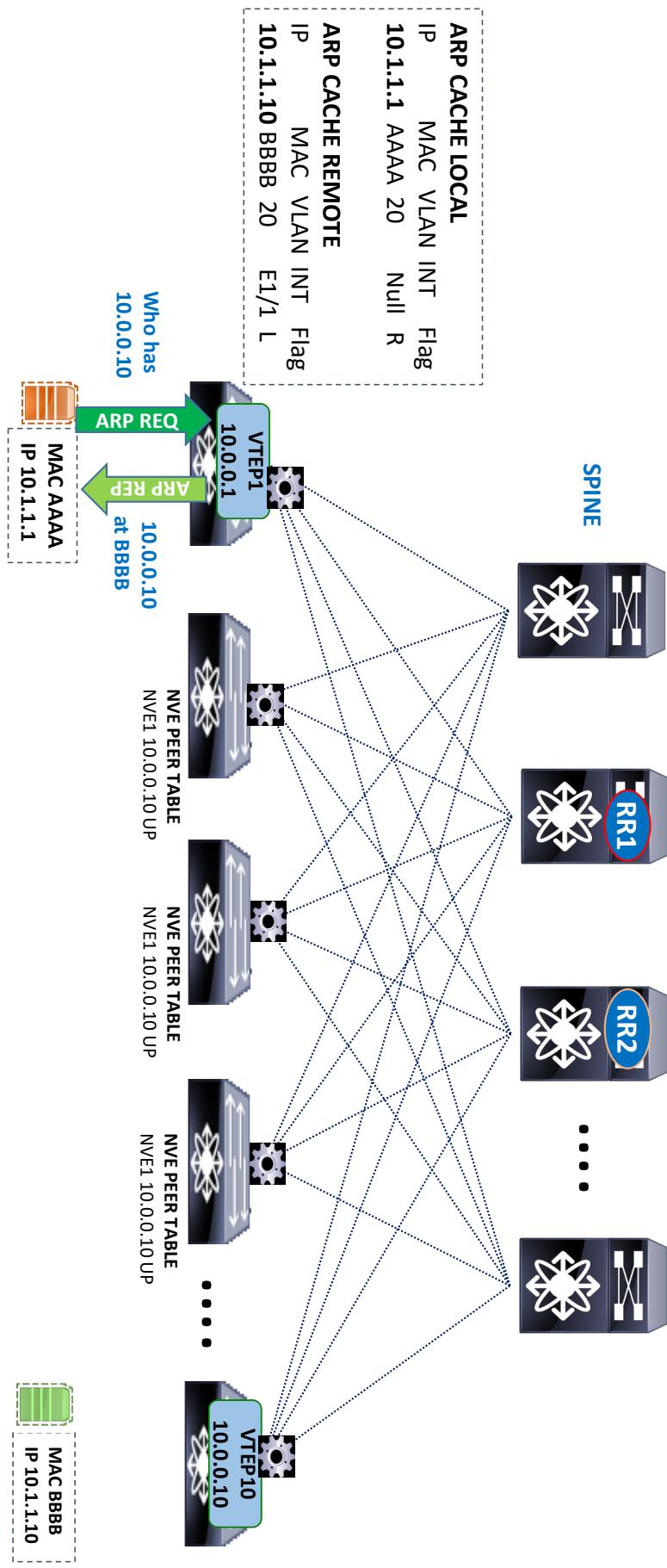


ARP Suppression in MP-BGP EVPN

ARP suppression reduces network flooding due to host learning from G-ARP



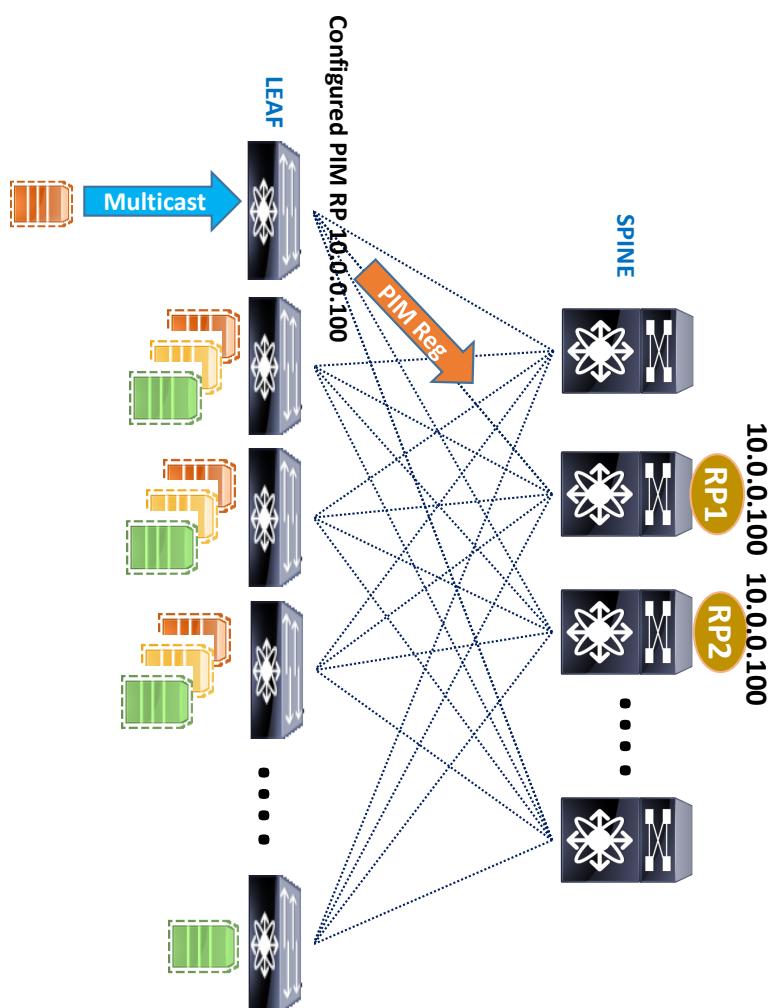
ARP Suppression in MP-BGP EVPN



If there is no local ARP CACHE, then FLOOD using multicast group or Ingress Replication

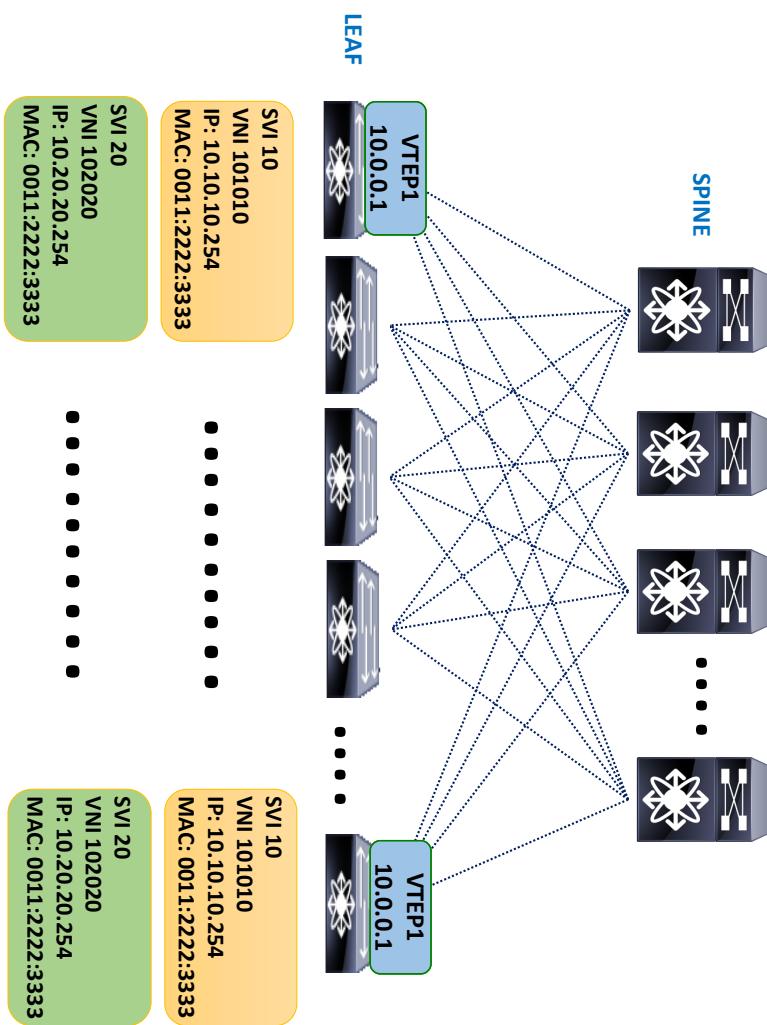
ARP suppression requires “hardware access-list tcam region arp-ether 256” and Reload

Anycast – One-to-Nearset Association



- a **network addressing and routing methodology**
- group of potential receivers, all identified by the **same destination address**
- datagrams sent from a single sender to the topologically **nearest node**
- **Example**
 - PIM ANYCAST RP
 - RP1 and RP2 has same IP Address
 - PIM Router send PIM registration to either RP1 or RP2 based on routing metric

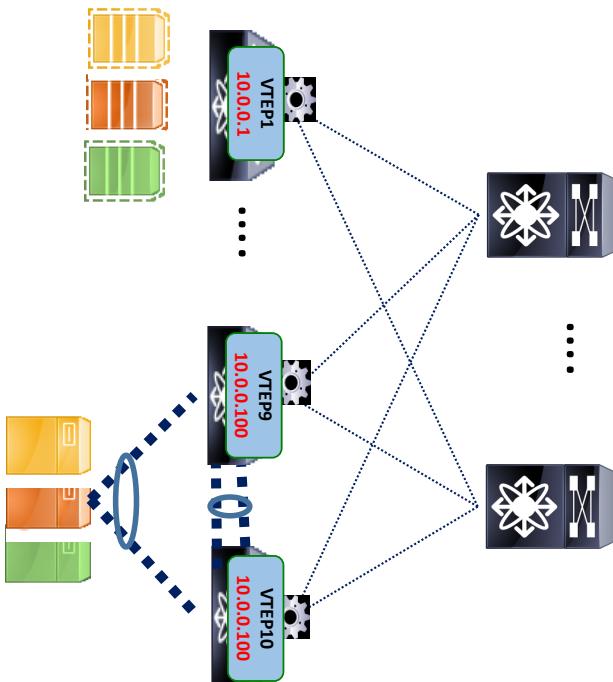
EVPN: Distributed IP Anycast Gateway



- All Leafs have same gateway IP and MAC for a Subnet
 - No need of FHRP any more
 - ARPs are terminated on Leafs
 - No Flooding beyond Leaf
 - Seamless Mobility
- Seamless Layer-2 or Layer-3 communication
 - Integrated Route & Bridge (IRB)
 - Route whenever you can, Bridge when needed
 - Optimal Scalability – Route Distributed & closest to the Host
- Reduced Failure Domain
 - Layer-2/Layer-3 Boundary at Leaf
- Consistent Configuration
 - Logical Configuration (VLAN, VRF, VNI) consistently instantiated on ALL Leafs
- Easy for Automation/Maintenance

VXLAN EVPN – vPC VTEP

- A pair of vPC VTEPS appears as one VTEP to other peer VTEP
- Use secondary IP of NVE Source Loopback as BGP update's Next-Hop
- vPC VTEPS establishes BGP session from BGP update source Loopback
- Bind NVE to the loopback interface that is not used for underlay routing protocols
 - VPC put NVE Source Loopback on VPC Secondary SW into Admin Shut when peer-link is shut
 - This will cause underlay routing issue if NVE uses underlay routing protocol loopback as source



```
VTEP9
interface loopback0
ip address 10.0.0.9/32
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

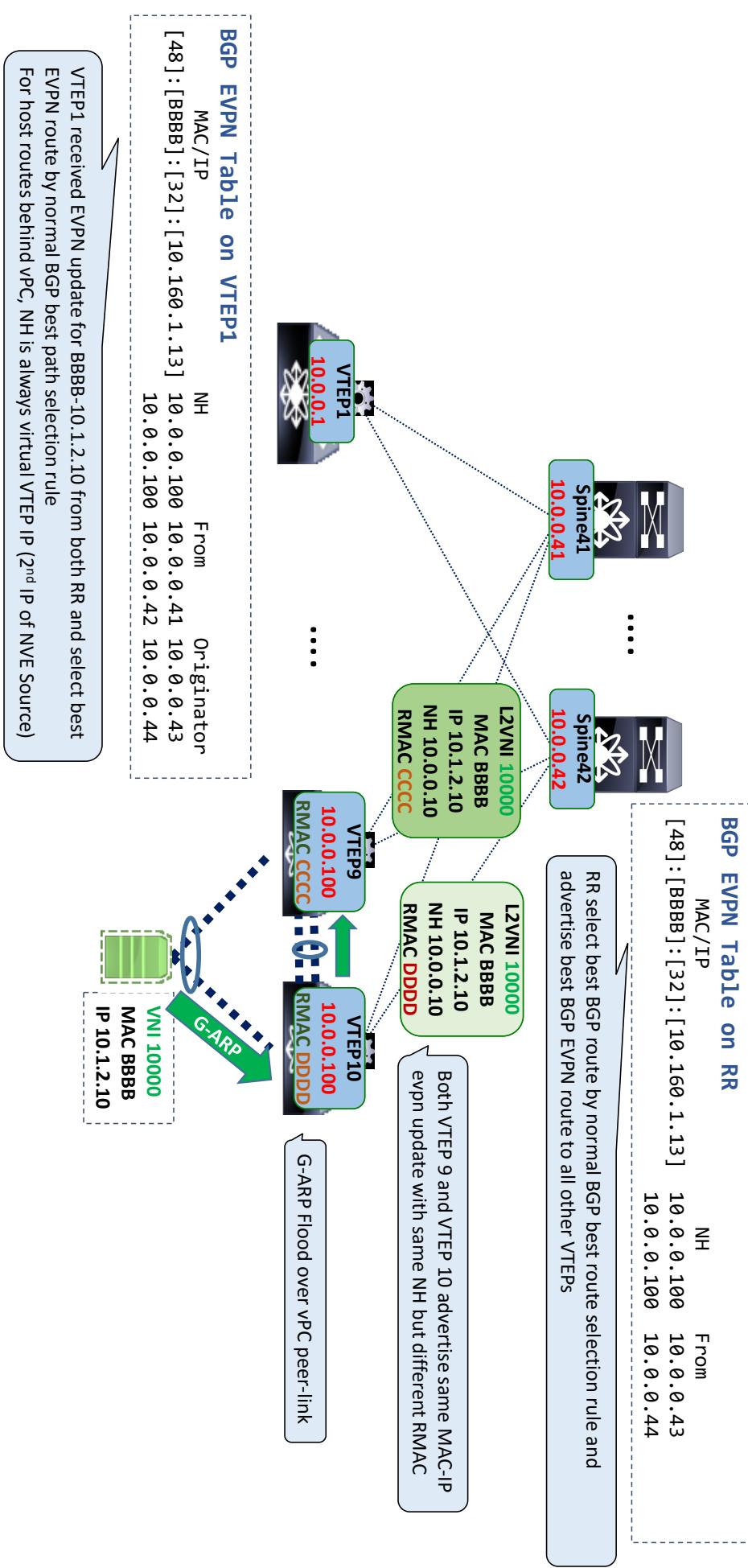
interface loopback1
ip address 10.0.1.9/32
ip address 10.0.0.100/32 secondary
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

VTEP10
interface loopback0
ip address 10.0.0.10/32
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

interface loopback1
ip address 10.0.1.9/32
ip address 10.0.0.100/32 secondary
ip router ospf 1 area 0.0.0.41
ip pim sparse-mode

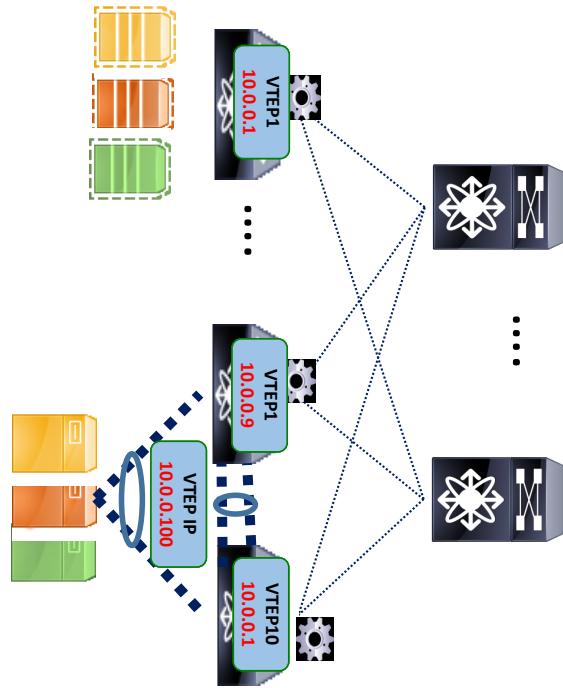
interface nve 1
source-interface loopback1
```

VXLAN EVPN – vPC VTEP



VXLAN EVPN – vPC VTEP Recommendation

- VPC Peer must have
 - Consistent VLAN to VN-seg mapping
 - Consistent NVE 1 binding to same loopback interface
 - Same secondary IP
 - Different Primary IP
 - Consistent VNI to multicast group mapping
- VPC Recommendation
 - peer-gw, peer-sw, ip arp sync
 - Increase STP hello timer to 4 sec
 - Underlay Backup SVI with PIM enabled
 - Overlay backup SVI with PIM enabled if DHCP relay is required

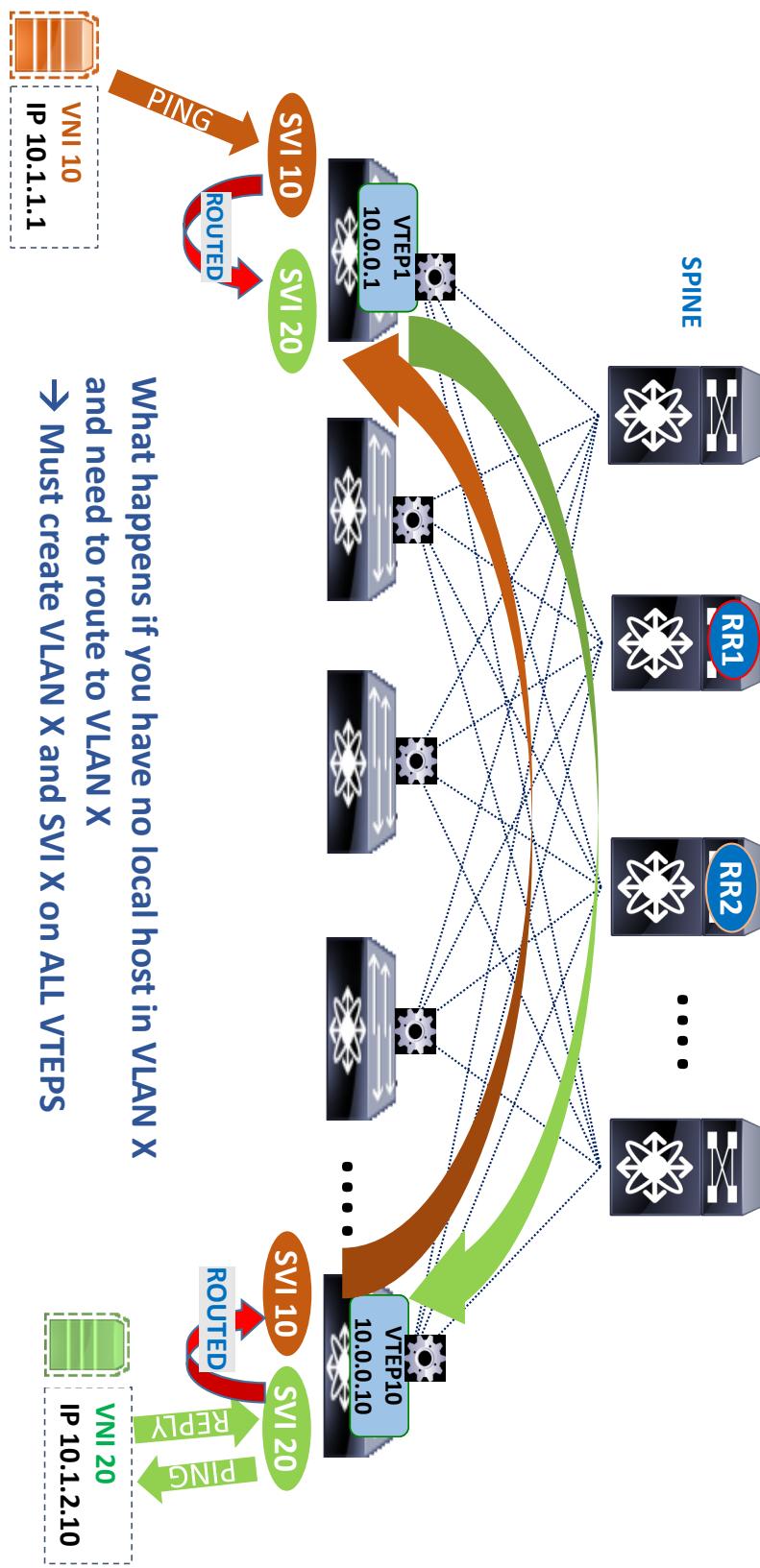


EVPN: Integrated Routing and Bridging (IRB)

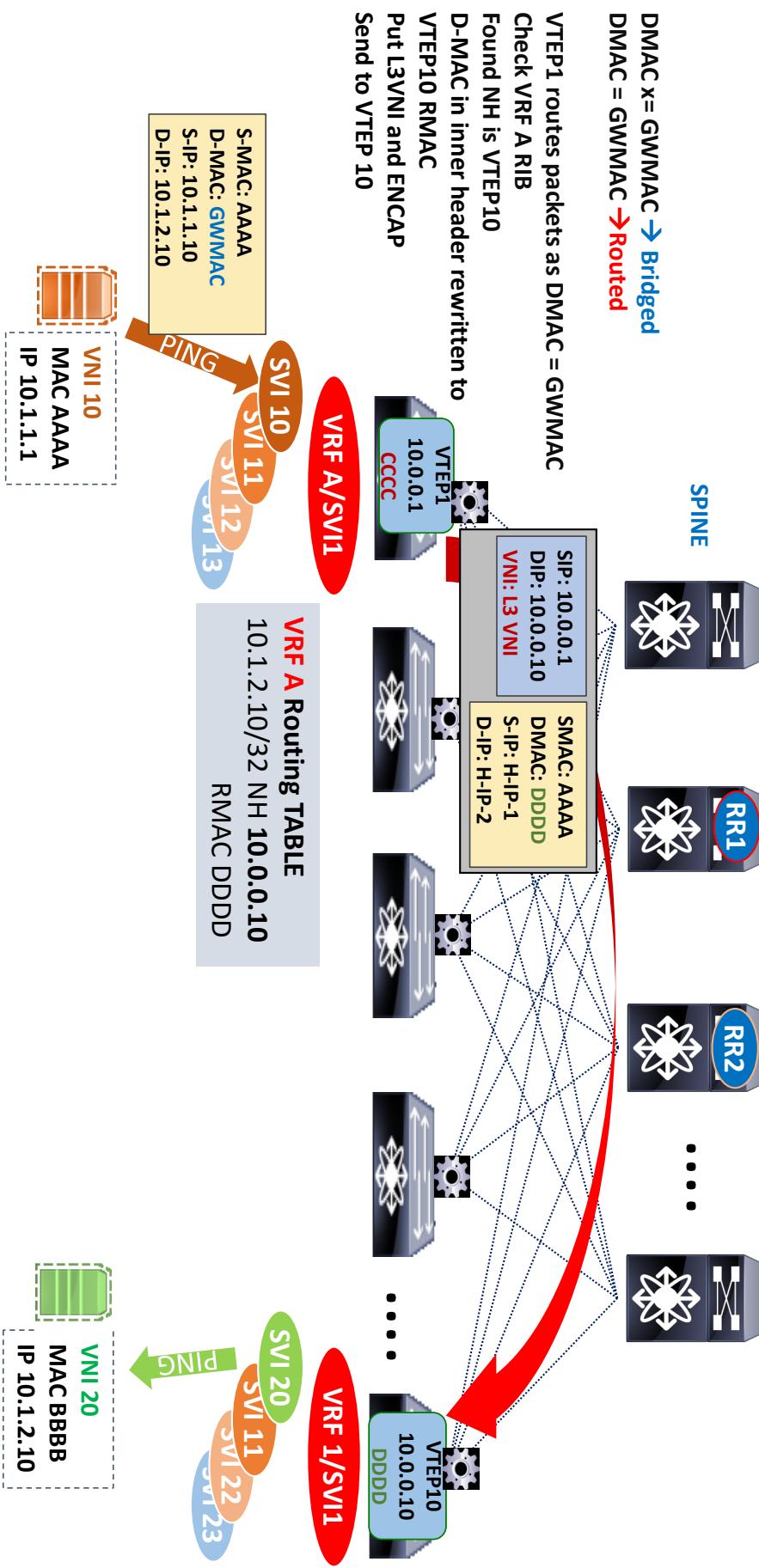
VXLAN/EVPN based overlays follow two slightly different Integrated Routing and Bridging (IRB) semantics

- Asymmetric
 - Uses an “asymmetric path” from the Host towards the egressing port of the VTEP vs. the way back
- Symmetric*
 - Uses an “symmetric path” from the Host towards the egressing port of the VTEP vs. the way back

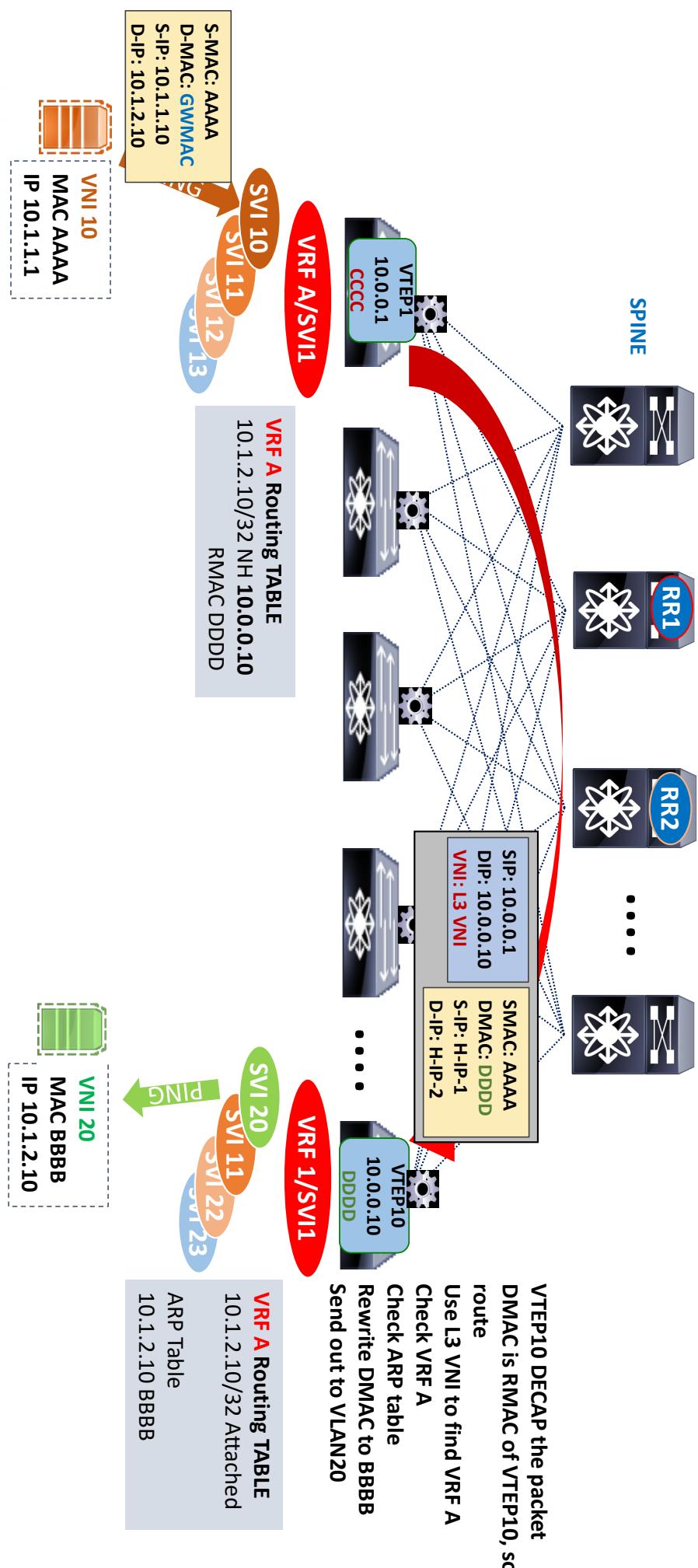
EVPN IRB: Asymmetric



EVPN IRB: Symmetric

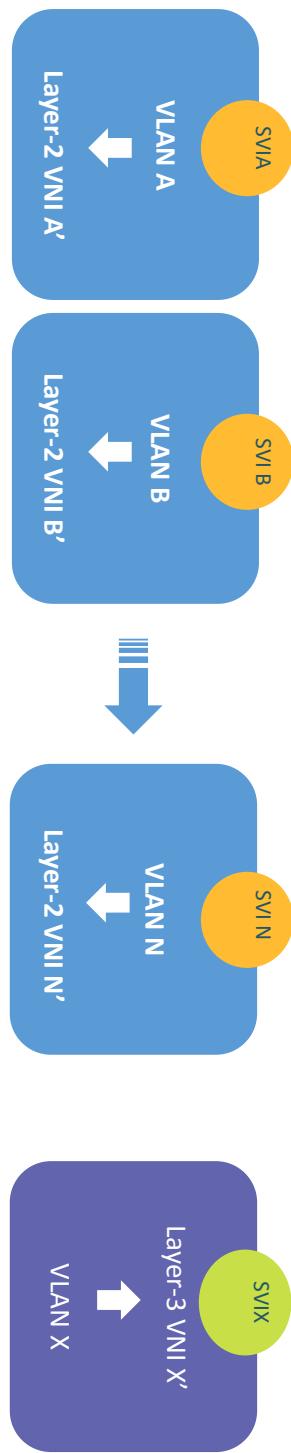


EVPN IRB: Symmetric



Logical Construct of Multi-Tenant VXLAN EVPN

Tenant X (VRF X)



- One VLAN maps to one Layer-2 VNI Layer-2 VNI per Layer-2 segment
- A Tenant can have multiple VLANs, therefore multiple Layer-2 VNIs
- Traffic within one Layer-2 VNI is bridged
 - Traffic between Layer-2 VNIs is routed
- 1 Layer-3 VNI per Tenant (VRF) for routing
 - VNI X' is used for routed packets

VXLAN EVPN Configuration

```
!Enable Features
feature nv overlay
feature vn-segment-vlan-based
feature bgp
nv overlay evpn
```

```
!Set distributed anycast GW MAC for all LEAFs
fabric forwarding anycast-gateway-mac 1600.1600.1600
```

```
!Create vlan and map vni for L3 VNI
vlan 1600
  name 13-vni-vlan-for-tenant-kk1
vn-segment 16000
```

```
!Create L2 VLAN and map VNI for L2 VNI
vlan 1601
  name 12-vni-BD1601
vn-segment 16010
```

```
vlan 1602
  name 12-vni-BD1601
vn-segment 16020
```

VXLAN EVPN Configuration

!Create L3 VRF and link to L2 VNI

vrf context evpn-tenant-kk1

vni 16000

rd 16000:1 or rd auto

address-family ipv4 unicast

route-target both 16000:1

route-target both 16000:1 evpn

Connect VNI to VRF

Import/Export for EVPN, too

!Create L3 VNI SVI and put it into L3 VRF

interface Vlan1600

description l3-vni-svi-for-tenant-kk1-routing

no shutdown

vrf member evpn-tenant-kk1

ip forward

L3 VLAN VNI in VRF

No need o IP address

VXLAN EVPN – L2 VNI Config

!Create Routing SVI for L2 VNI

```
interface Vlan1601
```

```
no shutdown
```

```
vrf member evpn-tenant-kk1
```

```
ip address 10.160.1.254/24
```

```
fabric forwarding mode anycast-gateway
```

interface **Vlan1602**

```
no shutdown
```

```
vrf member evpn-tenant-kk1
```

```
ip address 10.160.2.254/24
```

```
fabric forwarding mode anycast-gateway
```

This is how we connect L3 VRF to L2 VNI by putting L2 VNI SVI into L3 VRF

!Create NVE interface and map VNI

```
interface nve1
```

```
no shutdown
```

```
source-interface loopback0
```

```
host-reachability protocol bgp
```

```
member vni 16000 associate-vrf
```

```
member vni 16010
```

```
suppress-arp
```

```
mcast-group 230.0.0.160.1
```

```
member vni 16020
```

```
suppress-arp
```

```
mcast-group 230.0.0.160.2
```

L3 VNI is only for VRF

EVPN still can use Multicast CORE for BUM traffic

VXLAN EVPN Configuration

!Add VRF to BGP EVPN

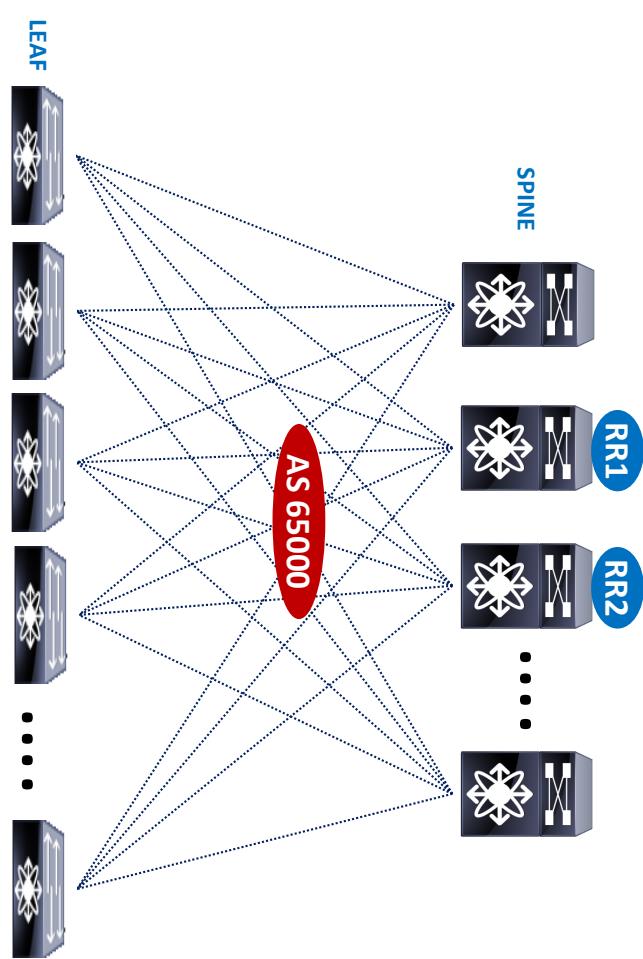
```
router bgp 65000
  vrf evpn-tenant-kk1
    address-family ipv4 unicast
      advertise l2vpn evpn
    address-family ipv6 unicast
      advertise l2vpn evpn
```

!Add L2 VNI to overlay with BGP control plane, enable ARP suppression

```
evpn
  vni 16010 12
    rd 16010:1 or rd auto
    route-target import 16010:1
    route-target export 16010:1

  vni 16020 12
    rd 16020:1
    route-target import 16020:1
    route-target export 16020:1
```

VXLAN EVPN – iBGP



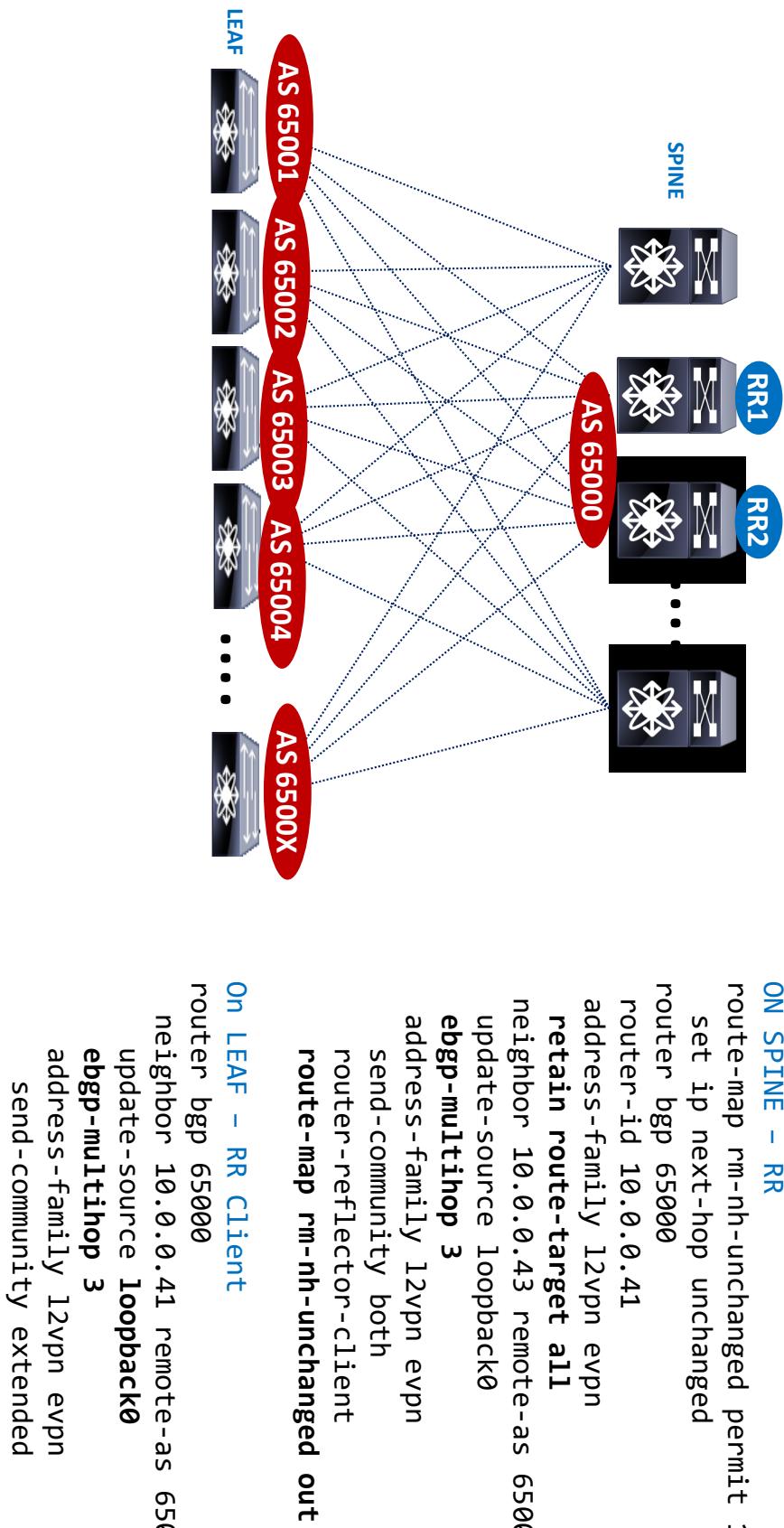
ON SPINE - RR

```
router bgp 65000
  route-id 10.0.0.41
  neighbor 10.0.0.43 remote-as 65000
  update-source loopback0
  address-family l2vpn evpn
  send-community both
  router-reflector-client
```

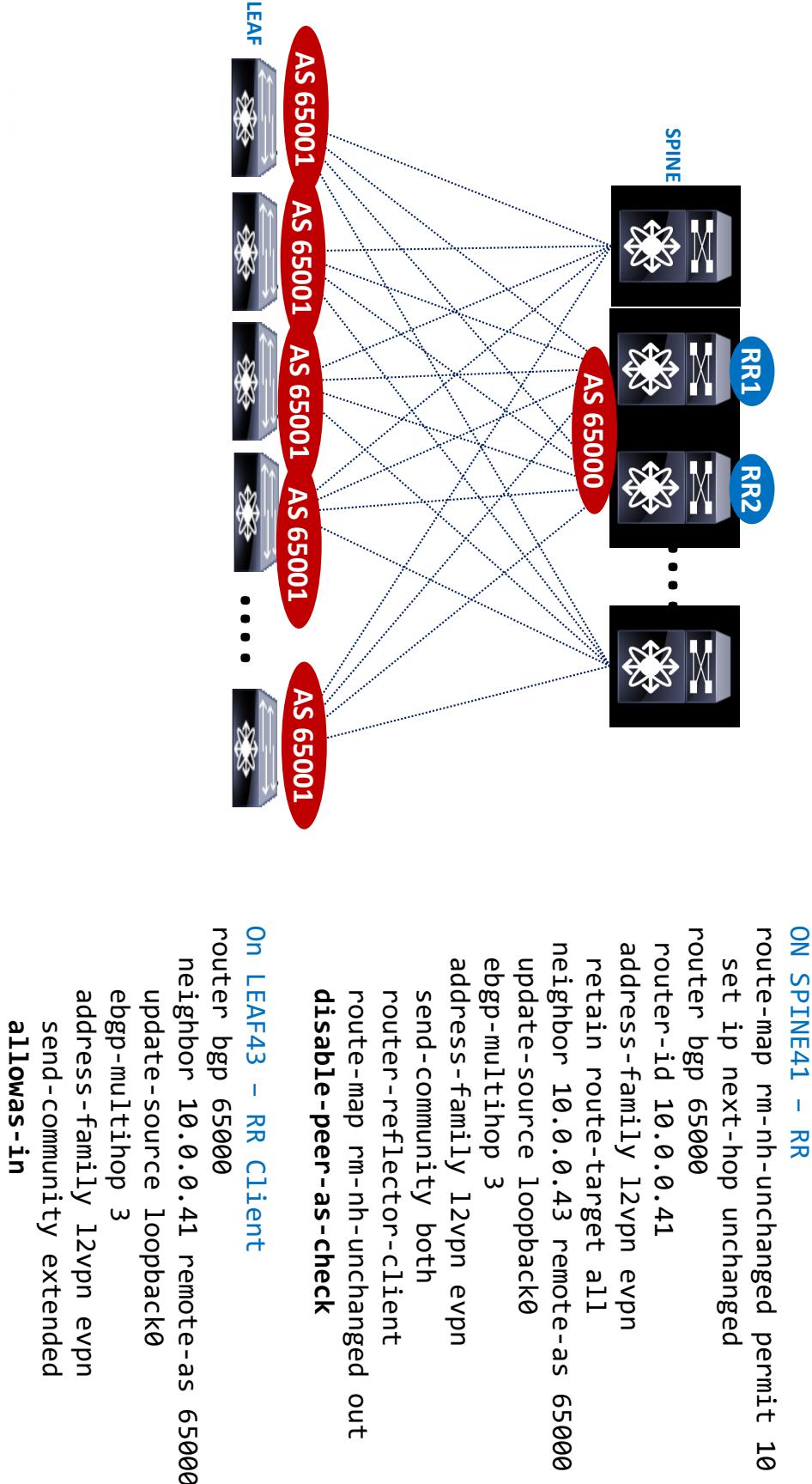
On LEAF - RR Client

```
router bgp 65000
  route-id 10.0.0.43
  neighbor 10.0.0.41 remote-as 65000
  update-source loopback0
  address-family l2vpn evpn
  send-community both
```

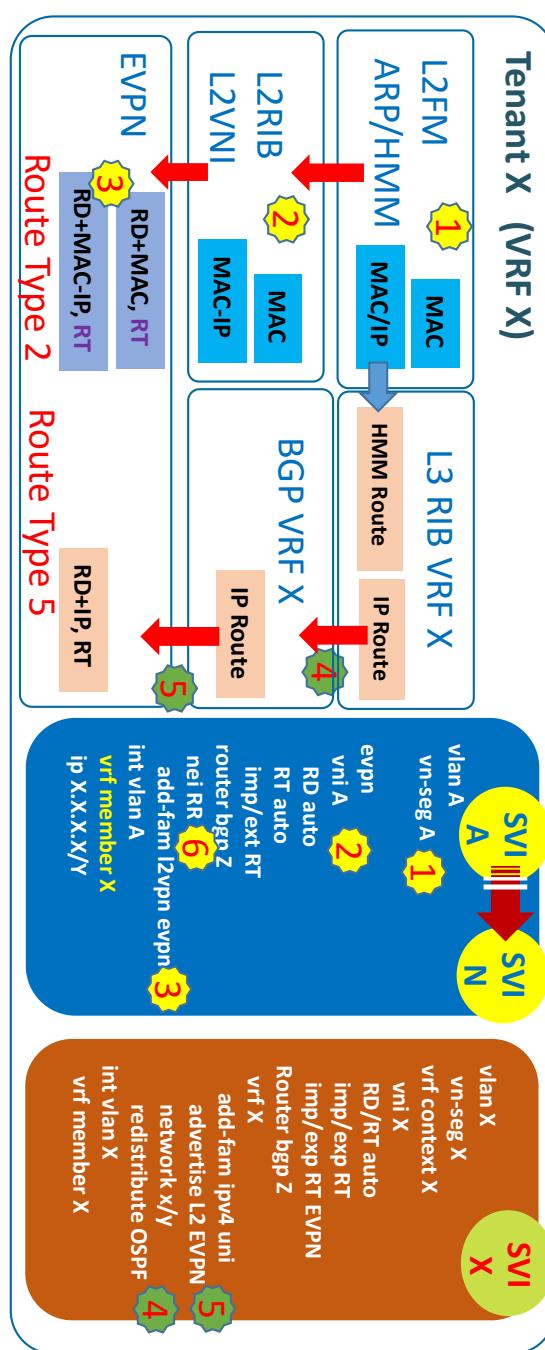
VXLAN EVPN – eBGP with Unique AS on LEAF



VXLAN EVPN – eBGP with Same AS on LEAF



How L2/L3 VNI works with BGP L2VPN EVPN



1 LEARN MAC from L2FM and MAC-IP from ARP

```
N9372TX-43# sh mac add vlan 1601
+ 1601 001a.6c70.3697 dyn 0 F F Po1743
N9372TX-43# sh ip arp vrf evpn-tenant-kk1
Address Age MAC Address Interface
10.160.1.13 00:00:20 001a.6c70.3697 Vlan1601
```

2 Install MAC and MAC-IP into L2RIB per EVPN ID

```
N9372TX-43# sh l2route evpn mac evi 1601
Mac Address Prod Next Hop (s)
001a.6c70.3697 Local Po1743
Mac Address Prod Host IP Next Hop (s)
001a.6c70.3697 HMM 10.160.1.13 N/A
```

3 L2RIB exports to L2VPN EVPN

```
N9372TX-43# sh bgp 12vpn evpn vni-id 16010
Route Distinguisher: 16010:1 (L2VNI 16010)
Network
*>[1][2]:[0]:[0]:[48]:[001a.6c70.3697]:[0]:[0.0.0.0]/216
*->[1][2]:[0]:[0]:[48]:[001a.6c70.3697]:[32]:[10.160.1.13]/272
10.0.0.100 100 32768 i
```

4 BGP VRF's network/redistribute creates IPv4 BGP routes

```
interface Loopback601
vrf member evpn-tenant-kk1
ip address 160.1.0.43/32
router bgp 2
vrf evpn-tenant-kk1
address-family ipv4 unicast
network 160.1.0.43/32
```

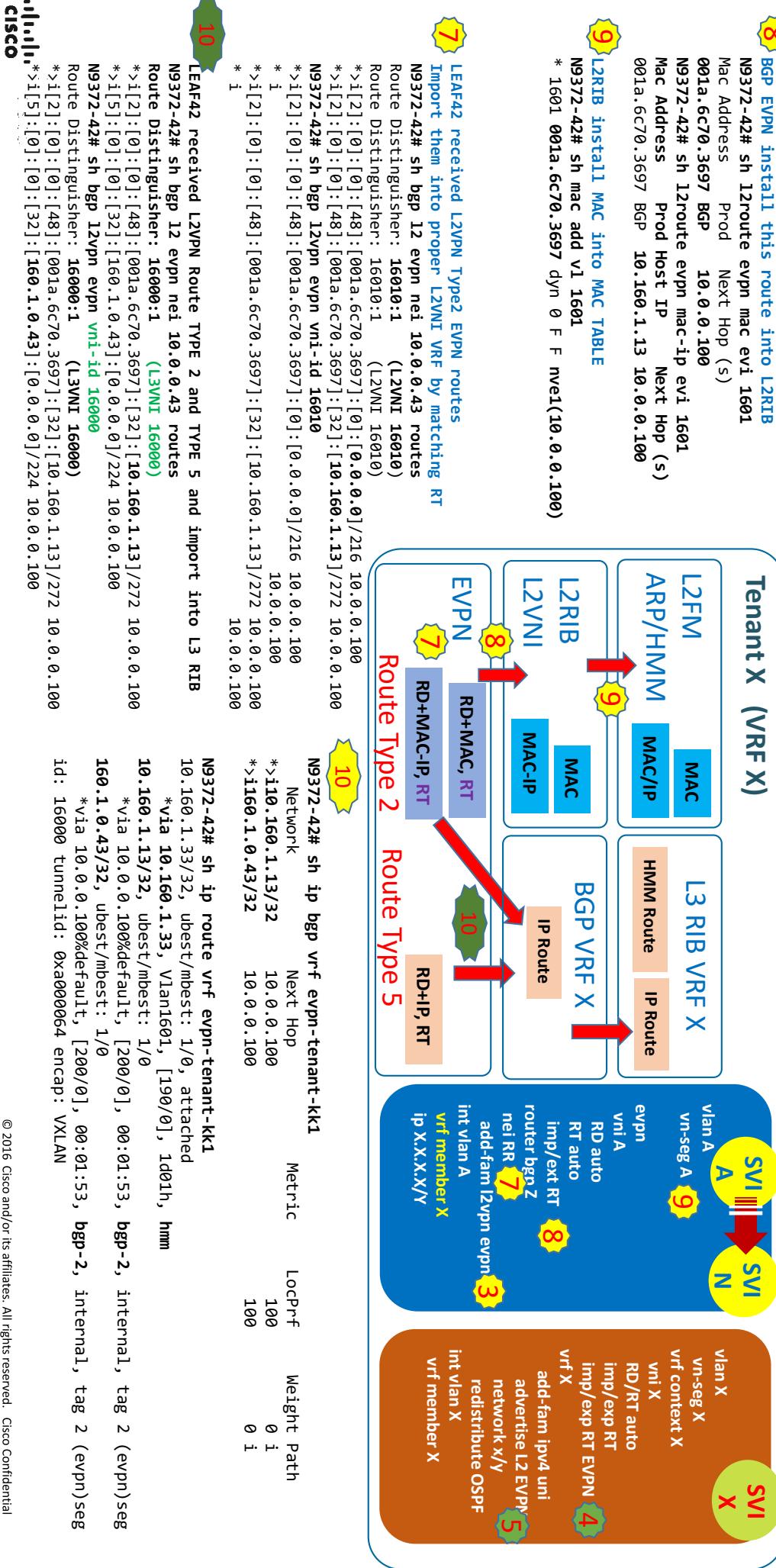
5 Route Type 5 is created in L2 EVPN L3 VNI from L3 VRF network statement

```
N9372TX-43# sh bgp 12 evpn vni-id 16000
*>1[5]:[0]:[0]:[32]:[160.1.0.43]:[0.0.0.0]/224 10.0.0.100
Route Distinguisher: 16010:1 (L2VNI 16010)
int vlan A
vrf member X
ip X.X.X.X/Y
```

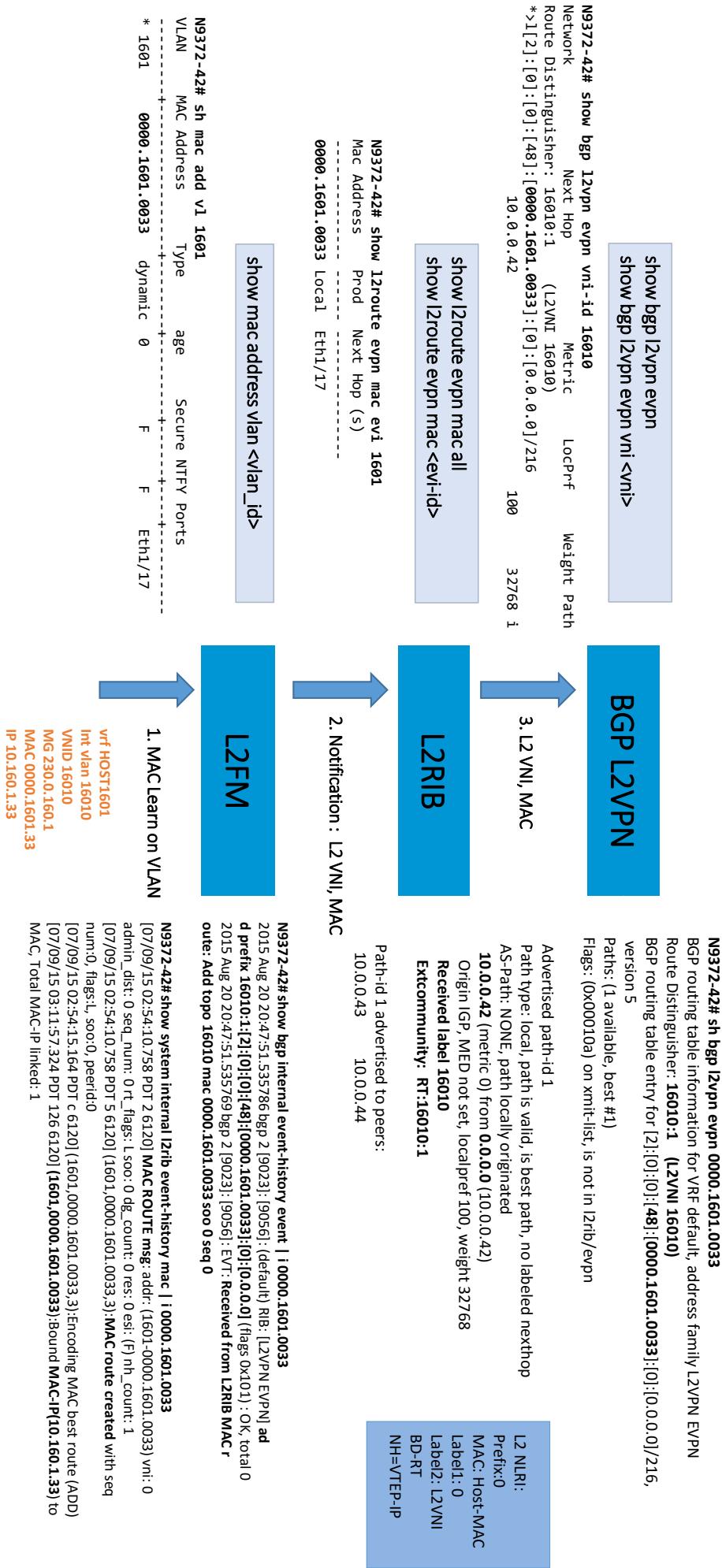
6 Leaf43 advertise L2/L3 VNI routes to LEAF42

```
N9372TX-43# sh bgp 12 evpn nei 10.0.0.42 advertised-routes
Route Distinguisher: 16010:1 (L2VNI 16010)
*>1[2]:[0]:[48]:[001a.6c70.3697]:[0]:[0.0.0.0]/216 10.0.0.100
*>1[2]:[0]:[48]:[001a.6c70.3697]:[32]:[10.160.1.13]/272 10.0.0.100
```

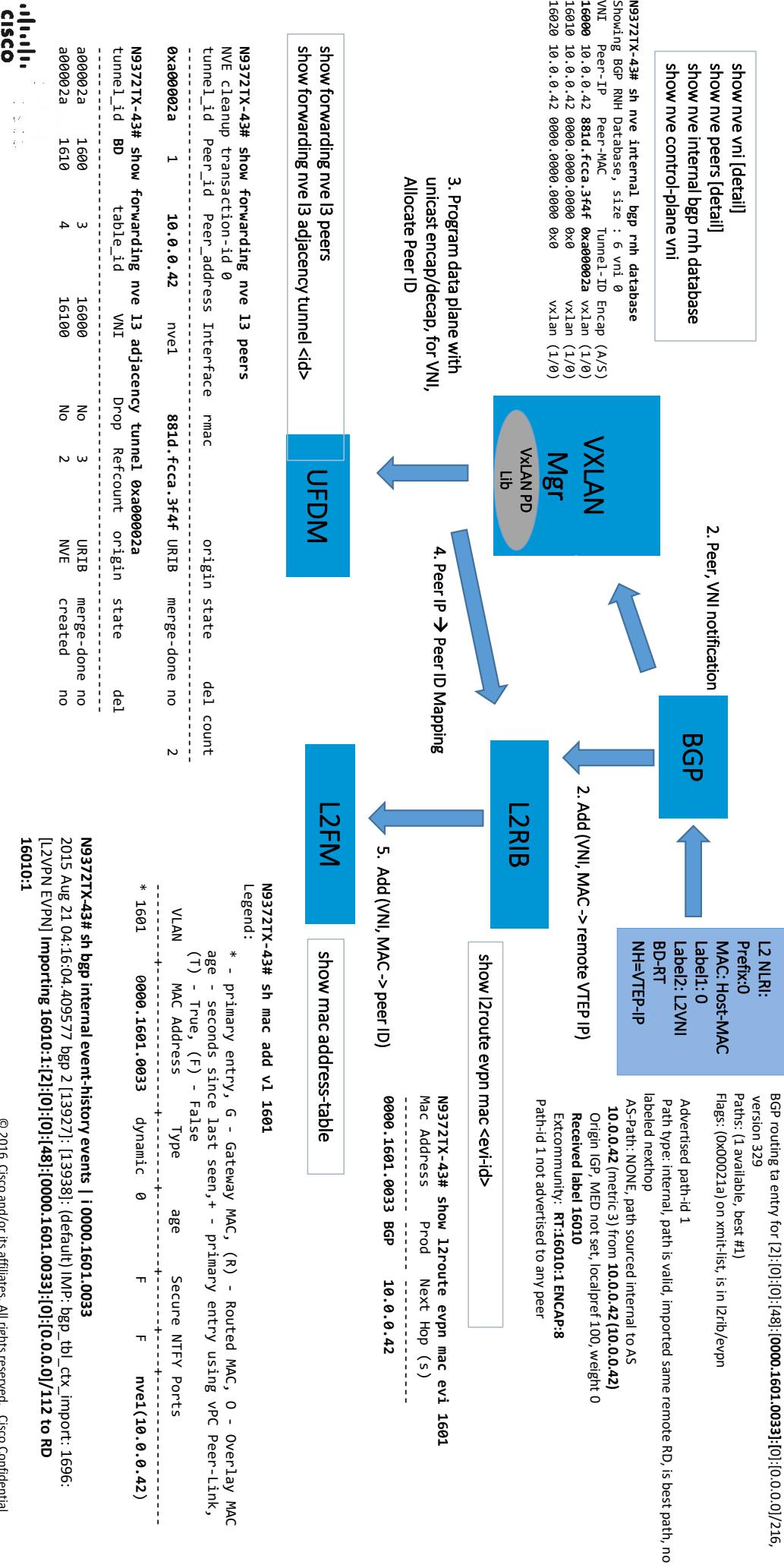
How L2/L3 VNI works with BGP L2VPN EVPN



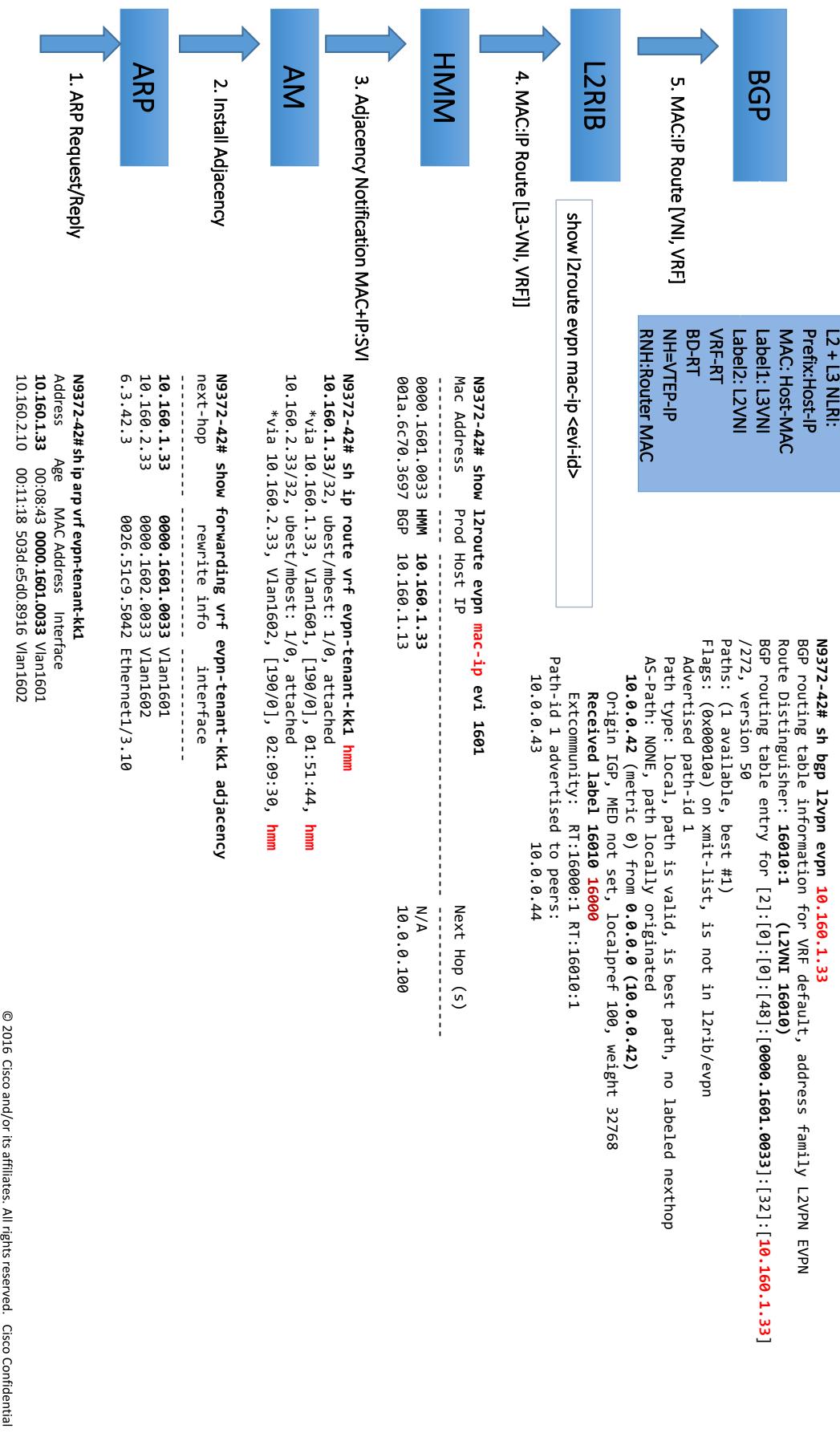
Local MAC routes advertisement via BGP EVPN



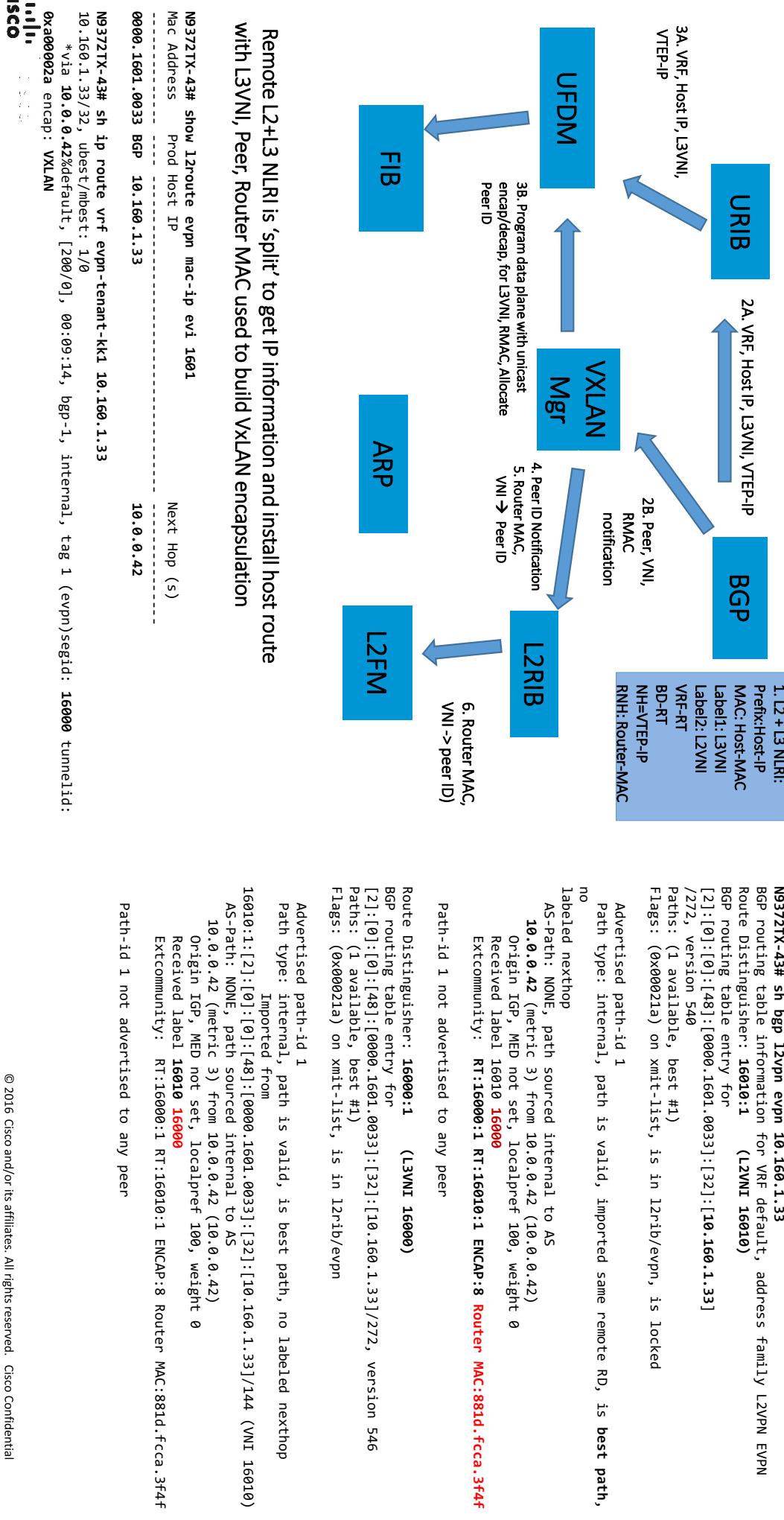
Remote L2 MAC Route Installation via BGP EVPN



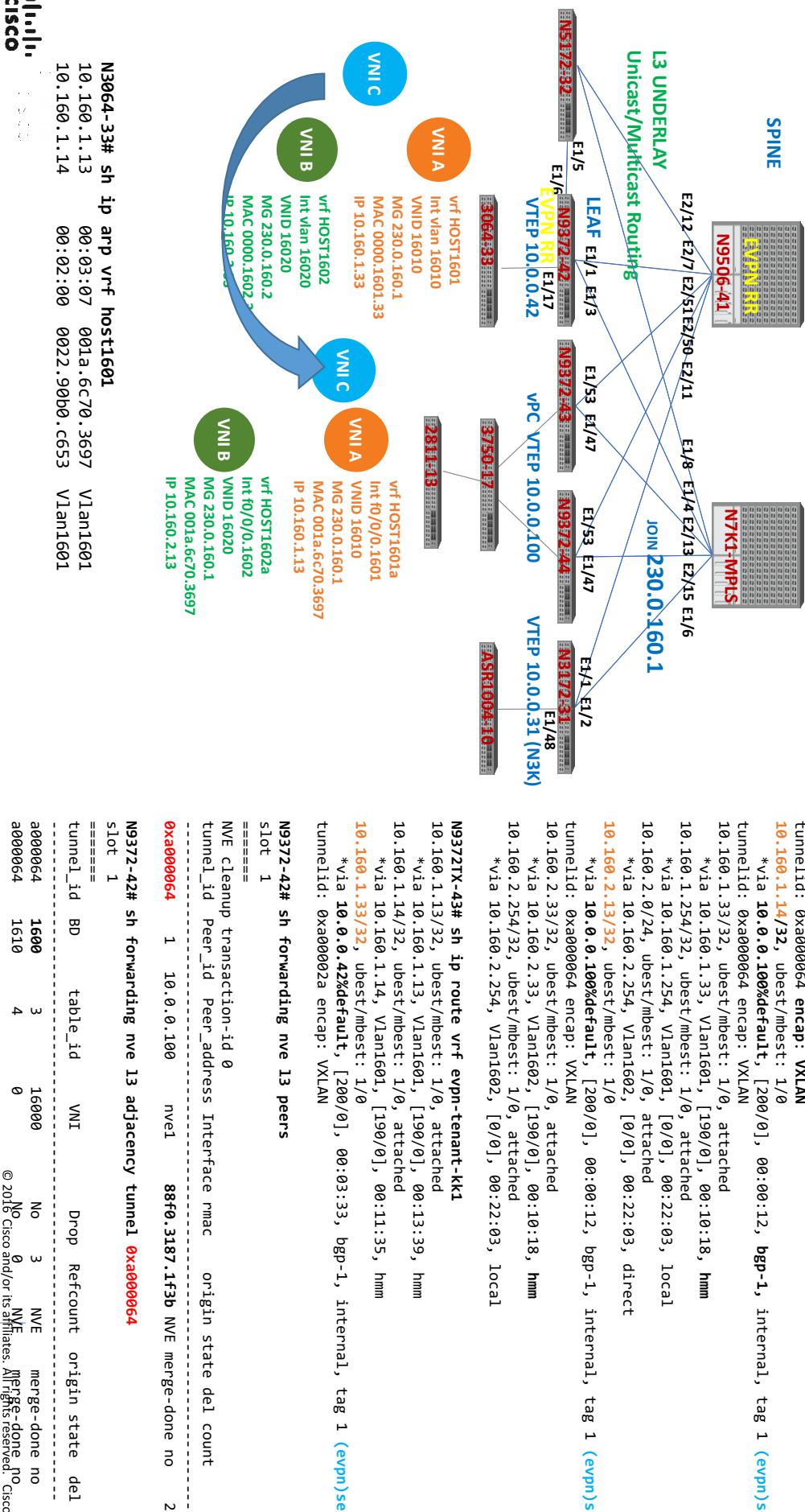
Local IP, MAC Route Advertisement via BGP EVPN



Remote IP Route Installation via BGP EVPN



VXLAN EVPN – BGP



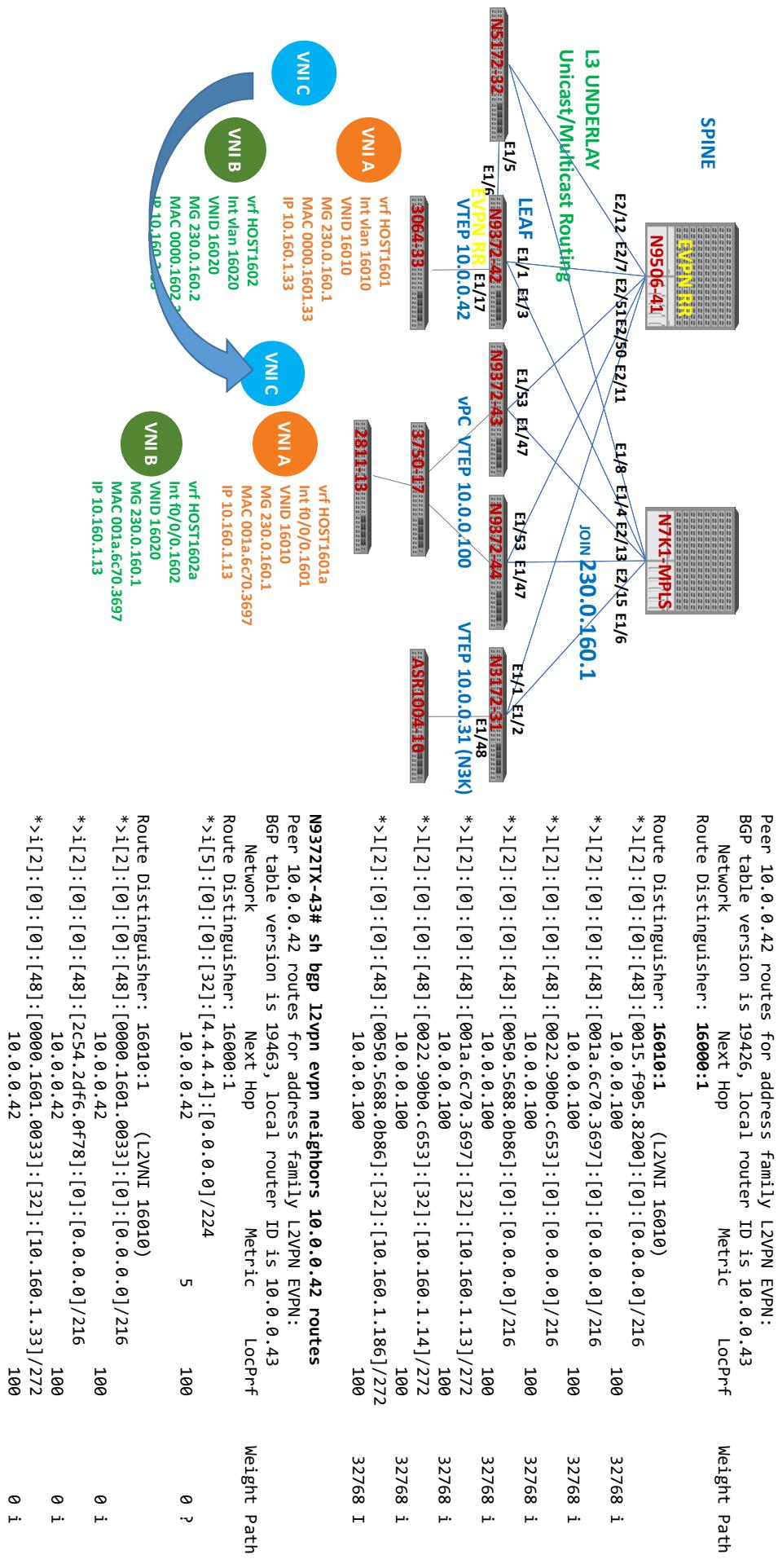
© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

•••••
cisco

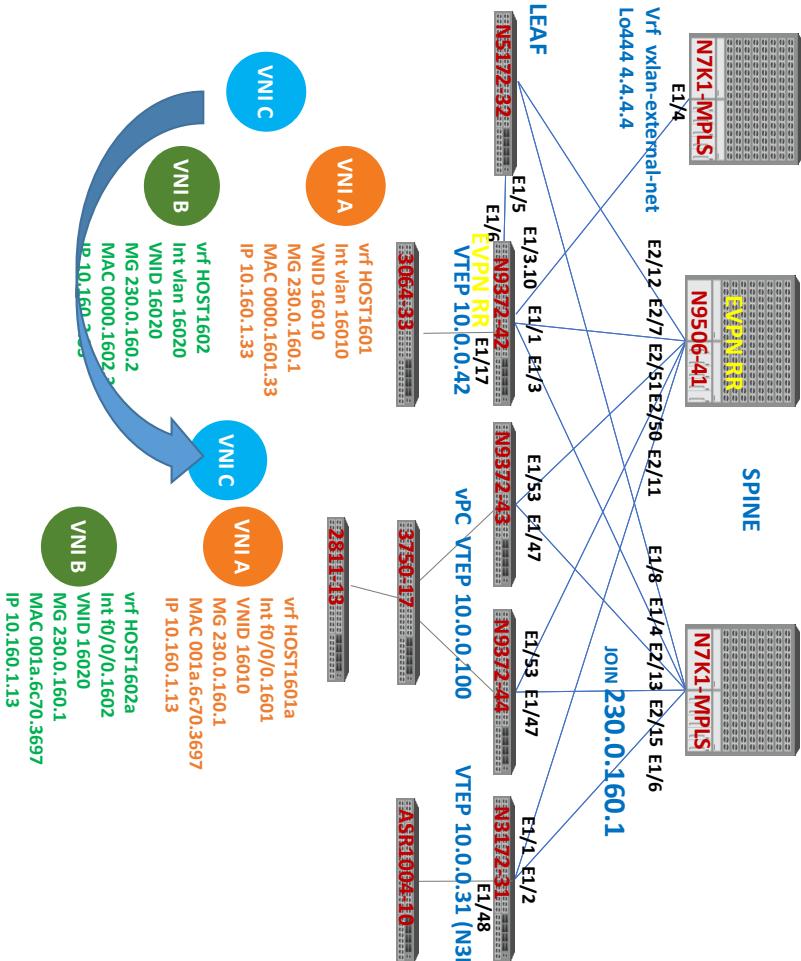
EVPN Summary

- MP-BGP IPv4 Unicast AF and L2VPN EVPN SAFI peering between LEAFs and/or SPINES
- Utilize RR to reduce iBGP Mesh
- Utilize VRF-LITE BGP for L3 VRF
- L2 VNI(Address-family l2vpn evpn) carries MAC and MAC-IP, EVPN Route Type 2
 - router bgp 2
 - router-id 10.0.0.43
 - address-family l2vpn evpn
 - neighbor 10.0.0.41
 - address-family l2vpn evpn
 - send-community extended
- MAC Learning on Leaf creates MAC entries
- ARP creates ADJ and MAC-IP entries
- L2 VNI put EVPN learned /32 into L3 VNI
- L3 VNI only carries EVPN Type5 Routes
 - vrf evpn-tenant-kk2
 - address-family ipv4 unicast
 - advertise l2vpn evpn
 - network X/Y
 - redistribute ospf 2 route-map rm-permit-all
- L3 VRF RIB receives routes from HMM, L2VNI, and L3VNI
- If Dest MAC = X RMAC → BRIDGE, look up MAC TABLE
- If Dest MAC = RMAC → ROUTE, look up L3 VNI Route TABLE

VXLAN EVPN – L2 VNI



VXLAN EVPN – External



Create **Border LEAF** that connects to external routing peers with iBGP or OSPF

Think it as just normal MPLS L3 VPN

Border Leaf already has L3 VNIs(VRFs)

Put external interface into proper VRF(L3 VNI) and configure VRF routing and redistribute OSPF2BGP and BGP2OSPF if you use OSPF as external routing protocol

For eBGP, just configure proper filtering

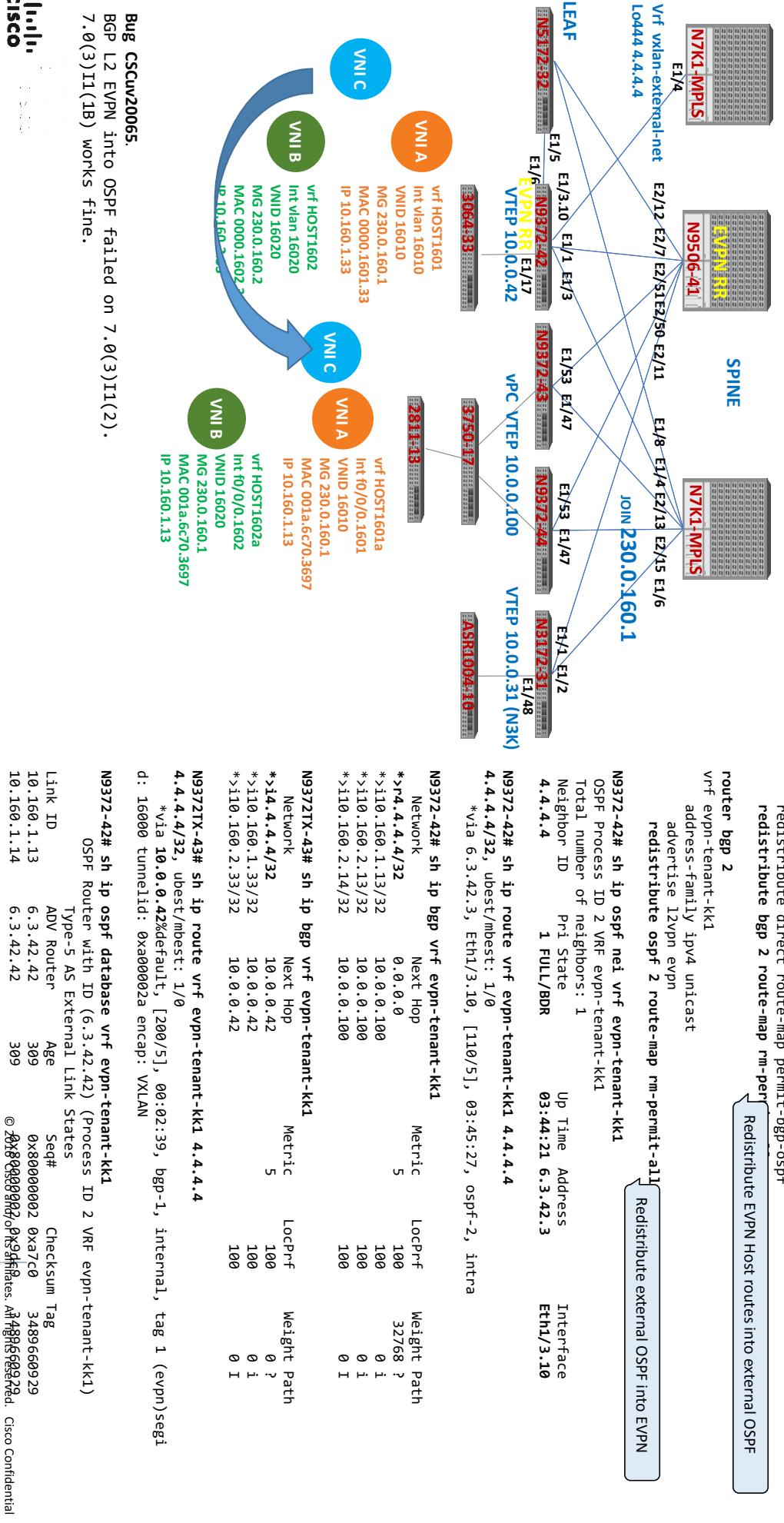
If you want hosts behind border leafs to reach external, configure redistribute direct under OSPF or BGP

By default, only EVPN hosts routes are redistributed

Summarize address on Border router if you don't want to advertise /32 to External peers

- Bug CSCuv20065. BGP L2 EVPN into OSPF failed on 7.0(3)I1(2). 7.0(3)I1(1B) works fine.
- Workaround is “match internal” in route-map

VXLAN EVPN – External



VXLAN EVPN – DHCP

Supported DHCP Relay in EVPN after 7.0(3)I1(1b)

- DHCP Server in default VRF; DHCP Client in Tenant VRF
- DHCP Server in non-default VRF; DHCP Client in Tenant VRF.
- DHCP Server and Client in same tenant VRF but different SVI (using Option-82 and use-vrf keyword)
- DHCP Relay and Client in Different Tenant VRF (Using Option-82 and use-vrf keyword)

http://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/7-x/vxlan/configuration/guide/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x/b_Cisco_Nexus_9000_Series_NX-OS_VXLAN_Configuration_Guide_7x_appendix_0110.html

Why is it so complex?

- DHCP server can't reply back to the original leaf that relayed DHCP DISC as all EVPN LEAFs have same IP for same SVI
 - Overcome this by using special loopback X and **DHCP option 82**
 - DHCP Relay on LEAF uses loopback X in vrf X as source IP for DHCP relay
 - DHCP Server must find proper DHCP pool by not using DHCP relay source IP(GI Addr), but using DHCP option 82, like Circuit ID/Server ID
 - Override/Link selection
- **vPC VTEPs add another complexity as both VTEPs uses virtual VTEP IP as NH**

HOST, Windows XP, is connected to vPC VTEP, Leaf43/44 in vlan 1601. SVI 1601 on LEAF43/44 is configured as DHCP Relay

CSR61 works as DHCP Server in vlan 1602 and connected to LEAF42

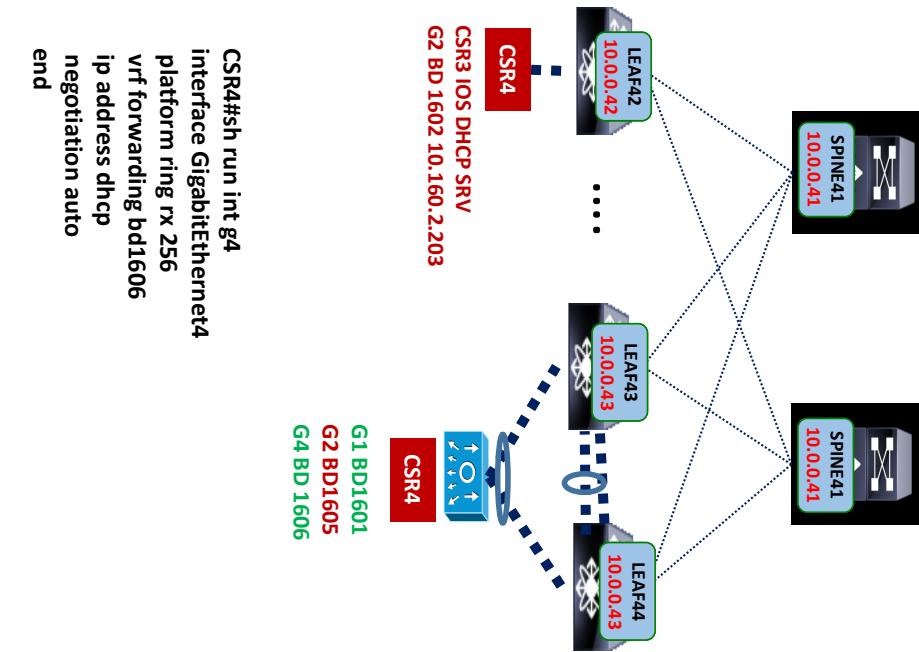
On IOS DHCP SERVER

```
ip dhcp pool dhcp-vxlan-1601
vrf evpn-tenant-k1
network 10.160.1.0 255.255.255.0
default-router 10.160.1.254
```

DHCP POOL's VRF NAME must match DHCP Client's VRF name
IOS DHCP server doesn't require to configure separate OPTION 82 match policy

VXLAN EVPN

- DHCP SRV and Client in same VRF



Detail Configuration STEP (Green color means only required for vpc vtep)

1. Enable DHCP Service and configure dhcp options

```
service dhcp
ip dhcp relay
ip dhcp relay information option
ip dhcp relay sub-option type cisco
ip dhcp relay information option vpn
```

Populating DHCP option 82

Put Cisco's Server ID Override and Link Selection Option

2. Create special loopback under VRF that you

vrf member evpn-tenant-kk1

ip address 160.1.0.43/32

ip router ospf 1 area 0

Include VRF info in option 82 field

3. Advertise Lox into L3 VRF BGP

```
Router bgp 2
vrf X
network 160.1.0.43/32
```

4. Configure DHCP Relay on SVI un

```
interface Vlan1606
no shutdown
vrf member evpn-tenant-kk1
ip address 10.160.6.254/24
fabric forwarding mode anycast-gateway
ip dhcp relay address 10.160.2.103
ip dhcp relay source-interface loopback601
```

Without IP Fwd, PKT that is encapsulated with VXLAN header and destined to router's own IP will be dropped

DHCP Relay uses Lox as source IP

5. Make sure L3 VNI SVI is configured with "IP Forwarder"

```
interface Vlan1600
vrf member evpn-tenant-kk1
ip forward
```

Without IP Fwd, PKT that is encapsulated with VXLAN header and destined to router's own IP will be dropped

6. Create Routing vlan/svi for vpc vrf

```
Vlan 99
interface Vlan99
vrf member evpn-tenant-kk1
ip address 10.0.99.43/24
ip router ospf 1 area 0.0.0.41
ip pim sp
```

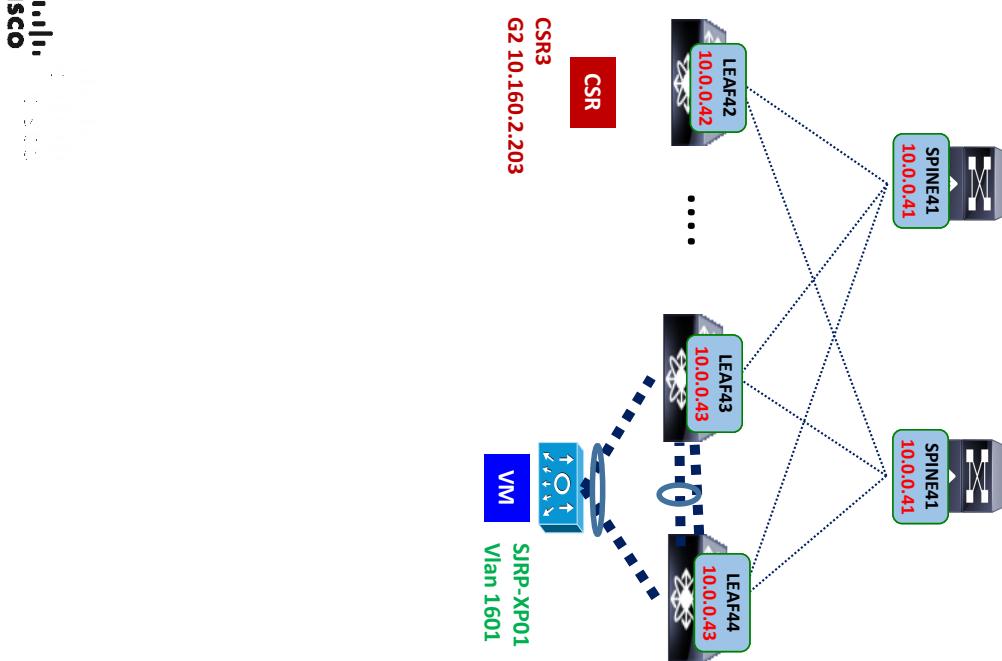
7. Create VRF Routing Process

```
router ospf 1
```

vr-f evpn-tenant-kk1

router-id 10.0.99.43

VXLAN EVPN – DHCP



```

2016 Mar 23 15:09:49.820358 dhcp_snoop: (PKT)Packet received on L3 interface
2016 Mar 23 15:09:49.820969 dhcp_snoop: (PKT)Initial if_index from netstack Vlan1601
2016 Mar 23 15:09:49.821017 dhcp_snoop: (PKT)RECEIVED packet type DHCP RELAY
2016 Mar 23 15:09:49.821201 dhcp_snoop: (PKT) DHCPDISCOVER on Intf Vlan1601(84), phy port-channel444(124), vlan
1601, pvlan 1601, vdc 0, vrf evpn-tenant-kk1
2016 Mar 23 15:09:49.821213 dhcp_snoop: (PKT)Packet: Len 342, l2_HDR 14, DHCP data 300, hdr flag 0x36
2016 Mar 23 15:09:49.821239 dhcp_snoop: (PKT)L2: DMAC ff:ff:ff:ff:ff:ff, SMAC 00:50:56:88:0b:86, PROTO 0x8
2016 Mar 23 15:09:49.821254 dhcp_snoop: (PKT)IP: DIP 255.255.255.255, SIP 0.0.0.0, t1 328, protocol 17, ttl 128
2016 Mar 23 15:09:49.821267 dhcp_snoop: (PKT)UDP: sport 68, dport 67, len 308
2016 Mar 23 15:09:49.821284 dhcp_snoop: (PKT)DHCP HDR: op 1, htype 1, hlen 6, hops 0, xid 0xcd77351, secs 0, flags
0x0, ciaddr 0.0.0.0, yiaddr 0.0.0.0, siaddr 0.0.0.0, giaddr 0.0.0.0, chaddr 00:50:56:88:0b:86, sname: file
2016 Mar 23 15:09:49.821315 dhcp_snoop: (PKT)DHCP Options: list 12,43,53,55,60-61,116 [ Code 53, len 1, val 01 Code
116, len 1, val 01 Code 61, len 7, val 01 00:50:56:88:0b:86 Code 12, len 11, val 73:6a:72:70:2d:6b:78:70:30:31 Code
60, len 8, val 4d:53:46:54:20:35:2e:30 Code 55, len 11, val
2016 Mar 23 15:09:49.821443 dhcp_snoop: (PKT)source-intf is loopback601 vrf evpn-tenant-kk1
2016 Mar 23 15:09:49.821511 dhcp_snoop: (PKT)Got src intf address 160.1.0.44
2016 Mar 23 15:09:49.821534 dhcp_snoop: (PKT)Source interface context id is 0xc
2016 Mar 23 15:09:49.821580 dhcp_snoop: (PKT) Adding option82 suboptions
2016 Mar 23 15:09:49.821594 dhcp_snoop: (PKT)Mac addr is 88:f0:31:87:1:f3b
2016 Mar 23 15:09:49.821607 dhcp_snoop: (PKT)Add cktid suboption, ifindex 0x9010641 (svi) phy_intf 0x160001bb
2016 Mar 23 15:09:49.821632 dhcp_snoop: (PKT)Could not get slot for ifindex 0x160001bb
2016 Mar 23 15:09:49.821644 dhcp_snoop: (PKT)Could not get port for ifindex 0x160001bb
2016 Mar 23 15:09:49.821654 dhcp_snoop: (PKT)this is in segment 16010 slot 0 port 0 in 160001bb
2016 Mar 23 15:09:49.821665 dhcp_snoop: (PKT)Circuit Id and Remote Id suboptions are added
2016 Mar 23 15:09:49.821729 dhcp_snoop: (PKT) VPN suboptions are added
2016 Mar 23 15:09:49.821768 dhcp_snoop: (PKT) Option82 Hex Dump = [F 0x52 L 0x32 V [F 0x1 L 0xa V 0x1 0x8 0x0 0x6
0x0 0x0 0x3e 0x8a 0x0 0x0 ]|T 0x2 L 0x6 V 0x88 0xf0 0x31 0x87 0x1f 0x3b ]|T 0x97 L 0x10 V 0x0 0x65 0x76 0x70 0x6e
0x2d 0x74 0x65 0x6e 0x61 0x6e 0x74 0x2d 0x6b 0x6b 0x31 ]|T 0x98 L 0x4 V 0
2016 Mar 23 15:09:49.821782 dhcp_snoop: (PKT) RELAYING to server 10.160.2.103
2016 Mar 23 15:09:49.821796 dhcp_snoop: (PKT)IP: DIP 10.160.2.103, SIP 160.1.0.44, t1 31745, protocol 17, ttl 255
2016 Mar 23 15:09:49.821808 dhcp_snoop: (PKT)UDP: sport 68, dport 67, len 360
2016 Mar 23 15:09:49.821827 dhcp_snoop: (PKT)DHCP HDR: op 1, htype 1, hlen 6, hops 1, xid 0xcd77351, secs 0, flags
0x0, ciaddr 0.0.0.0, yiaddr 0.0.0.0, siaddr 0.0.0.0, giaddr 160.1.0.44, chaddr 00:50
56:88:0b:86, sname: file
2016 Mar 23 15:09:49.821886 dhcp_snoop: (PKT)DHCP Options: list 12,43,53,55,60-61,116 [ Code 53, len 1, val 01
Code 116, len 1, val 01 Code 61, len 7, val 01 00:50:56:88:0b:86 Code 12, len 11
, val 73:6a:72:70:2d:6b:6b:78:70:30:31 Code 60, len 8, val 4d:53:46:54:20:35:2e:30 Code 55, len 11, val

```

VXLAN EVPN – DHCP

```
15:09:49.820858 dhcp_snoop: (PKT)Packet received on L3 interface
15:09:49.820969 dhcp_snoop: (PKT)Initial if_index from netstack Vlan1601
2016 Mar 23 15:09:49.821017 dhcp_snoop: (PKT)RECEIVED packet type DHCP RELAY
2016 Mar 23 15:09:49.821201 dhcp_snoop: (PKT)DHCPDISCOVER on Intf Vlan1601(84), phy
port-channel4/44(124), vlan 1601, pvid 0, vrf evpn-tenant-kk1
2016 Mar 23 15:09:49.821213 dhcp_snoop: (PKT)Packet: Len 342, I2 HDR 14, DHCP data 300,
hdr flag 0x36
2016 Mar 23 15:09:49.821239 dhcp_snoop: (PKT)I2: DMAc ff:ff:ff:ff:ff, SMAC 00:50:56:88
0b:86, PROTO 0x8
2016 Mar 23 15:09:49.821254 dhcp_snoop: (PKT)IP: DIP 255.255.255.255, SIP 0.0.0.0, tl1328,
protocol 17, ttl 128
2016 Mar 23 15:09:49.821267 dhcp_snoop: (PKT)UDP: sport 68, dport 67, len 308
2016 Mar 23 15:09:49.821284 dhcp_snoop: (PKT)DHCP HDR:op 1, htype 1, hlen 6, hops 0, xid
0xcd7c7351, secs 0, flags 0x0, ciaddr 0.0.0.0, siaddr 0.0.0.0, giaddr 0.0.0.0,
chaddr 00:50:56:88:0b:86, sname , file
2016 Mar 23 15:09:49.821315 dhcp_snoop: (PKT)DHCP Options: list 12,43,53,55,60-61,116 [
Code 53, len 1, val 01 Code 116, len 1, val 01 Code 61, len 7, val 01 00:50:56:88:0b:86 Code
12, len 11, val 73:6a:72:70:2d:6b:6b:78:70:30:31 Code 60, len 8, val 4d 53:46:54:20:35:2e:30
Code 55, len 11, val
2016 Mar 23 15:09:49.821443 dhcp_snoop: (PKT)source-intf is loopback601 vrf evpn-tenant-
kk1
2016 Mar 23 15:09:49.821511 dhcp_snoop: (PKT)Got src intf address 160.1.0.44
2016 Mar 23 15:09:49.821534 dhcp_snoop: (PKT)Source interface context id is 0xc
2016 Mar 23 15:09:49.821580 dhcp_snoop: (PKT)Adding option82 suboptions
2016 Mar 23 15:09:49.821594 dhcp_snoop: (PKT)Mac addr is 88:0f:31:87:1f:3b
2016 Mar 23 15:09:49.821607 dhcp_snoop: (PKT)Add cktid suboption, ifindex 0x9010641 (svi)
phy_intf 0x160001bb
2016 Mar 23 15:09:49.821632 dhcp_snoop: (PKT)Could not get slot for ifindex 0x160001bb
2016 Mar 23 15:09:49.821644 dhcp_snoop: (PKT)Could not get port for ifindex 0x160001bb
```

2016 Mar 23 15:09:49.821654 dhcp_snoop: (PKT)this is in segment 16010 slot 0 port
0 in 160001bb

2016 Mar 23 15:09:49.821665 dhcp_snoop: (PKT)Circuit Id and Remote Id suboptions
are added

2016 Mar 23 15:09:49.821729 dhcp_snoop: (PKT)VPN suboptions are added

2016 Mar 23 15:09:49.821768 dhcp_snoop: (PKT) Option&2 Hex Dump = [T 0x52 L

0x32 V [T 0x1 L 0xa V 0x1 0x8 0x0 0x6 0x0 0x0 0x3e 0x8a 0x0 0x0]][T 0x2 L 0x6 V

0x88 0x0 0x31 0x87 0x1f 0x3b][T 0x97 L 0x10 V 0x0 0x65 0x76 0x70 0x6e 0x2d 0x74

0x65 0x6e 0x61 0x6e 0x74 0x2d 0x6b 0x6b 0x31][T 0x98 L 0x4 V 0

2016 Mar 23 15:09:49.821782 dhcp_snoop: (PKT)RELAYING to server 10.160.2.103

2016 Mar 23 15:09:49.821796 dhcp_snoop: (PKT)IP: DIP 10.160.2.103, SIP 160.1.0.44,

tl 31745, protocol 17, ttl 255

2016 Mar 23 15:09:49.821808 dhcp_snoop: (PKT)UDP: sport 68, dport 67, len 360

2016 Mar 23 15:09:49.821827 dhcp_snoop: (PKT)DHCP HDR:op 1, htype 1, hlen 6,
hops 1, xid 0xcd7c7351, secs 0, flags 0x0, ciaddr 0.0.0.0, siaddr 0.0.0.0, giaddr 0.0.0.0,
giaddr 160.1.0.44, chaddr 00:50:56:88:0b:86, sname , file

56 88 0b:86, sname , file

2016 Mar 23 15:09:49.821886 dhcp_snoop: (PKT)DHCP Options: list 12,43,53,55,60-

61,82,116 [Code 53, len 1, val 01 Code 116, len 1, val 01 Code 61, len 7, val 01 00:

50:56:88:0b:86 Code 12, len 11

, val 73:6a:72:70:2d:6b:6b:78:70:30:31 Code 60, len 8, val 4d 53:46:54:20:35:2e:30

Code 55, len 11, v

VXLAN EVPN

- DHCP SRV in DEF VRF, Client in VRF

Detail Configuration STEP

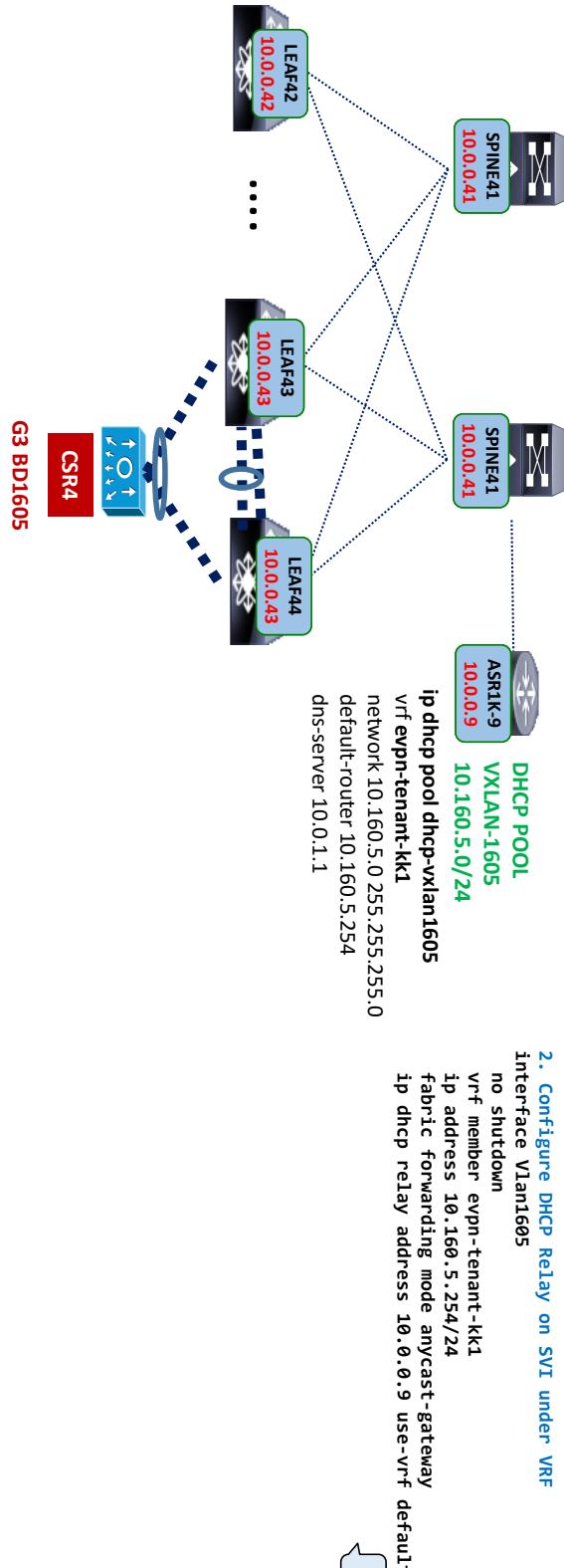
1. Enable DHCP Service and configure dhcp options

```
service dhcp  
ip dhcp relay  
ip dhcp relay information option vpn
```

2. Configure DHCP Relay on SVI under VRF

```
interface Vlan1605  
no shutdown  
vrf member evpn-tenant-kk1  
ip address 10.160.5.254/24  
fabric forwarding mode anycast-gateway  
ip dhcp relay address 10.0.0.9 use-vrf default  
network 10.160.5.0 255.255.255.0  
default-router 10.160.5.254  
dns-server 10.0.1.1
```

Just configure "use-vrf default"



```
CSR4#sh run int g3  
interface GigabitEthernet3  
platform ring rx 256  
vrf forwarding bd1605  
ip address dhcp  
negotiation auto
```

VXLAN EVPN

- DHCP SRV in External , Client in VRF

Detail Configuration STEP
 1. Enable DHCP Service and configure dhcp options
 service dhcp
 ip dhcp relay

2. Configure DHCP Relay only BORDER LEAF 42's SVI

On LEAF42
 interface Vlan1601
 no shutdown
 mtu 9216

vrf member evpn-tenant-kk1
 ip address 10.160.1.254/24

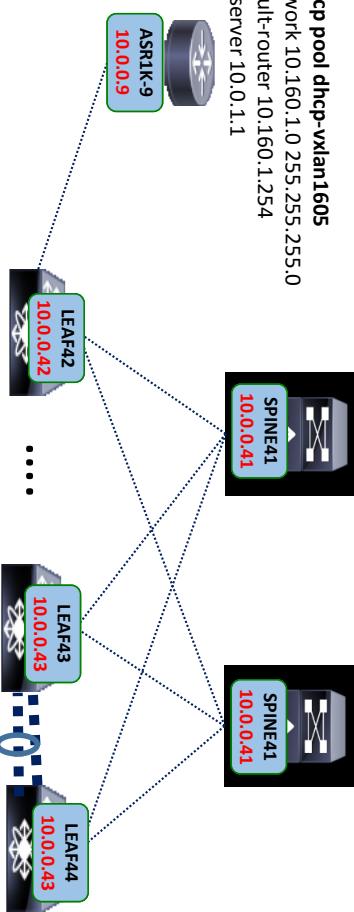
fabric forwarding mode anycast-gateway
 ip dhcp relay address 10.0.0.9

On LEAF43/44, no ID DHCP RELAY configured
 interface Vlan1601
 no shutdown
 mtu 9216

vrf member evpn-tenant-kk1
 ip address 10.160.1.254/24

fabric forwarding mode anycast-gateway

Just configure "use-vrf default"

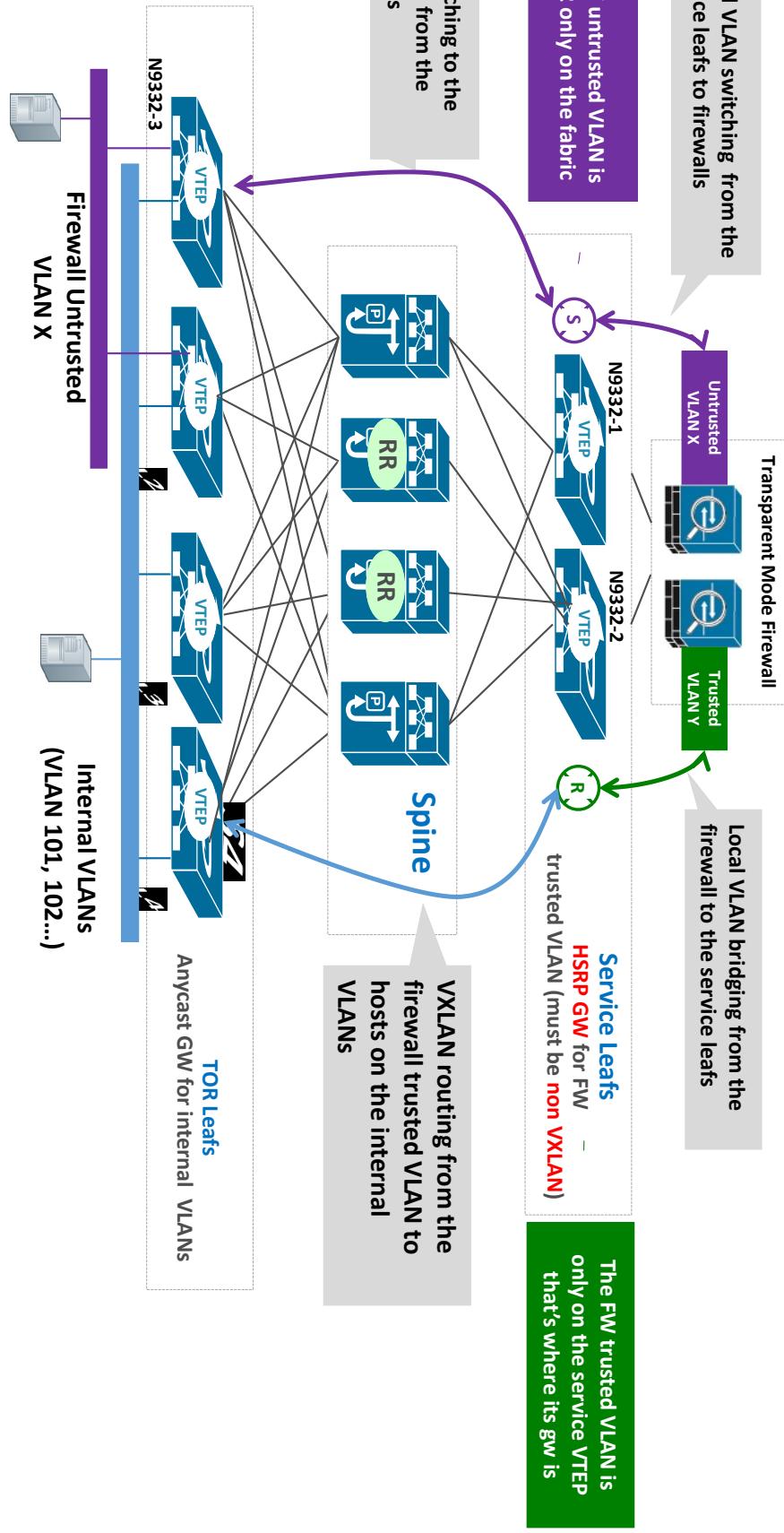


```
CSR4#sh run int g3
interface GigabitEthernet3
platform ring rx 256
vrf forwarding bd1605
ip address dhcp
negotiation auto
```

DHCP Relay Configuration Links

- http://www.cisco.com/c/en/us/td/docs/sw/7-x/vxlan/configuration/guide/bCisco_Nexus_9000_Series_NX-OS_VxLAN_Configuration_Guide_7x/bCisco_Nexus_9000_Series_NX-OS_VxLAN_Configuration_Guide_7x_appendix_0110.html
- <https://techzone.cisco.com/t5/Nexus-9300/DHCP-Relay-in-VxLAN-EVPN/ta-p/842437>
- <https://techzone.cisco.com/t5/Nexus-9500/VxLAN-DHCP-relay-Microsoft-DHCP-server-config/ta-p/844198?attachment-id=56286>

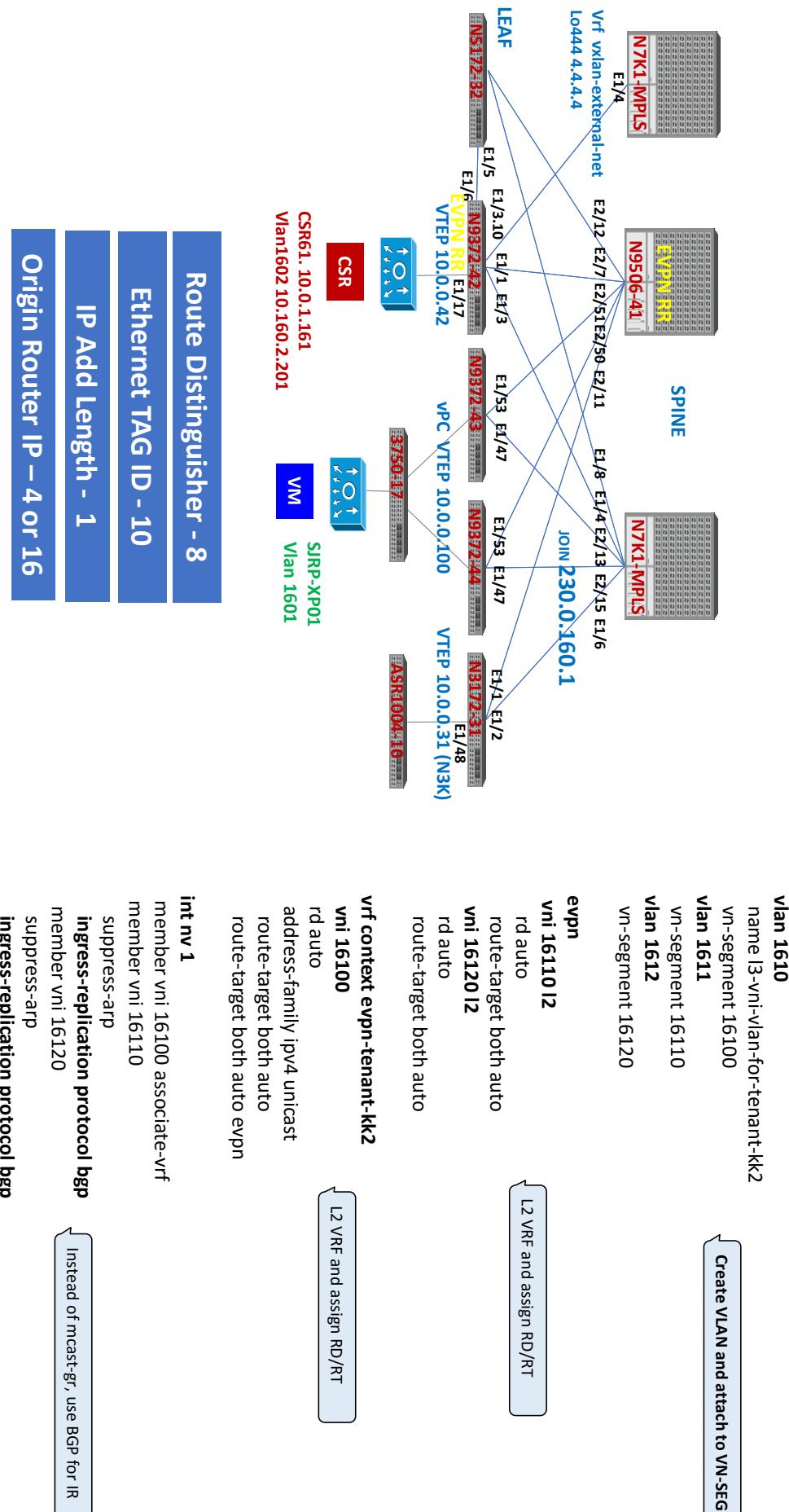
Transparent Firewall Insertion to VXLAN EVPN Fabric



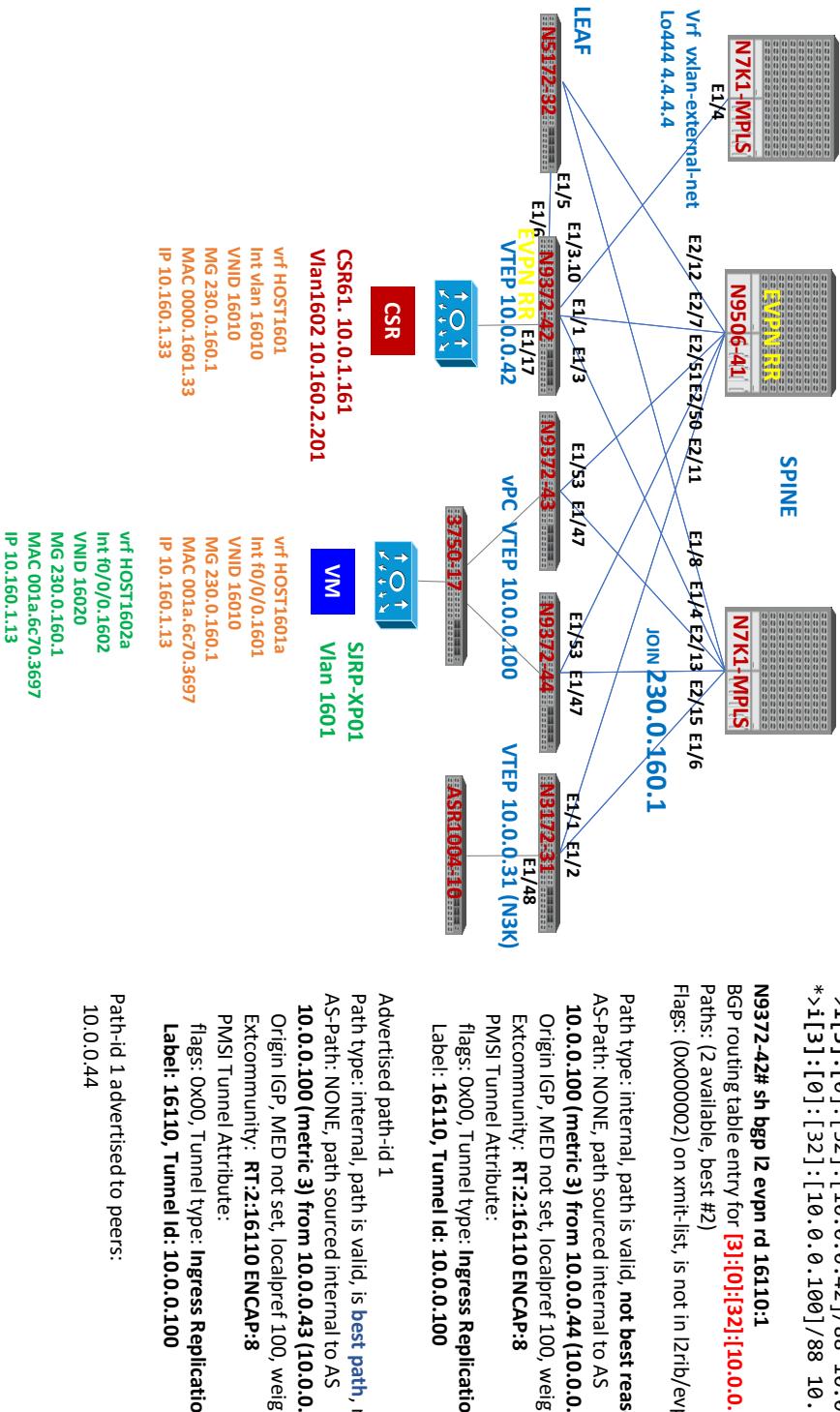
VXLAN Peer and Host Learning Options

Host Learning	Data-Plane	Control-Plane
CORE		
Multicast		
Flood and Learn		EVPN-Multicast
Peer Learning: DP		Peer Learning: BGP-RnH
	<pre>Vlan 2 vn-segment 4098 Interface nve 1 member vni 10000 mcast-group 225.1.1.1</pre>	<pre>Vlan 2 vn-segment 10000 Interface nve 1 host-reachability protocol bgp member vni 4098 mcast-group 225.1.1.1</pre>
Unicast		
Static Ingress-Replication		EVPN Ingress-Replication
Peer Learning: CLI		Peer Learning: BGP-IMET
	<pre>Vlan 2 vn-segment 4098 Interface nve 1 member vni 4098 ingress-replication protocol static peer x.x.x</pre>	<pre>Vlan 2 vn-segment 4098 Interface nve 1 host-reachability protocol bgp member vni 4098 ingress-replication protocol bgp</pre>

VXLAN EVPN – Ing Replication



VXLAN EVPN – Ingr Replication



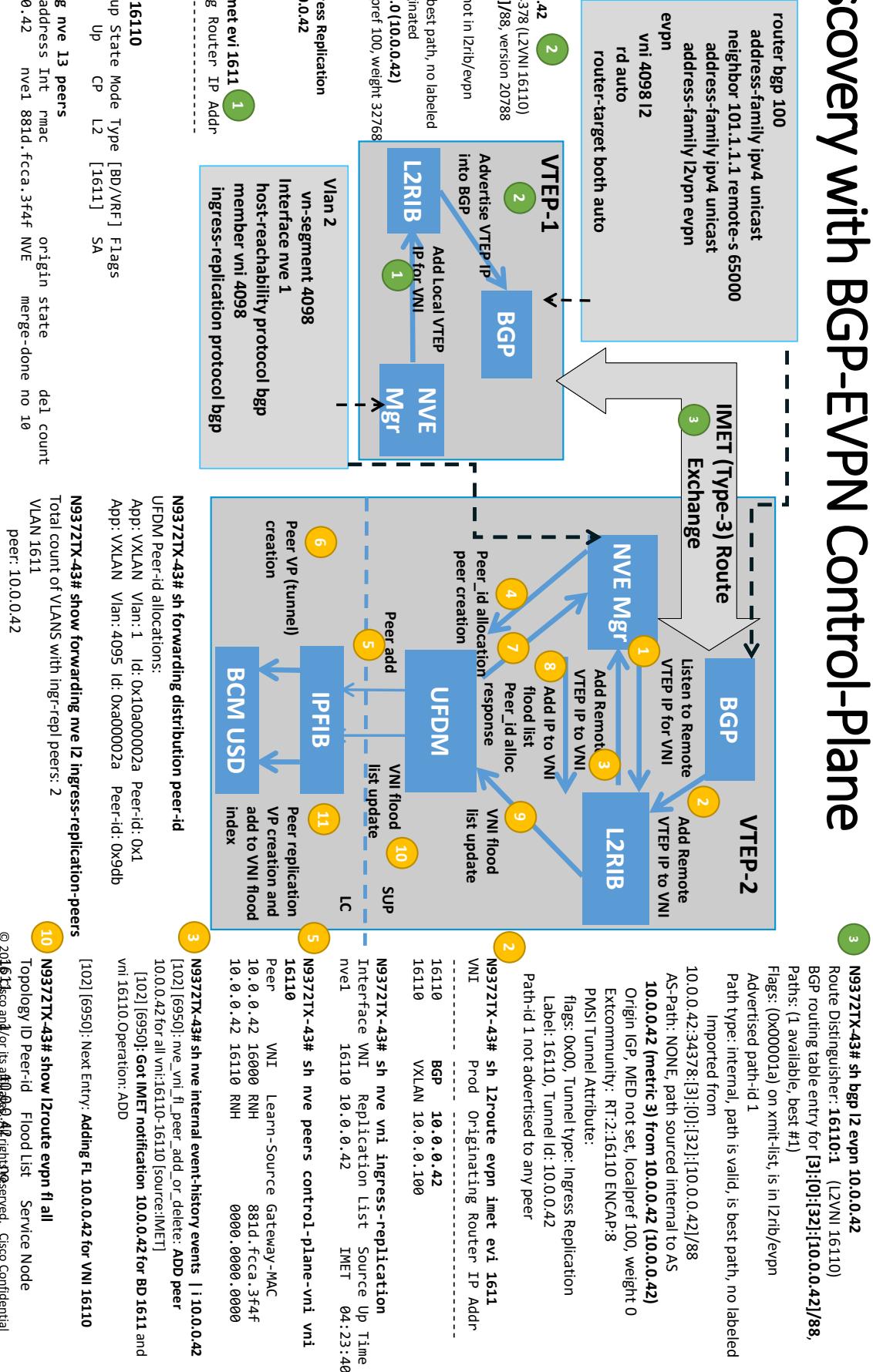
```
N9372-42# sh bgp 12 evpn vni-id 16110
Network  Next Hop      Metric LocPrf Weight Path
          Distinguisher: 10.0.0.42:34378  (L2VNI 16110)
        *->[3]:[0]:[32]:[10.0.0.42]/88 10.0.0.42 100 32768 i
        *>i[3]:[0]:[32]:[10.0.0.100]/88 10.0.0.100 100 0 I
```

Network	Next Hop	Metric	LocPrf	Weight	Path
Route Distinguisher: 10.0.0.42:34378 (L2VNI 16110)					
*->[3]:[0]:[32]:[10.0.0.42]/88 10.0.0.42 100 32768 i					
*->i[3]:[0]:[32]:[10.0.0.100]/88 10.0.0.100 100 0 I					

N9372-42# sh bgp 12 evpn rd 16110:1
BGP routing table entry for [3]:[10]:[32]:[10.0.0.100]/88, version 33057
Paths: (2 available, best #2)
Flags: (0x0000002) on xmit-list, is not in 1rib/evpn, is locked

Path type: internal, path is valid, **not best** reason: Neighbor Address, no labeled nexthop
AS-Path: NONE, path sourced internal to AS
10.0.0.100 (metric 3) from 10.0.0.44 (10.0.0.43)
Origin IGP, MED not set, localpref 100, weight 0
Extcommunity: **RT:2:16110 ENCAP:8**
PMI Tunnel Attribute:
flags: 0x00, Tunnel type: Ingress Replication
Label: 16110, Tunnel Id: 10.0.0.100

Peer discovery with BGP-EVPN Control-Plane



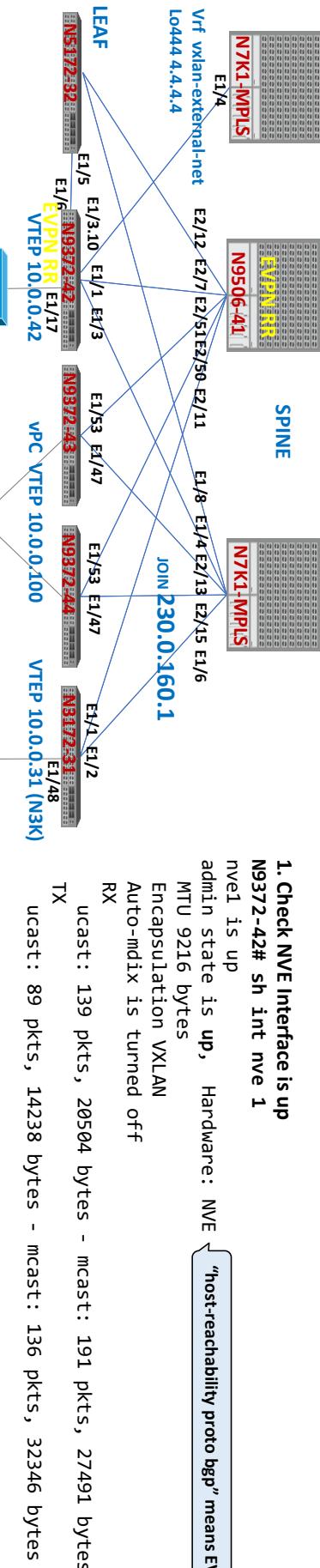
VXLAN EVPN – TS

Troubleshooting for Infrastructure

Host **10.160.1.33** on LEAF42 can't ping host **10.160.1.13** on vpc LEAF 43/44.
Ingress LEAF is LEAF42 and Egress LEAF is vPC LEAF43/44.

1. Check NVE Interface is up
N9372-42# sh int nve 1
 nve1 is up
 admin state is up, Hardware: NVE
 MTU 9216 bytes
 Encapsulation VXLAN
 Auto-mdix is turned off

"host-reachability proto bgp" means EVPN



If down, make sure you didn't no shut

2. Check NVE State

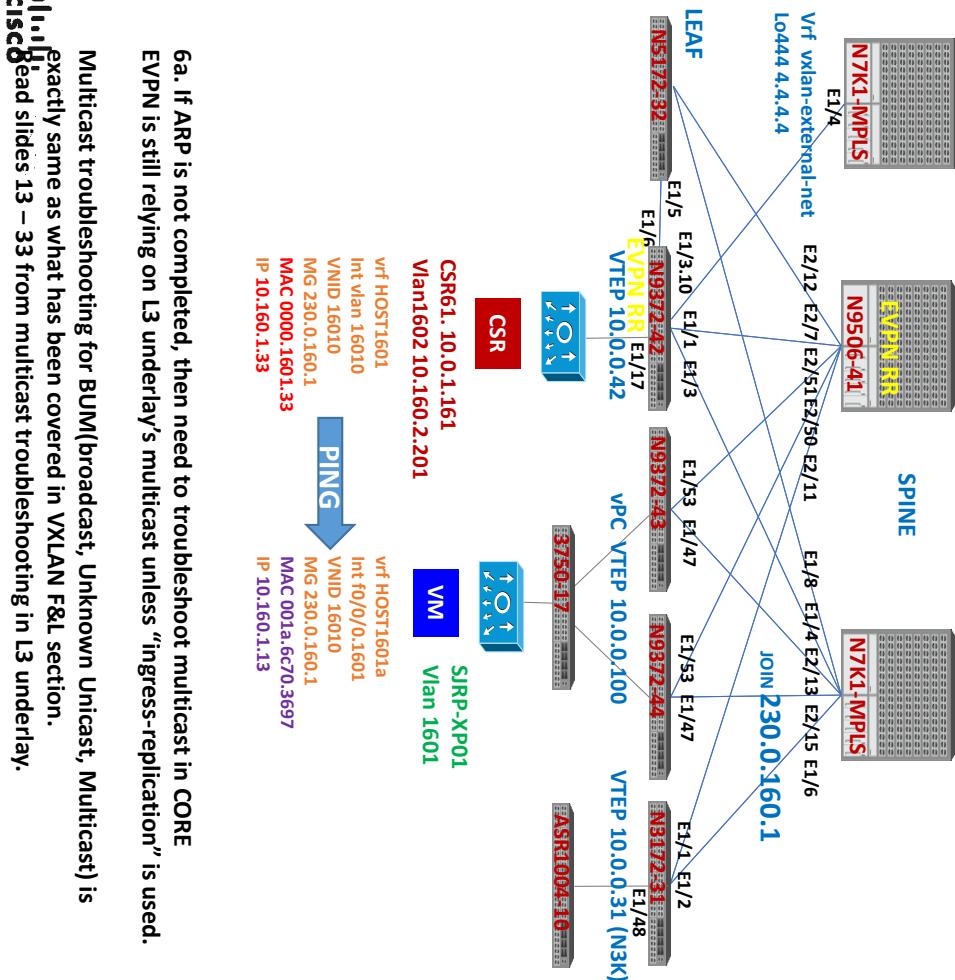
```

N9372-42# sh nve interface nve 1 det
Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VTP-Only [not-notified]
Local Router MAC: 881d.fcca.3f4f
Host Learning Mode: Control-Plane
Source-Interface: loopback0 (primary: 10.0.0.42, secondary: 0.0.0.0)
Source Interface State: Up
NVE Flags: No Request Pending
Interface Handle: 0x49000001
  
```

- a. If ARP is not completed, then need to troubleshoot multicast in CORE
 EVPN is still relying on L3 underlay's multicast unless "ingress-replication" is used.

- Multicast troubleshooting for BUM(broadcast, Unknown Unicast, Multicast) is exactly same as what has been covered in VXLAN F&L section.
- cisco8dead slides 13 – 33 from multicast troubleshooting in L3 underlay.

VXLAN EVPN – TS



```
N9372-42# sh bgp internal evi
BGP L2VPN/EVPN RD Information for 16010:1
```

L2VNI ID : 16010 (evi_16010)

#Prefixes Local/BRIB : 0 / 8

#Paths L3VPN->EVPN/EVPN->L3VPN : 0 / 3

```
BGP L2VPN/EVPN RD Information for 16000:1
```

L3VNI ID : 16000 (evi_16000)

#Prefixes Local/BRIB : 1 / 6

#Paths L3VPN->EVPN/EVPN->L3VPN : 10 / 5

```
N9372-42# sh bgp internal evi 16010
*****
```

L2RIB bound / VNI Req to L2RIB : Yes / 1

VNI Adds / Dels from L2RIB : 24 / 0

L2RIB is up/registered/local-req: 1/1

L2RIB down: in-prg/up-defer: 0/0

L2RIB register/failures: 1/0

L2RIB deregister/failures: 0/0

L2RIB flow control (#enabled/#disabled): Disabled (0/0)

```
*****
```

```
BGP L2VPN/EVPN RD Information for 16010:1
L2VNI ID : 16010 (evi_16010)
```

#Prefixes Local/BRIB : 0 / 8

#Paths L3VPN->EVPN/EVPN->L3VPN : 0 / 3

```
*****
BGP Configured VNI Information:
VNI ID (Index) : 16010 (0)
RD : 16010:1
```

Export RTs	Export RT cfg list: 16010:1
Import RTs	Import RTs : 1
Import RT cfg list:	16010:1

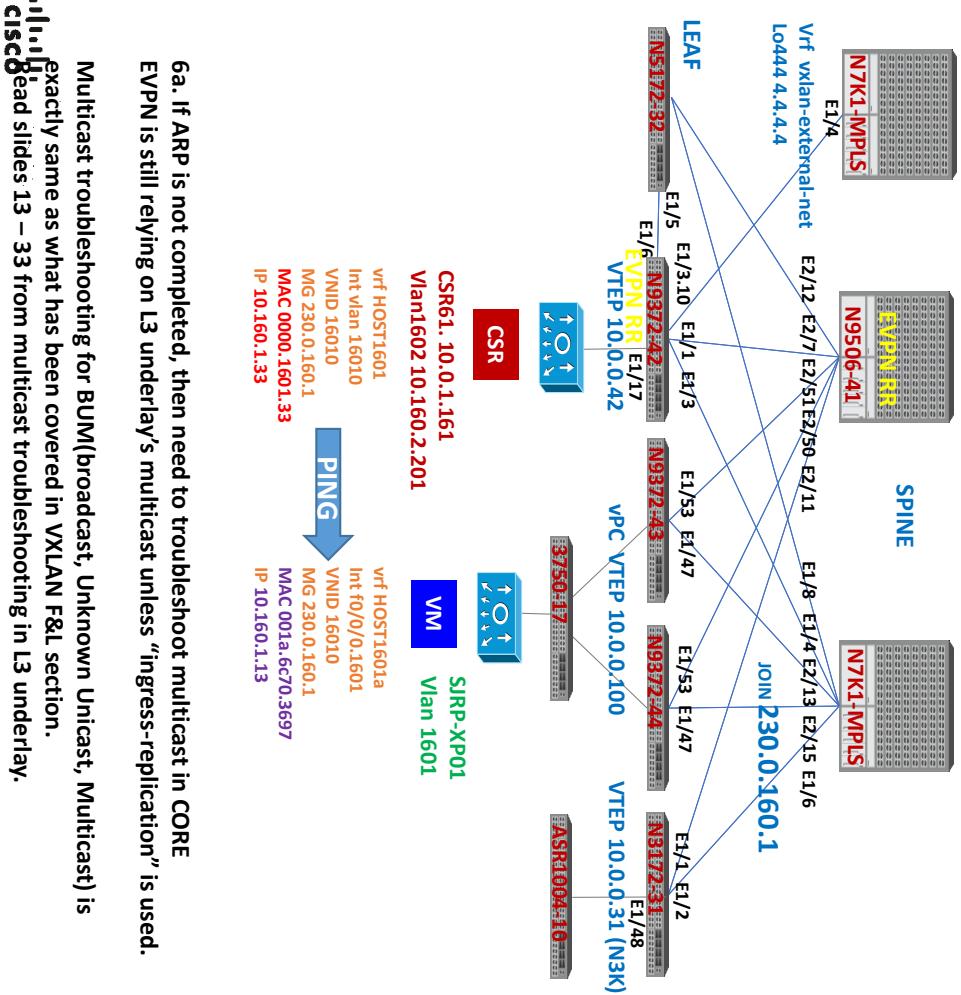
Topo Id	1601
VTEP IP	10.0.0.42
VTEP VPC IP	0.0.0.0

If no, make sure you configured VNI in NVE interface

- 6a. If ARP is not completed, then need to troubleshoot multicast in CORE
- EVPN is still relying on L3 underlay's multicast unless "ingress-replication" is used.

- Multicast troubleshooting for BUM(broadcast, Unknown Unicast, Multicast) is exactly same as what has been covered in VXLAN F&L section.
- cisco dead slides 13 – 33 from multicast troubleshooting in L3 underlay.

VXLAN EVPN – TS



- 6a. If ARP is not completed, then need to troubleshoot multicast in CORE**
EVPN is still relying on L3 underlay's multicast unless "ingress-replication" is used.
- Multicast troubleshooting for BUM(broadcast, Unknown Unicast, Multicast) is exactly same as what has been covered in VXLAN F&L section.
- cisco** Dead slides 13 – 33 from multicast troubleshooting in L3 underlay.

```
+++++
BGP VNI Information for evi_16010
L2VNI ID : 16010 (evi_16010)
RD : 16010.1
```

Prefixes (local/total)	0/8
Enabled	yes
Delete pending	0
Import pending	0
Import in progress	0
Encap	VXLAN
ToD Id	1601
VTEP IP	10.0.0.42
VTEP VPC IP	0.0.0.0
Active Export RTs	1
Config Export RTs	16010.1
Export RT cfg list: 16010:1	0/0
Import RT chg/chg-pending	0/0
IMET Reg/Unreg from L2RIB	1/0
Active Import RTs	1
Config Import RTs	16010.1
Import RT cfg list: 16010:1	0/0
IMET Reg/Unreg from L2RIB	1/0
MAC Reg/Unreg from L2RIB	1/0
MAC IP Reg/Unreg from L2RIB	1/0
IMET Add/Del from L2RIB	0/0
MAC Add/Del from L2RIB	1/2
MAC IP Add/Del from L2RIB	0/0
IMET Dnld/Wdraw to L2RIB	0/0
MAC Dnld/Wdraw to L2RIB	58/0
MAC IP Dnld/Wdraw to L2RIB	42/0

“host-reachability proto bgp” means EVPN

```
N9372TX-43# sh bgp internal nve-peer-vni
PeerAddress VNI VrfID GatewayMAC TunnelID Encap EgressVNI F
10.0.0.42 16000 1 881d.fccca.3f4f 0xa00002a 1 0 0
10.0.0.42 16010 1 0000.0000.0000 0x0 1 0 0
10.0.0.42 16020 1 0000.0000.0000 0x0 1 0 0
```

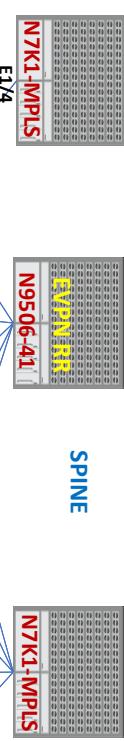
VXLAN EVPN – TS

Troubleshooting Steps for BRIDGE Traffic

Host **10.160.1.33** on LEAF42 can't ping host **10.160.1.13** on vpc LEAF 43/44.
Ingress LEAF is LEAF42 and Egress LEAF is vPC LEAF43/44.

- Check if it is EVPN or VXLAN
N9372TX-43# sh run int nv 1

"host-reachability proto bgp" means EVPN



N9372-42# sh run int nv 1
vrf vxlan-extn-net
Lo444 4.4.4.4
E1/4
E2/12 E2/7 E2/51 E2/50 E2/11
E1/8 E1/4 E2/13 E2/15 E1/6
JOIN 230.0.160.1

- Check if it is EVPN or VXLAN
N9372TX-43# sh run int nv 1
- Check local reachability from ingress LEAF42 to local **HOST, 10.160.1.33**
You can only ping local host from any LEAF and can't ping any remote HOST due to Dist Anycast GW made all SVI on all Leafs have same IP.
If you need to test remote connectivity from LEAF, configure Loopback X on a LEAF and put it into proper VRF and redistribute DIRECT2BGP under BGP VRF

N9372-42# ping **10.160.1.33** vrf evpn-tenant-kk1
PING 10.160.1.33 (10.160.1.33): 56 data bytes
64 bytes from 10.160.1.33: icmp_seq=0 ttl=254 time=0.942 ms

- Check if local/remote MAC address is installed in MAC table

- If ping to Local host failed, then check if LEAF learned Local MAC

```
N9372-42# sh mac add v1 1601
VLAN      MAC Address    Type      age      Secure  NTFY Ports
-----+-----+-----+-----+-----+-----+
      +-----+-----+-----+-----+-----+
 1601  0000.1601.0033  dynamic   0        F        F      Eth1/17
```

- If no MAC found, check spanning tree and L2 related configuration

```
N9372-42# sh spanning-tree vlan 1601 | i 1/17
Eth1/17          Root  FWD 2      128.17  P2p
```

- If local PING works on both ingress/egress LEAF but host can't ping each other, make sure host can ping GW and check subnet mask

- Check ARP on local HOST if ARP is completed.

```
N9372-42# sh ip arp vrf host1601 | i 1.13
10.160.1.13  00:06:03  001a.6c70.3697  Vlan1601
```

- If ARP is not completed, then need to troubleshoot multicast in CORE

EVPN is still relying on L3 underlay's multicast unless "ingress-replication" is used.

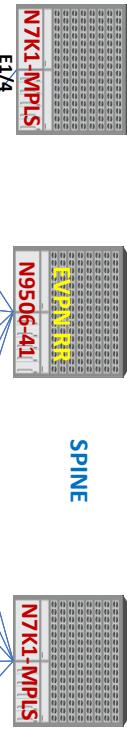
- Multicast troubleshooting for BUM(broadcast, Unknown Unicast, Multicast) is exactly same as what has been covered in VXLAN F&L section.
- cisco8dead slides 13 – 33 from multicast troubleshooting in L3 underlay.

VXLAN EVPN – TS

7. Check Remote MAC on ingress LEAF42

```
N9372-42# sh l2route evpn mac evi 1601
* 1601 001a.6c70.3697 dynamic 0
```

F F nve1(10.0.0.100)



8. Check L2 Route to remote MAC on ingress LEAF42

```
N9372-42# sh l2route evpn mac evi 1601
Mac Address Prod Next Hop (s)
-----
```

0000.1601.0033	Local	Eth1/17
0015.f905.8214	BGP	10.0.0.100
001a.6c70.3697	BGP	10.0.0.100

L2RIB shows that ingress LEAF42 learned 10.0.0.100 (vPC VTEP 4344) from remote MAC 3697 from remote VTEP 10.0.0.100 (vPC VTEP 4344)

9. If no L2 route to remote MAC, check egress LEAF's has local MAC installed for .3697

```
N9372TX-43# sh mac add via 1601 | i 36
+ 1601 001a.6c70.3697 dynamic 0 F F F Po1743
```

N9372TX-43# sh l2route evpn mac evi 1601 | i 3697

001a.6c70.3697 Local Po1743

Check Egress LEAF's MAC table and L2RIB table

10. If L2 RIB is good, then check EVPN route for MAC is installed on Egress LEAF.

```
N9372TX-43# sh bgp l2vpn evpn 001a.6c70.3697
Route Distinguisher: 16010:1 (L2VNI 16010)
```

BGP routing table entry for [2]:[0]:[0]:[48]:[001a.6c70.3697]:[0]:[0.0.0.0]/216,

version 16

Paths: (1 available, **best #1**)

Flags: (0x00010a) on xmit-list, is not in l2rib/evpn

Advertised path-id 1

Path type: local, path is valid, is **best path**, no labeled next-hop

AS-Path: NONE, path locally originated

10.0.0.100 (metric 0) from 0.0.0.0(10.0.0.43)

Origin (GP, MED not set, localpref 100, weight 32768)

Received label **16010**

Extcommunity: RT:16010:1 S00:10.0.0.100:0

RT from EVPN is attached to S00, site of origina, is egress VTEP's IP

Only one label is showed up and it is derived from L2 VNI of DMAC

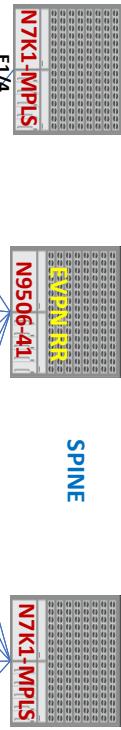
Path-id 1 advertised to peers:

10.0.0.41 10.0.0.42

- These troubleshooting steps are very similar to usual MPLS L3 VPN TS
- 1. Check ingress/egress PE is getting traffic/information about local host
- 2. Check ingress PE received egress PE's route and installed in RIB
- 3. Check egress PE if destination route is installed local and advertised to MP-BGP
- cisco: Check ingress PE received correct routes from egress PE

VXLAN EVPN – TS

11. After we verified that egress LEAF advertised destination route into BGP, now we check ingress LEAF got proper EVPN routes from RR



LEAF
N9372-42
E1/4
Lo444 4.4.4.4
Vrf vxlan-extenal-net
E2/12 E2/7 E2/5 E2/50 E2/11
E2/13 E2/15 E1/6
E1/5 E1/3.10 E1/1 E1/3
E1/17 E1/42
E1/53 E1/47
E1/53 E1/47
E1/1 E1/2
E1/48

JOIN 230.0.160.1

LEAF
N9372-43
E1/4
E1/5 E1/3.10 E1/1 E1/3
E1/17 E1/42
VTEP 10.0.0.42
VPC VTEP 10.0.0.100
VTEP 10.0.0.31 (N3K)

Advertised path-id 1
Path type: internal, path is valid, imported same remote RD, is best path, no labeled nexthop

AS-Path: NONE, path sourced internal to AS
10.0.0.100 (metric 3) from 10.0.0.43 (10.0.0.43)
Origin IGP, MED not set, localpref 100, weight 0

Received label 16010

Extcommunity: RT:16010:1 SOO:10.0.0.100:0 ENCAP:8
Path type: internal, path is valid, imported same remote RD, not best reason:
Neighbor Address, no labeled nexthop
AS-Path: NONE, path sourced internal to AS
10.0.0.100 (metric 3) from **10.0.0.44** (10.0.0.43)
Origin IGP, MED not set, localpref 100, weight 0
Received label 16010
Extcommunity: RT:16010:1 SOO:10.0.0.100:0 ENCAP:8

RT is correct and now we see ENCAP info
This is L2 MAC route so no router mac info is there.

CSR
CSR61. 10.0.1.161
Vlan1602 10.160.2.201
VM SRP-XP01
Vlan 1601

vrf HOST1601
Int f0/0/0.1601
VNID 16010

vrf HOST1601a
Int f0/0/0.1601
VNID 16010

MG 230.0.160.1
MAC 0000.1601.33
IP 10.160.1.13

vrf HOST1602a
Int f0/0/0.1602
VNID 16020
MG 230.0.160.1
MAC 001a.6c70.3697
IP 10.160.1.13

10. If EVPN route is correct but not installed in L2Route, read slide 46 and follow all step

VXLAN EVPN – TS

Troubleshooting Steps for ROUTED traffic

Host **10.160.1.33** in vlan 1601 on LEAF42 can't ping host **10.160.2.13** in vlan 1602 on vpc

LEAF 43/44. → InterVXLAN traffic

Ingress LEAF is LEAF42 and Egress LEAF is vPC LEAF43/44.

1. Assume all L2 is good.
2. Always check L3 VRF RIB first for any L3 issue as DMAC is router mac when traffic is L3 and ingress LEAF route packet based on L3 VNI VRF

```
N9372-42# sh ip route vrf evpn-tenant-kk1 10.160.2.13
IP Route Table for VRF "evpn-tenant-kk1"
10.160.2.13/32, ubest/mbest: 1/0
  *via 10.0.0.100%default, [200/0], 00:00:44, bgp-2, internal, tag 2
(evpn) segid: 16000 tunnelid: 0xa000064 encap: VXLAN
```

3. If not there, check L2RIB on dest VXLAN. (src vlan is 1601 and dest vlan is 1602)

	Mac Address	Prod Host IP	Next Hop (s)
0000.1602.0033	HMM	10.160.2.33	N/A
001a.6c70.3697	BGP	10.160.2.13	10.0.0.100
0022.90b0.c653	BGP	10.160.2.14	10.0.0.100

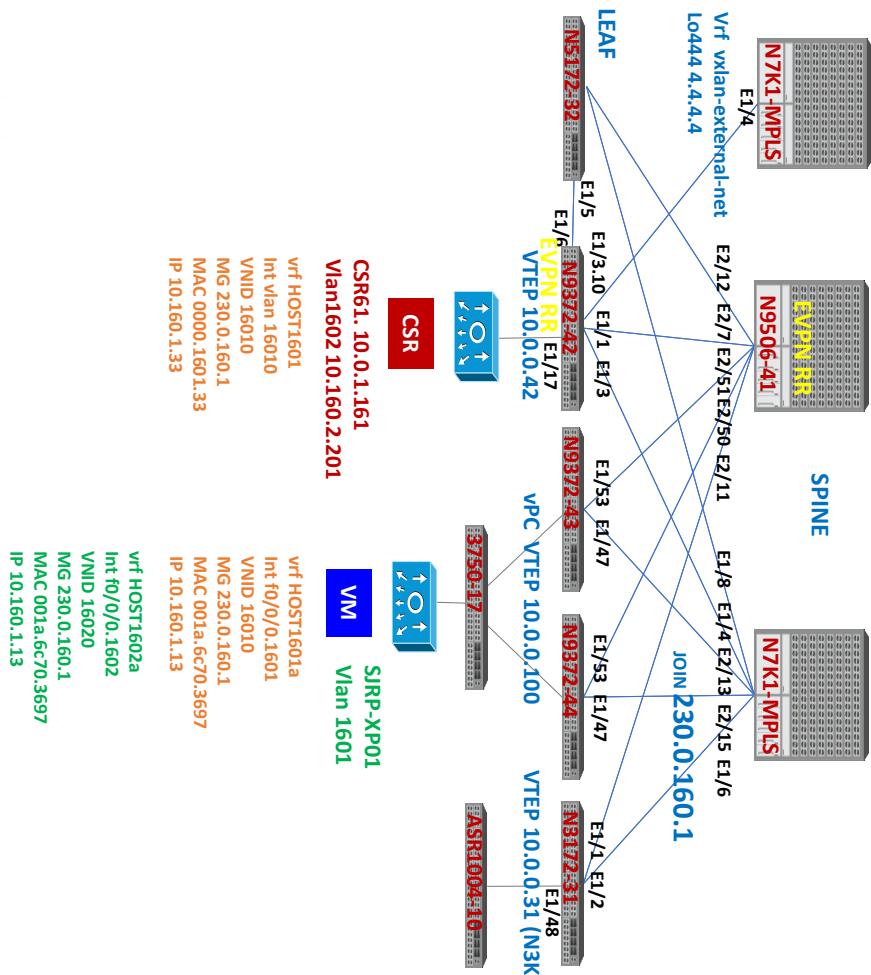
CSR
CSR61. 10.0.1.161
Vlan1602 10.160.2.201

```
vrf HOST1601
int f0/0/0.1601
vnid 16010
vniD 16010
MAC 0000.1601.33
MG 230.0.160.1
MAC 001a.6c70.3697
IP 10.160.1.13
vrf HOST1602a
int f0/0/0.1602
vnID 16020
MG 230.0.160.1
MAC 001a.6c70.3697
IP 10.160.2.13
```

4. If dest IP not in L2RIB on dest VLAN, then check IPv4 VRF BGP if/32 host route is learned from egress LEAF

```
N9372-42# sh ip bgp vrf evpn-tenant-kk1
      Network          Next Hop            Metric     LocPrf    Weight Path
        *>r4.4.4.4/32    0.0.0.0           5          100      32768 ?
        *>i10.0.0.143/32 10.0.0.100       0          100      0 ?
        *>i10.160.1.0/24 10.0.0.100       0          100      0 ?
        *>i10.160.1.13/32 10.0.0.100       0          100      0 i
        *>i10.160.1.14/32 10.0.0.100       0          100      0 i
        *>i10.160.2.0/24 10.0.0.100       0          100      0 ?
        *>i10.160.2.13/32 10.0.0.100       0          100      0 i
        *>i10.160.2.14/32 10.0.0.100       0          100      0 I
```

VXLAN EVPN – TS



5. Check egress LEAF L2 MAC-IP to make sure EGRESS LEAF learned dest IP in dest VLAN

N9372TX-43# sh l2route evpn mac-ip evi 1602	Mac Address	Prod Host IP	Next Hop (s)
0000.1602.0033 BGP	10.160.2.33	10.0.0.42	
001a.6c70.3697 HMM	10.160.2.13	N/A	
0022.90b0.c653 HMM	10.160.2.14	N/A	

6. If MAC-IP for dest IP is not learned, follow slides 47 for local IP learn/advertise

7. Check egress LEAF advertise MAC-IP into L2VPN BGP

N9372TX-43# sh bgp l2vpn evpn 10.160.2.13
BGP routing table information for VRF default, address family L2VPN EVPN

Route Distinguisher: 16020:1 (L2VNI 16020)

BGP routing table entry for [2]:[0]:[0]:[48]:[001a.6c70.3697]:[32]:[10.160.2.13]/272, version

3

Paths: (1 available, best #1)
Flags: (0x00010a) on xmit-list, is not in l2rib/evpn

MAC-IP

Advertised path-id 1
Path type: local, path is valid, is best path, no labeled nexthop

AS-Path: NONE, path locally originated

10.0.0.100 (metric 0) from 0.0.0.0 (10.0.0.43)

0.0.0.0 as this LEAF is originating route

Origin IGP, MED not set, localpref 100, weight 32768

Received label 16020 **16000**

Extcommunity: RT:16000:1 RT:16020:1 SOO:10.0.0.100:0

vrf HOST1601a
Int f0/0/0.1601
VNID 16010

MG 230.0.160.1
MAC 000a.6c70.3697

IP 10.160.1.13

vrf HOST1602a
Int f0/0/0.1602
VNID 16020

MG 230.0.160.1
MAC 001a.6c70.3697

IP 10.160.1.13

Advertised to two RR, LEAF42
and SPINE41

VXLAN/EVPN HW FWD

Nexus 9500 Switch Line Cards

X9600 Series Line Cards (NFE only)

High Performance 40G Aggregation

- N9K-X9636PQ

Merchant only

NX-OS Mode Only

X9500 Series Line Cards (NFE + ALE)

Performance 10G/40G Access/Aggregation

- N9K-X9564PX
- N9K-X9564TX
- N9K-X9536PQ

Merchant +

NX-OS Mode &
ACI Leaf Ready
Wlan/EVPN

X9400 Series Line Cards (NFE only)

Entry Level 10GE/40GE Access and Aggregation

- N9K-X9432PQ
- N9K-X9464PX
- N94-X9464TX

Merchant only

NX-OS Mode Only

X9700 Series Line Cards (ASE only)

High Performance 40G ACI Spine

- N9K-X9736PQ

Merchant +

ACI Spine Only



Nexus 9500 NX-OS Line Cards

Merchant+



9564 (48x10G+4x40G)



9564 (48x10G+4x40G)



9536 (36x40G)

- Additional buffer
- **VXLAN routing***
- ACI Enabled*

Merchant



9464TX (48x10G+4x40G)



9464PX (48x10G+4x40G)



9636PQ (36x40G)



9432PQ (32x40G)

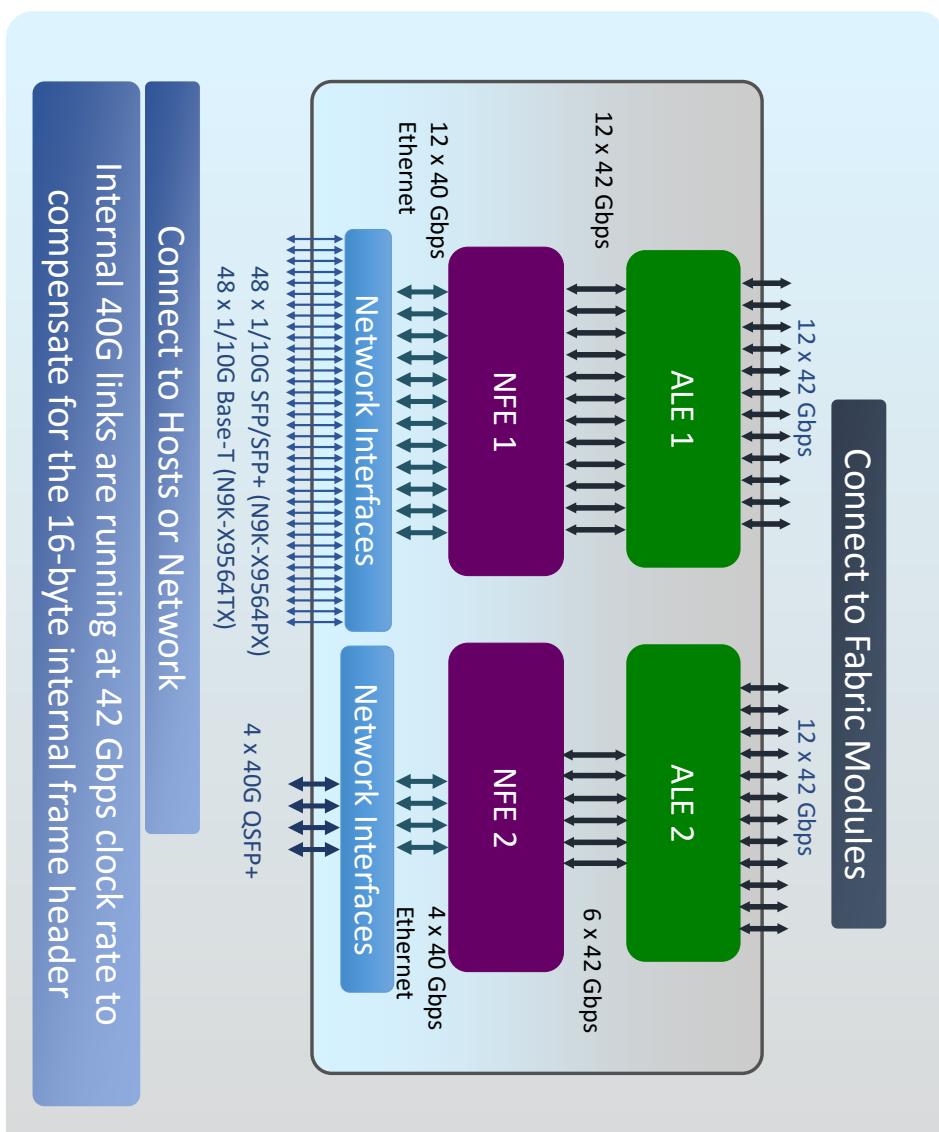
10G Copper

10G Fiber

40G Fiber

Nexus 9500 N9K-X9500 Series Line Cards

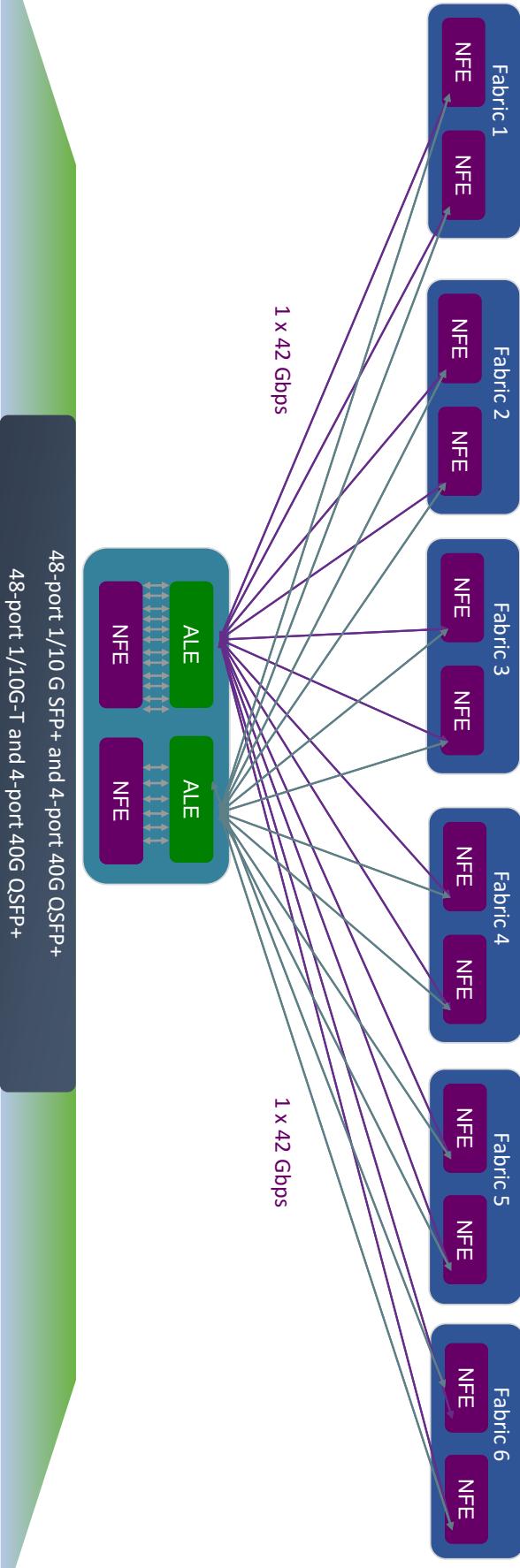
- 2 network forwarding engines (NFEs)
NFE ASICs act as main forwarding engines for standalone mode.
- 2 application leaf engines (ALEs) for additional buffering and packet handling
- Works in 4, 8 and 16 slot chassis
- Line rate performance on all ports and all packet sizes with 3 or 6 fabric modules



Nexus 9500 N9K-X9500 Series Line Cards

N9K-X9564PX & N9K-X9564TX Fabric Connectivity

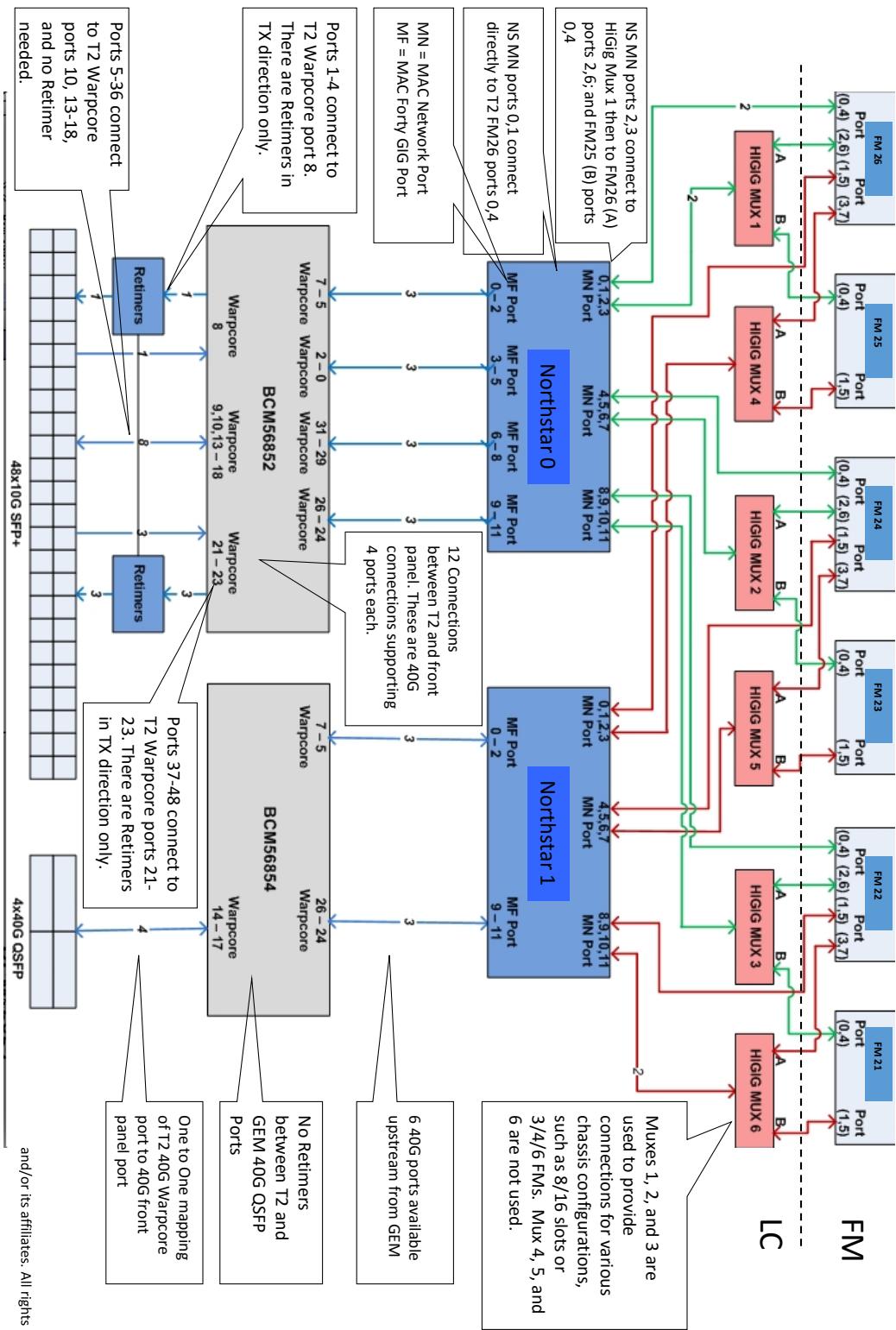
Fabric Connectivity with 6 Fabric Modules in Nexus 9508 Chassis



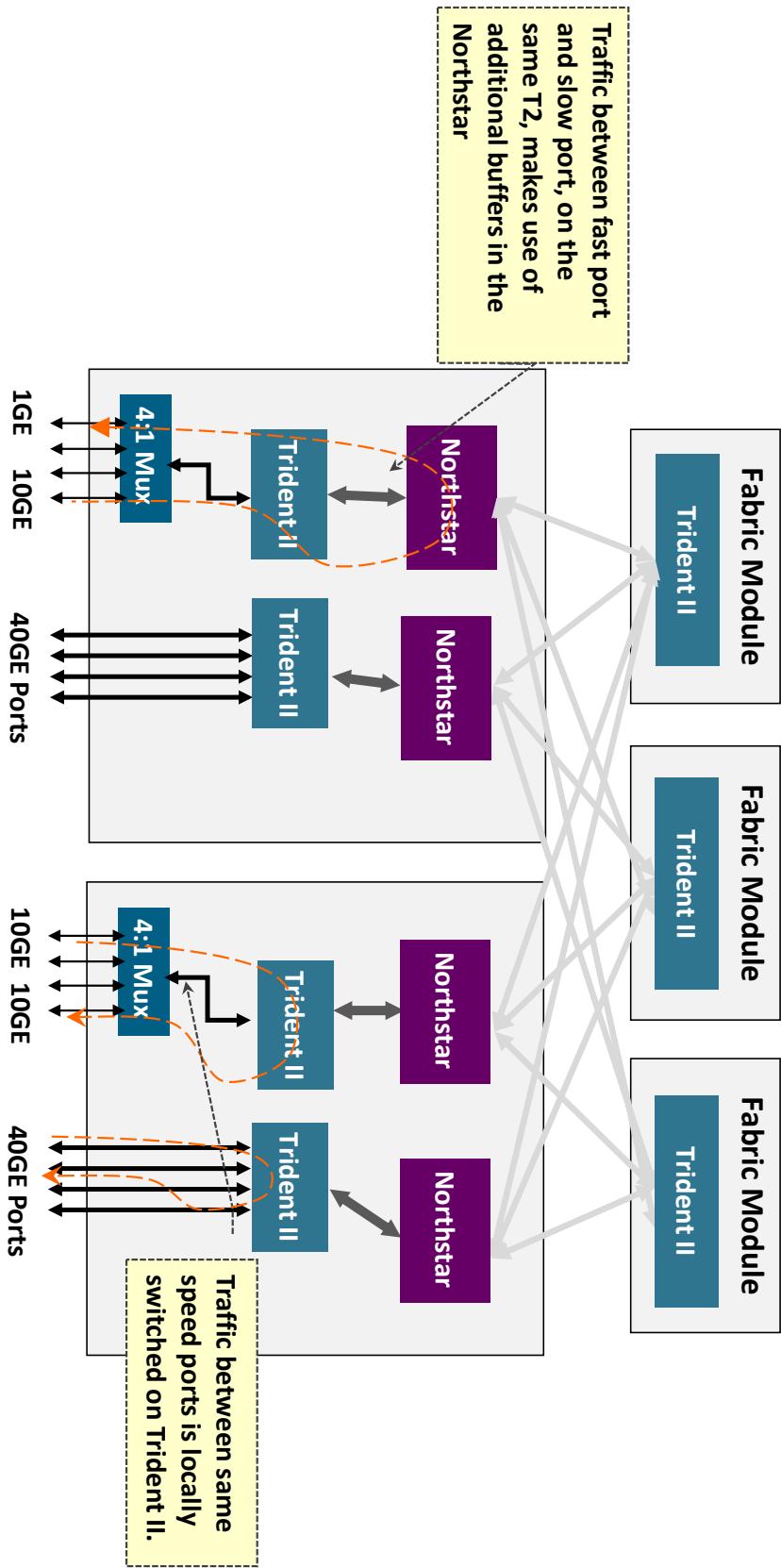
When a chassis has 6 fabric modules, N9K-X9564PX and N9K-X9564TX are connected to and utilize all 6 fabric modules

The above diagram shows the line card fabric connectivity in a Nexus 9508 chassis as an example

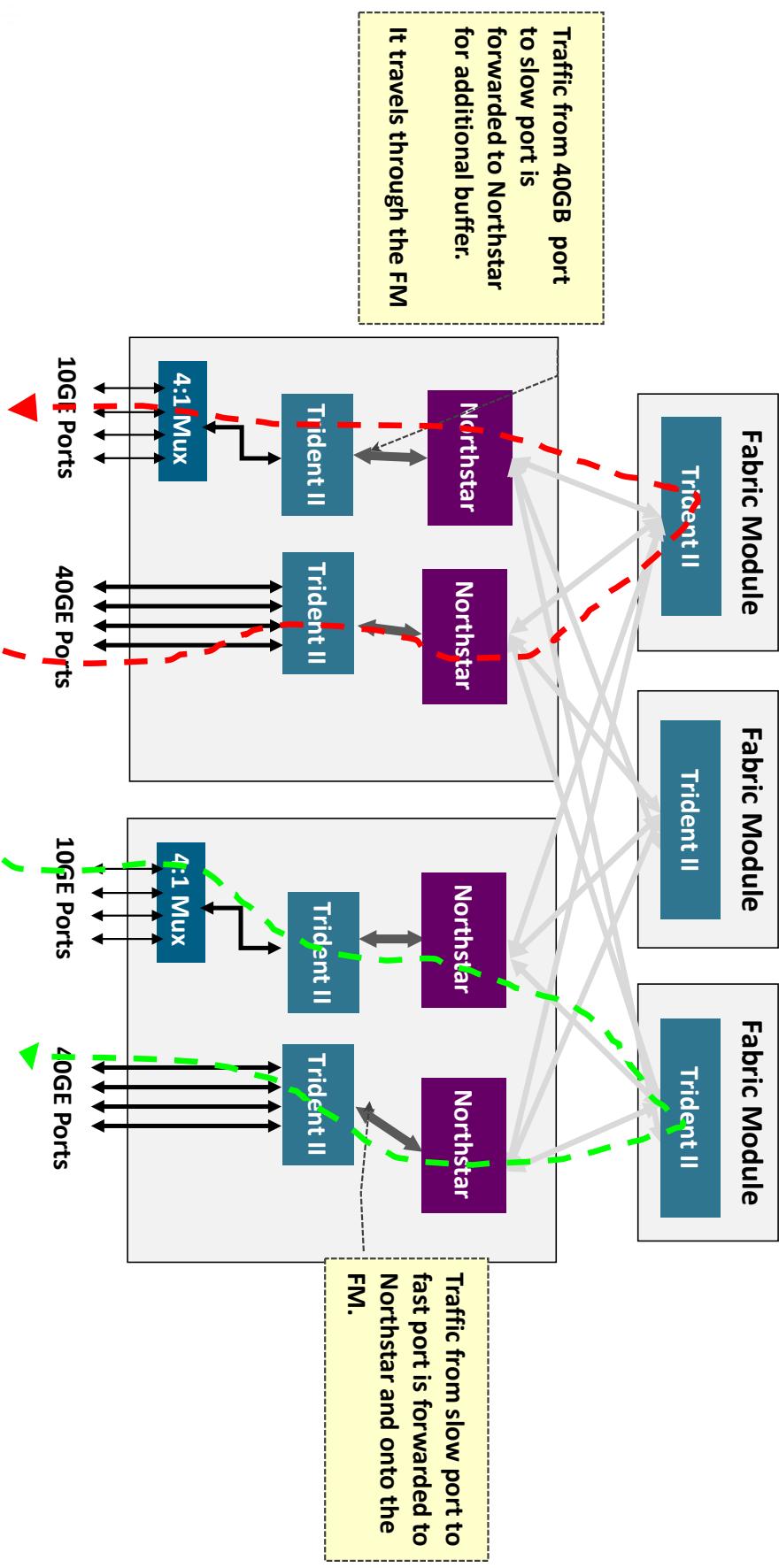
N9K-X9564PX Block Diagram (Blackcomb)



Local Switching 48p 1G/10GE + 4p QSFP Module



Fabric Switching 48p 1G/10GE + 4p QSFP Module



Nexus 9300 Series Switches Portfolio

N9K-C93120TX

New!

N9K-C9332PQ

N9K-C9372PX

N9K-C9372TX



N9K-C9396PX



N9K-C9396TX



N9K-C93128TX



Nexus 93128TX/ 93128PX

3 RU with 1 GEM module slot

1,280Gbps

96-port 1/10 G-T on Nexus 93128TX

96-port 1/10 SFP+ on Nexus 93128P

6 ports 40 Gb QSFP+ on N9K-M6PQ GEM

module

12 ports 40 Gb QSFP+ on N9K-M12PQ GEM

module

8 ports 40 Gb QSFP+ on N9K-M12PQ GEM

module

2 ports 100 Gb CFP2 on N9K-M4PC-CFP2 GEM

module

- 3 RU with 1 GEM module slot
- 1,280Gbps
- 96-port 1/10 G-T on Nexus 93128TX
- 96-port 1/10 SFP+ on Nexus 93128P
- 6 ports 40 Gb QSFP+ on N9K-M6PQ GEM module
- 12 ports 40 Gb QSFP+ on N9K-M12PQ GEM module
- 8 ports 40 Gb QSFP+ on N9K-M12PQ GEM module
- 2 ports 100 Gb CFP2 on N9K-M4PC-CFP2 GEM module

- Nexus® 9372PX/ 9372TX
 - 1 RU w/n GEM module slot
 - 720Gbps
 - 6-port 40 Gb QSFP+
 - 48-port 1/10 Gb SFP+ on Nexus 9372PX
 - 48-port 1/10 G-T on Nexus 9372TX
- Nexus 9332PQ
 - 1 RU w/n GEM module slot
 - 1,280Gbps
 - 32-port 40 Gb QSFP+
- Nexus 93120TX
 - 2 RU w/h GEM module slot
 - 1200Gbps
 - 6-port 40 Gb QSFP+
 - 96-port 1/10 G-T

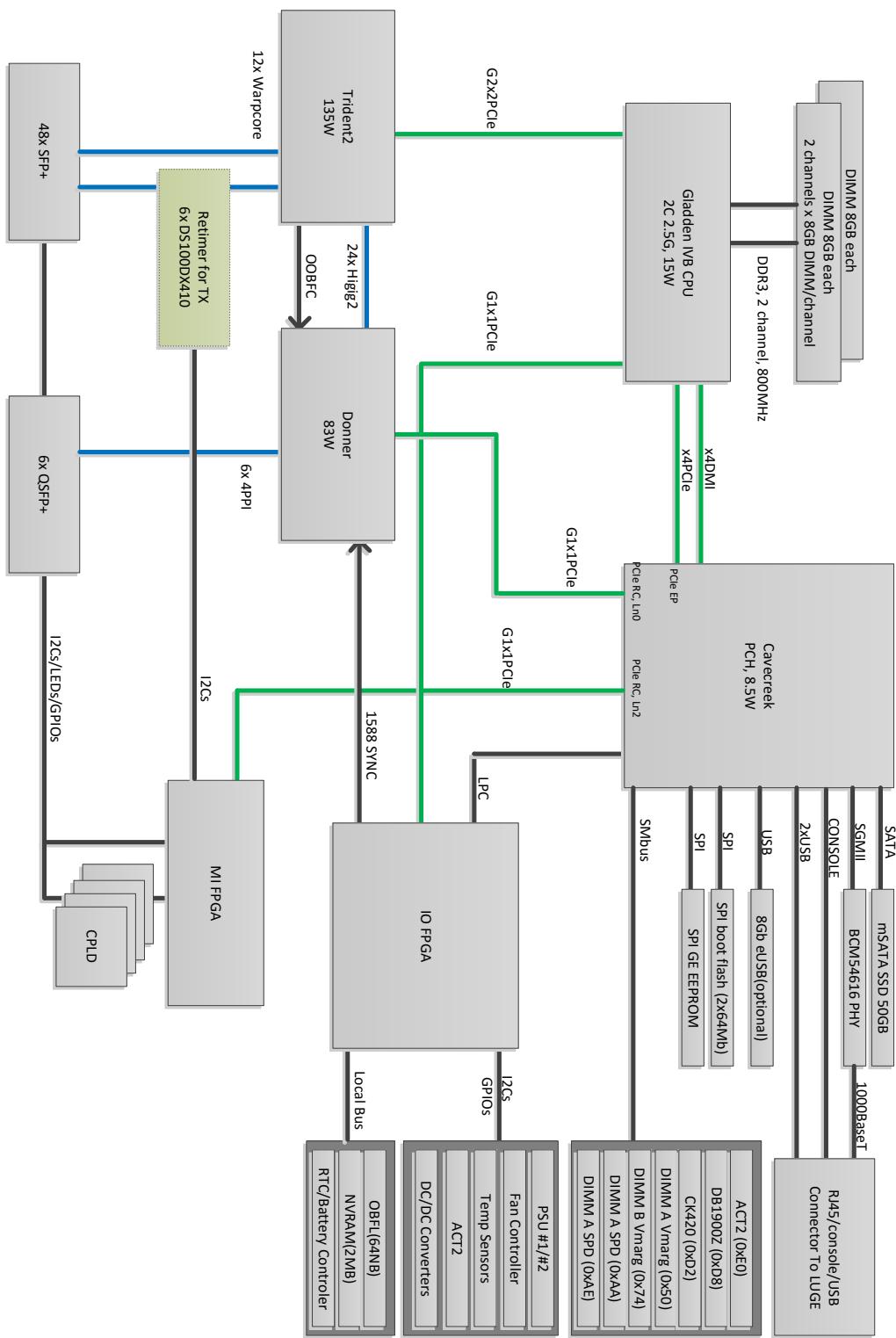
Nexus 9300 Series Switch Architecture

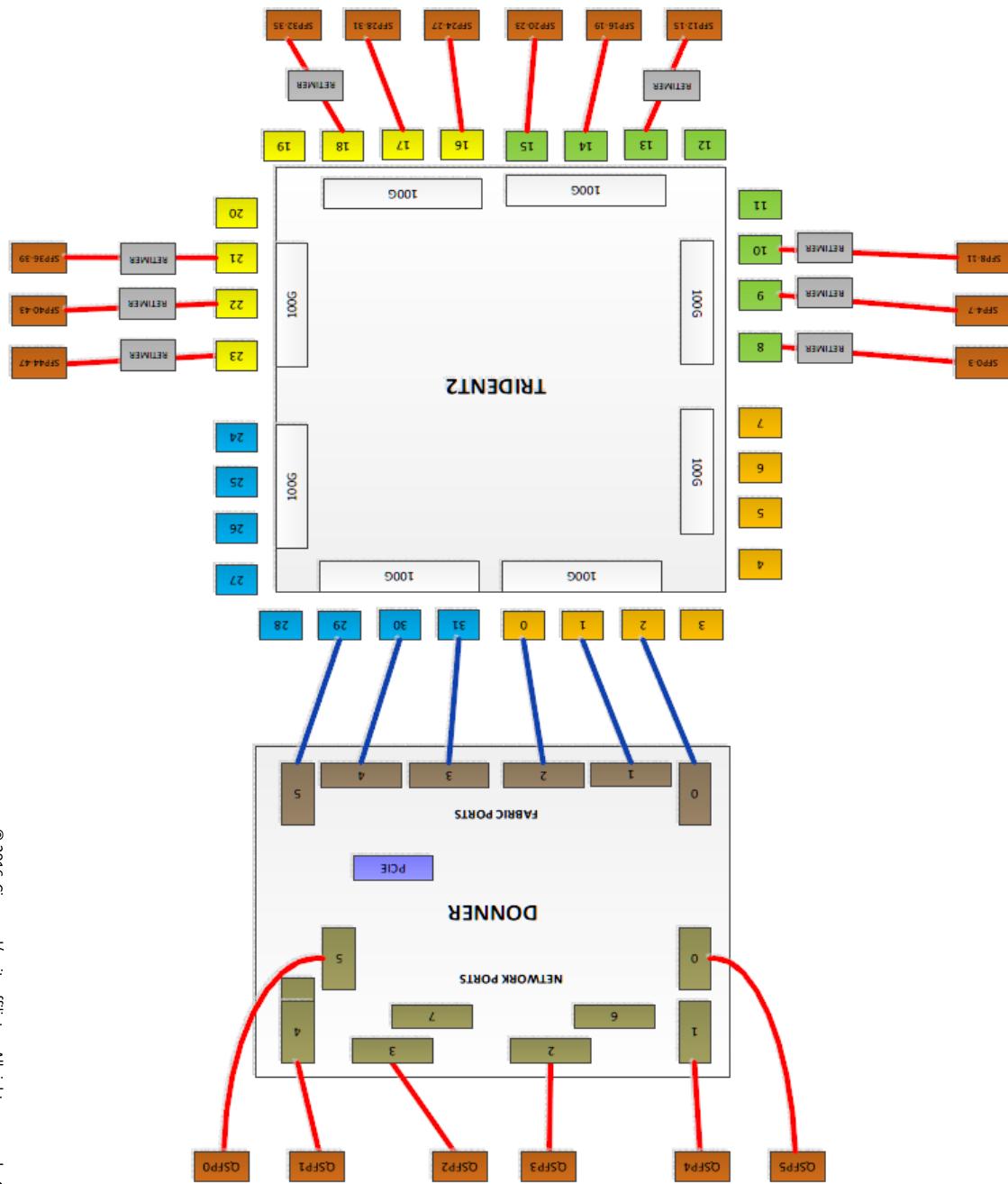
Nexus 9372PX/ Nexus 9372TX Block Diagram



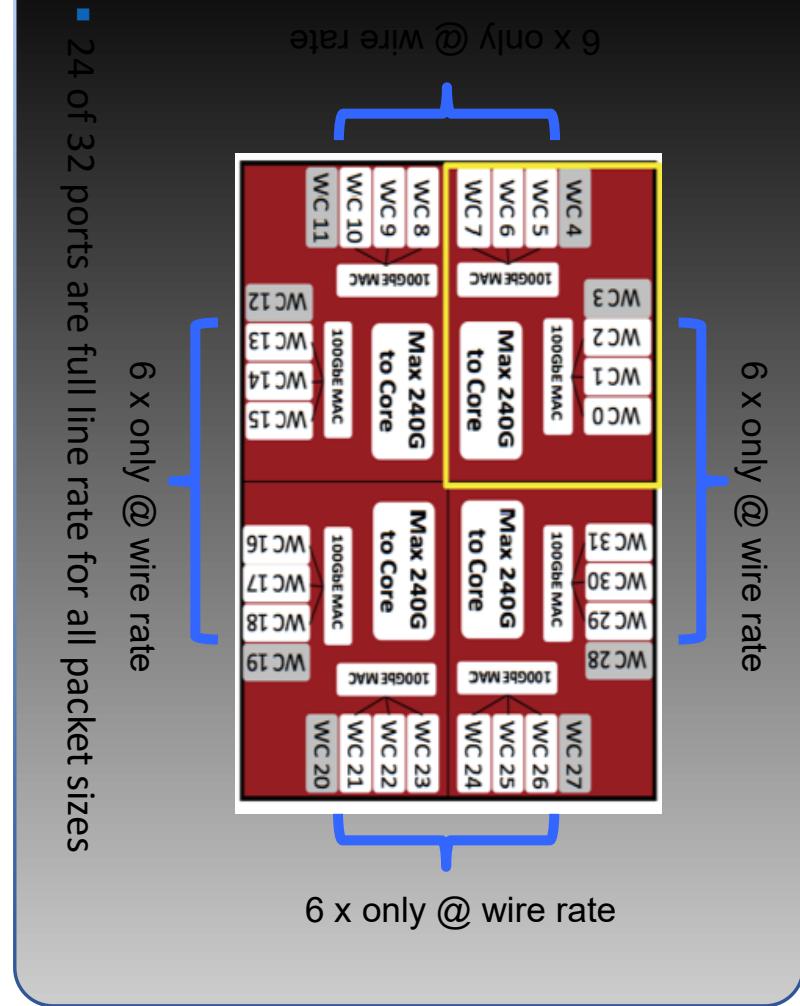
- The 6 40GE links between NFE and ALE-2 run at 42Gbps clock rate to accommodate the internal packet header
- Hardware is capable of VXLAN bridging and routing
- Hardware is capable of supporting both NX-OS and ACI modes
- Full line rate performance for all packet sizes

Oslo System Block Diagram (3/12/2014)





NFE Forwarding Capacity (BCM T2)



Two forwarding Modes on NFE

➤ Full Late-Rate Mode (FLM)

➤ Over-subscribed Mode (OSM)

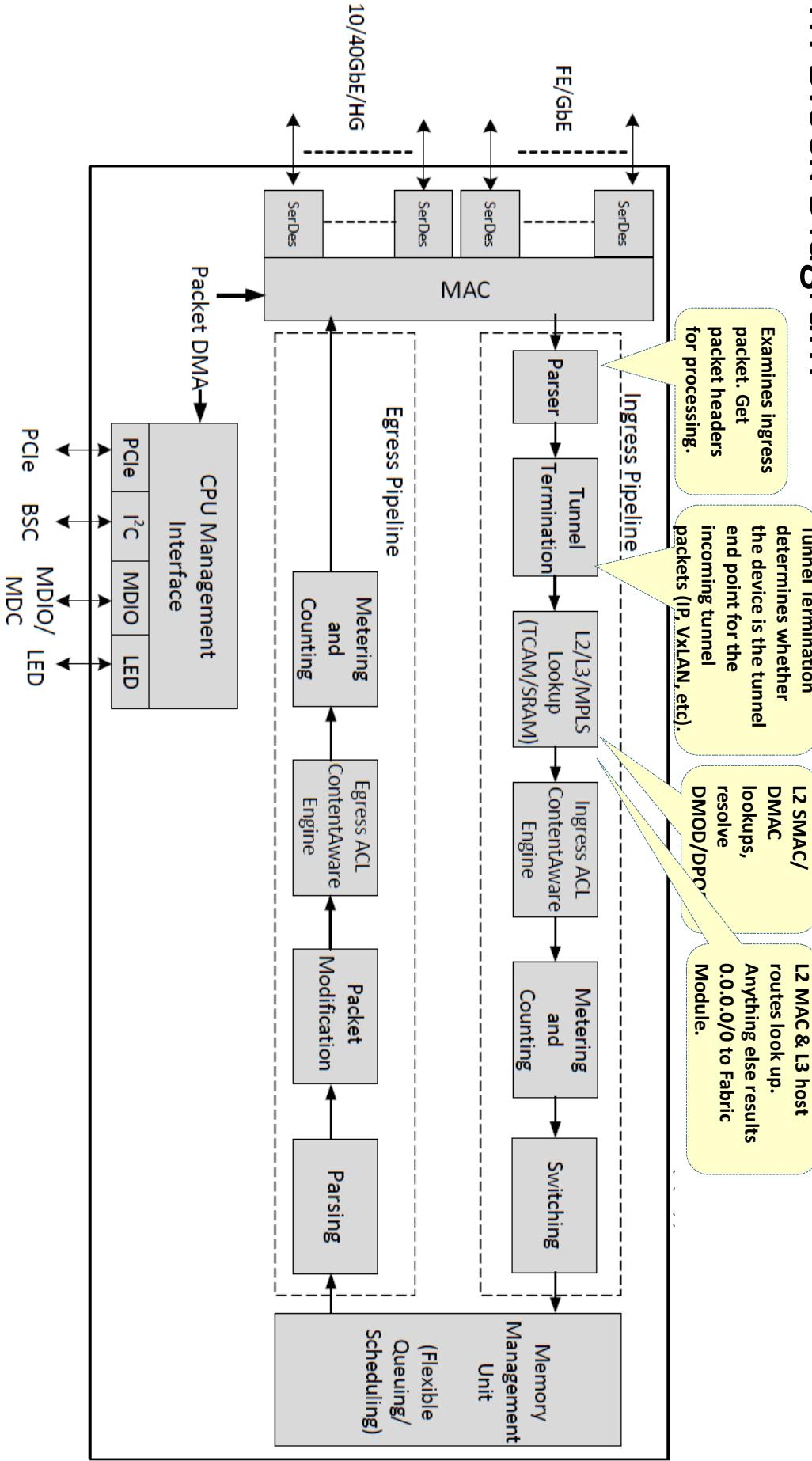
Full Late-Rate Mode (FLM):

- Only 24 40GE ports are used
- Every port is full line-rate for all packet sizes

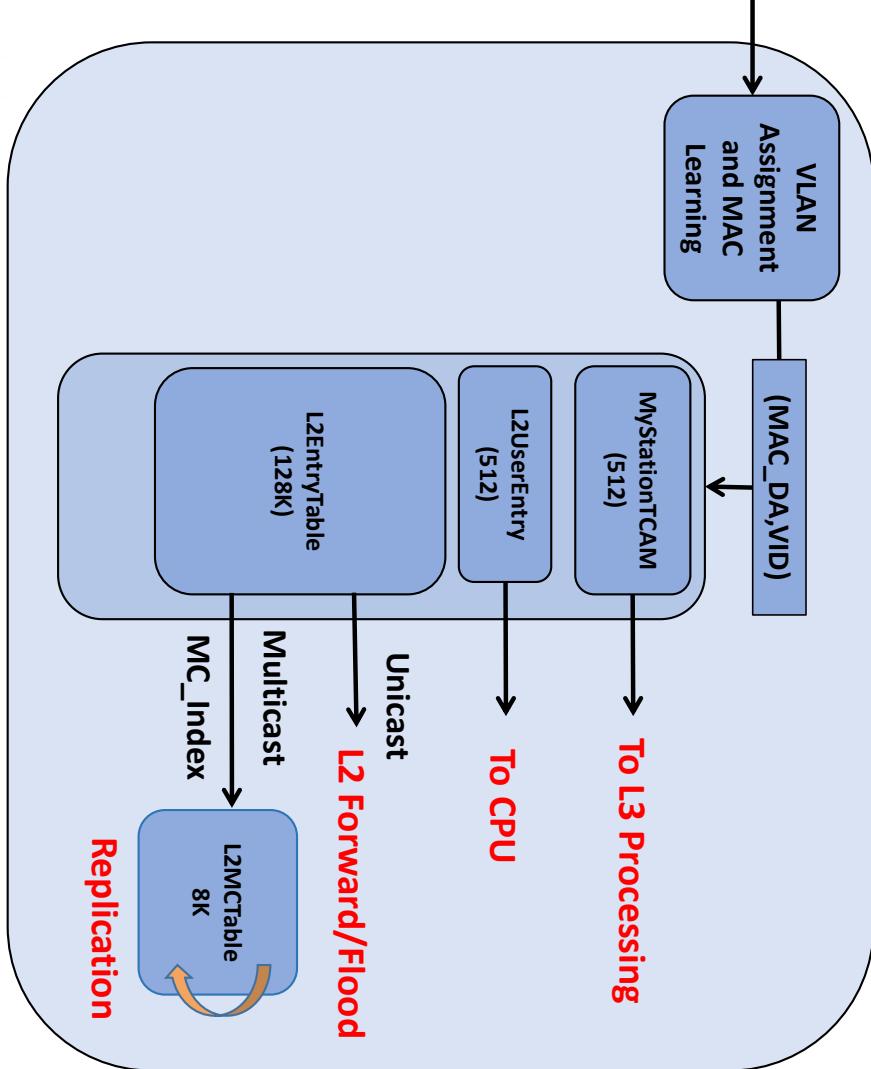
Over-subscribed Mode (OSM)

- All 32 40GE ports are used
- Every port is line-rate for packet sizes > 193 Bytes

BCM Block Diagram



BCM L2 Forwarding



- If packet is tagged, VLAN ID is picked from packet itself. For untagged packets we lookup Port table and get the VLAN information
- STP state is checked to see if this VLAN is allowed
- Packet is now at learning phase. If (MAC_SA, VLAN ID) does not exist in L2_Entry_Table, then HW assisted MAC learning happens
- The next lookup is based on (MAC_DA, VID).
- Match to MyStationTCAM(RMAC), Packet will be forwarded to L3 processing
- Match to L2Entry Table, FWD as L2 UNICAST
- In case of L2 Multicast packet when there is match in L2_Entry_Table, we index into the L2MCTable. The L2MCTable entry contains the port bitmap where the packet has to egress.

BCM L2 Forwarding - Tables

Table	Purpose	BCM SHELL
L2EntryTable	CAM table – 128K entries (16K buckets 8 entries each)	l2 show
L2UserEntry	Contains BPDU entries for CPU redirection. 512 entries	d chg l2_user_entry d chg my_station_tcam or l2 station show id=x
MyStationTCAM	Programmed with router and FHRP MAC. 512 entries	
L2MCTable	Used to select bitmap for L2 multicast replication. 8K entries	
SpanningTreeGroup	Contains STP state	stg stp 5
PortTable	Contains native VLAN information. 4096 MST 512 PVST	dump chg port 3 1 (E1/3)
VLAN Table	Maintains VLAN-Port membership	vlan show x
Port-Channel Table	16 ports per Port-Channel, 64 Port-Channels	trunk show

How to find all BCM tables if you don't know shell commands → use "list" command from bcm shell

```
N9372-42# bcm-shell mod 1
bcm-shell.0> list L2_user
Flags          Name           Entry/Copy  Description
--A-bC        L2_USER_ENTRY      512       Combined L2_ENTRY TCAM/Data RAM fo...
----b-        L2_USER_ENTRY_DATA_ONLY 512       Data SRAM for L2_USER_ENTRY TCAM.
--A-bC        L2_USER_ENTRY_ONLY   512       TCAM for guaranteed L2 entries and...
```

BCM L3 Forwarding - Tables

Table	Purpose	BCM SHELL
L3_Entry	Host /32 Route Table	l3 l3 table show l3 ip6host show
L3_DEFIP	LPM(Longest Prefix Match) Table	l3 defip show l3 ip6route show
L3_ECMPP	ECMP Table	l3 multipath show
L3_IPMC	Layer 3 multicast table	ipmc table show
EGR_L3_INTF	Egress Interface Information	l3 egress show

NFE Unified Forwarding Table

- NFE has a 16K traditional LPM TCAM table.
- Additionally NFE has the following Unified Forwarding Table for ALPM (Algorithm LPM) Mode

Dedicated L2 MAC Entries: 32k x 105 bits	4k x 420 bits	bank-0
	4k x 420 bits	bank-1

Mode	L2	L3 Hosts	LPM
0	288K	16K	0
1	224K	56K	0
2	160K	88K	0
3	96K	120K	0
4	32K	16K	128K

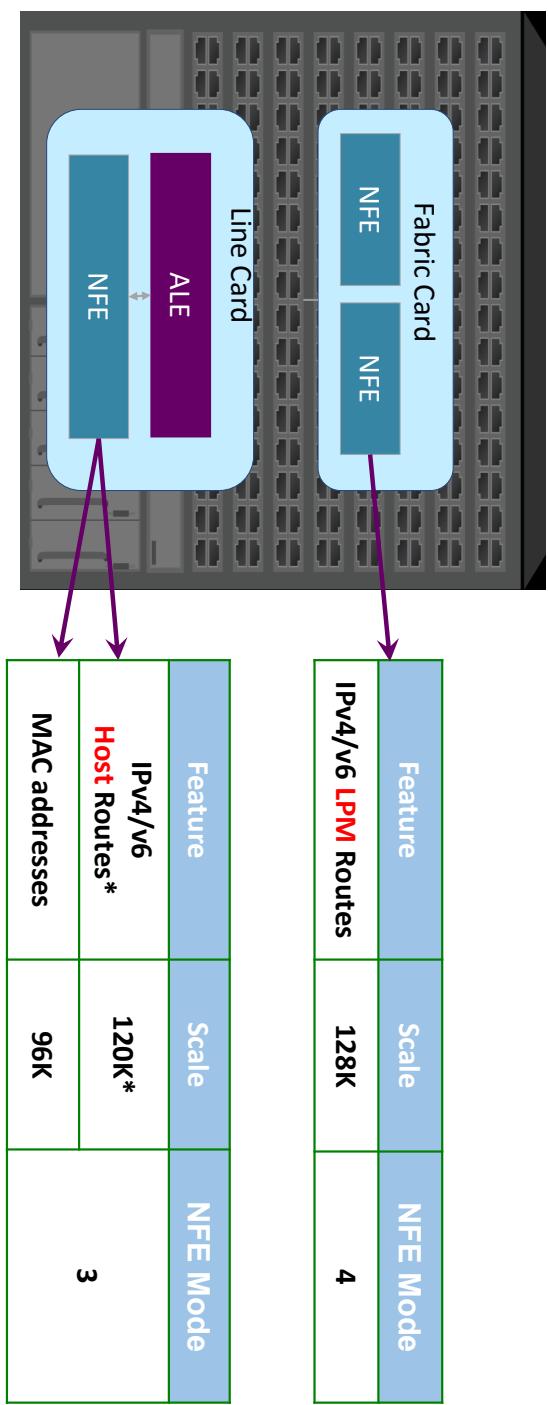
SUPPORTED COMBINATIONS

Dedicated L3 Host Entries: 16k x 105 bits	1k x 420 bits	bank-6
	1k x 420 bits	bank-7
	1k x 420 bits	bank-8

In default setting, N9500 line card NFE uses Mode 3, fabric module NFE uses Mode 4.

Nexus 9500 Forwarding Programming Mode

Hierarchical Routing Mode (Default)



From linecard, you can see only default route is installed in LPM

bcm-shell.0> 13 defip show

```
13 defip show
Unit 0, Total Number of DEFIP entries: 12288
#          VRF      Net addr           Next Hop Mac      INTF MODID PORT PRIO CLASS HIT VLAN
3072    Global   0.0.0.0/0        00:00:00:00:00:00  149149    0    0    0    0    y
```

Nexus 9500 Forwarding Programming Mode

	MAC Table		IPv4/IPv6 Host Table		IPv4/IPv6 LPM Route Table		Multicast Route Table	
	Location	NFE Mode	Location	NFE Mode	Location	NFE Mode	Location	NFE Mode
Hierarchical routing mode (default)	LC	3	LC	3	FM	4	LC+FM	3
Hierarchical 64-bit ALPM mode	LC	3	LC	3	FM	4	LC+FM	3
Hierarchical Max-host routing mode	LC	2	IPv4 on FM		IPv4 on FM	3	LC+FM	
Non-hierarchical routing mode	LC	3	LC	3	LC	3	LC	3
Non-hierarchical routing Max-L3 mode	LC	4	LC	4	LC	4	LC	4

Forwarding Programming Mode	Configuration Command
Default Hierarchical routing mode	Default
Hierarchical 64-bit ALPM mode	9508(config)# system routing hierarchical max-mode l3 64b-alpm
Hierarchical Max-host routing mode	9508(config)# system routing max-mode host
Non-hierarchical routing mode	9508(config)# system routing non-hierarchical
Non-hierarchical routing Max-L3 mode	9508(config)# system routing non-hierarchical max-mode l3

CLI to Show Forwarding Programming Mode

```
9508# sh system routing mode
Configured System Routing Mode: Non-Hierarchical (Default)
Applied System Routing Mode: Hierarchical (Default)
Configured SVI post-routed unknown-unicast hardware flood mode: enabled
US-DUR-LC01-9508#
```

```
9508# show forwarding route summary module 1
```

Module Type	:	Line-Card
Module Mode	:	Mode-3
Module Route	:	Host only
(IPv4+IPv6)	(1)	

IPv4 routes for table default/base

'***' denotes routes NOT programmed in hardware due to hierarchical routing

Cumulative route updates: 1005038
Cumulative route inserts: 1005005
Cumulative route deletes: 143
Total number of routes: 24
Total number of paths : 25

Number of routes per mask-length:
/32 : 24

```
9508# show forwarding route summary module 26
```

Module Type	:	Fabric-Module
Module Mode	:	ALPM (Mode-4)
Module Route	:	LPM only (IPv4+IPv6)
(2)		

IPv4 routes for table default/base

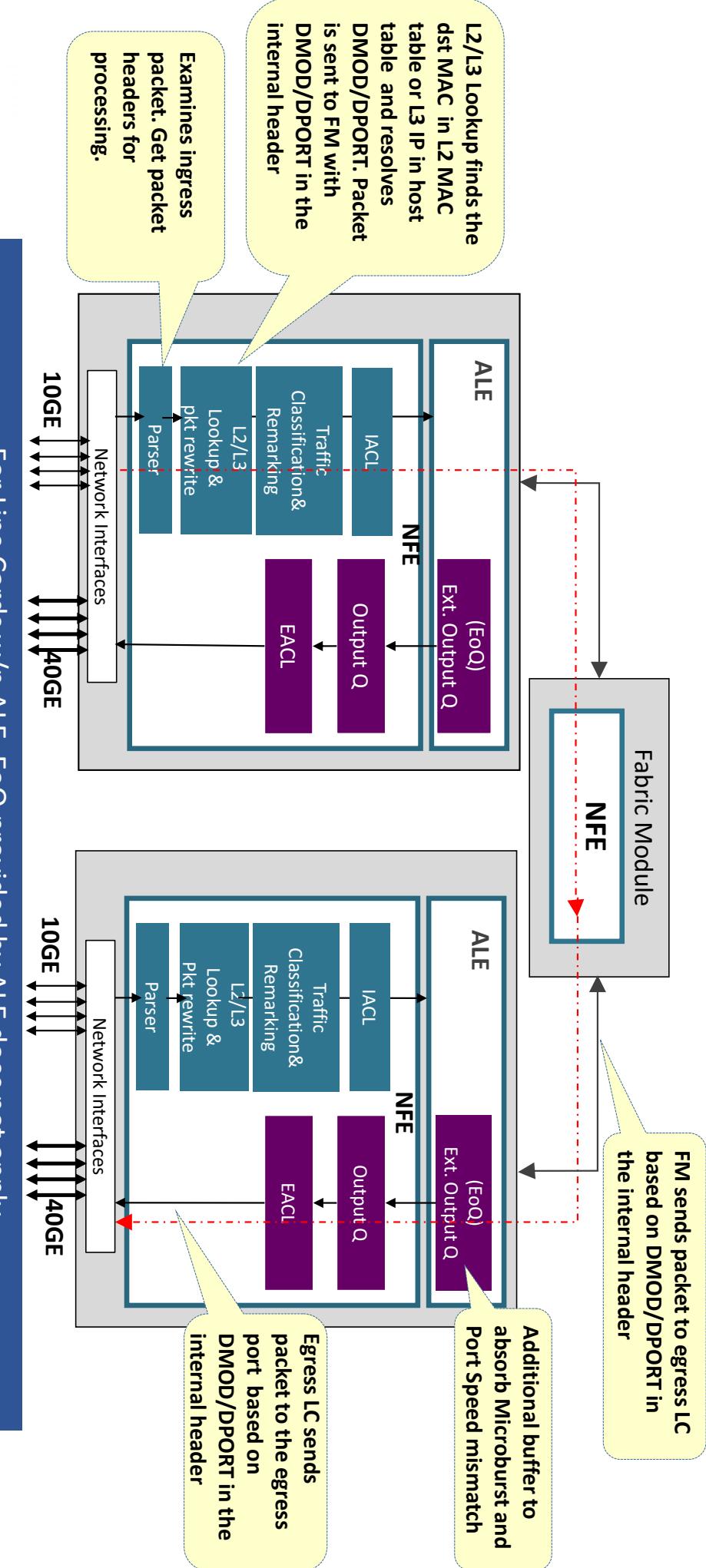
'***' denotes routes NOT programmed in hardware due to hierarchical routing

Cumulative route updates: 1005043
Cumulative route inserts: 1004930
Cumulative route deletes: 54
Total number of routes: 8
Total number of paths : 8

Number of routes per mask-length:
/8 : 1 /30 : 5

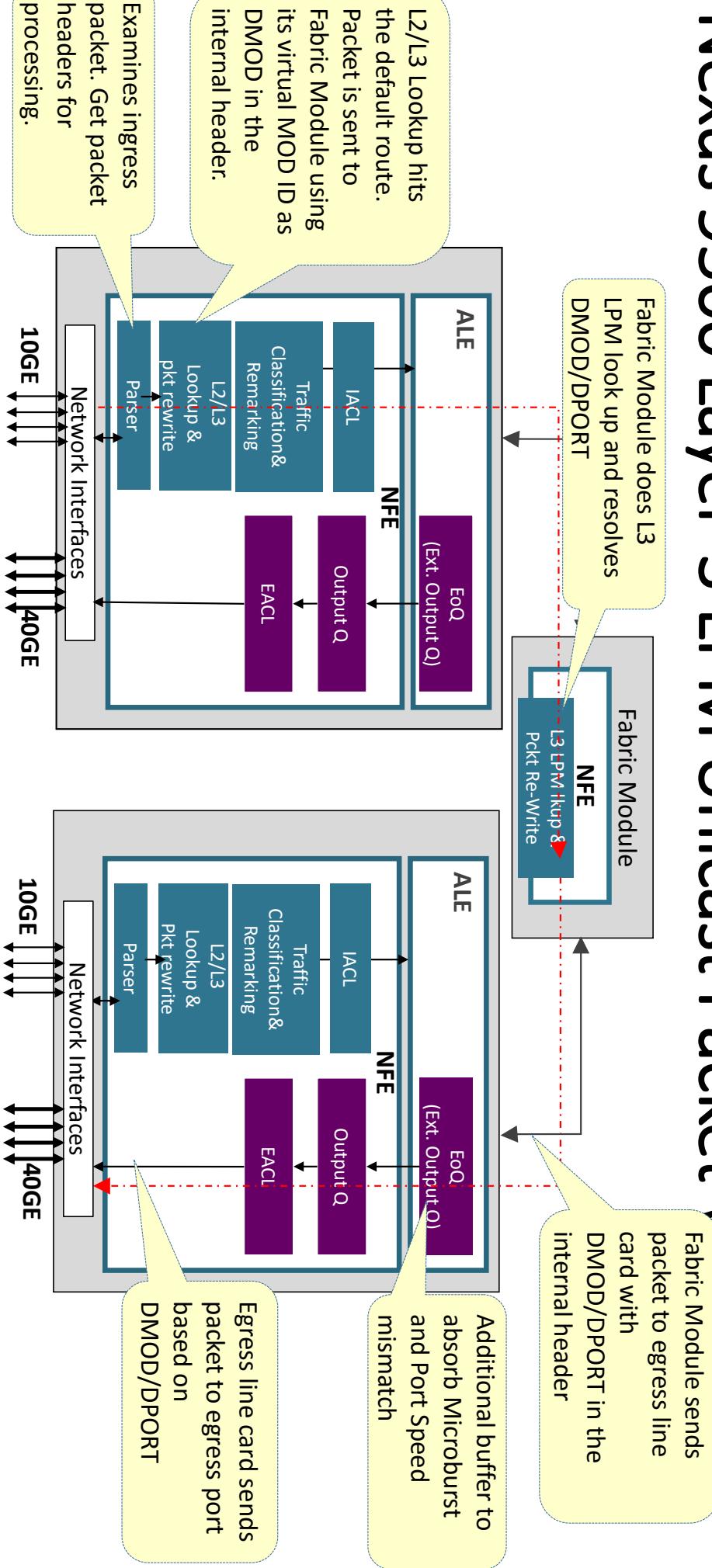
US-DUR-LC01-9508#

Nexus 9500 Unicast Packet Walk

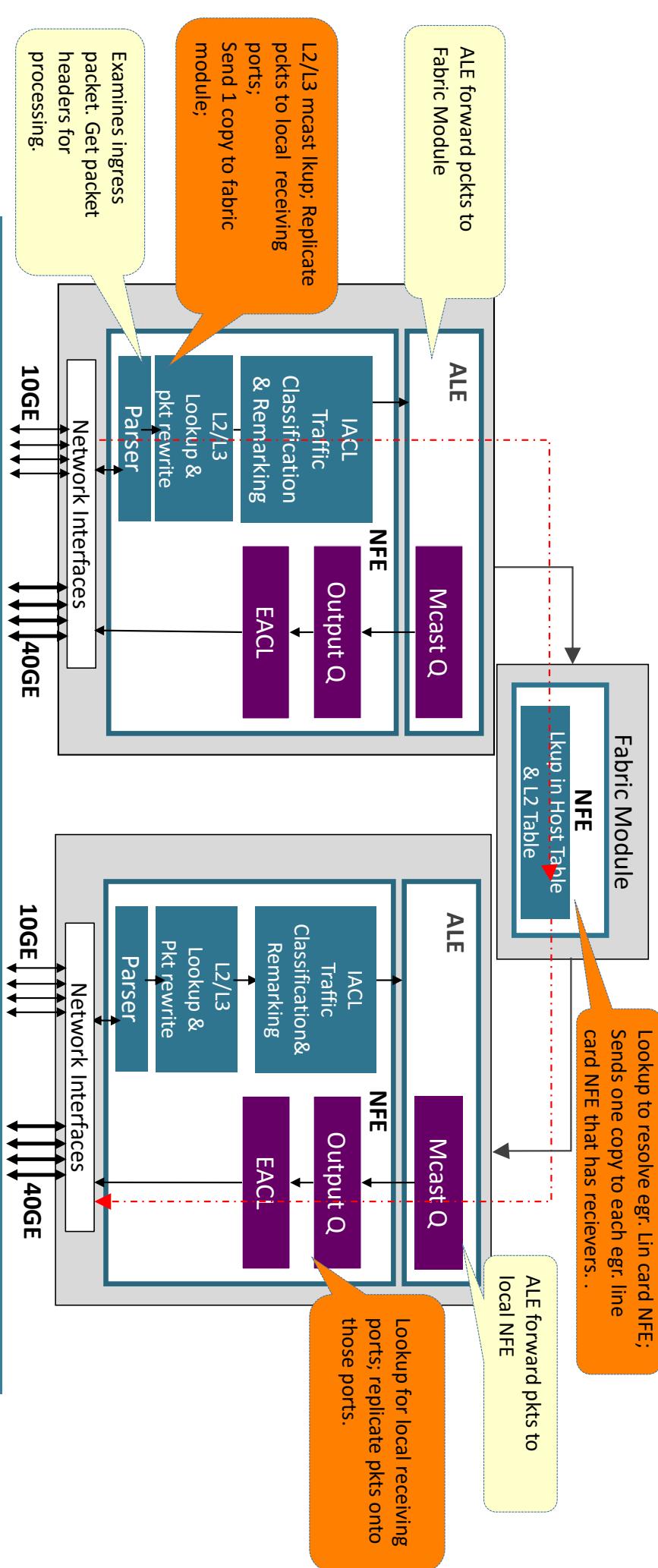


For Line Cards w/o ALE, EoQ provided by ALE does not apply.

Nexus 9500 Layer-3 LPM Unicast Packet Walk

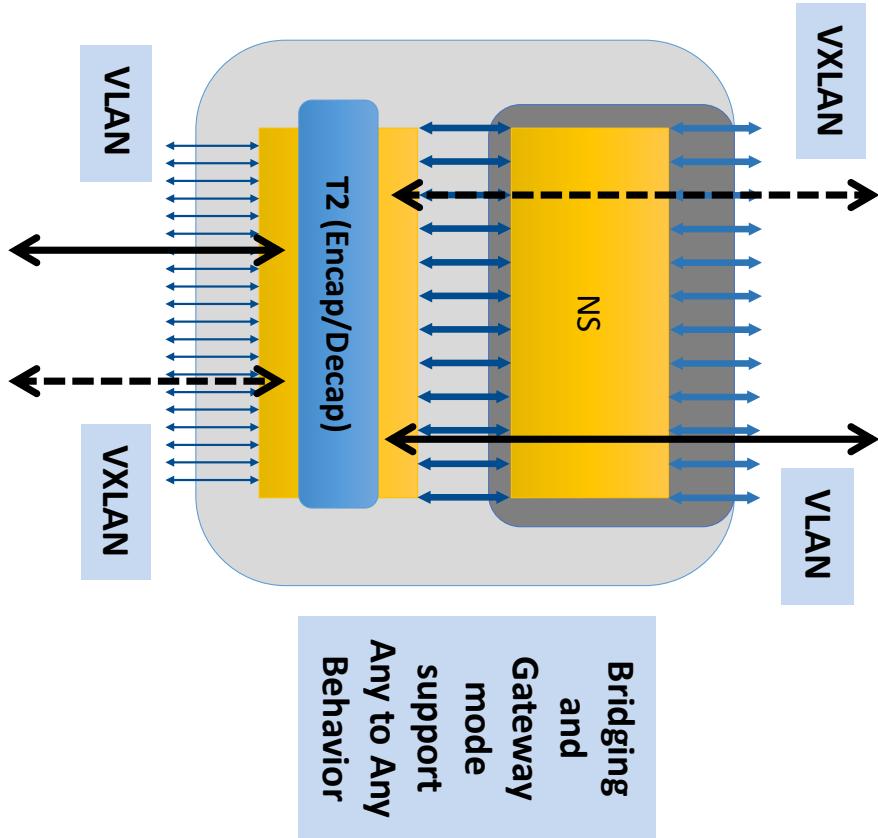


Nexus 9500 L2/L3 Multicast Packet Walk

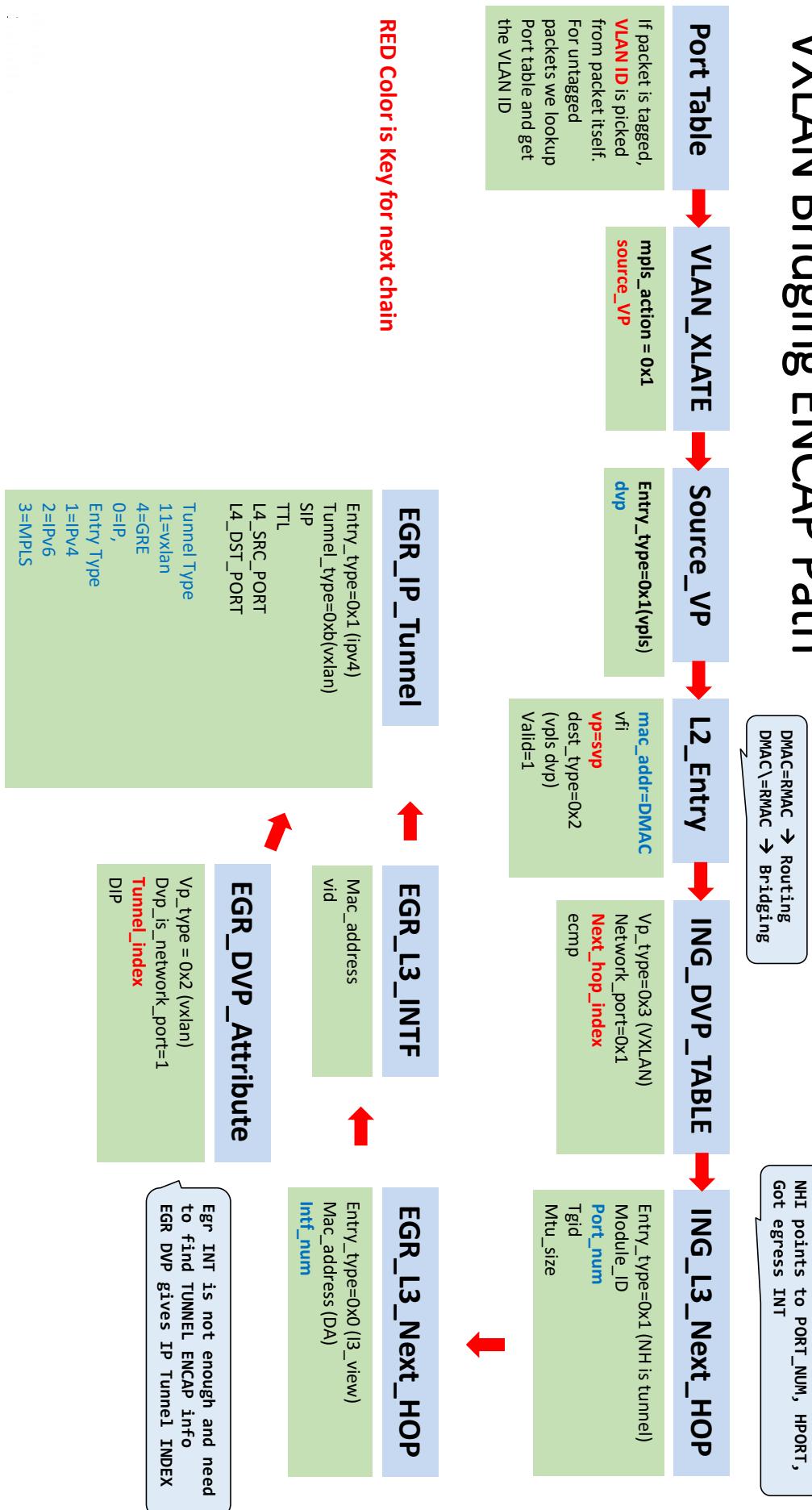


VXLAN Bridging packet path

- VXLAN Encapsulation and Decapsulation occur on T2
 - VLAN -> VXLAN (Access to Network)
 - VXLAN – VLAN (Network to Access)
 - VXLAN – VXLAN (Transit – Native IP)
- Bridging and Gateway are independent of the port type (1/10/40G ports)



VXLAN Bridging ENCAP Path



BCM VXLAN Forwarding - Tables

Table	Purpose	Important Fields	BCM SHELL
VLAN_Xlate	BCAM for VLAN trans of tagged pkts.	Mpls_Action=0x1 Source_VP	d chg vlan_xlate
MPLS_Entry	VXLAN/MPLS/TRILL label lookup table	VXLAN_VN_ID:VN_ID VXLAN_VN_ID:VFI VXLAN_VN_ID:SIP VXLAN_SIP:SIP	d chg mpls_entry
ING_DVP_TABLE	Used to generate a NHI from the DVP value and other DVP properties.	NEXT_HOP_INDEX ECMP=1	d chg ing_dvp_table
ING_L3_NEXT_HOP	Reduced version of L3_NEXT_HOP Used to provide just mod and port/TGID	Entry_Type=0x1(NH=Tunnel) PORT_NUM	d chg ing_dvp_l3_next_hop
EGR_L3_NEXT_HOP			d chg egr_l3_next-Hop
EGR_DVP_ATTRIBUTE	Egress Destination Virtual Port Table From Next Hop Table	VXLAN:TUNNEL_INDEX VXLAN:DIP	d chg egr_dvp_attribute
EGR_IP_TUNNEL	Make new Tunnel Header	Tunnel Type SIP L4_DEST_PORT Entry Type	d chg egr_ip_tunnel

VXLAN EVPN ENCAP

Troubleshooting Steps for Bridging HW ENCAP

Host 10.160.1.33 in vlan 1601 on LEAF42 can't ping host 10.160.1.13 in vlan 1602 on vpc LEAF 43/44.
Ingress LEAF is LEAF42 and Egress LEAF is VPC LEAF43/44.

1. Assume control plane is good and RIB/FIB has correct entries

```
N3064-33# sh ip arp vrf host1601
10.160.1.13  00:09:10  001a.6c70.3697  v1an1601
```



LEAF42 installed remote MAC properly

N9372-42# sh mac add vl 1601 | i 3697

* 1601 001a.6c70.3697 dynamic 0

N9372-42# sh nve peers

Interface	Peer-IP	State	LearnType	Uptime	Router-Mac
nve1	10.0.0.100	Up	CP	1d23h	88f0.3187.8c4b

N9372-42# show forwarding nve 13 peers
NVE cleanup transaction-id 0

tunnel_id	Peer_id	Peer_address	Interface	rmac	origin state	del count		
-	0xa000064	2	10.0.0.100	nve1	88f0.3187.8c4b	NVE	merge-done no	13

2. Ingress INT is E1/17, Trunk Port and ingress traffic is tagged with 1601, Ensure E1/17 is FWD state

```
N9372-42# bcm-shell1 mo 1 'stg show'
```

STG 4 : contains 591 VLANs (1,10,1113-1199,1500-1999,3001,3100)

Block: xe0-xe1,xe3-xe5-xe9,xe11-xe15,xe17-xe19,xe27-xe28,xe30-xe47

Forward: xe2,xe4,xe10,**xe16**,xe20-xe26,xe29,hg

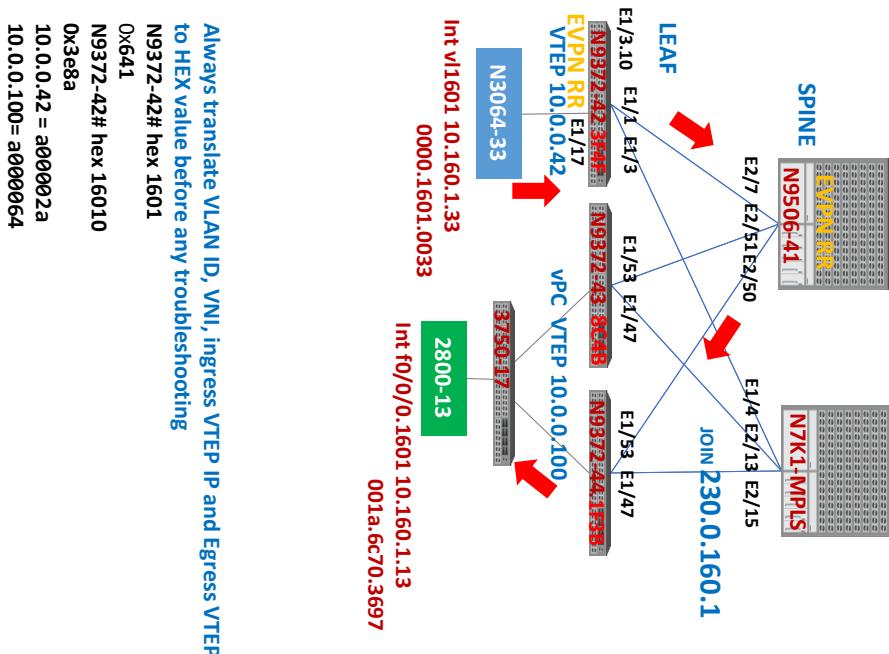
3. Check VLAN_XLATE for VLAN 1601 = 0x641

N9372-42# bcm-shell1 mo 1 'd chg vlan_xlate' | i **0x641**

VLAN_XLATE.ipipe[5760]:

```
<XLATE:VLAN_ID=0x641,XLATE:VLAN_ACTION_VALID=1,XLATE:TGID=0x97,XLATE:TAG_ACTION_PROFILE_PTR=2,XLATE:SV
P_VALID=1,XLATE:SOURCE_VP=0x810,XLATE:SOURCE_FIELD=0x97,XLATE:PORT_NUM=0x17,XLATE:OVID=0x641,XLATE:OTA
G=0x641,XLATE:OLD_VLAN_ID=0x641,XLATE:NEW_VLAN_ID=0x641,XLATE:NEXT_OVID=0x641,XLATE:MPLS_ACTION=1,XLATE
:MODULE_ID=1,XLATE:KEY=0x000000032080ba4,XLATE:ITAG=0x641,XLATE:INCOMING_VIDS=0x641,XLATE:HASH LSB=0x
641,XLATE:GLP=0x97,XLATE:DATA=0x000004080000000020000001905,VLAN_ID=0x641,VLAN_ACTION_VALID=1,VALID=1
,TGID=0x97,TAG_ACTION_PROFILE_PTR=2,SVP_VALID=1,SOURCE_VP=0x810,SOURCE_SOURCE_TYPE=1,SOURCE_FIELD=0x97,PORT_N
UM=0x17,OVID=0x641,OTAG=0x641,OLD_VLAN_ID=0x641,NEW_OVID=0x641,MPLS_ACTION=1,MODULE_ID=1,KEY_TYPE=4,KEY=0x000000032080ba4,ITAG=0x641,INCOMING_VIDS=0x641,HASH LSB=0x641,GLP=0x97,DATA=0x0
00004084000000020000001905>
```

VXLAN EVPN



4. Check Source_VP table and L2_Entry table for DMAC

```
N9372-42# bcm-shell module 1 'd chg source_vp 2064'
0x810==>2064
N9372-42# ipipe0[2064]:
<VFI=0x641,TPID_SOURCE=2,TPID_ENABLE=0xf,SD_TAG_MODE=1,IPV6L3_ENABLE=1,IPV4L3_ENABLE=1,FLEX_CTR_POOL_N
NUMBER=2,FLEX_CTR_OFFSET_MODE=1,FLEX_CTR_BASE_COUNTER_IDX=0x70,EXP_PVLAN_VID=0x641,ENTRY_TYPE=1,DVP=0x6
41,DISABLE_VLAN_CHECKS=1,DEFAULT_VLAN_TAG=0x190,DEFAULT_VID=0x641,CML_FLAGS_NEW=8,CML_FLAGS_MOVE=8,>
```

```
N9372-42# bcm-shell mo 1 'd chg l2_entry' | i 3697
```

```
L2_ENTRY.ipipe0[91780]:
<VPG_TYPE=1,VPG_1=0xd,VPG=0x2db8,VLAN_ID=0x641,VIF:L2MC_PTR=0x2db8,VFI=0x641,VALIDD=1,T_1=1,TGID_1=0xd,
TGD=0x1b8,RESERVED_69=68=1,PORT_NUM_1=0xd,PORT_NUM=0x38,OVID=0x641,MODULE_ID=0x5b,MAC_ADDR=0x001a6c70
3697,LOCAL_SA=1,L2MC_PTR=0x2db8,L2:VPG_TYPE=1,L2:VPG=0x2db8,L2:VLAN_ID=0x641,L2:VFI=0x641,L2:TGID=0x1b
8,L2:RESERVED_69=68=1,L2:PORT_NUM=0x38,L2:MODULE_ID=0x5b,L2:MAC_ADDR=0x001a6c703697,L2:L3MC_PTR=0xdb8,
L2:L2MC_PTR=0x2db8,L2:KEY=0x00069b1c0da5c6413,L2:HASH_LSB=0x697,L2:DEST_TYPE=2,L2:DESTINATION=0x2db8
,L2:DATA=0x00000012db8,L2:ASSOCIATED_DATA=0x00000012db8,KEY_TYPE=3,KEY=0x00069b1c0da5c6413,VID=0xa5c,H
TSA=1,DEST_TYPE_1=1,DEST_TYPE=2,DESTINATION_1=0xd,DESTINATION=0x2db8,DATA=0x00000012db8,ASSOCIATED_DAT
A=0x00000012db8>
```

Int v1601 10.160.1.33
0000.1601.0033
Int f0/0/0.1601.10.160.1.13
001a.6c70.3697

5. Verify SVP for NVE Peer from MPLS_Entry

```
N9372-42# bcm-shell module 1 'd chg mpls_entry' | grep 0x2db8
MPLS_ENTRY.ipipe0[4336]:VXLAN_SIP:SVP=0x2db8,VXLAN_SIP:SIP=0x00000064,VXLAN_SIP:KEY=0x00a000648,VXLAN_
SIP:HASH_LSB=0x64,VXLAN_SIP:DATA=0x2db8,VALID=1,KEY_TYPE=8,EVEN_PARITY=1,>
```

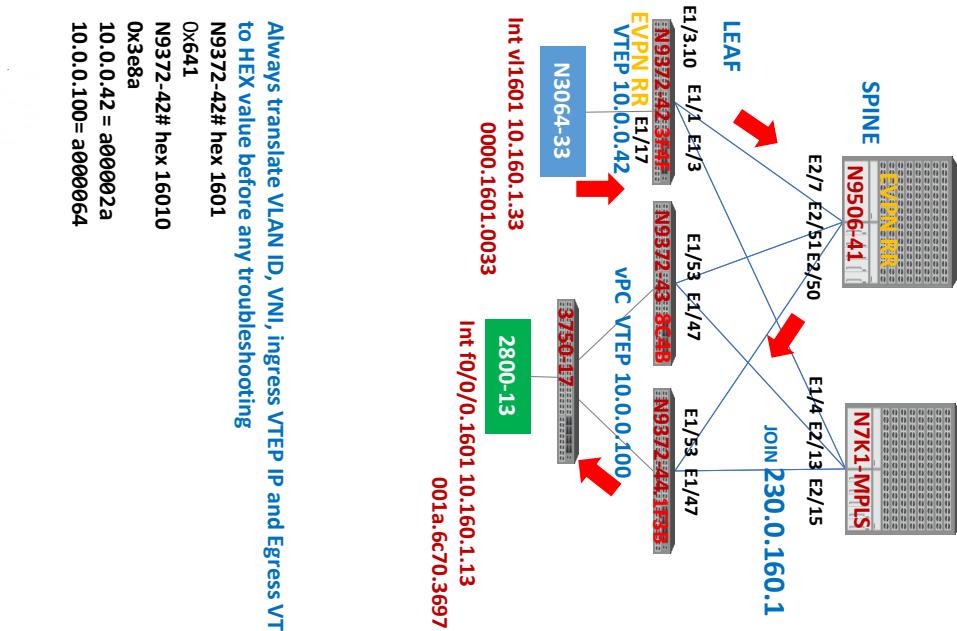
6. Verify MAC in HW pointing to correct Virtual Port

```
N9372-42# sh hardware mac address-table 1 vlan 1601 | i 3697 |MAC
FE | PI | VLAN | MAC | Trunk | TGID | Mod | Port | Virtual | Static | Hit | Hit | CPU | Pending
0 1 30273 001a.6c70.3697 0 0 0 2db9 0 1 1 0 0
```

HW VLAN ID = 28672+VXLAN ID = 28672+1601 = 30273, so we look for vlan 30273

Verify MAC in HW pointing to Virtual Port in BCM shell
N9372-42# bcm-shell mod 1 '12 show' | grep 36:97
mac=00:1a:6c:70:36:97 v1an=30273 GPORT=0x80002db9 port=0x80002db9(vxlan) Hit

VXLAN EVPN – HW



7. Using the **SVP** entry, get the **NEXT_HOP_INDEX** for the peer.

If **ECMP=1**, it means the route towards the peer is an ECMP route.

N9372-42# bcm-shell module 1 'd chg ing_dvp_table 0x2db9'.

```
<VLAN_ID=0xffff, Tgid=0x89, PORT_NUM=9, MTU_SIZE=0x3fff, MODULE_ID=1, L3_OIF=0x1fff, ENTRY_TYPE=2, ENTRY_INFO_>
<VP_TYPE=3, NEXT_HOP_INDEX=0xa6, NETWORK_PORT=1, ECMP_PTR=0xa6, DVP_GROUP_PTR=0xa6,>
```

8. Using **NH index**, get egress port num

```
N9372-42# bcm-shell module 1 "dump chg ing_l3_next_hop 0xa6 "
ING_L3_NEXT_HOP.ipipe[166]:
<VLAN_ID=0xffff, Tgid=0x89, PORT_NUM=9, MTU_SIZE=0x3fff, MODULE_ID=1, L3_OIF=0x1fff, ENTRY_TYPE=2, ENTRY_INFO_>
UPPER=3, DVP_RES_INFO=0x7f,>
```

ECMP Case

```
N9372-42# bcm-shell module 1 'd chg ing_dvp_table 0x2db9',
ING_DVP_TABLE.ipipe[11705]:<VP_TYPE=3,NEXT_HOP_INDEX=0x103,NETWORK_PORT=1,ECMP_PTR=0x103,ECMP=1,DVP_G
ROUP_PTR=0x103,>
```

0x103=259

N9372-42# bcm-shell mo 1 '13 multipath show' | b 259

Multipath Egress Object 200259

Interfaces: 100110 100180

```
N9372-42# bcm-shell module 1 '13 egress show' | i 100110
100110 7c:0e:ce:48:ef:bf 4095 4112 7 1 -1 no no
100180 00:26:51:c9:50:42 4095 4114 9 1 -1 no no
```

Port_Num is Hardware port no.

```
N9372-42# bcm-shell module 1 "phy info" | i port|9
          port  id0  id1  addr  iaddr           name      timeout
xe0( 7 )  600d  8770  ad   ad           TSC-A2/08/0  250000
xe2( 9 )  600d  8770  af   af           TSC-A2/08/2  250000
```

Port_Num is Hardware port no.

N9372-42# sh interface hardware-mappings | i 9|port

HPort - Hardware Port Number or Hardware Trunk Id:

FPort - Fabric facing port number. 255 means N/A

NPort - Front panel port number

VPort - Virtual Port Number. -1 means N/A

Name	Index	Siad Unit	HPort	FPort	NPort	VPort
Eth1/1	1a000000 1	0	7	255	0	-1
Eth1/3	1a000400 1	0	9	255	2	-1

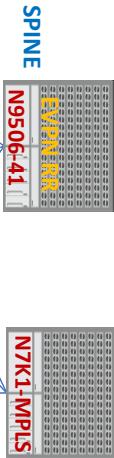
VXLAN EVPN – HW

Tunnel Verification

8. Using the SVP entry, 0x2db9, get egr_dvp_attribute to verify TUNNEL ID which has the SIP programmed.

Also, check the DIP points to peer ip.

```
N9372-42# bcm-shell module 1 'd chg egr_dvp_attribute 0x2db9'
EGR_DVP_ATTRIBUTE:epipe0[11:705]:
<VXLAN:TUNNEL_INDEX=1,VXLAN:DTP=0xa000064,VXLAN:DELETE_VNTAG=1,VP_TYPE=2,>
```



Using the TUNNEL_ID verify the SIP

```
N9372-42# bcm-shell module 1 'd chg egr_ip_tunnel 1'
```

EGR_IP_TUNNEL.epipe0[1]:

```
<TUNNEL_TYPE=0xb, TTL=0xff, SIP=0xa00002a, L4_DEST_PORT=0x12b5, ENTRY_TYPE=1, DSSCP_SEL=1,>
Tunnel Type 11=vlan 4=GRE 0=IP, SIP= 10.0.0.42,dst port=4789, Entry Type 1=IPv4 2=IPv6 3=MPLS
```

Verify the vlan associated with the vni. Here you can see the BD[VLAN] associated with the VNI.

Interface	VNI	Multicast-group	State	Mode	Type [BD/VRF]	Flags
nve1	16010	230.0.160.1	Up	CP	L2 [1601]	SA

Verify the VFI is created for Vlan 1601. 1601=0x641, 16010=0x3e8a

```
N9372-42# bcm-shell module 1 'd chg mpls_entry' | grep 0x641
MP_LS_ENTRY.ipipe0[8312]:
```

```
<VXLAN_VN_ID:VN_ID=0x3e8a, VXLAN_VN_ID:VFI=0x641, VXLAN_VN_ID:KEY=0x0000000000003e8a9, VXLAN_VN_ID:HASH_LB
=0x68a, VXLAN_VN_ID:DATA=0x641, VALID=1, KEY_TYPE=9,>
```

From VFI, we can find Unknown Unicast/Unknown Multicast/Broadcast flooding ports

```
N9372-42# bcm-shell module 1 'd chg vfi 1601'
VFI.ipipe0[1601]:
```

```
<VP_1=0xfe8, VP_0=0x1fd0, UUC_INDEX=0x1fd0, UMC_INDEX=0x1fd0, PROTOCOL_PKT_INDEX=2, FLEX_CTR_POOL_NUMBER=4,
FLEX_CTR_BASE_COUNTER_IDX=0x74, BC_INDEX=0x1fd0>
```

We can see BUM traffic will be forwarded to correct ports.

```
N9372-42# bcm-shell module 1 "mc show" | b 1fd0
Group 0xc001fd0 (VXLAN)
  port xe0,  encap id 400192   E1/1
  port xe2,  encap id 400197   E1/3
  port xe10, encap id 400068   E1/11
  port xe16, encap id 400027   E1/17
  port xe24, encap id 400071   E1/25
```

Interface	Role	Sts	Cost	Prio	Nbr	Type
Eth1/11	Desg	FWD	20000	128.11	P2p	
Eth1/17	Desg	FWD	20000	128.17	P2p	
Eth1/25	Desg	FWD	20000	128.25	P2p	

VXLAN EVPN – HW

How to find which one of vPC pair received ICMP PING traffic

There are two ECMP paths from LEAF42 to SPINES and two ECMP form SPINES to LEAF43/44

N9372-42# sh ip route 10.0.0.100

10.0.0.100/32, ubest/mbest: 2/0

*via 10.3.42.3, Eth1/3, [110/3], 3d02h, ospf-1, intra

*via 10.41.42.41, Eth1/1, [110/3], 3d02h, ospf-1, intra

N9504-41# sh ip route 10.0.0.100

10.0.0.100/32, ubest/mbest: 2/0

*via 10.41.43.43, Eth2/49, [110/2], 1w0d, ospf-1, intra

*via 10.41.44.44, Eth2/51, [110/2], 1w0d, ospf-1, intra

Easiest way is to utilize NVE interface counter and sending icmp ping

N9372TX-43# clear count

N9372TX-43# sh int nve 1

nve1 is up

admin state is up, Hardware: NVE

MTU 9216 bytes

Encapsulation VXLAN

Auto-mdix is turned off

RX

 icast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes

TX

 icast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes

N9372TX-43# sh nve vni 16010 counters

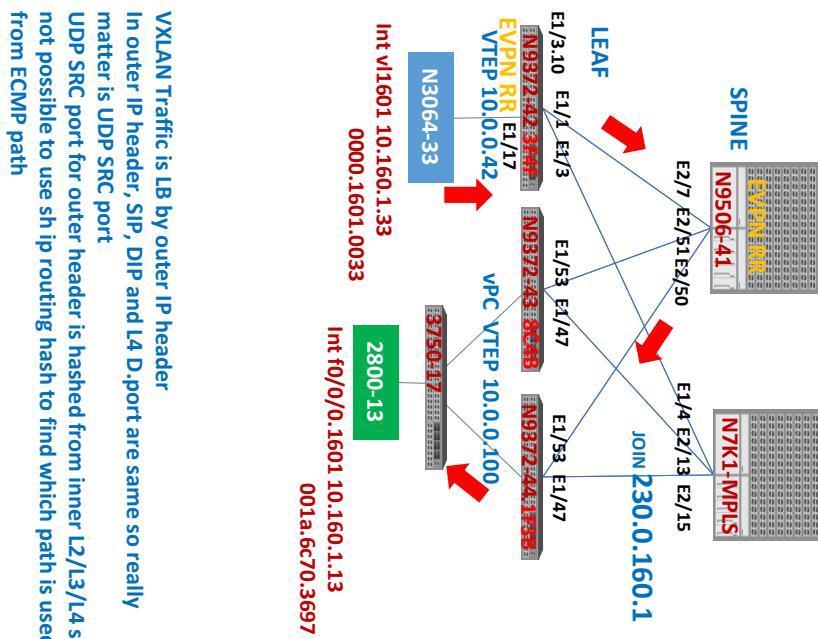
VNI: 16010

TX Use VNI level counter
N9372TX-43# clear nve vni 16100 counters
Cleared counters.

176395 unicast packets 30221175 unicast bytes

128069 multicast packets 1595552 multicast bytes

RX
126461 unicast packets 21189856 unicast bytes
10417 multicast packets 1824716 multicast byte



VXLAN EVPN – Utilize ERSPAN for Troubleshooting

Most powerful and easy way to troubleshooting EVPN is using ERSPAN

You just need IP connectivity from LEAF to LAPTOP or VM with wireshark

1. Configure ERSPAN as below

```
N9372TX-44(config-erspan-src)# sh run mon
monitor session 1 type erspan-source
erspan-id 2
vrf default
destination ip 10.0.0.1.2
source interface Ethernet1/47 both
no shut
monitor erspan origin ip-address 10.0.0.44 global
```

2. Run Wireshark on LAPTOP or VM

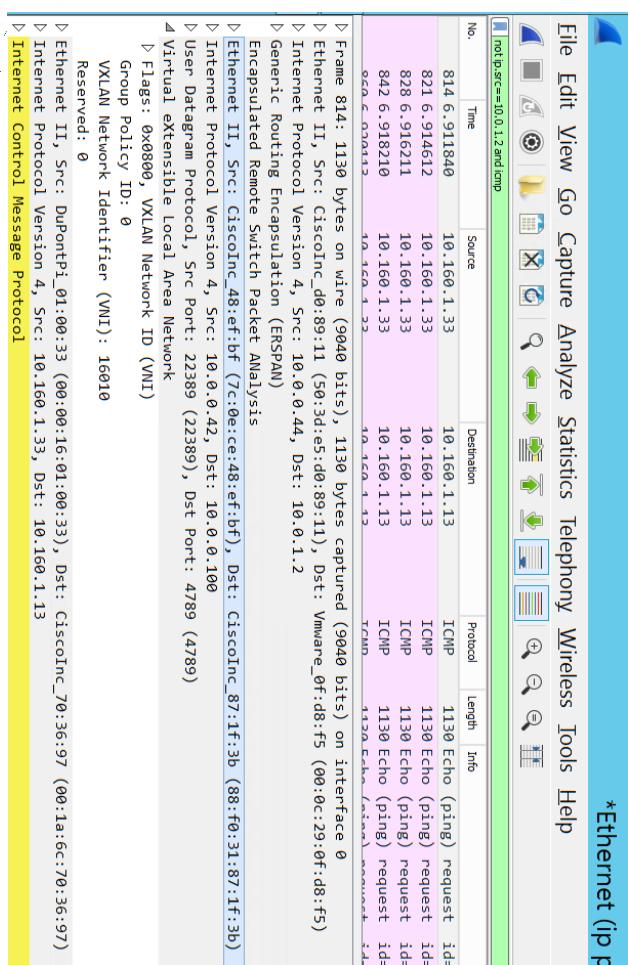
Go to Preference → Protocol → ERSPAN → "Force decode FAKE ERSPAN"

Use "ip proto 0x2f" as capture filter to get only ERSPAN packets

Use "not ip.src==10.0.1.2 and icmp" as display filter all ICMP Unreachable from LAPTOP or VM

You can ERSPAN on egress port to VM to verify PKT is leaving egress VTEP

Can't use VRF management for erspan
40G port can only do RX, not TX direction



N9504-41# sh system internal forwarding route 10.0.0.100/32 det mod 2

```
10.0.0.100/32 , Ethernet2/49
Dev: 1, RPF Enabled: N , VPN: 1
AdjId: 0x30e66, LIF: Ethernet2/49 (0x1083),
DMAC: 8800.3187.8c4b SMAC: 7c0e.ce48.efbf
AdjId: 0x30e66, LIF: Ethernet2/51 (0x1085),
DMAC: 88f0.3137.113b SMAC: 7c0e.ce48.efbf
```

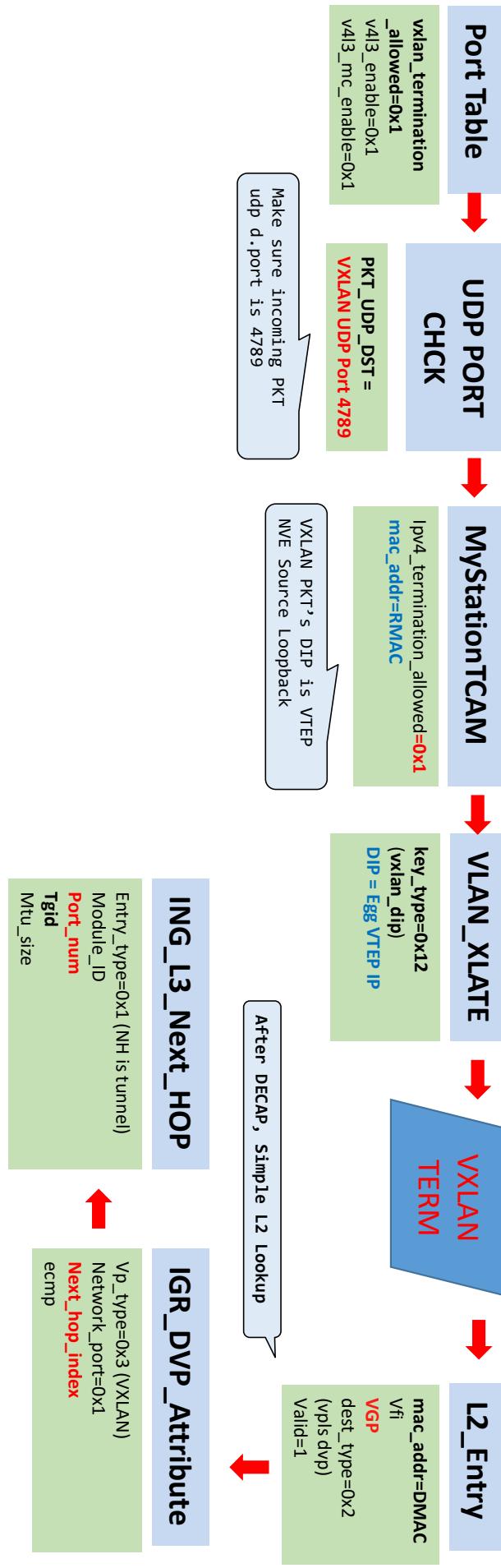
Outer DMAC is RMAC of LEAF44
Outer SIP is ingress LEAF42 VTEP IP and DIP is Leaf43/44
virtual VTEP IP

- ▷ Frame 814: 1130 bytes on wire (9040 bits), 1130 bytes captured (9040 bits) on interface 0
- ▷ Ethernet II, Src: CiscoInc d8:91:11 (50:3d:e5:d8:91:11), Dst: VMware_0f:d8:f5 (00:0c:29:0f:d8:f5)
- ▷ Internet Protocol Version 4, Src: 10.0.0.44, Dst: 10.0.1.2
- ▷ Internet Protocol Version 4, Src: 10.0.0.44, Dst: 10.0.1.2
- ▷ Generic Routing Encapsulation (ERSPAN)
- ▷ Encapsulated Remote Switch Packet Analysis
- ▷ Ethernet II, Src: CiscoInc_48:ef:bf (7c:0e:ce:48:ef:bf), Dst: CiscoInc_87:1f:3b (88:f0:31:87:1f:3b)
- ▷ Internet Protocol Version 4, Src: 10.0.0.42, Dst: 10.0.0.100
- ▷ User Datagram Protocol, Src Port: 22389 (22389), Det Port: 4789 (4789)
- ▷ Virtual extensible Local Area Network
 - ▷ Flags: 0x0800, VXLAN Network ID (VNI): 16010
 - ▷ VXLAN Network Identifier (VNI): 16010

VXLAN VNI is 16010, LVNI of vlan 1601

```
Reserve: 0
▷ Ethernet II, Src: DuPontPl_01:00:33 (00:00:16:01:00:33), Dst: CiscoInc_70:36:97 (00:1a:6c:70:36:97)
▷ Internet Protocol Version 4, Src: 10.160.1.33, Dst: 10.160.1.13
▷ Internet Control Message Protocol
```

VXLAN Bridging Decap Path



VXLAN EVPN – HW

Troubleshooting Steps for Bridging HW DECAP

Traffic from Host 10.160.1.33 to host 10.160.1.13 reaches LEAF44.

When VXLAN encapsulated packet is received by LEAF43, outer DIP is VTEP IP and DPoRT 4789.

This packet gets decapped and sent out using the inner dmac on the VXLAN access port.

1. Verify core facing INT has VXLAN TERM is allowed

```
N937TX-44# bcm-shell module 1 'd chg port 51 1'
PORT.ipipe[51]:VXLAN TERMINATION ALLOWED=1, VT_KEY_TYPE=4, VLAN_PROTOCOL_DATA_INDEX=0x3
3,VFP_ENABLE=1, V6L3_ENABLE=1, V6IPMC_ENABLE=1, V4L3_ENABLE=1, V4IPMC_ENABLE=1, TRUST_INCOMING_VID=1, TRUST_
DOT1P_PTR=1, SUBNET_BASED_VID_ENABLE=1, PRI_MAPPING=0xfac688, PORT_VID=1, OVID=1, OUTER_TPID_ENABLE=1, MY_MO
DID=1, MPLS_ENABLE=1, MM_TERM_ENABLE=1, MAC_BASED_VID_ENABLE=1, IPRI_MAPPING=0xfac688, IPMC_DO_VLAN=1, IEEE
_802_1AS_ENABLE=1, ICFI_1_MAPPING=1, FP_PORT_FIELD_SEL_INDEX=0x33, FILTER_ENABLE=1, EN_IFILTER=1, DISCARD_I
F_VNTAG_PRESENT=1, DATA_4=0x0000000000000000, DATA_3=0x0000000000000001f58c10800, DATA_2=0x80004000000000
000666668, DATA_1=0x0000040003763444148001, DATA_0=0x0100100000100000070ce021, CML_FLAGS_NEW=8, CML_FLAG
_S_MOVE=8, CFI_1_MAPPING=1, >
```

2. Verify that VTEP SRC loopback IP and VXLAN UDP port match on ingress/egress VTEP

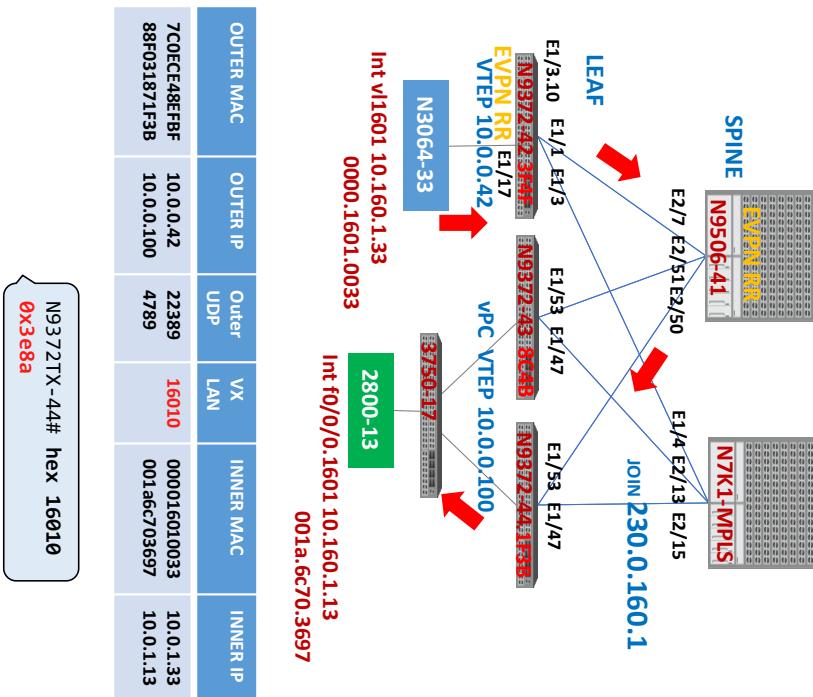
```
N937TX-44# sh nve interface
Interface: nve1, State: Up, encapsulation: VXLAN
VPC Capability: VPC-VIP-Only [notified]
Local Router MAC: 88f0.3187.1f3b
Host Learning Mode: Control-Plane
```

Source-Interface: Loopback0 (primary: 10.0.0.44, secondary: 10.0.0.100)

```
N937TX-44# sh nve vxlan-params
VXLAN Dest. UDP Port: 4789
```

3. Check My_Station_TCAM for DMAC **88f0.3187.1f3b**

```
N937TX-44# bcm-shell module 1 'd chg my_station_tcam' | i 1f3b
MY_STATION_TCAM.ipipe[0]:<VALID=1, MPLS_TERMINATION_ALLOWED=1, MASK=0x00000000ffffffffff, MAC_ADDR_MAC
K=0xffffffffffff, MAC_ADDR=0x88f031871f3b, KEY=0x0000000088f031871f3b, IPV6_TERMINATION_ALLOWED=1, IPV4_TE
RMINATION_ALL_ORED=1, DATA=0x3a, ARP_RARP_TERMINATION_ALLOWED=1>
```

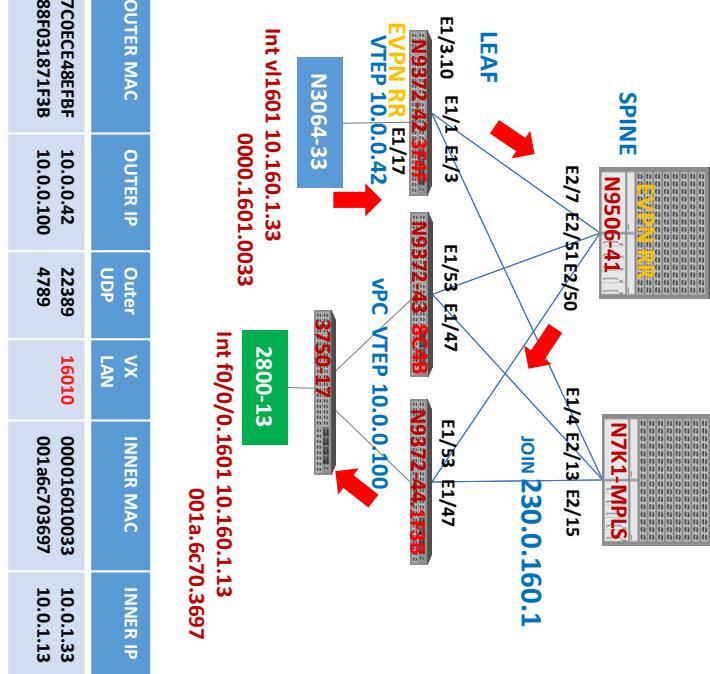


OUTER MAC	OUTER IP	Outer UDP LAN	VX LAN	INNER MAC	INNER IP
7C0ECE48EFBF 88F031871F3B	10.0.0.42 10.0.0.100	22389 4789	16010	000016010033 001a6c703697	10.0.1.33 10.0.1.13

```
Int v1601 10.160.1.33
0000.1601.0033
```

```
Int f0/0/0.1601 10.160.1.13
001a.6c70.3697
```

VXLAN EVPN – HW



4. Verify VLAN_XLATE with DIP as 10.0.0.100 exist, then VTEP decap VXLAN traffic

```
N9372TX-44# bcm-shell module 1 'd chg vlan_xlate' | i 0xa0000064
VLAN_XLATE.ipipe0[1048]: <VXLAN_DIP:KEY=0x0500003212,VXLAN_DIP:IGNORE_UDP_CHECKS
UM=1,VXLAN_DIP:HASH LSB=0x64, VXLAN_DIP:DIP=0xa0000064,VXLAN_DIP:DATA=0x0000000000000000,VALID=1,KEY_TYPE=0x12,>
```

5. Verify MPLS_Entry table has proper SIP and VFI

```
N9372TX-44# bcm-shell module 1 'd chg mpls_entry' | i 0xa00002a
MPLS_ENTRY.ipipe0[4368]:VXLAN_SIP:SIP=0x0000002a,VXLAN_SIP:KEY=0x00a0002a8,V
XLAN_SIP:HASH_LSB=0x2a,VXLAN_SIP:DATA=0x2db8,VALID=1,KEY_TYPE=8,EVEN_PARITY=1,>
MPLS_ENTRY.ipipe0[8312]:<VXLAN_VN_ID:VN_ID=0x3e8a,VXLAN_VN_ID:VFI=0x641,VXLAN_VN_ID:DATA=0x641,VAL_ID=1,KEY_TYPE=9,>
```

```
N9372TX-44# bcm-shell module 1 'd chg mpls_entry' | i 0x641 Vxlan_sip
MPLS_ENTRY.ipipe0[8312]:<VXLAN_VN_ID:VN_ID=0x3e8a,VXLAN_VN_ID:VFI=0x641,VXLAN_VN_ID:DATA=0x641,VAL_ID=1,KEY_TYPE=9,> Vxlan_vn_id
```

6. After DECAP, just usual L2 Lookup on based inner DMAC and vlan tag

```
N9372TX-44# bcm-shell module 1 'd chg l2_entry' | i 3697
L2_ENTRY.ipipe0[10328]:<VPG_TYPE=1,VPG_1=0xd,VPG=0x86e,VLAN_ID=0x641,VIF:L2MC_PTR=0x86e,VFI=0x641
,VALID=1,T_1=1,TGID_1=0xd,TGID=0x6e,RESERVED_69_68=1,PORT_NUM_1=0xd,PORT_NUM=0x6e,OVID=0x641,MODU
LE_ID=0x10,MAC_ADDR=0x001a6c703697,LOCAL_SA=1,L2MC_PTR=0x86e,L2:VPG_TYPE=1,L2_VPG=0x86e,L2:VLAN_I
D=0x641,L2:VFID=0x641,L2:TGID=0x6e,L2:RESERVED_69_68=1,L2:PORT_NUM=0x6e,L2:MODULE_ID=0x10,L2:MAC_A
DDR=0x001a6c703697,L2:L3MC_PTR=0x86e,L2:L2MC_PTR=0x86e,L2:KEY=0x0000069b1c0da5c6413,L2:HASH_LSB=0x
3697,L2:DEST_TYPE=2,L2:DESTINATION=0x86e,L2:DATA=0x00001086e,L2:ASSOCIATED_DATA=0x000001086e,KEY
_TYPE=3,KEY=0x0000069b1c0da5c6413,IVID=0xa5c,HITSA=1,HITDA=1,DEST_TYPE_1=1,DEST_TYPE=2,DESTINATION
_1=0xd,DESTINATION=0x86e,DATA=0x000001086e,ASSOCIATED_DATA=0x000001086e>
```

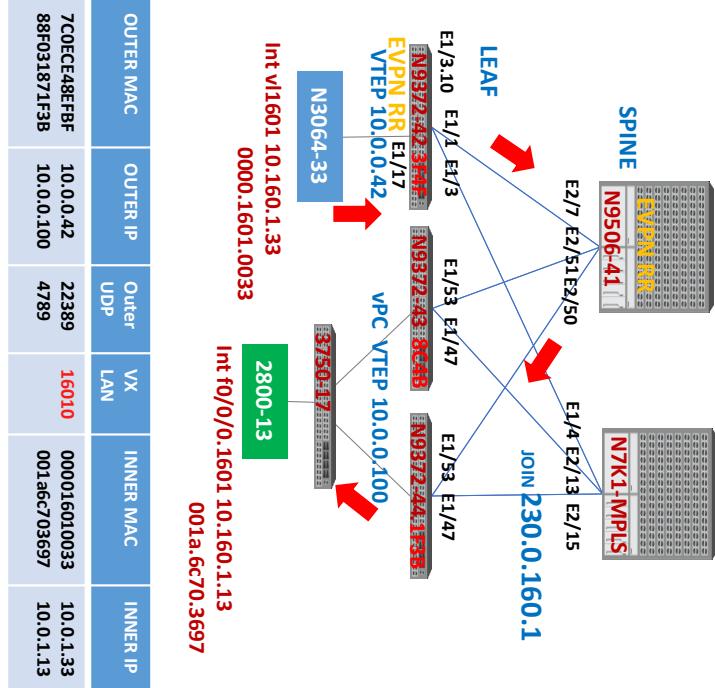
7. Check MAC in HW for VPG

```
N9372TX-44# sh mac address-table v1 1601 | i 3697
+ 1601 001a.6c70.3697 dynamic 0 F F P01744
```

N9372TX-44# show hardware mac address-table 1 vlan 1601 | i 3697|MAC

FE	PI	VLAN	MAC	Trunk	TGID	Mod	Port	Virtual	Static	Hit	Hit	CPU	Pending
0	1	30273	001a.6c70.3697	0	0	0	0	86e	0	1	1	0	0

VXLAN EVPN – HW



8. Use SVP and find NHI

```
N9372TX-44# bcm mo 1 'd chg ing_dvp_table 0x86e'
ING_DVP_TABLE.ipipe0[2158]: <NEXT_HOP_INDEX=0x9b,ECMP_PTR=0x9b,DVP_GROUP_PTR=0x9b,>
```

```
N9372TX-44# bcm-shell module 1 'd chg ing_13_next_hop 0x9b'
ING_L3_NEXT_HOP.ipipe0[155]:VLAN_ID=0xffff,TGID=5,T=1,PORT_NUM=5,MTU_SIZE=0x1ffe,L3_OIF=0x1ffe,ENTRY_IN
FO_UPPER=1,DVP_RES_INFO=0x7e,>
```

N9372TX-44# bcm-shell module 1 'trunk show id=5'

```
Device supports 1072 trunk groups:
1024 front panel trunks (0..1023), 256 ports/trunk
48 fabric trunks (1024..1071), 64 ports/trunk
```

```
trunk 5: (front panel, 2 ports)=xe4,xe5 dlf=any mc=any ipmc=any psc=portflow (0x9)
```

```
N9372TX-44# sh port-ch sum | i 1744
1744 Po1744(SU) Eth LACP Eth1/5(P) Eth1/6(P)
```

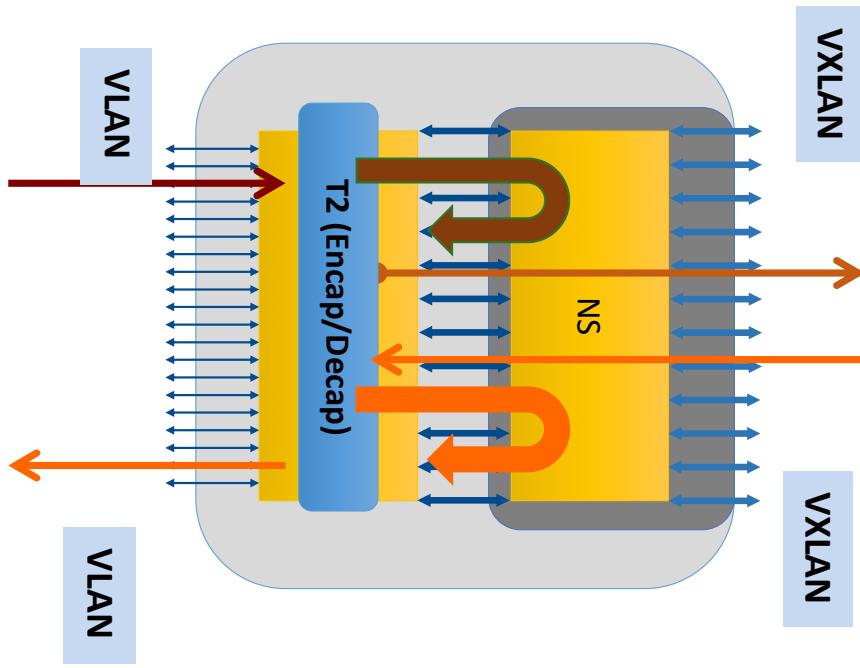
```
N9372TX-44# sh mac address-table vlan 1601 | i 3697
+ 1601 001a.6c70.3697 dynamic 0 F F Po1744
```

OUTER MAC	OUTER IP	Outer UDP	VX LAN	INNER MAC	INNER IP
7C0ECE48EFBF 88F031871F3B	10.0.0.42 10.0.0.100	22389 4789	16010 001a.6c70.3697	000016010033 10.0.1.33 10.0.1.13	001a.6c70.3697

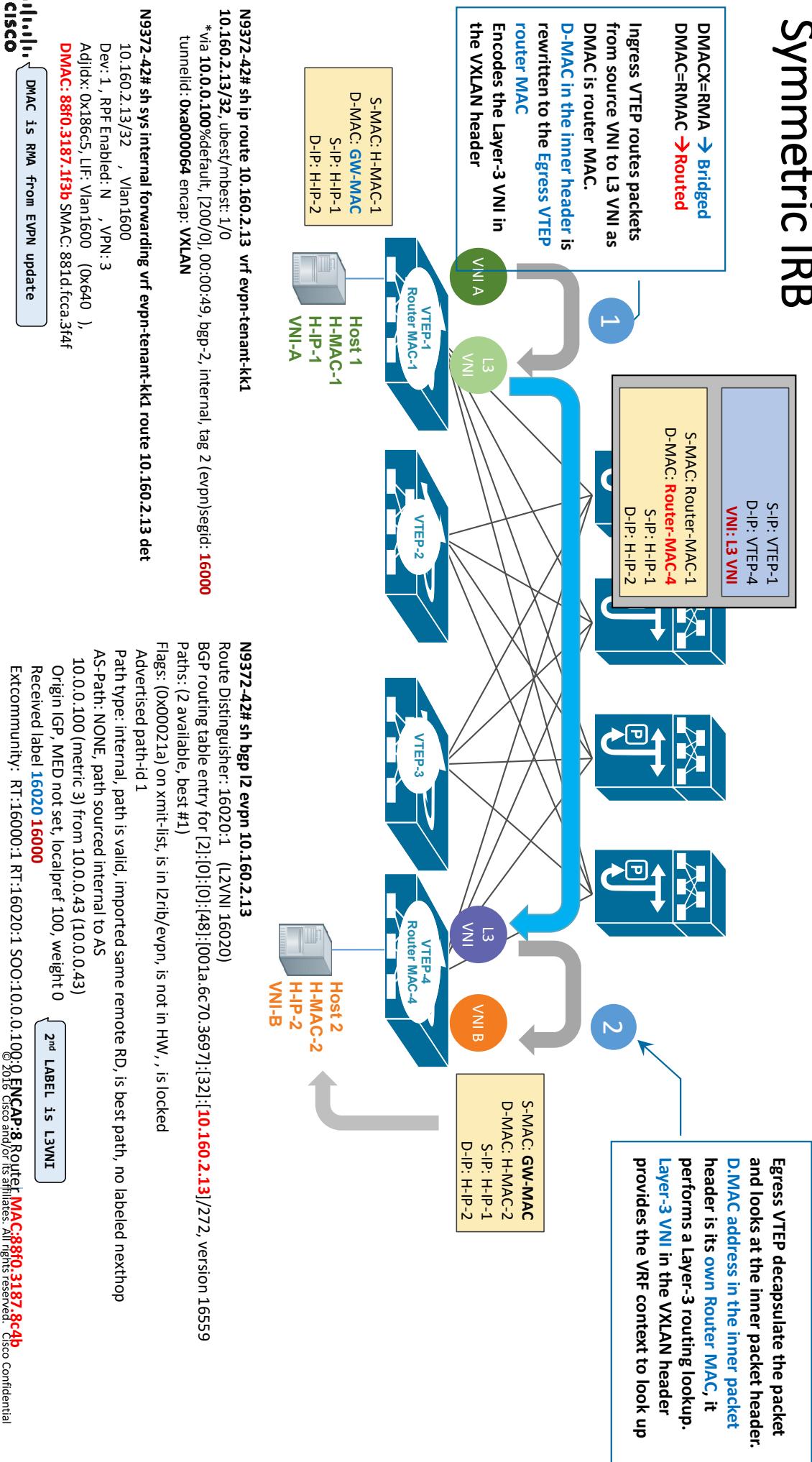
```
N9372TX-44# hex 16010
0x3e8a
```

VXLAN Routing packet path

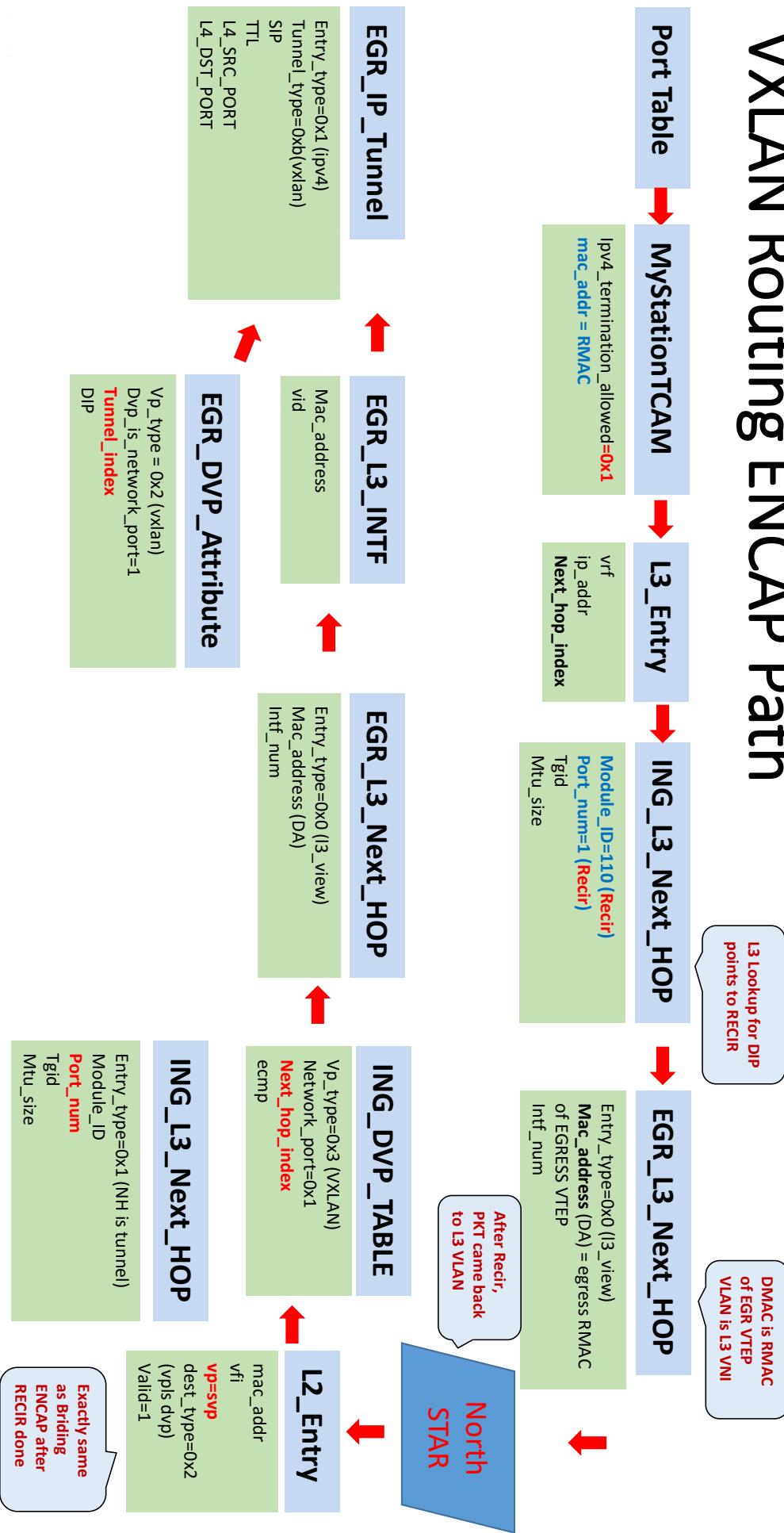
- VXLAN Encapsulation and Decapsulation occur on T2 (in Bridging pipeline)
- VLAN → VXLAN (Access to Network)
 - Packet from T2 (**after routing lookup**) recirculate via NS and **bridging lookup** is triggered.
- VXLAN → VLAN (Network to access)
 - Packet from T2 (**after bridging lookup**) recirculate via NS and **routing lookup** is triggered.
- VXLAN → VXLAN (Transit – Native IP)
- Routing is independent of the port type (1/10/40G ports)



Symmetric IRB



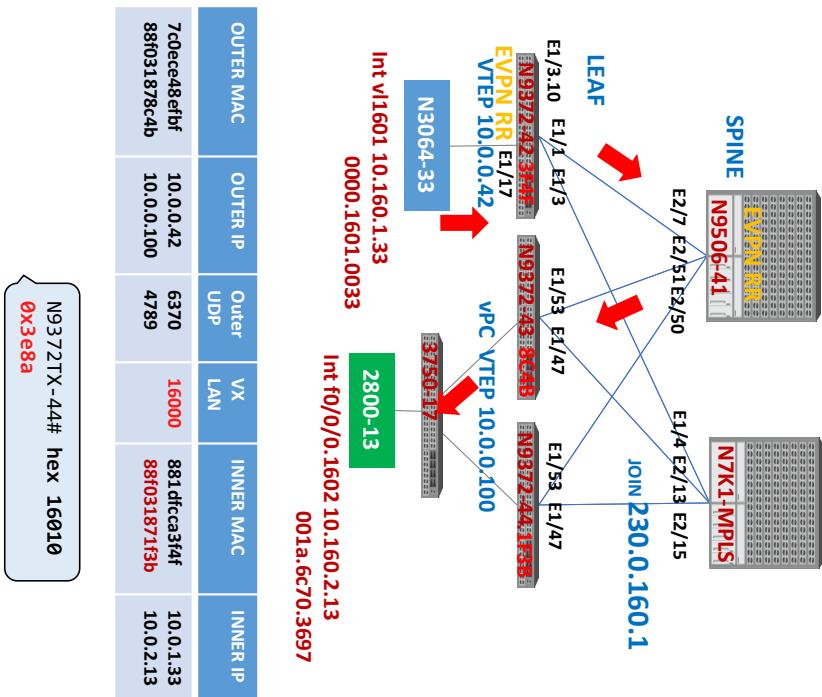
VXLAN Routing ENCAP Path



VXLA EVPN Routing ENCAP

Troubleshooting Steps for Routing HW ENCAP

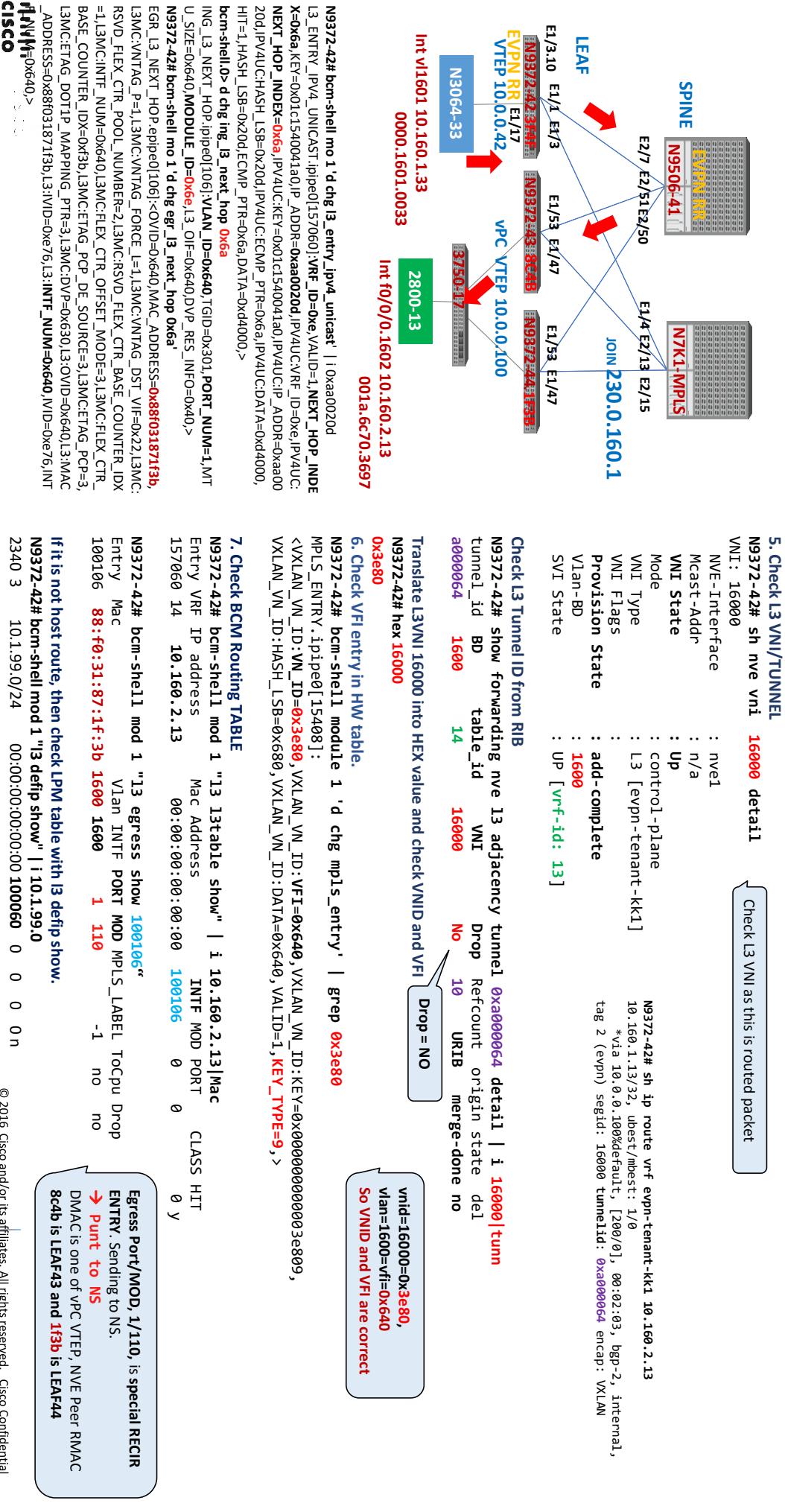
Host 10.160.1.33 in vlan 1601 on LEAF42 can't ping host 10.160.2.13 in vlan 1602 on vpc LEAF 43/44. Ingress LEAF is LEAF42 and Egress LEAF is vPC LEAF43/44.



- Assume control plane is good and RIB/FIB has correct entries
N9372-42# sh ip route vrf evpn-tenant-kk1 10.160.2.13
10.160.2.13/32, ubest/mbest: 1/0
*via 10.0.0.100%default, [200/0], 00:25:32, bgp-2, internal, tag 2 (evpn) segid: 16000 tunnelid:
0xa000064 encap: VXLAN
N9372-42# sh vrf peer
nve1 10.0.0.100 up CP 6d16h 88f0.3187.1f3b
N9372-42# sh vrf evpn-tenant-kk1 detail
VRF-Name: evpn-tenant-kk1, VRF-ID: 13, State: Up
VPNID: unknown
RD: 16000:1
VNI: 16000
Max Routes: 0 Mid-Threshold: 0
Table-ID: 0x8000000d, AF: IPv6, Fwd-ID: 0x8000000d, State: Up
Table-ID: 0x0000000d, AF: IPv4, Fwd-ID: 0x0000000d, State: Up
- Check ingress INT is in correct VRF.**
N9372-42# sh vrf | i kk1
evpn-tenant-kk1
N9372-42# bcm-shell mod 1 "d chg l3_if 1601" 13 up --
L3_If:ipipe0[1601]:<VRF=0xd,l3_if_PROFILE_INDEX=2,IPMC_L3_IF=0x641,CLASS_ID=0xc00,>
- Intervlan packet comes with DST MAC as ingress LEAF's router mac, so check distribute anycast GW MAC
N9372-42# sh run fabric forwarding | i mac
fabric forwarding anycast-gateway-mac 1600.1600.1600
interface Vlan1601
fabric forwarding mode anycast-gateway
- Let's check if RMAC is proper installed in HW TCAM, my_station_team**
N9372-42# bcm-shell module 1 "dump chg my_station_tcam"
MY_STATION_TCAM.ipipe0[1]: <VAL_ID=1,MPLS_TERMINATION_ALLOWED=1,MASK=0x00000000ff
ffffffffff,MAC_ADDR_MASK=0xffffffffffff,MAC_ADDR=0x160016001600,KEY=0x0000000016
0016001600,IPV6_TERMINATION_ALLOWED=1,IPV4_TERMINATION_ALLOWED=1,DATA=0x3a,ARP_R
ARP_TERMINATION_ALLOWED=1>

cisco

VXLA EVPN Routing ENCAP



VXLA EVPN Routing ENCAP

8. Packet is sent out of T2 to NS and received back on T2 again on the L3 VRF VLAN, in this case 1600, VFI_ID=640.

Return traffic hits the FP entry for VRF VLAN. This points to VFI under which the dmac lookup has to happen.

```
N9372-42# bcm-shell module 1 'fp show' | grep -B 6 -A 12 640
EID 0x00004fc8: gid=0x38,
```

```
    slice=0, slice_idx=0x3, part =0 prio=0x2000, flags=0x10202, Installed, Enabled
```

```
    tcam: color_indep=0,
```

```
StageLookup OuterVlanId Offset: 177 Width: 12
```

```
DATA=0x000000640
```

```
MASK=0x000000ffff
```

```
HiGig Offset: 176 Width: 1
```

```
DATA=0x00000001
```

```
MASK=0x00000001
```

```
SrcMac Offset: 68 Width: 48
```

```
DATA=0x0000881d fccaa3f4f
```

```
MASK=0x0000ffff ffffffff
```

```
action={act=IncomingMplsPortSet, param0=402668440(0x18003b98), param1=0(0), param2=0(0),
```

```
param3=0(0)}
```

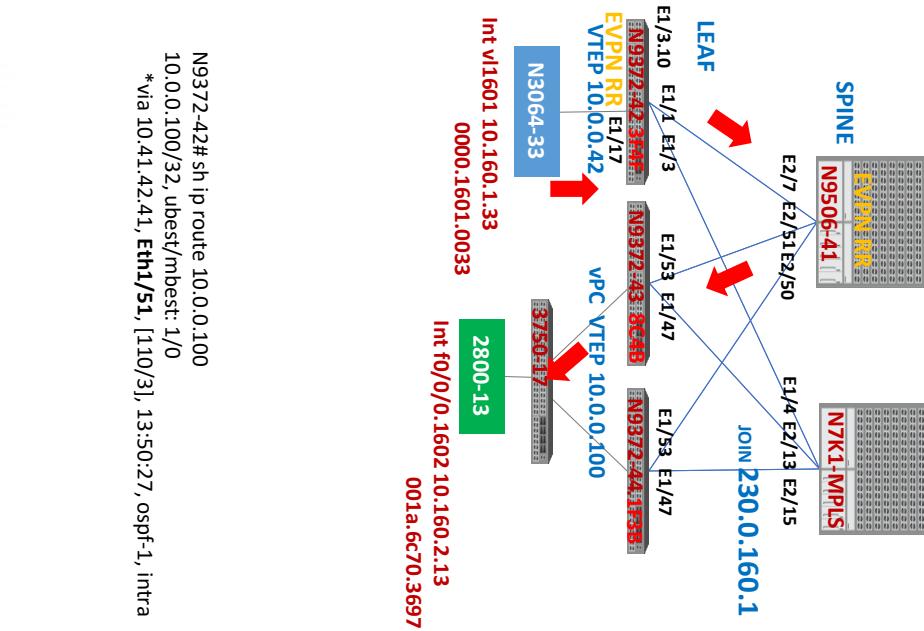
```
policer= statistics={stat id 20424 slice = 1 idx=15 entries=1}{Packets}
```

Get the VP, virtual port, entry from param0[Here it is 3B98]. Check the source_vp entry for 0x3B98

```
N9372-42# bcm-shell module 1 'd chg source_vp 0x3b98'
SOURCE_VP.ipipe0[15256]:
<VFI=0x640,TPID_ENABLE=0xf,SD_TAG_MODE=1,IPv6L3_ENABLE=1,FLEX_CTR_POOL_NUMBER=2,FLEX_C
TR_OFFSET_MODE=1,FLEX_C
TR_BASE_COUNTER_IDX=0x20,EXP_PVLAN_VID=0x640,ENTRY_TYPE=1,DVP=0x640,DISABLE_VLAN_CHECKS=1,DEFAULT_VLAN
_TAGS=0x190,DEFAULT_VID=0x640,>
```

```
N9372-42# bcm-shell mod 1 "13 egress show 100106"
Entry Mac Vlan INTF PORT MOD MPLS_LABEL ToCpu Drop
100106 88:f0:31:87:1f:3b 1600 1600 1 110 -1 no no
```

VXLA EVPN Routing ENCAP



9. Now we are doing normal VXLAN MAC lookup in vlan 1600, VFI, for DMAC 8C4B, from L3 egress table

```
N9372-42# sh mac address-table vlan 1600 | i 1f3b
* 1600    88f0.3187.1f3b    static    -      F      F      nve1(10.0.0.100)
```

HW VLAN ID = 28672+VXLAN ID = 28672+1600 = 30272, so we look for VLAN 30272

Find DST MAC 1F3B from L2 Table

```
N9372-42# bcm-shell module 1 'l2 show' | grep 1f:3b
mac=88:f0:31:87:1f:3b vlan=30272 GPORT=0x80002db8 port=0x80002db8 (vxlan) Static Hit
N9372-42# bcm-shell module 1 'd chg ing_dvp_table 0x2db8'
ING_DVP_TABLE.ipipe[11704]: <VTP_TYPE=3,NEXT_HOP_INDEX=0x73,NETWORK_PORT=1,ECMP_PTR=0x73,DVP_GROUP_PTR=0x73,>
```

N9372-42# bcm-shell module 1 'd chg ing_l3_next_hop 0x73'

```
ING_L3_NEXT_HOP.ipipe[115]: <VLAN_ID=0xffff,TGID=0x83,PORT_NUM=3,MTU_SIZE=0x3fff
,MODULE_ID=1,L3_OIF=0x1ffff,ENTRY_TYPE=2,ENTRY_INFO_UPPER=3,DVP_RES_INFO=0x7f,>
```

If T=1 is present, then TGID will be the trunk id in hex. Use *bcm-shell module 1 'trunk show' to get the trunk members.

```
N9372-42# bcm-shell module 1 'phy info' | gre '3'
hg2( 3) 600d 8770 81 81          TSC-A2/00/4 250000
```

Egress port is HG2, E1/51

```
N9372-42# sh interface hardware-mappings | i 1/51|Name
Name   Ifindex Smod Unit HPort FPort Nport Vport
Eth1/51 1a006400 1 0 3 2 50 -1
```

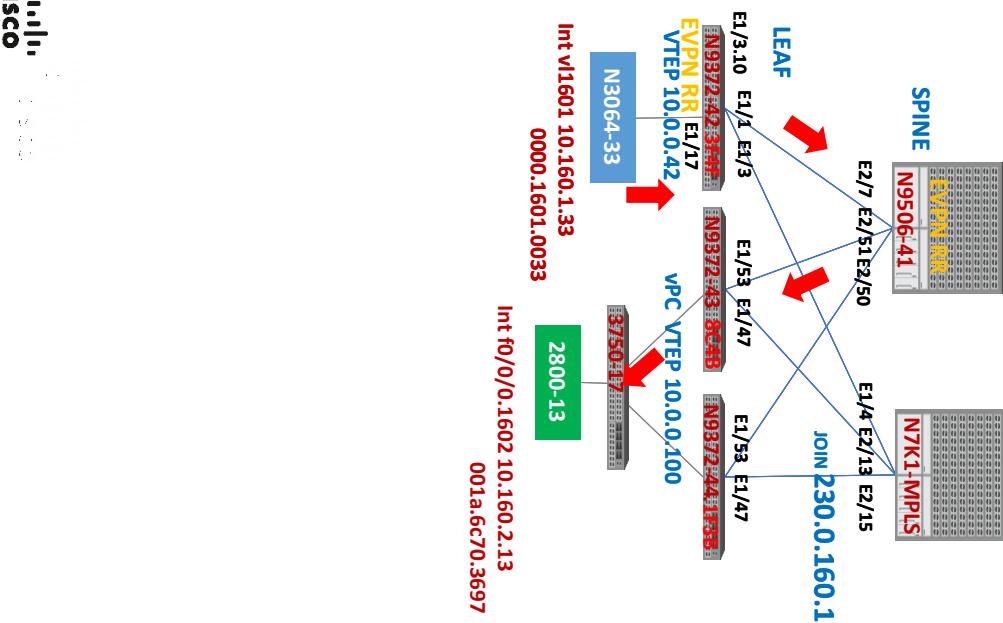
11. Check Tunnel DIP from SVF

We got GPORT=0x80002db9 from NVE DMAC and VNI 16000 so use it to find tunnel encaps

```
N9372-42# bcm-shell module 1 'd chg mpls_entry' | i 2db8
MPLS_ENTRY.ipipe[4336]: <VXLAN_SIP:SVI=0x2db8,VXLAN_SIP:DATA=0xa000064,VXLAN_SIP:KEY=0x00a000648,VXLAN_SIP:HASH_LSB=0x64,VXLAN_SIP:VALID=1,KEY_TYPE=8,EVEN_PARITY=1,>
```

```
DEST IP 0xa000064 = 10.0.0.100
```

VXLA EVPN Routing ENCAP



12. Verify tunnel attributes for SVP peer by using GPORT info and get TUNNEL ID

```
N9372-42# bcm-shell module 1 'd chg egr_dvp_attribute 0x2d18'
EGR_DVP_ATTRIBUTE.epipe0[11704]: <VXLAN:TUNNEL_INDEX=1,VXLAN:DVP_IS_NETWORK_PORT
=1,VXLAN:DIP=0xa000064,VXLAN:DELETE_VNTAG=1,VP_TYPE=2,>
```

13. Check Tunnel ID

```
N9372-42# bcm-shell module 1 'd chg egr_ip_tunnel 1'
EGR_IP_TUNNEL.epipe0[1]: <TUNNEL_TYPE=0xb,TTL=0xff,SIP=0xa00002a,L4_DEST_PORT=0x
12b5,ENTRY_TYPE=1,DSCP_SEL=1,>
```

SIP a00002a = 10.0.0.42

```
14. Now, check NH for this TUNNEL, SIP a00002a = 10.0.0.42 DIP 0xa000064 = 10.0.0.100
N9372-42#bcm-shell module 1 '13 l3table show' | egre 10.0.0.100|Mac
Entry VRF IP address Mac Address INTF Mod PORT CLASS HIT
168792 1 10.0.0.100 00:00:00:00:00:00 100066 0 0 0 n
N9372-42# bcm-shell module 1 '13 egress show 100066'
Entry Mac Vlan INTF PORT MOD MPLS_LABEL ToCpu Drop
100066 7c:0e:ce:48:ef:bf 4095 4162 3 1 -1 no no
```

Int v1601 10.160.1.33
0000.1601.0033

Int fo/0/0.1602 10.160.2.13
001a.6c70.3697

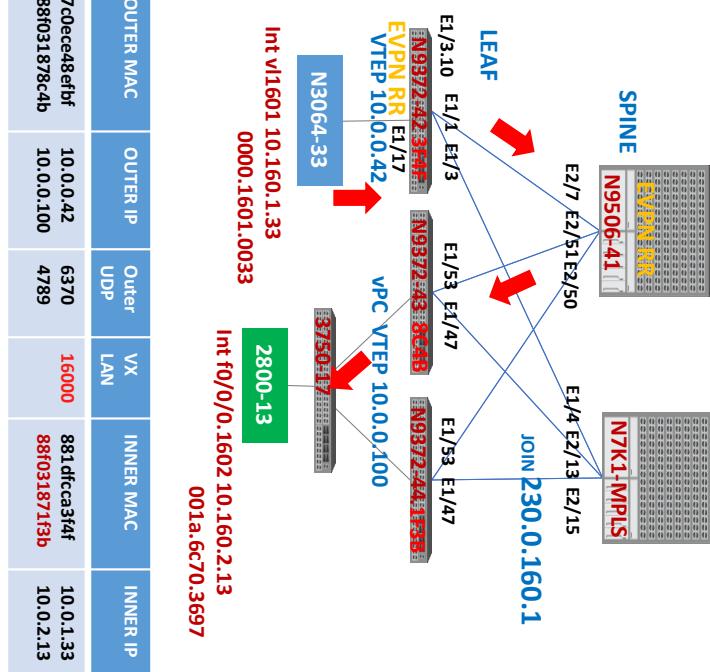
```
N9372-42# sh system internal forwarding route 10.0.0.100/32 det
10.0.0.100/32 , Ethernet1/51
Dev: 1 , RPF Enabled: N , VPN: 1
AdjIdx: 0x186e2, LIF: Ethernet1/51 (0x1042),
DMAc: 7c0e.ce48.efbf SMAC: 881d.fcca.3f4f
```

N9372-42# sh interface hardware-mappings | egre Eth1/1|Port

```
HPort - Hardware Port Number or Hardware Trunk Id:
Front - Fabric facing port number. 255 means N/A
NPort - Front panel port number
VPort - Virtual Port Number. -1 means N/A
Name Ifindex Smod Unit HPort FPort NPort VPort
Eth1/51 1a006400 1 0 3 2 50 -1
```

```
N9372-42# sh ip adjacency e1/51
Address MAC Address Pref Source Interface
10.41.42.41 7c0e.ce48.efbf 50 arp Ethernet1/51
Leaf42 sends ICMP ECHO to SPINE41 out of E1/1
```

VXLA EVPN Routing ENCAP



```
N9504-41# received ICMP ECHO on int E2/7
N9504-41# sh ip route 10.0.0.100
10.0.0.100/32, ubest/mbest: 2/0
*via 10.41.43.43, Eth2/47, [110/5], 3d04h, ospf-1, intra
*via 10.41.44.44, Eth2/48, [110/5], 3d04h, ospf-1, intra
```

```
N9504-41# bcm-shell module 2 "13 13table show" | i 10.0.0.100
Entry VRF IP address Mac Address INTF MOD PORT CLASS HIT
168792 1 10.0.0.100 00:00:00:00:00:00 200281 0 0 (ECMP)
```

```
N9504-41# bcm-shell module 2 "13 multipath show 200281"
```

```
Multipath Egress Object 200281
Interfaces: 100025 100028
```

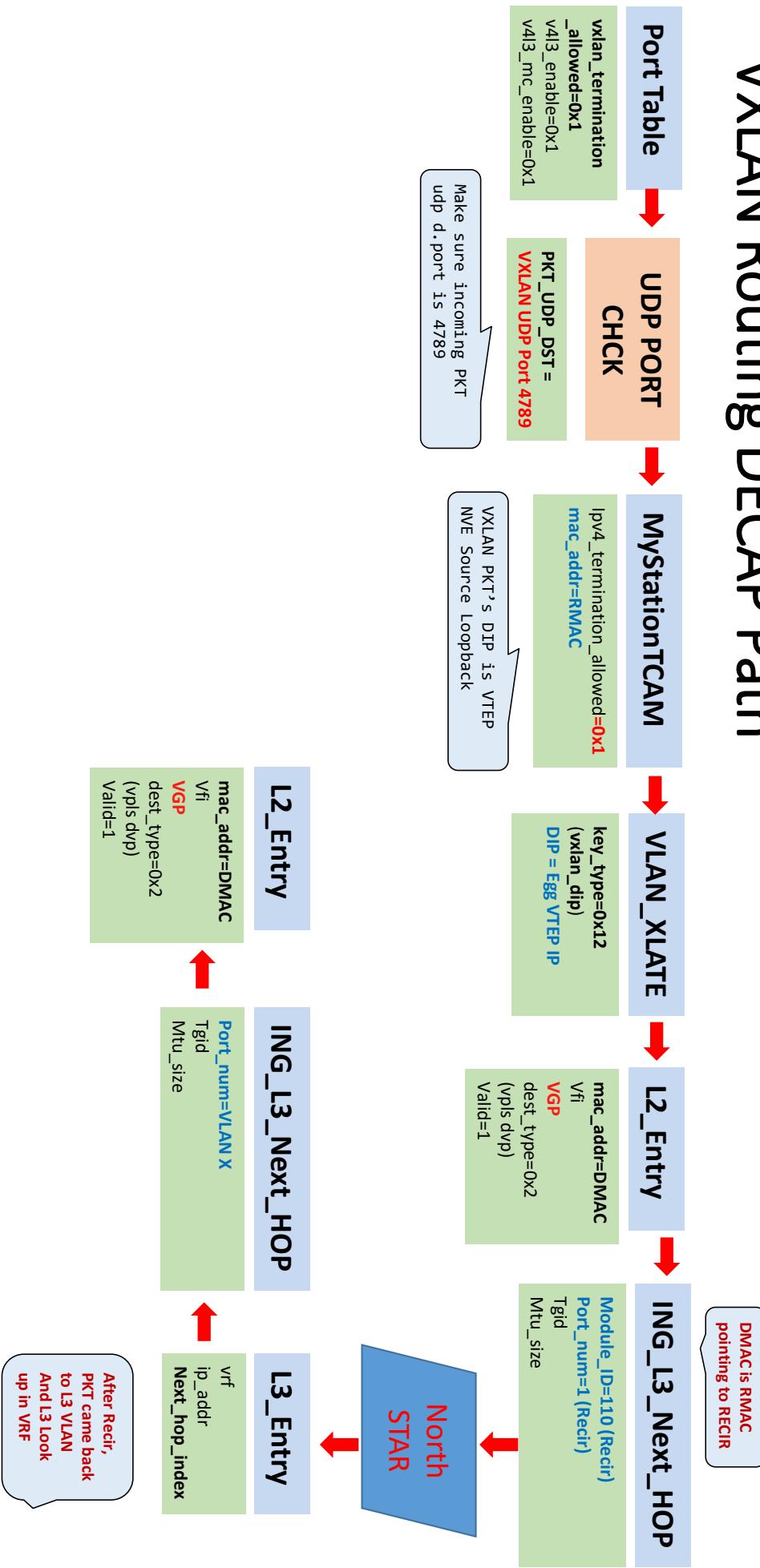
```
N9504-41# bcm-shell mod 2 "13 egress show 100025"
Entry Mac Vlan INTF PORT MOD MPLS_LABEL ToCpu Drop
100025 88:f0:31:87:8c:4b 4095 4225 59 4 -1 no no
```

```
N9504-41# bcm-shell mod 2 "13 egress show 100028"
Entry Mac Vlan INTF PORT MOD MPLS_LABEL ToCpu Drop
100028 88:f0:31:87:1f:3b 4095 4226 60 4 -1 no no
```

```
N9504-41# sh routing hash 10.0.0.42 10.0.0.100 54568 4798 mod 2
```

```
Load-share parameters used for software forwarding:
Load-share mode: address source-destination
No IPv4 protocol specified, defaulting to UDP
Hash for VRF "default"
Hashing to path *10.41.44.44
Out Interface: Eth2/48
For route:
10.0.0.100/32, ubest/mbest: 2/0
*via 10.41.43.43, Eth2/47, [110/5], 3d04h, ospf-1, intra
*via 10.41.44.44, Eth2/48, [110/5], 3d04h, ospf-1, intra
```

VXLAN Routing DECAF Path

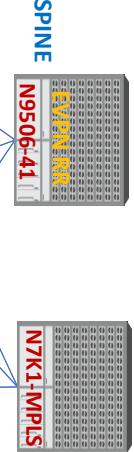


VXLA EVPN Routing DECAP

How do we find which VPC VTEP receive PKT? SPINE41 has two ECMP to 10.0.0.100

Clear counter on NVE int and use ping with timeout 0. You might see ping failure as COPP will drop it.
 N9372TX-43# ping 10.160.1.13 vrf host1601 count 1000 timeout 0
 N9372TX-43# sh int nv 1

nve1 is up
 admin state is up, Hardware: NVE
 MTU 9216 bytes
 Encapsulation VXLAN



Auto-midx is turned off
 RX
 ucast: 3975 pkts, 658424 bytes - mcast: 0 pkts, 0 bytes
 TX
 ucast: 3395 pkts, 570360 bytes - mcast: 38 pkts, 4432 bytes

When LEAF43 received ICMP ECHO, outer DIP is 10.0.0.100, outer DMAC is for E1/51, and inner DMAC is RMAC of VTEP.
 Remember LEAF42 had RMAC of LEAF44 but ECHO can reach LEAF43 or LEAF44 due to ECMP

So, peer-gw must be configured for any of vPC VTEP to process incoming VXLAN ENCAP traffic
 From previous ENCAP TS on LEAF42, we saw DMAC for 10.160.2.13 is 8C4B.

```
N9372-42# bcm-shell mod 1 "13 13table show" | i 10.160.2.13|Mac
Entry VRF IP address      Mac Address          INTF MOD PORT      CLASS HIT
157060 14 10.160.2.13   00:00:00:00:00:00  100106      0 0 0 y
N9372-42# bcm-shell mod 1 "13 egress show 100106"
Entry Mac          Vlan INTF PORT MOD MPLS_LABEL ToCpu Drop
100106 88:f0:31:87:1f:3b 1600 1600 1 110 -1 no no
```

OUTER MAC	OUTER IP	Outer UDP	VX LAN	INNER MAC	INNER IP
7c0ce48efbf 88f031871f3b	10.0.0.42 10.0.0.100	6370 4789	16000 88f031871f3b	881dfccaa314f 10.0.1.33 10.0.2.13	

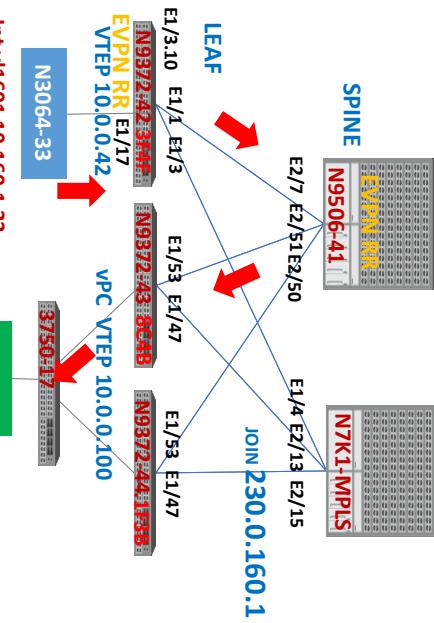
You can see 1f3b is RMAC of LEAF44 but G flag is set due to peer-gateway so LEAF44 will route packet.
 N9372TX-43# sh mac add v1 1600 | i sup vpc
 age - seconds since last seen,+ - primary entry using vPC Peer-Link,
 G 1600 88:f0:31:87:1f3b static - F F vPC Peer-Link(R)

OUTER MAC	OUTER IP	Outer UDP	VX LAN	INNER MAC	INNER IP
7c0ce48efbf 88f031871f3b	10.0.0.42 10.0.0.100	6370 4789	16000 88f031871f3b	881dfccaa314f 10.0.1.33 10.0.2.13	

Check mystationtcam for DMAC

```
N9372TX-43# bcm-shell mo 1 'd chg my_station_tcam' | i 1f3b
MY_STATION_TCAM.ipipe[1]:<VALID_ID=1, MASK=0x000000ffffffffff, MAC_ADDR_MASK=0xffffffffffff, MAC_A
DDR=0x88f031871f3b, KEY=0x000000088f031871f3b, IPV6_TERMINATION_ALLOWED=1, IPv4_TERMINATION_ALLOWED
=1, DATA=0x38, ARP_RARP_TERMINATION_ALLOWED=1>
```

VXLA EVPN Routing DECAP



Verify MPLS_ENTRY for VLAN 1600 exist

```
N9372TX-43# bcm-shell module 1 'd chg mpls_entry' | grep 0x3e80
MPLS_ENTRY.ipipe0[15408]: <VLAN_VN_ID:VN_ID=0x3e80,VXLAN_VN_ID:KEY=0x000000000003e809,VXLAN_VN_ID:HASH_L
SB=0x680,VXLAN_VN_ID:DATA=0x640,VALIDID=1,KEY_TYPE=9,>
```

Verify VLAN Xlate table has entry for DIP, 10.0.0.10=0xa0000064

```
N9372TX-43# bcm-shell mo 1 'd vlan_xlate' | i 0xa0000064
VLAN_XLATE.ipipe0[1048]: <VXLAN_DIP:USE_OUTER_HEADER, PHB=0, VXLAN_DIP:RSVD_FLEX_CTR_POOL_NUMBER=0, VXLAN_DIP:RSVD_FLEX_CTR_BASE_COUNTER_IDX=0, VXLAN_DIP:RESERVED_0=0x00000000000000000000000000000000, VXLAN_DIP:RSVD_FLEX_CTR_POOL_NUMBER=0, VXLAN_DIP:IGNORE_UDP_CHECKSUM=1, VXLAN_DIP:HASH_NETWORK_RECEIVERS_PRESENT=0, VXLAN_DIP:KEY=0x050003212, VXLAN_DIP:FLEX_CTR_OFFSET_MODE=0, VXLAN_DIP:FLEX_CTR_BASE_LSB=0x64, VXLAN_DIP:DIP=FLEX_CTR_POOL_NUMBER=0, VXLAN_DIP:DATA=0x00000000000000000000000000000000, VALIDID=1, SOURCE_COUNTER_IDX=0, VXLAN_DIP:DIP=0xa0000064, VXLAN_DIP:DATA=0x00000000000000000000000000000000, VALIDID=1, SOURCE_CE_TYPE=0, KEY_TYPE=0x12, EVEN_PARITY=0,>
```

Find SVP from HW MAC table for DMAC

```
N9372TX-43# sh hardware mac address-table 1 vlan 1600 | i 3b|MAC
FE |PI| VLAN | MAC |Trunk| TGID |Mod|Port|Virtual|Static|Hit|Hit|CPU|Pending
0 1 30272 88f0.3187.1f3b 0 0 0 3b98 1 1 1 0 0
HW VLAN ID = 28672+VXLAN ID = 28672+1600 = 30272, so we look for vlan 30272
```

Use SVP to check ing_dvp_table for NH Index

```
N9372TX-43# bcm-shell module 1 'd chg ing_dvp_table 0x3b98'
ING_DVP_TABLE.ipipe0[15256]: <NEXT_HOP_INDEX=0x134,ECMP_PTR=0x134,DVP_GROUP_PTR=0x134,>
N9372TX-43# bcm-shell module 1 'd chg ing_l3_next_hop 0x134'
ING_L3_NEXT_HOP.ipipe0[308]: <VLAN_ID=0xffe,TGID=0x301,PORT_NUM=1,MTU_SIZE=0x1ffe,MODULE_ID=0x6e,L3_OIF=0x1ffe,
ENTRY_INFO_UPPER=1,DVP_RES_INFO=0x7e,>
```

port1/mod110 is special RECIR port to NS and Decapped

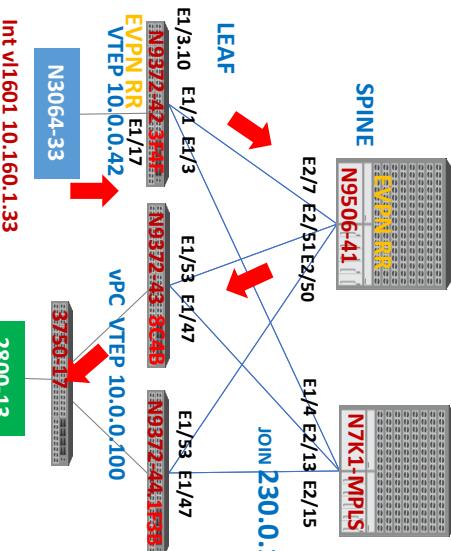
PKT is recirculated to NS and come back to BCM as decapped, so use DIP, 10.160.2.13 in VLAN 1600 for L3 Lookup

```
N9372TX-43# sh vrf evpn-tenant-kk1
VRF-Name          VRF-ID  State   Reason
evpn-tenant-kk1    12      Up     --

```

```
N9372TX-43# bcm-shell module 1 'd chg l3_if 1600'
L3_IIF.ipipe0[1600]: <VRF=0xc,L3_PROFILE_INDEX=2,IPMC_L3_IIF=0x640,CLASS_ID=0xc00,>
```

VXLA EVPN Routing DECAP



Check VRF Routing Table for DIP, 10.160.2.13

```
N9372TX-43# bcm-shell module 1 '13 l3table show' | i 10.160.2.14|Mac
Entry VRF IP address Mac Address INTF MOD PORT CLASS HIT
177140 12 10.160.2.14 00:00:00:00:00:00 100339 0 0 0 y
```

```
N9372TX-43# bcm-shell module 1 '13 egress show 100339'
Entry Mac Vlan INTF PORT MOD MPLS_LABEL ToCpu Drop
100339 00:22:90:b0:c6:53 1602 1602 5t -1 -1 no no
```

```
N9372TX-43# sh mac address-table | i c653
* 1602 0022.90b0.c653 dynamic 0 F F Po1843
```

```
N9372TX-43# sh ip arp vrf evpn-tenant-kk1
10.160.2.14 00:14:37 0022.90b0.c653 Vlan1602
```

```
N9372TX-43# show hardware mac address-table 1 | i c653
FE |PI| VLAN | MAC |Trunk| TGID |Mod|Port|Virtual|Static|Hit|Hit|CPU|Pending
0 1 30274 0022.90b0.c653 0 0 0 0 893 1 0 0 0 0
```

```
N9372TX-43# bcm-shell module 1 'trunk show id=5'
Device supports 1072 trunk groups:
```

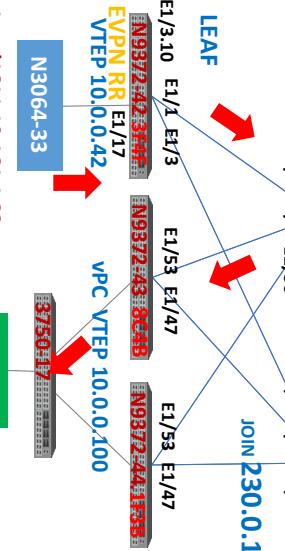
1024 front panel trunks (0..1023), 256 ports/trunk
48 fabric trunks (1024..1071), 64 ports/trunk

trunk 5: (front panel, 2 ports)=xe44,xe45 dlf=any mc=any ipmc=any psc=portflow (0x9)

```
N9372TX-43# sh port-c sum | i 1843
1843 Po1843(SU) Eth LACP Eth1/45(P) Eth1/46(P)
```

VXLA EVPN Ingress Replication TS

VRF EVPN-Tenant-KK2 uses BGP Ingress Replication for VLAN 1611 and VLAN 1612
Let's check if Ing Rep is working fine in control plane



```
N9372-42# sh nve peers
Interface Peer-IP State LearnType Uptime Router-Mac
nve1 10.0.0.100 Up CP 00:12:02 88f0.3187.8c4b
```

```
N9372-42# sh nve vni ingress-replication 16120
Interface VNI Replication List Source Up Time
nve1 16120 10.0.0.100 IMET 00:09:22
```

Check VNI 16120 set for Ingress Replication

```
N9372-42# sh 12route evpn imet evi 1612
VNI Prod Originating Router IP Addr
16120 BGP 10.0.0.100
16120 VXLAN 10.0.0.42
```

```
N9372-42# sh forwarding nve 12 ingress-replication-peers
slot 1
Total count of VLANS with ingr-repl peers: 2
VLAN 3100
```

peer: 10.0.0.100

VLAN 1612
peer: 10.0.0.100

```
N9372-42# sh forwarding distribution peer-id
```

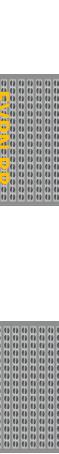
```
UDFM Peer-id allocations:
App: VXLAN Vlan: 1 Id: 0x10a000064 Peer-id: 0x2
App: VXLAN Vlan: 4095 Id: 0xa000064 Peer-id: 0xdead
We see two Peer ID, one for IR Encap
```

```
N9372-42# sh forwarding distribution internal ingress-replication-peer
Total Ingress-replication Peers: 1
Peer-0x9db Vlan list:
VL-1612 VL-3100
```

VXLA EVPN Ingress Replication TS

2. Check ING REP in HW.

Start with VFI for 1612 to find flood INDEX for BUM



SPINE

N9506-41

N7K1-MRSL

E2/7
E2/51
E2/42

E2/14
E2/13
E2/15

E1/3,10
E1/1
E1/3

E1/53
E1/47

E1/47

E1/53
E1/47

```
N9372-42# bcm-shell module 1 'mc show' | b 1fb5
Group 0xc001fb5 (VXLAN)
    port hg2, encap id 400029
    port xe16, encap id 400029
Uses NHI to find DVP to verify tunnel SIP/DIP
N9372-42# bcm-shell module 1 'd chg egr_13_next_hop 156'
EGR_L3_NEXT_HOP .epipe0[156]:
<L3MC:MAC_ADDRESS=0x7c0ecce48efbf, L3MC:L3MC_USE_CONFIGURED_MAC=1, L3MC:L2_MC_VLAN_DISABLE=1, L3MC:L2_MC_SA_DISABLE=1, L3MC:L2_MC_DA_DISABLE=1, L3MC:INTF_NUM=0x1041, L3MC:DVP_VALID=1, L3MC:DVP=0x3792, ENTRY_TYPE=7>
Int v1611.10.1611.13
0000.1611.0033
Int f0/0/0.1611.10.1611.13
001a.6c70.3697
N9372-42# bcm-shell module 1 'd chg ing_13_next_hop 0x9c'
ING_L3_NEXT_HOP .ipipe0[156]:<VLAN_ID=0xffff, TID=0x83, PORT_NUM=3, MTU_SIZE=0x3fff, MODULE_ID=1, L3_OIF=0x1ffff, ENTR_Y_TYPE=2, ENTRY_INFO_UPPER=3, DVP_RES_INFO=0x7f,>
```

```
Name      Ifindex Smod Unit HPort FPort NPort VPort
Eth1/51   1a006400 1    0     3    2    50    -1
```

```
N9372-42# bcm-shell module 1 'd chg egr_dvp_attribute 0x3792'
EGR_DVP_ATTRIBUTE.epipe0[14226]:<VXLAN:TUNNEL_INDEX=1, VXLAN:DVP_IS_NETWORK_PORT=1, VXLAN:DIP=0xa0000064, VXLAN:DELETE_VNTAG=1, VP_TYPE=2,>
```

UUC-unknown unicast UMC -unknown multicast
BC-Broadcast

ENCAP ID's last digits are NHI
HG2 is E1/51XE16 is E1/17 so correct
Also, Ing REP uses/install only one of ECMP path
We will see multiple Encap ID for same port if we do have multiple Ing Rep peers

```
N9372-42# bcm-shell module 1 'd chg ing_dvp_table 0x3792'
ING_DVP_TABLE.ipipe0[14226]: <VP_TYPE=3, NEXT_HOP_INDEX=0x9c, NETWORK_PORT=1, ECMP_PTR=0x9c, DVP_GROUP_PTR=0x9c,>
N9372-42# bcm-shell module 1 'd chg ing_13_next_hop 0x9c'
ING_L3_NEXT_HOP .ipipe0[156]:<VLAN_ID=0xffff, TID=0x83, PORT_NUM=3, MTU_SIZE=0x3fff, MODULE_ID=1, L3_OIF=0x1ffff, ENTR_Y_TYPE=2, ENTRY_INFO_UPPER=3, DVP_RES_INFO=0x7f,>
```

```
Name      Ifindex Smod Unit HPort FPort NPort VPort
Eth1/51   1a006400 1    0     3    2    50    -1
```

```
N9372-42# bcm-shell module 1 'd chg egr_ip_tunnel 1'
EGR_IP_TUNNEL.epipe0[1]: <TUNNEL_TYPE=0xb, TTL=0xff, SIP=0xa000002a, L4_DEST_PORT=0x12b5, ENTRY_TYPE=1, DSCP_SEL=1,>
```

Verified correct TUNNEL SRC 10.0.0.42(0x2a) DST 10.0.0.100(0x64)

DECAP would be just same as UNICAST DECAP

VXLA EVPN Multicast Encap

Multicast HW Encap is very similar to Ing Rep as we just check UUC/UMC/BC index from VFI



Interface	VNI	Multicast-group	State	Mode	Type	[BD/VRF]	Flags
nve1	16010	230.0.160.1	Up	CP	L2	[1601]	SA

```
N9372-42# sh nve vni 16010
Interface VNI      Multicast-group      State Mode Type [BD/VRF]      Flags
-----  -----  -----  -----  -----  -----  -----  -----
nve1      16010    230.0.160.1      Up    CP    L2 [1601]      SA

N9372-42# sh ip mroute 230.0.160.1 10.0.0.42
IP Multicast Routing Table for VRF "default"
(10.0.0.42/32, 230.0.160.1/32), uptime: 00:44:17, nve pim mrib 1p
  Incoming interface: Loopback0, RPF nbr: 10.0.0.42
  Outgoing interface list: (count: 1)
    Ethernet1/51, uptime: 00:43:56, pim
```

```
N9372-42# bcm-shell module 1 'ipmc table show'
SRC IP ADDRESS MC IP ADDRESS MC GROUP VID VRF COS HWMIDX CLASS HIT
10.0.0.42      230.0.160.1      0x20000009      0      1      0      85538      1      no
```

```
N9372-42# bcm-shell module 1 'd chg ipmc 0x9'
L3_IPMC.pipe0[9]: <VALID=1,REPL_HEAD_BASE_PTR=0xa9,L3_BITMAP=0x00000000000000000000000000000008,>
```

```
N9372-42# bcm-shell module 1 'pbmp 0x8'
```

```
0x00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000008 ==> hg2
```

```
module-1# show system internal iftmc info vlan 1601 | i vni
vni = 16010, ipmc_idx = 0x001fd0 ipmc_remap_idx = 0xc001fcf dvif: 0eid[0] = 0, alt_eid[0] = 0eid[1] = 0,
alt_eid[1] = 0eid[2] = 0, alt_eid[2] = 0eid[3] = 0, alt_eid[3] = 0eid[4] = 0, alt_eid[4] = 0eid[5] = 0,
alt_eid[5] = 0eid[6] = 0, alt_eid[6] = 0eid[7] = 0, alt_eid[7] = 0
```

```
N9372-42# bcm-shell module 1 'mc show' | b 1fd0
```

```
Group 0xc001fd0 (VXLAN)
  port hg2, encap id 400087
  port xe16, encap id 400020
```

XE16=E1/17 and interface to HOST

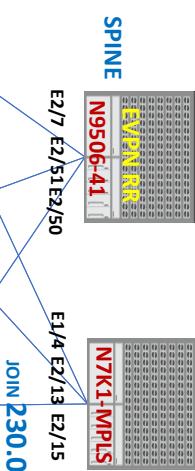
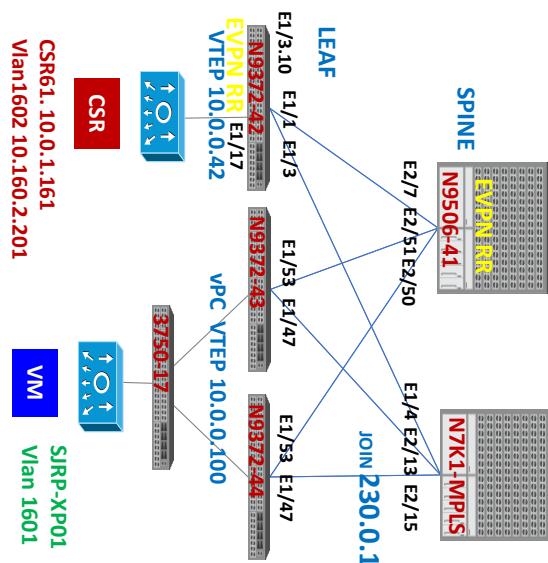
MAC to MC address translation → 230.0.160.1

```
N9372-42# bcm-shell module 1 "d chg egr_l3 next_hop 87"
EGR_L3_NEXT_HOP.epipe0[87]:<L3MC:MAC_ADDRESS=0x0005e00a001,L3MC:L3MC_USE_CONFIGURED_MAC=1,L3MC:L2_MC_VLAN_DISABLE=1,L3MC:L2_MC_SA_DISABLE=1,L3MC:L2_MC_DA_DISABLE=1,L3MC:INTF_NUM=0xffff,L3MC:DNP_VALID=1,L3MC:DNP_TYPE=0x37ba,ENTRY_TYPE=7>
```

VXLA EVPN Multicast DECAF

LEAF43/44 received multicast so how we check?

In VPC, only one DECAPPER is elected by vPC RPF source election, like normal L3 MU source in VPC



```
N9372TX-43# sh ip pim internal vpc rpf-source
Source: 10.0.0.42
Pref/Metric: 110/3
Source role: primary
Forwarding state: Tie (forwarding)
MRIB Forwarding state: forwarding
```

```
N9372TX-44# sh ip pim internal vpc rpf-source
Source: 10.0.0.42
Pref/Metric: 110/3
Source role: secondary
Forwarding state: Tie (not forwarding)
MRIB Forwarding state: not forwarding
```

(10.0.0.42/32, 230.0.160.1/32), uptime: 5w6d, ip(0) mrib(0) pim(0) nve(1)

Data Created: Yes

VXLAN Flags

VXLAN Decap

VPC Flags

RPF-Source Forwarder

Stats: 27446/2884636 [Packets/Bytes], 27.333 bps

Stats: Active Flow

Incoming interface: Ethernet1/51, RPF nbr: 10.41.44.41

Outgoing interface list: (count: 1)

nve1, uptime: 5w6d, nve

```
N9372TX-44# sh ip mr 230.0.160.1 10.0.0.42 detail
(10.0.0.42/32, 230.0.160.1/32), uptime: 5w6d, ip(0) mrib(0) pim(1) nve(1)
```

Data Created: Yes

Stats: 27706/2916868 [Packets/Bytes], 27.333 bps

Stats: Active Flow

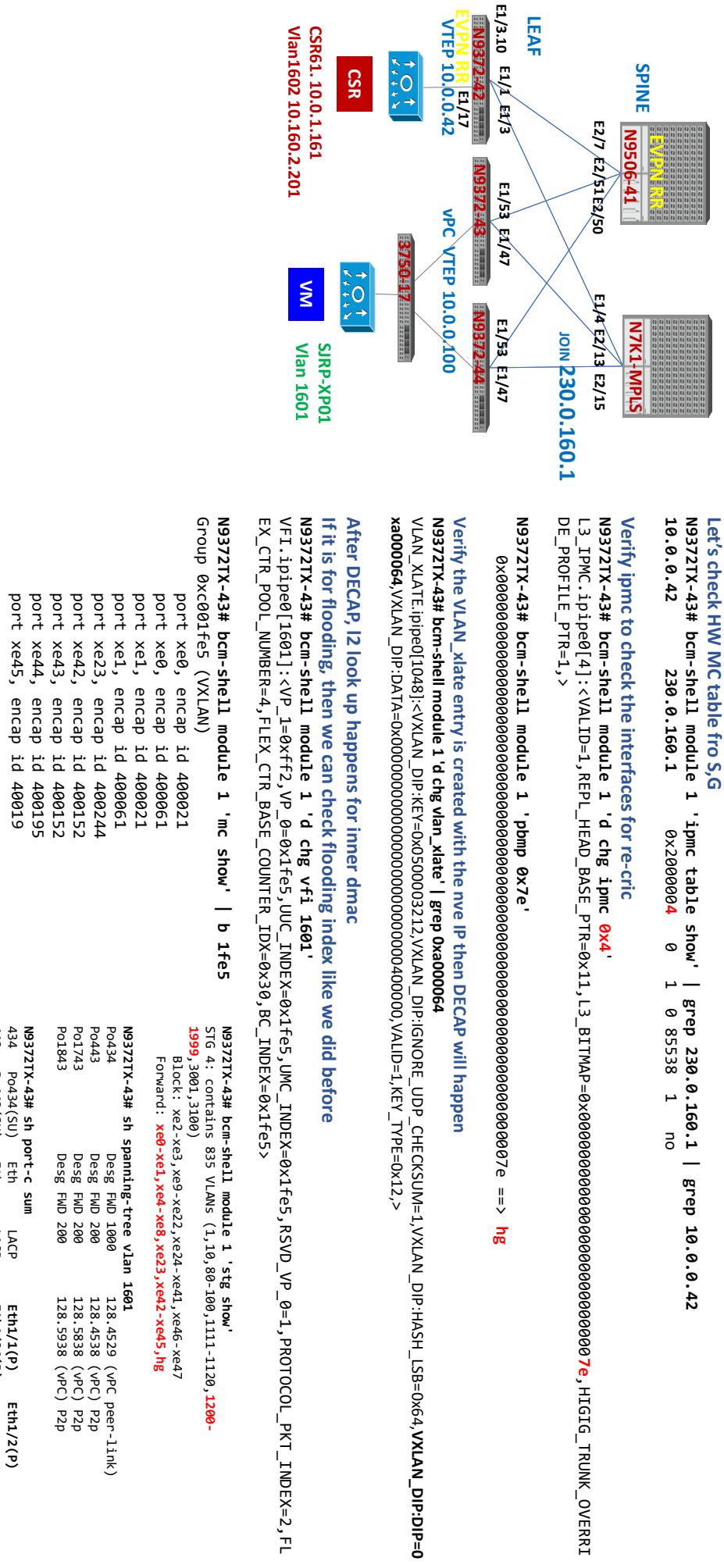
Incoming interface: Ethernet1/51, RPF nbr: 10.41.44.41

Outgoing interface list: (count: 2)

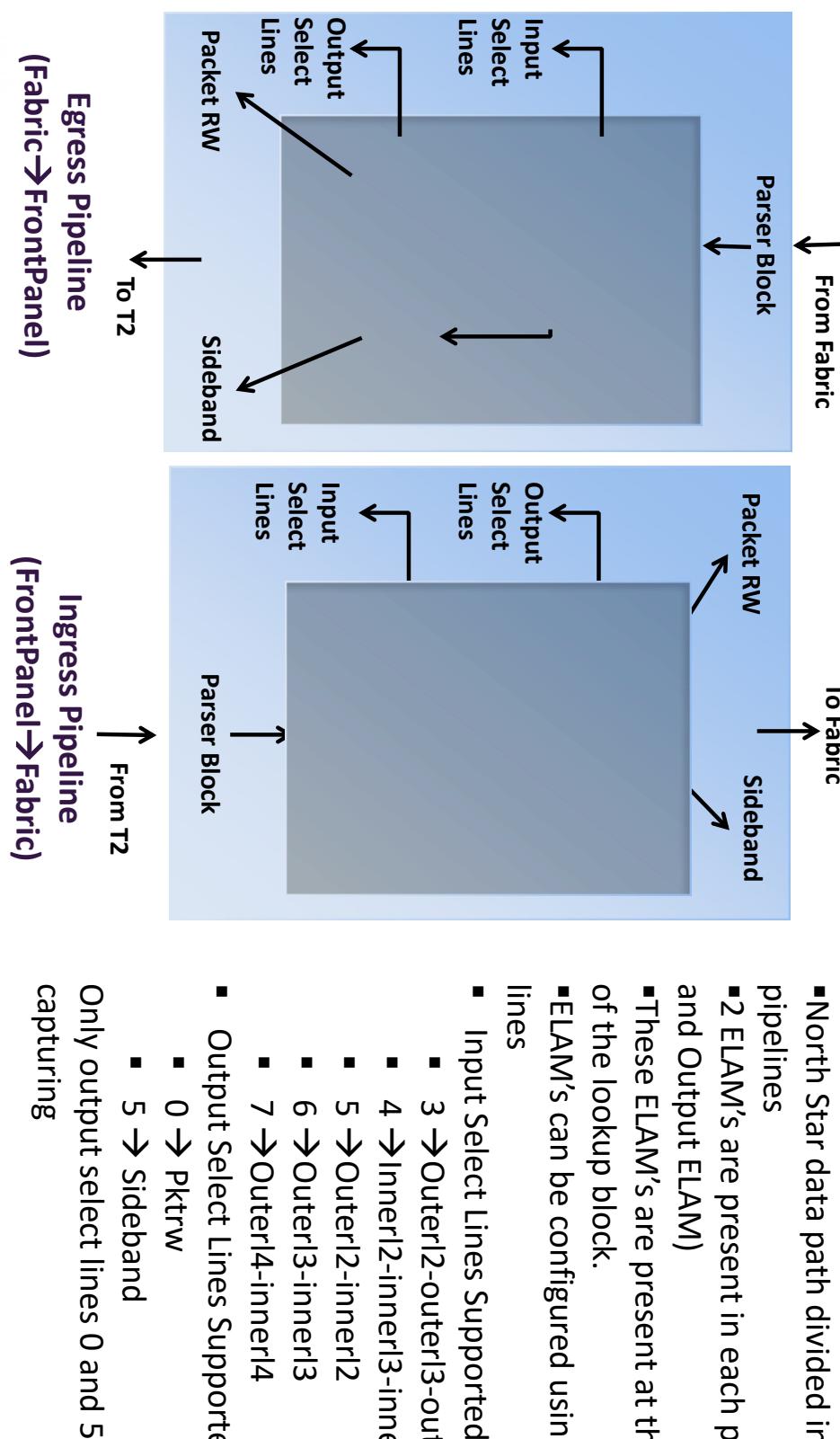
```
Ethernet1/47, uptime: 00:06:46, pim
nve1, uptime: 5w6d, nve
```

So we only care on LEAF43

VXLA EVPN Multicast DECAF



ELAM Support in North Star

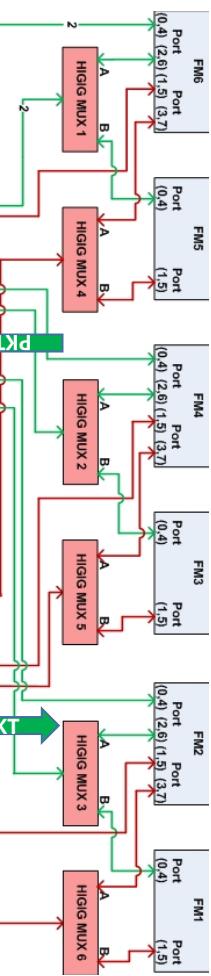


- North Star data path divided into ingress and egress pipelines
- 2 ELAM's are present in each pipeline (Input ELAM and Output ELAM)
- These ELAM's are present at the beginning and end of the lookup block.
- ELAM's can be configured using the available select lines
 - Input Select Lines Supported
 - 3 → Outerl2-outerl3-outerl4
 - 4 → Innerl2-innerl3-inner l4
 - 5 → Outerl2-innerl2
 - 6 → Outerl3-innerl3
 - 7 → Outerl4-innerl4
 - Output Select Lines Supported
 - 0 → Pktrw
 - 5 → Sideband

Only output select lines 0 and 5 are supported for capturing

ELAM on N9K-X95XX LINE

You can do it from SUP or from MOD
Also, you can use elam-wrap that will capture from all LC and all ASIC instances at one time



```
N9504-41# debug platform internal ns elam-wrap ing in-select 3 out-select 5 outer
ip4 src_ip 10.160.1.33 dst_ip 10.160.1.28 13-type 2
N9504-41# debug platform internal ns elam-wrap star
N9504-41# debug platform internal ns elam-wrap stat
=====
N9504-41# attach mod 2
module-2# debug platform internal ns elam asic 1
module-2(ns-elam-insel13)# trigger init ingress in-select 3 out-select 5
module-2(ns-elam-insel13)# set outer ip4 src_ip 10.160.1.33 dst_i 10.160.1.13 13-
type 2
module-2(ns-elam-insel13)# start/stat/reset → don't forget reset when u retry ELAM
=====
I capture ICMP ECHO from 10.160.1.33 to 10.160.1.28.
Ingress port is E2/50, NS instance 1
N9504-41# debug platform internal ns elam-wrap report
ELAM REPORT FOR MODULE 2 INST 0
=====
ELAM not triggered. No report available
ELAM REPORT FOR MODULE 2 INST 1
=====
```

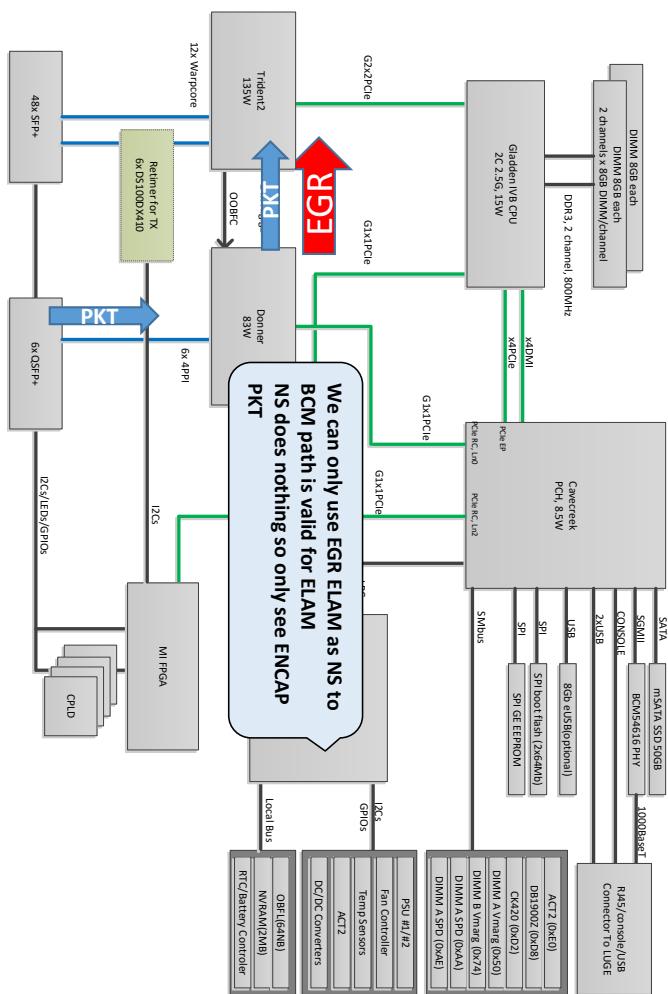
Status: Triggered

Snip...

```
GLB_C++: [MSG] - 12vec0 is complete
GLB_C++: [INFO] Decoded Dump of a Binary structure assumed to be of type: LU_VE
C_L2: [INFO] ce_da: 00077DB1f443 Address 10.160.1.28 Age 00:16:01 00077db1f443 Interface Vlan1601
GLB_C++: [INFO] ce_sa: 0000016010033
GLB_C++: [INFO] qtag0_vlan: 641
GLB_C++: [INFO] - 13vec0 is complete
GLB_C++: [INFO] Decoded Dump of a Binary structure assumed to be of type: LU_VE
C_L3: [INFO] ip_ttl: FF
GLB_C++: [INFO] ip_prot: 01 → ICMP
GLB_C++: [INFO] ip_csum: F01A
GLB_C++: [INFO] ip_da: 00000000000000000000000000000000 → 10.160.1.28
GLB_C++: [INFO] ip_sa: 00000000000000000000000000000000 → 10.160.1.33
```

ELAM on N9372

Osl System Block Diagram (3/12/2014)



I capture ICMP ECHO from 10.160.1.33 to 10.160.1.28 on LEAF42.

Ingress port is E1/51, NS instance 0, Egress port is E1/17

On TOR, only EGR ELAM is possible and see ENCAPPED PKT so see Outer IP and Outer MAC

N9372-42# debug platform internal ns elam-wrap egress in-select 3 out-select 5 outer
N9372-42# debug platform internal ns elam-wrap star
N9372-42# debug platform internal ns elam-wrap report
ELAM REPORT FOR MODULE 1 INST 0
=====

GLB_C++: [MSG] - 12vec0 is complete

GLB_C++: [INFO] Decoded Dump of a Binary structure assumed to be of type: LU_VE

GLB_C++: [INFO] 13_type: 2

GLB_C++: [INFO] ip_ttl: FF

GLB_C++: [INFO] ip_prot: 11 → 17 UDP

GLB_C++: [INFO] ip_pklen: 0054

GLB_C++: [INFO] ip_da: 0000000000000000A000002A

GLB_C++: [INFO] ip_sa: 00000000000000000000000000000029

GLB_C++: [MSG] - 14vec0 is complete

GLB_C++: [INFO] Decoded Dump of a Binary structure assumed to be of type: LU_VE

C_L4: [INFO]

GLB_C++: [INFO] 14_type: 1

GLB_C++: [INFO] 14_sp: 63B8

GLB_C++: [INFO] 14_dp: 12B5

GLB_C++: [INFO] 14_len: 0082

Ping from IOS adds 50 bytes but
NXOS adds 78 bytes

How to match interested packet → Trick is matching ping data size+50 byte in ELAM pkt len
3750X-28#ping vrf host1601 10.160.1.33 size 200 repeat 100000000

N9372-42# debug platform internal ns elam-wrap egress in-select 3 out-select 5 outer
N9372-42# debug platform internal ns elam-wrap report | i len
GLB_C++: [INFO] ip_pklen: 00FA → 250=200+50 (50 byte VXLAN overhead)

IOS send 20B IP+8B ICMP+172B Data = 200B

NXOS send 20B IP+8B ICMP+200B Data = 228B

VXLAN ENCAP Header = 50 or 54 Bytes depends on Dot1Q tag

= 14B Eth + 4B DOT1Q + 20B Outer IP+ 8B UDP+ 8byte VXLAN

So, it can be +50, +54, +78 or +82B with 200 ping size option

cisco

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Dynamic Fabric Automation -IP Fabric (EVPN)

Cisco Data Center Networks- Choices in Automation and Programmability

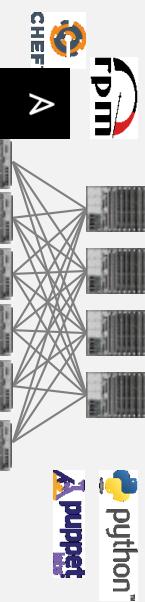
Application Centric Infrastructure



Programmable Fabric



Programmable Network



Turnkey integrated solution with security, centralized management, compliance and scale

FabricPath w/BGP
VxLAN-BGP EVPN
standard-based

3rd party controller support

Cisco Controller for software
overlay provisioning and
management across N2K-N9K

Modern NX-OS with enhanced NX-APIs

DevOps toolset used for Network Management

(Puppet, Chef, Ansible etc.)

Broad and deep ecosystem

Automation, API's, Controllers and Tool-chain's

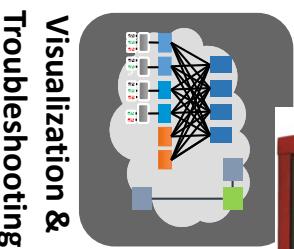
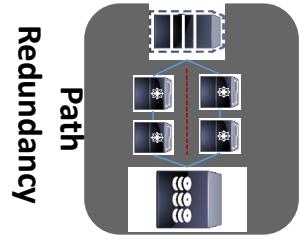
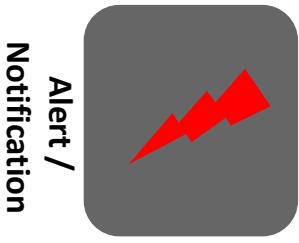
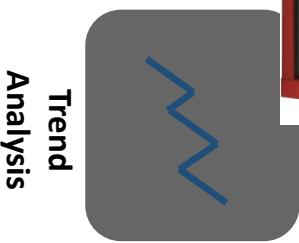
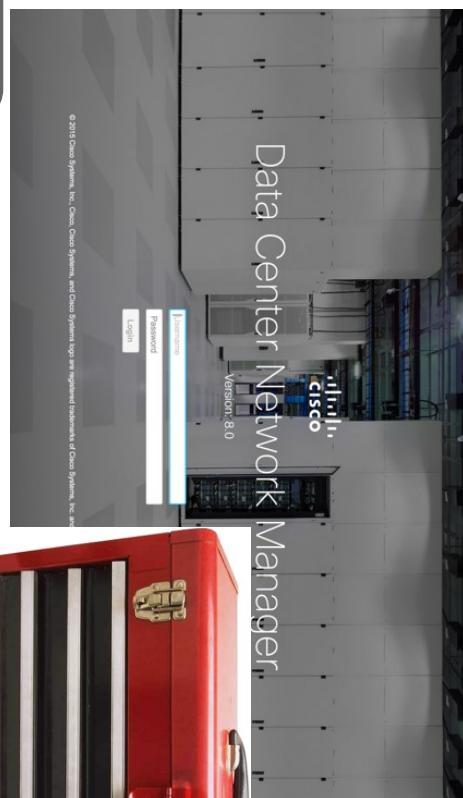
3 Automation Solutions from CISCO

- Simplifies cloud operations and automates network configuration for lower Opex in next-generation data center fabrics

	Dynamic Fabric Automation	Nexus Fabric Manager	Virtual Topology System
Overlay Support	FabricPath with BGP VXLAN F&L VXLAN MP-BGP EVPN	VXLAN MP-BGP EVPN	VXLAN F&L (only on N3K) VXLAN MP-BGP EVPN
Platform Support	N2K/N5K/N6K/N7K/N9K	N2K/N9K (N7K in future)	N2K/N3K/N5K/N6K/N7K/N9K
Day 0 Support	- Built-in POAP Server - Underlay and BGP EVPN is configured from POAP Template	- POAP is supported by external server - Automated VTEP discovery and underlay configuration	Need to manually configure or use NX-API - Underlay IGP and Multicast - NVE, BGP EVPN
Auto Provision Overlay Network	- GUI or REST/NX-API	- GUI or REST/NX-API	- GUI, Openstack, VMWARE
Host Trigger	- DOT1Q, VDP, VMTRACKER, LLDP, Auto-Pull (GUI)	- GUI or REST/NX-API - Vmtracker in future	- OPENSTACK Plugin - VMWARE Plugin

What is DCNM

Multi-Device, Multi-Fabric, Multi-Tenant LAN, SAN, Fabric Management TOOL for Nexus & MDS



Automation
Health Monitoring,
Inventory & Diagnostics

Visualization &
Troubleshooting

Configuration

Trend Analysis

Alert / Notification

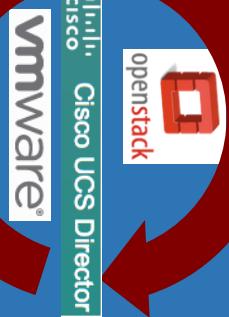
Path Redundancy

Automation for Fabric

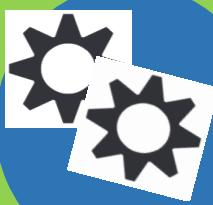
Power-On Auto Provisioning [POAP]



Auto-Configuration



REST /SMI-S APIs



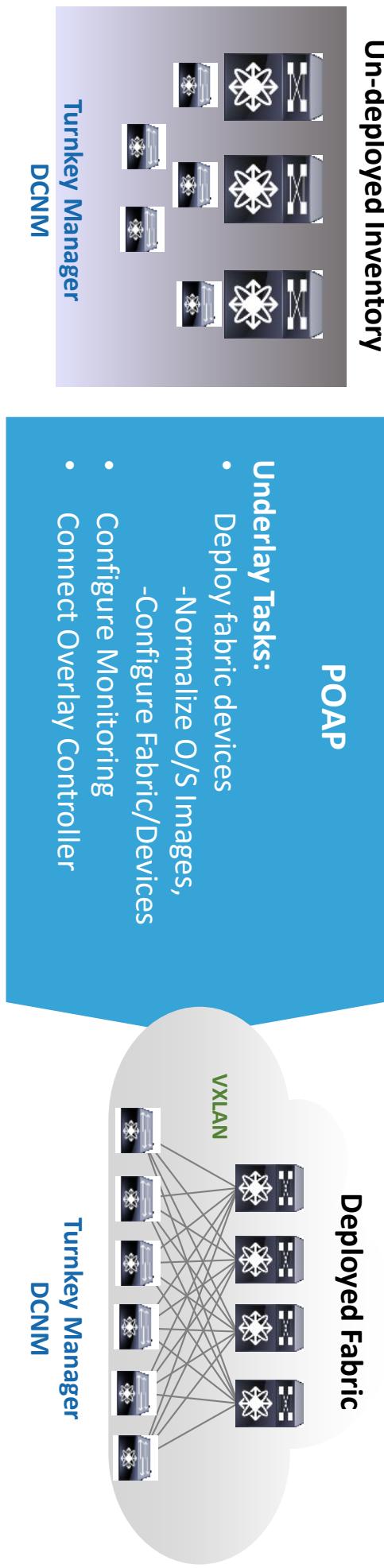
Day 0 bring up of new Fabrics

Multi-tenant VM & PM* Network Provisioning

Customizable Automation via REST

Day 0 Operations: Install, Provision, Deploy

Challenge: Deploy Fabric with many devices with minimal OPEX and errors

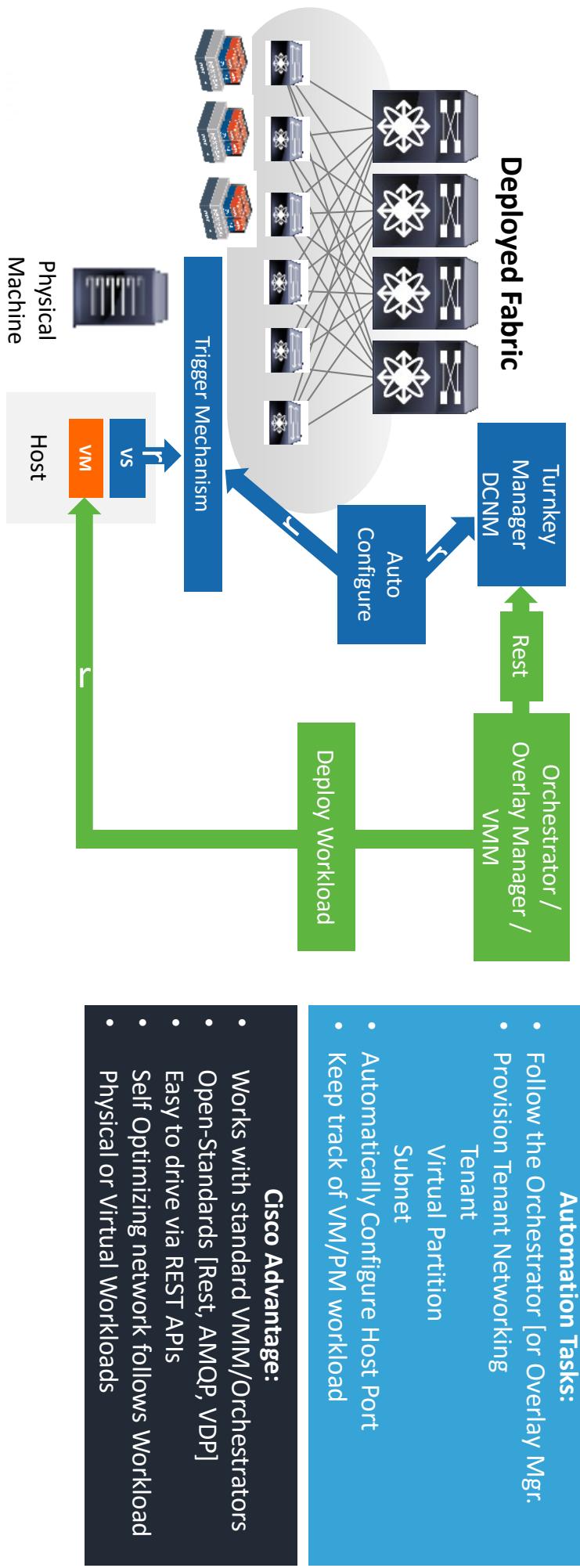


Cisco Advantage:

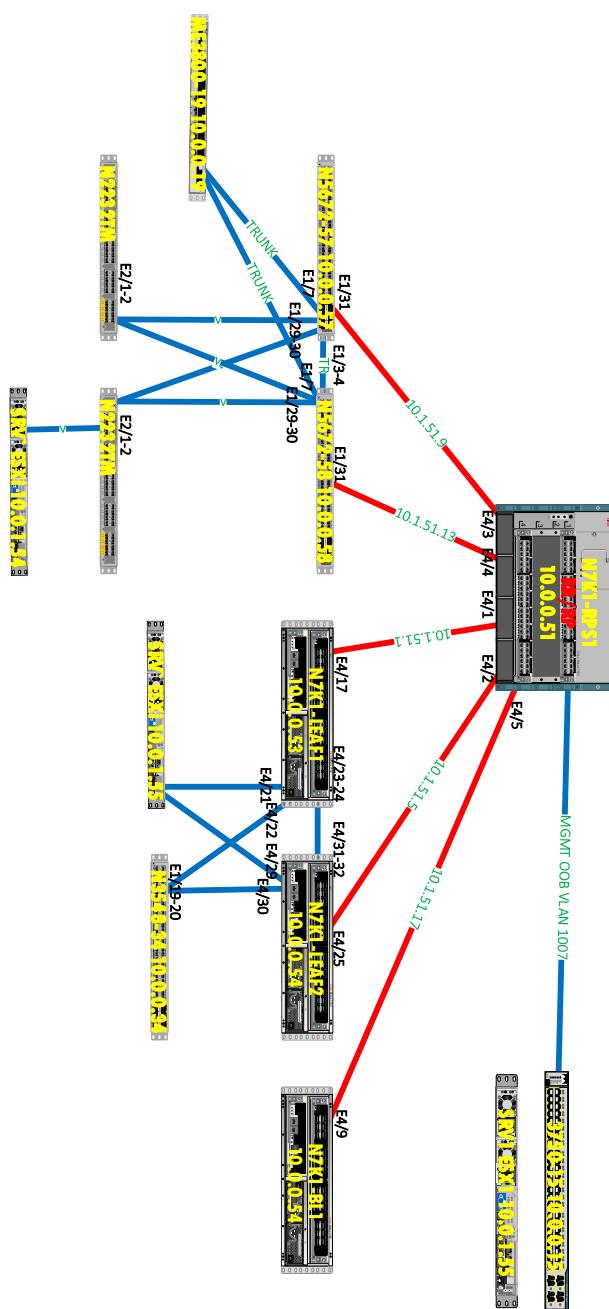
- Turnkey management infrastructure
- Makes large fabrics easy to deploy with reduced Errors

Automate – Fabric Automation with VXLAN/FP

Challenge: Automatically deploy Tenant Overlay Network upon Workload/VM deployment



SJRP DFA EVPN LAB

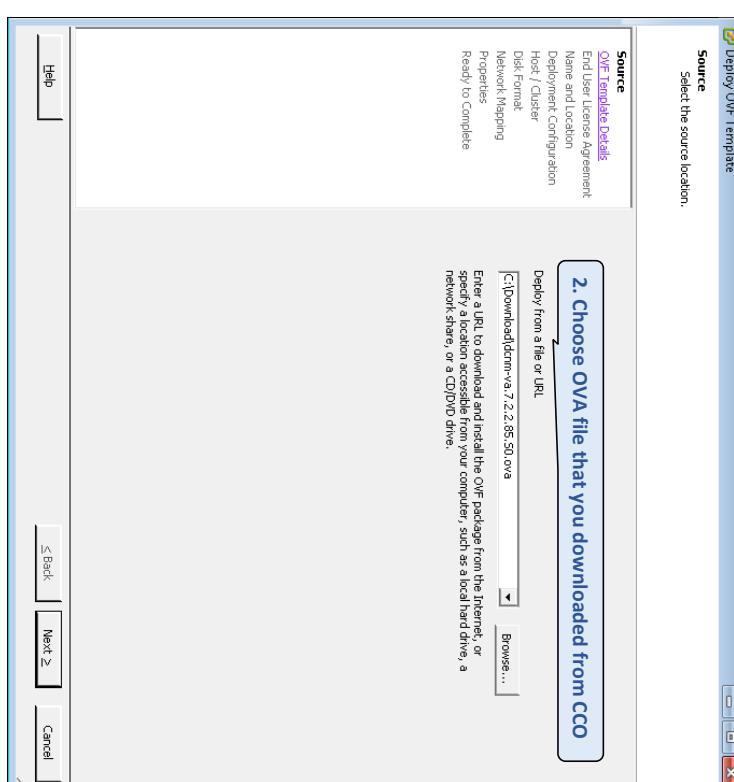


DCNM OVA Deployment on VMWARE

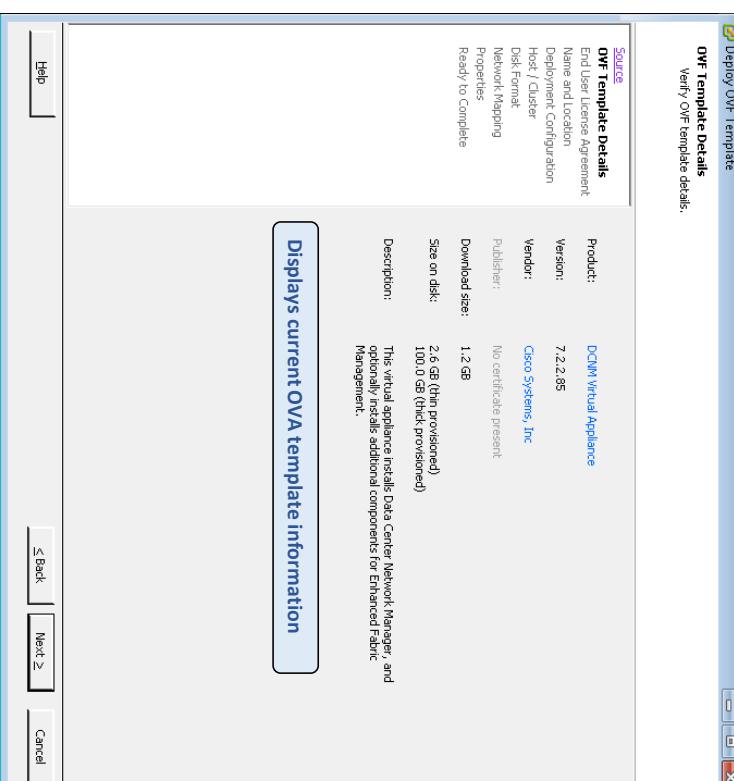
1. From Vcenter, click "File" → "Deploy OVF Template"



2. Choose OVA file that you downloaded from CCO



3. Displays current OVA template information



DCNM OVA Deployment on VMWARE

Deploy OVF Template

Name and Location
Specify a name and location for the deployed template

Source: Deploy, OVF Template
OVF Template Details: jprodmc
End User License Agreement:

Name and Location
Deployment Configuration
Host / Cluster
Disk Format
Network Mapping
Properties
Ready to Complete

3. Type VM Name and choose VC Datacenter to deploy OVA

Name: jprodmc
The name can contain up to 60 characters and it must be unique within the inventory folder.

Inventory Location: SRRP-DC1

Deployment Configuration
Select a deployment configuration.

Source: Deploy, OVF Template
OVF Template Details: jprodmc
End User License Agreement:

Name and Location
Deployment Configuration
Host / Cluster
Disk Format
Network Mapping
Properties
Ready to Complete

4. Choose deployment size

Configuration: Small

Use this deployment option to configure a small version of appliance with 2vCPUs and 8g RAM.

Deployment Configuration
Configuration: Small

Source: Deploy, OVF Template
OVF Template Details: jprodmc
End User License Agreement:

Name and Location
Deployment Configuration
Host / Cluster
Disk Format
Network Mapping
Properties
Ready to Complete

DCNM OVA Deployment on VMWARE

The image displays two consecutive screens from the VMware 'Deploy OVF Template' wizard.

Screenshot 1: Step 5 - Choose Which ESXi Host to deploy OVA

This screen shows the 'Host / Cluster' selection step. The user has chosen 'SRP-DC1' as the host. The interface includes tabs for 'Source', 'OVF Template Details', 'End User License Agreement', 'Name and Location', and 'Deployment Configuration'. A large blue box highlights the '5. Choose Which ESXi Host to deploy OVA' step. Navigation buttons at the bottom are labeled 'Help', '< Back', 'Next >', and 'Cancel'.

Screenshot 2: Step 6 - Choose Disk Format

This screen shows the 'Disk Format' selection step. It lists 'Datastore: datastore1' and 'Available Space (GB): 587.4'. Under 'Disk Format', the 'Thick Provision Lazy Zeroed' option is selected. Other options include 'Thick Provision Eager Zeroed' and 'Thin Provision'. A large blue box highlights the '6. Choose Disk Format' step. Navigation buttons at the bottom are labeled 'Help', '< Back', 'Next >', and 'Cancel'.

DCNM OVA Deployment on VMWARE

Deploy OVF Template

Network Mapping
What networks should the deployed template use?

Source
OVF Template Details
End User License Agreement
Name and Location
Deployment Configuration
Host / Cluster
Storage
Disk Format
Network Mapping
Properties
Ready to Complete

Map the networks used in this OVF template to networks in your inventory

Source Networks
dcmn-mgmt
enhanced-fabric-mgmt
DFAMGMT
VM Network
Destination Networks
VM Network
Storage
Disk Format
Network Mapping

7. Map network to NIC
DCNM-MGMT provides connectivity(ssh,scp,http,https) to the DCNM server.
Enhanced fabric management provides POAP connectivity and OOB mgmt access to N9K/N7K/N5K switches

You can use same interface for both networks if you want

Properties
Ready to Complete

Deploy OVF Template

Properties
Customize the software solution for this deployment.

Source
OVF Template Details
End User License Agreement
Name and Location
Deployment Configuration
Host / Cluster
Storage
Disk Format
Network Mapping
 Properties
Ready to Complete

Application Management
Select this option to install the applications for Enhanced Fabric Management along with DCNM. This includes implementations for the following protocols: XNMP, LDAP, DHCP, AMQP. There are repositories that could be accessed via the following protocols: TFTP, SFTP, SCP.

Management Properties

Administration, Password
Enter the root password for SSH, Console and all the selected applications. Must be at least 8 characters in length and must contain at least an alphabet and a number. It can contain only the special characters .(dot), +(plus), -(underscore) and -(hyphen)

Enter password

Confirm password

DCNM Network (All fields are required)

Hostname
Enter the hostname as a Fully Qualified Domain Name, for example: myhost.cisco.com
spdcnm1.sjpt.local

IP Address
10 . 0 . 1 . 10

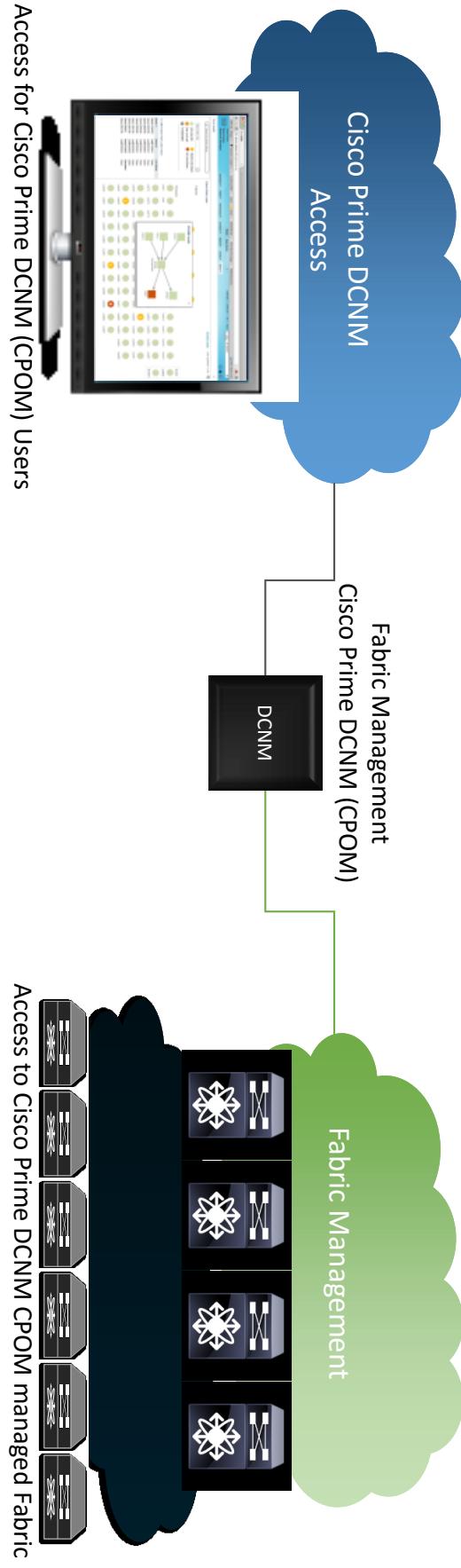
Properties with invalid values will be left unassigned. The vApp will not be able to power on until properties have valid values.

Help < Back Next > Cancel

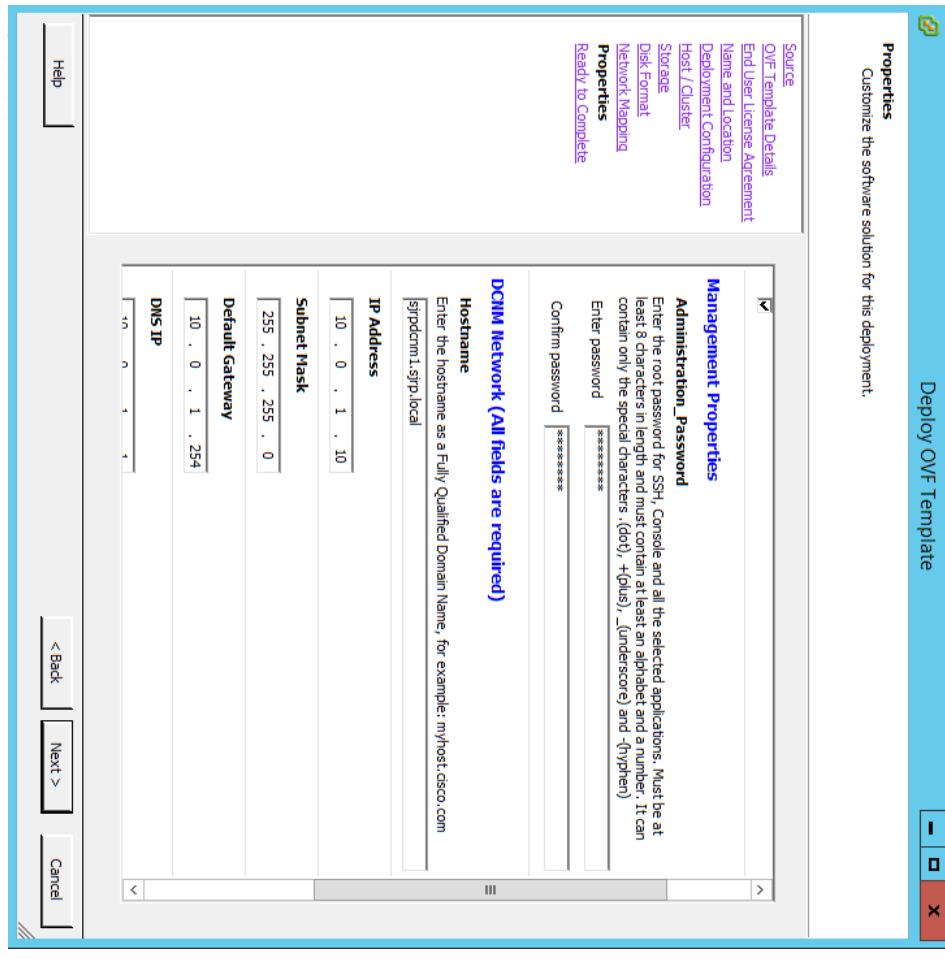
Help < Back Next > Cancel

Connecting DCNM (CPOM) for DFA

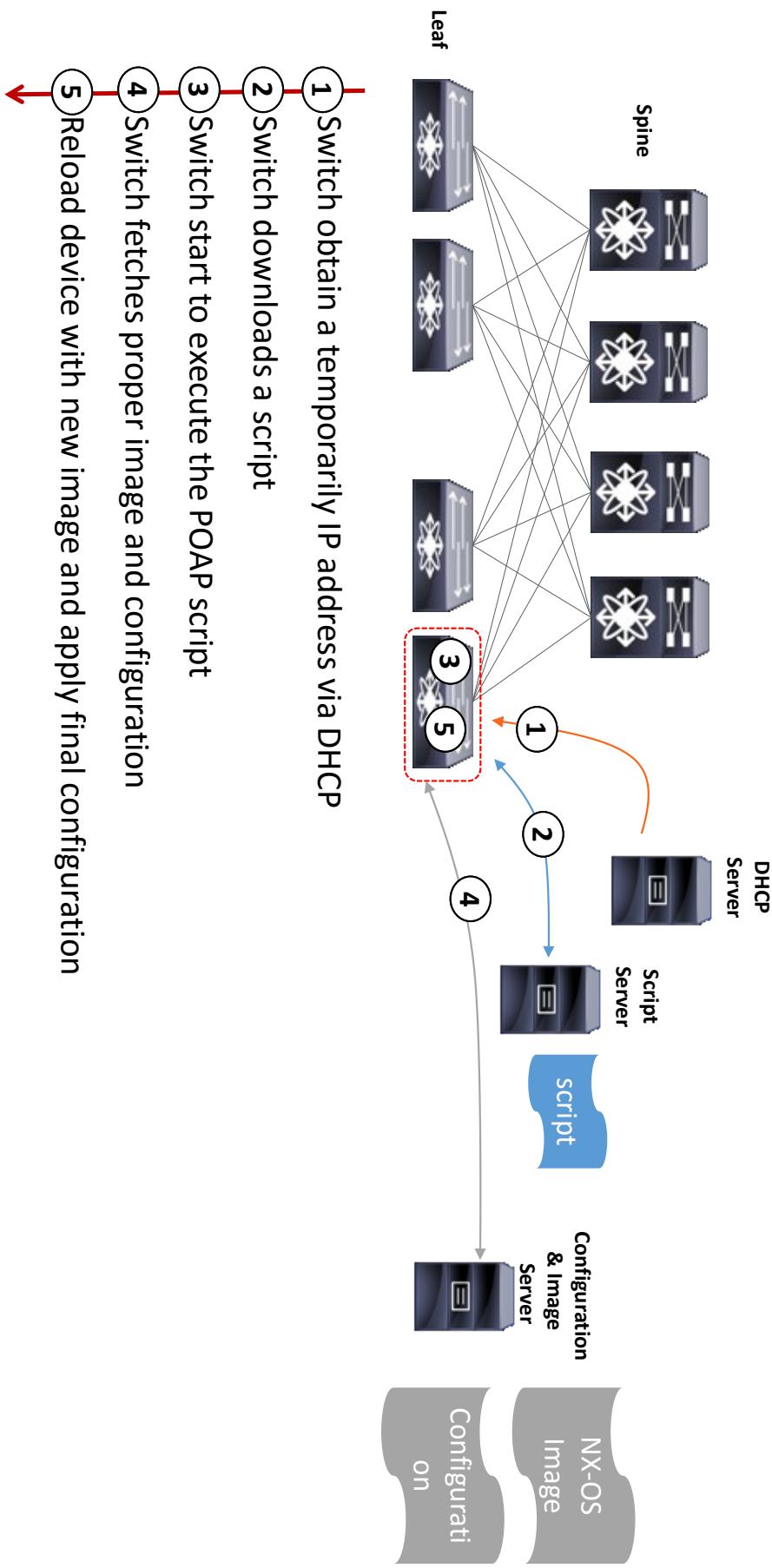
- **DCNM Access** for User-Access to the DCNM (CPOM) WebUI or DCNM via Fat-Client
 - Configuration of this Interfaces requires:
 - IP Address
 - Netmask
 - Gateway
 - DNS-Server
- **Fabric Management** for Access to the Network Switch Out-of-Band Management (mgmt0)
 - Configuration of this interface requires:
 - IP Address
 - Netmask
 - DNS-Server



DCNM OVA Deployment on VMWARE



Power on Auto-Provisioning (POAP)

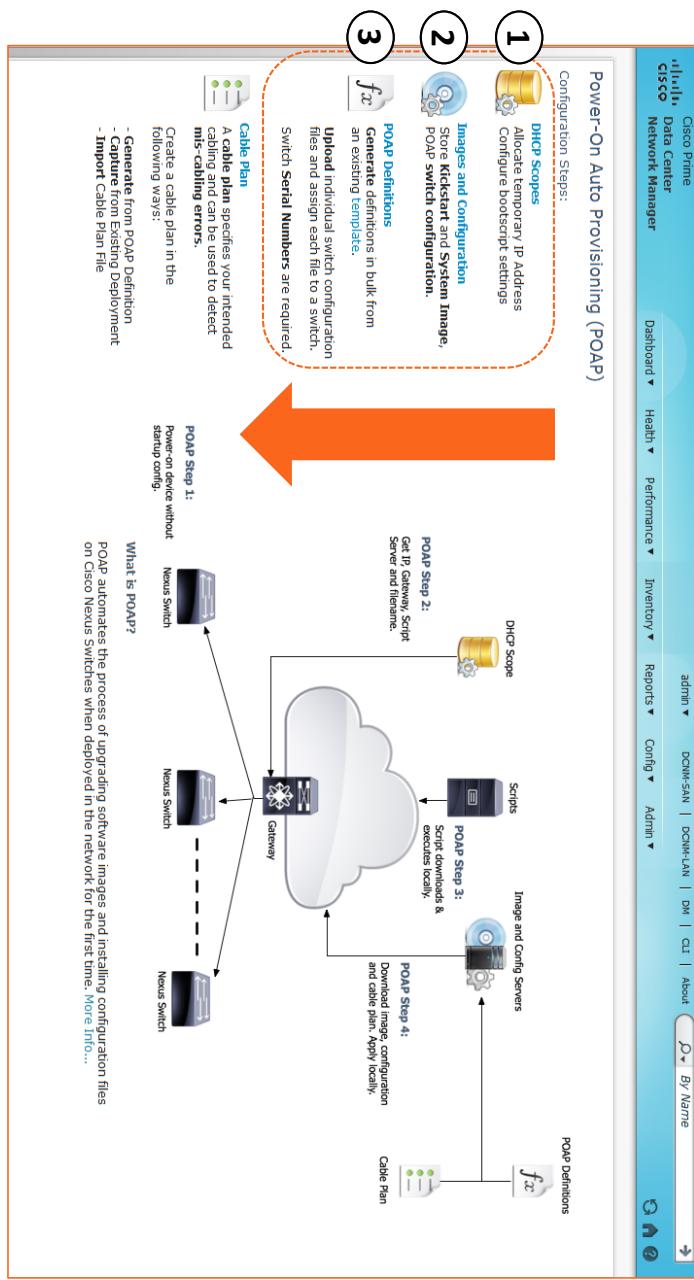


POAP with DCNM

How to Setup

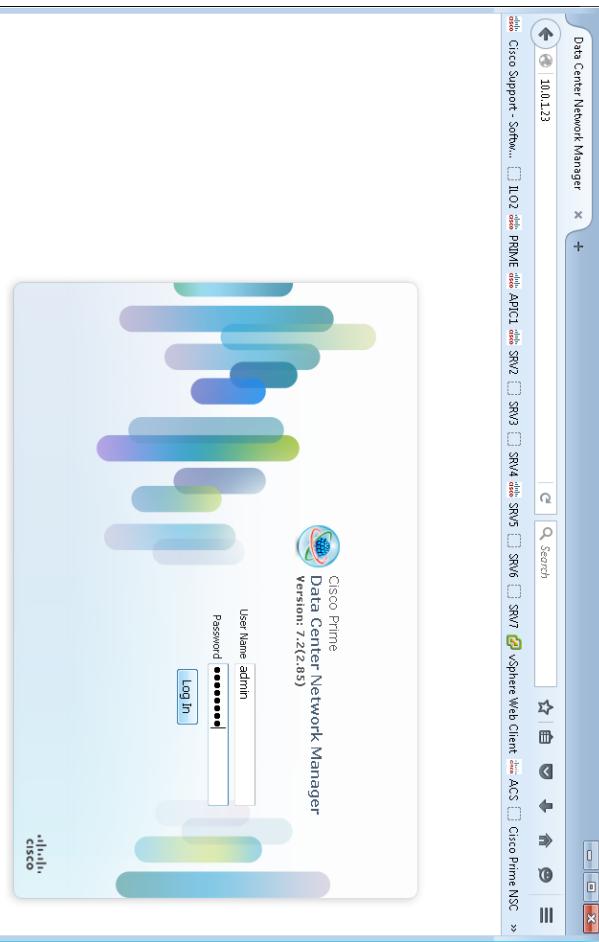
3 Major Steps

- DHCP Scope
- Images and Configuration
- POAP Definition



DCNM POAP

Let's login into DCNM GUI



DCNM - 10.1.13

Cisco Support Software: 10.0.2 PRIME APICL SRV2 SRV3 SRV4 SRV5 SRV6 SRV7 vSphere Web Client ACS Cisco Prime NSC

Welcome to DCNM

DCNM Licenses

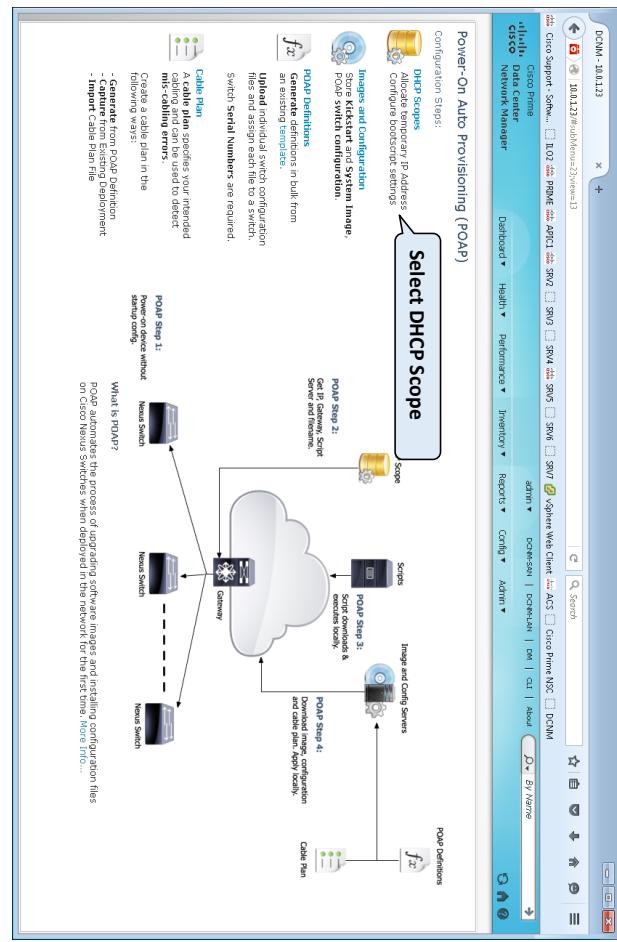
Power-On Auto Provisioning

Performance Collection

Documentation

DCNM POAP – DHCP SCOPE

Before we do use POAP, need to modify default DHCP scope



Edit DHCP Scope

Scope Name:	enhanced_fabric_mgmt_scope
Scope Subnet:	10.0.7.0/24 (E.g.:1.1.1.1.1.0/24)
IP Address Range:	10.0.7.200-10.0.7.230 (E.g.:1.1.1.1.1.0-1.1.1.1.1.5.1.1.1.1.9)
Default Gateway:	10.0.7.254
Domain Name Servers:	10.0.1.1
Lease time(seconds):	3600
Bootscript Name:	poap_dcnm.py
TFTP Server:	10.0.7.254

Select DHCP Scope and Edit

This IP range is only used for POAP stage, not OOB mgmt IP of SW.
MGMT IP is assigned manually by POAP definition

DCNM POAP- Image Upload

Let's copy NXOS images into DCNM directory

Cisco Prime
Data Center
Network Manager

Dashboard ▾ Health ▾ Performance ▾ Inventory ▾ Reports ▾ Config ▾ Admin ▾

admin ▾ DCNM-SAN | DCNM-LAN

Image and Configuration Servers (1 rows)

Add Edit Delete

Name	URL	Username
Default_SCP_Repository	scp://10.0.7.10/var/lib/dcnm	root

NXOS Image location

WinSCP - DOWN - admin@10.0.1.10 - WinSCP

New Session

Local Mark Files Commands Session Options Remote Help

Session: admin@10.0.1.10

Transfer Settings Default

C:\DOWN

Name	Size	Changed
mpif.cfg	12 K	2/29/2016 3:1
n7k1-sll.cfg	108 K	FOC1836R2C06
n7000-s2-kickstart73.001.1.bin		2/29/2016 7:1
n7000-uk9-72.ova		2/29/2016 7:1
n7000-uk973.N1.1.bin		2/29/2016 7:1
n7000-s2-dl9.73.D1.1.bin		2/29/2016 7:1
n7000-s2-kickstart-npe73.001.1.bin		2/29/2016 7:1
n7000-d9.6.2.0.5.5.bin		2/29/2016 7:1
n3000-uk9-kickstart6.0.2.Ub5.5.bin		2/23/2016 3:5
voap.py	2 KB	2/23/2016 3:5
	3 KB	2/18/2016 2:33

File operations: Open, Edit With, Upload..., Upload and Delete..., Delete, Rename

Use SCP program and upload all NX switch images to this location.
Make sure you change DCNM dir to /var/lib/dcnm

You can SSH to DCNM1, 172.16.166.53

```
[root@sjrpdcnm1 ~]# pwd
/root
[root@sjrpdcnm1 ~]# ls
[root@sjrpdcnm1 dcnm]# ls
cableplan.xml      FOC1902R0E3      n7000-s2-dk9.7.3.0.D1.1.bin    SAL19017CEU
cacert.pem         FOX1844GP9Z      n7000-s2-kickstart.7.3.0.D1.1.bin  TBM13343273
dcnm-dfa.conf     golden-dcnm-dfa.conf  nxos.7.0.3.13.1.bin          TBM13343273:LEAF1
dcnm-server-list.cfg goldendhcpd.conf poap_dcnm.py            TBM13343273:LEAF2
dhcpd.conf        licenses           poap_global.settings       TBM13343273:RSPINE1
escape-lines      n6000-uk9.7.3.0.N1.1.bin   SAL18495M25
FOC1836R2Q6      n6000-uk9-kickstart.7.3.0.N1.1.bin  SAL18516QBO
voap.py
```

DCNM POAP – Fabric Encapsulation Setting

Before creating POAP definition, Need to choose either Fabric Path Encap or EVPN Enacap with PIM ASM or BIDIR mode

Admin->Fabric>Fabric Encapsulation Settings

Fabric Encapsulation

Enable Fabric Path Encapsulation

Enable VXLAN Encapsulation

Choose here either FP or EVPN

N5600 Leaf Network

N7K Leaf with PIM Bidir Network

N7K Leaf with PIM ASM Network

NSX Leaf Network

N5K Leaf Network

N5K & N9K Combined Leaf Network

Edge Router / Border Leaf / BorderPE Settings

POAP Settings

Fabric Encapsulation Settings **Click Here**

L2 Segment ID Range Manager

Multicast Group Subnet: 227.2.2.0/24

RP Subnet: 10.254.254.0 /24

Number of Rendezvous Points (RPs): 1

RP1 Multicast Group Subnet: 227.2.2.0/24

RP1 Phantom IP Address: 10.254.254.1

RP2 Multicast Group Subnet: 227.2.2.0/25

RP2 Phantom IP Address: 10.254.254.1

RP3 Multicast Group Subnet: 227.2.2.128/25

RP3 Phantom IP Address: 10.254.254.65

RP4 Multicast Group Subnet: 227.2.2.192/26

RP4 Phantom IP Address: 10.254.254.129

Number of Rendezvous Points (RPs): 4

RP1 Multicast Group Subnet: 227.2.2.0/26

RP2 Multicast Group Subnet: 227.2.2.64/26

RP3 Multicast Group Subnet: 227.2.2.128/26

RP4 Multicast Group Subnet: 227.2.2.192/26

RP1 Phantom IP Address: 10.254.254.1

RP2 Phantom IP Address: 10.254.254.65

RP3 Phantom IP Address: 10.254.254.129

RP4 Phantom IP Address: 10.254.254.193

Number of Rendezvous Points (RPs): 2

RP1 Multicast Group Subnet: 227.2.2.0/25

RP2 Multicast Group Subnet: 227.2.2.128/25

RP1 Phantom IP Address: 10.254.254.1

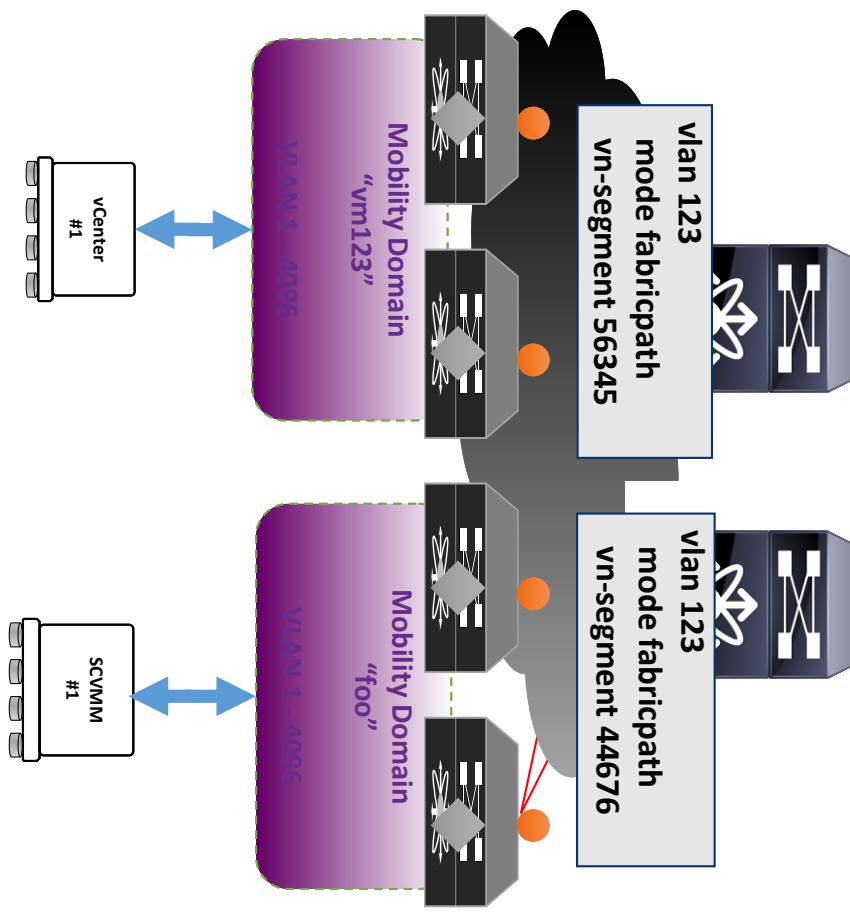
RP2 Phantom IP Address: 10.254.254.65

RP3 Phantom IP Address: 10.254.254.129

RP4 Phantom IP Address: 10.254.254.193

Mobility Domain (MD)

- A Mobility Domain (MD) defines a unique Layer-2 name-space (e.g. VLAN domain/range)
 - A Orchestrator or Virtual Machine Manager could represent a management entity for a Mobility Domain
- VLAN-IDs are shared within a Mobility Domain (MD) across multiple DFA Leafs
 - A given Leaf can only belong to one (1) Mobility Domain
 - N7K supports Multi MD on a leaf using vlan translation
- Multiple Mobility Domains can use overlapping VLAN-IDs mapped to different Segment-Ids. This all on top of the same DFA Fabric
- Mobility Domain "vm123": VLAN 123 -> segment-id 56345
- Mobility Domain "foo": VLAN 123 -> segment-id 44676
- Mobility Domain (MD) requires 802.1q header transport between Host (Server) and the Leaf
 - Mobility Domain uses only MAC learning as network auto-configuration trigger



Vinci Multi-Mobility Domain Auto-config Illustration

Switch configuration

dynamic server vlan range = 100-110
 translation range = 10-111
 Global mobility domain = MD0

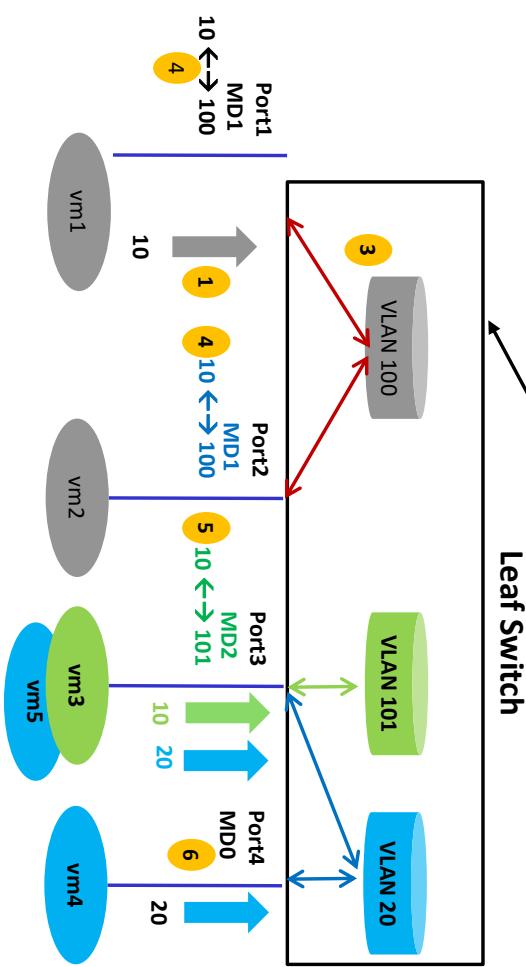
Mobility domain	Interfaces	Detectable Range
MD0	Port4	12-20
MD1	Port1, Port2	10
MD2	Port3	10,20

Auto-configuration entries in remote DB (LDAP)

Mobility domain	VLAN	Segment ID	Org/Partition
MD0	20	9999	Coke/Eng
MD1	10	10000	Pepsi/Eng
MD2	10	10001	IBM/Eng

Vlan Translation Auto-Configuration

1. Packet arrives on interface
2. Lookup vlan 10 + MD1 in remote DB
3. Determine the “to” vlan to use for translation. Pick “free” vlan from dynamic server vlan range
4. Program vlan translation (**switchport vlan mapping 10 100**) on all interfaces of the mobility domain.
5. Similar process as above for vlan 10 packet arriving on MD2 interface. Note that the “to” vlan is different, to provide the traffic segregation
6. Auto-configuration of a global vlan, 20. Vlan is available on all interfaces, except on MD1 where it is not part of the detectable range



DCNM POAP – Mobility Domain for N7K

N6K uses global default md0 but N7K needs to define MD
For N6K/N7K compatibility, create md0 for N7K

The screenshot shows the Cisco DCNM interface for a Cisco Prime Data Center Network Manager. The top navigation bar includes links for Admin, Dashboard, Health, Performance, Inventory, Reports, Config, and Admin. A search bar at the top right is set to 'By Name'.

The main content area displays a table titled 'Mobility Domains (1 rows)'. A single row is listed under 'Mobility Domain' with the value 'md0'. Below the table are buttons for 'Add', 'Edit', 'Delete', and 'Import/Export'.

A callout box labeled 'Click ADD' points to the 'Add' button. Another callout box labeled 'Click this' points to the 'Edit Mobility Domain' dialog box.

The 'Edit Mobility Domain' dialog box contains fields for 'Mobility Domain Name' (set to 'md0') and 'Detectable VLAN Range' (set to '1200-2499'). It also includes 'L2 Segment ID Range Management' and 'Mobility Domains' sections. At the bottom are 'OK' and 'Cancel' buttons.

A large callout box at the bottom right contains the following text:

- Must type "md0"
- This is only required for N7K.
- N6K/N9K uses default global mobility domain "md0" automatically

DCNM POAP – General Settings

Define system and core dynamic vlan in general setting

The screenshot shows the DCNM General Settings page under Admin > Fabric > General Settings. The left sidebar lists various configuration categories like General, Performance, Event, Federation, Management Users, and Fabric. Under Fabric, 'General Settings' is selected and highlighted with a blue box. A callout bubble points to this selection with the text 'Select this'.

The main configuration area is for LDAP settings:

- LDAP Server: 10.0.7.10
- LDAP User Name: cn=admin,dc=cisco,dc=com
- LDAP Password: *****
- Segment ID Range: 30000-49999
- Partition ID Range: 50000-60000
- Use SSL for DCNM to LDAP communication

A callout bubble points to the Segment ID Range input field with the text 'Default values for VN Segment ID Range'.

Below the LDAP section, there's a 'DHCP' section with a checkbox for 'Use local DHCPd for Fabric (changes cannot be undone)'. A callout bubble points to this checkbox with the text 'Uncheck this if you want to use outside DHCP server'.

At the bottom, there are sections for 'L2 Segment ID Range Management' and 'Other Configs'. In 'Other Configs', there's a checkbox for 'Use Top-Down Approach for Delegation management'. A callout bubble points to this checkbox with the text 'Match with previous mdo0 setting'.

Core Dynamic Vlan – used for L3 VNI for VRF management
System Dynamic VLAN – super range of VLAN available to DFA for dynamic mapping
Every VLAN from 2500 to 2999 will be used for the Layer-3 (VRF) to Segment-ID mapping.
start using VLAN 3000 and above for Layer-2 VLAN to Segment-ID mapping

Global Mobility Domain Name: mdo
Global Mobility Domain Detectable VLAN Range: 1200-2499
Global Anycast Gateway MAC: 2020.0000.00AA

Can't overlap system vlangs

DCNM POAP

Let's create a POAP definition for N6K

The screenshot shows two stacked windows of the DCNM interface, both titled "DCNM - 10.0.1.23".

Top Window: The title bar shows "POAP Steps". A button labeled "Click this" is overlaid on the "Add" button. The main area lists "Serial" and "DHCP Scope" under "POAP Steps". Below this is a section for "Image and Config Servers" with a "POAP Definitions" button.

Bottom Window: The title bar shows "POAP Definitions". A button labeled "Click this" is overlaid on the "Generate or Upload Config" button. The main area has tabs for "Select an Option", "Select Switches", "Template Parameters", and "General Config". It includes sections for "Upload Startup Config" (radio button selected) and "Generate Definition" (radio button selected). A note says "Generate POAP definition from a template".

DCNM N5600 POAP

[Return to Launchpad](#)

POAP Steps

Select Switches

Enter comma separated Serial Numbers or Serial Number\DC Name identifiers. Examples: 123456, 123457 or 456795\DC1, 456795\DC2

Total 1

* Switches: FOC1902R0E3

Get serial no from show license host-id

POAP Definitions

fx

Image and Config Servers

Import From CSV File...

Note: The username\password is used by DCNM only to manage the switch and DCNM will not create the entered username\password in the switch.

* Switch Type: N5600

* Image Server: Default SCP Repository

* System Image: n6000-uk9-7.3.0.N1.0.254.bin

Cable Plan

* Kickstart Image: n6000-uk9-kickstart-7.3.0.N1.0.254.

* Config Server: Default SCP Repository

* Switch Password: cisco111

Show password in clear text

PW is used by DCNM to ssh to SWITCH

< Back | Next > | Cancel

DCNM N5600 POAP

POAP Definitions

✓ Generate or Upload Config
✓ Select Switches
Template Parameters

Select Switch Config Template

Template: IPFabric_N5600_Leaf_v02
EVPN → **IPFabric**
Saved Settings: LEAF1
Apply Manage...

Template Parameters - 1 Switches : FOC1902R0E3

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config	FCoE
Switch Name: <input type="text" value="N5672-57"/> * Administrative Username: <input type="text" value="admin"/> * Administrative Password: <input type="password" value="cisco11"/> * Management IP: <input type="text" value="10.0.7.57"/> * Management Prefix: <input type="text" value="24"/> * Default Gateway: <input type="text" value="10.0.7.254"/> * Management VRF: <input type="text" value="management"/> * Breakout Cables: <input type="text" value=""/> Console timeout: <input type="text" value="0"/> * Console Speed: <input type="text" value="9600"/> *	Plain text password of at least 8 characters Management address for this device [example: 10.10.10.20] Management Network Prefix Default Gateway IP address For Out-of-Band Management use management for In-Band use something else								

Choose proper template for switch role

IPFabric_N5600_Leaf_v02
EVPN → **IPFabric**
FP → Fabric
Leaf/Spine

< Back Next > Cancel

Template Parameters - 1 Switches : FOC1902R0E3

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config	FCoE
<input checked="" type="checkbox"/> Enable Secure LDAP									
<p>true if secure LDAP will be used</p> <p>DCNM IP Address</p> <p>LDAP IP Address <input type="text" value="10.0.1.23"/> Hostname of the Auto-config LDAP Server [Input format must match the server certificate CN]</p> <p>LDAP2 Hostname <input type="text" value="sjpdcmn2.sjrp.local"/> Hostname of the Secondary Auto-config LDAP Server</p> <p>LDAP Username <input type="text" value="reader"/> LDAP query user [needed for secure LDAP]</p> <p>LDAP Password <input type="password" value="fabr1c"/> Password to access LDAP for Auto-Config Queries [needed for secure LDAP]</p> <p>Primary LDAP server <input type="radio" value="none"/> none Use LDAP / LDAP2 as primary server or none if no primary server</p> <p>Enable XMPP <input checked="" type="checkbox"/> true if XMPP will be used</p> <p>Xmpp Server IP <input type="text" value="10.0.1.23"/> IP Address of the XMPP Server</p> <p>DCNM IP Address</p> <p>Xmpp Server Name <input type="text" value="sjpdcmn2.sjrp.local"/> FQDN of the XMPP Server [name entered for DCNM OVA deployment]</p> <p>Xmpp Groups <input type="text" value="dcnm-leaf"/> Space Separated XMPP Leaf Group Names</p> <p>Xmpp Password <input type="password"/> Password for device in XMPP Server</p> <p>AAA Server Type <input type="button" value="tacacs"/> TACACS Setting</p> <p>AAA Server IP <input type="text" value="10.0.1.8"/> Name or IP Address of AAA Server if used</p> <p>AAA Shared Secret <input type="text" value="cisco"/> Shared secret if AAA Server is used</p> <p>DNS Server IP <input type="text" value="10.0.1.1"/> IP Address of DNS Server if used</p> <p>NTP Server IP <input type="text" value="10.0.1.254"/> IP Address of NTP Server if used</p>									

ip host sjpdcmn1.sjrp.local 10.0.1.10

How to access LDAP DB

fabric database type network

server protocol ldap host sjpdcmn1.sjrp.local vrf management

db-table ou=networks,dc=cisco,dc=com key-type 1

db-security user cn=reader,dc=cisco,dc=com password 7 iwf1c

fabric database type profile

server protocol ldap host sjpdcmn1.sjrp.local vrf management

db-table ou=profiles,pFabric,dc=cisco,dc=com

db-security user cn=reader,dc=cisco,dc=com password 7 iwf1c

fabric database type partition

server protocol ldap host sjpdcmn1.sjrp.local vrf management

db-table ou=partitions,dc=cisco,dc=com

server protocol ldap host sjpdcmn1.sjrp.local vrf management

db-table ou=hosts,dc=cisco,dc=com

server protocol ldap host sjpdcmn1.sjrp.local vrf management

db-table ou=hosts,dc=cisco,dc=com

server protocol ldap host sjpdcmn1.sjrp.local vrf management

db-security user cn=reader,dc=cisco,dc=com password 7 iwf1c

fabric access server sjpdcmn1.sjrp.local vrf management password 7

fabric access group dcmn-leaf

aaa authentication login default group **AAA local**

aaa authentication login console local

aaa accounting default group AAA

aaa authentication login error-enable

tacacs-server host 10.0.1.8 key 7 "fewhg"

aaa group server tacacs+ AAA

server 10.0.1.8

use-vrf management

ntp server 10.0.1.254 prefer use-vrf management

vrf context management

ip name-server 10.0.1.1

cisco

DCNM N5600 POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config	FCoE
<p>Switch Role <input type="text" value="leaf"/> * The role of the switch, e.g. leaf, spine</p> <p>Link-State Routing Protocol <input checked="" type="radio" value="ospf"/> * Select link-state routing protocol Either ISIS or OSPF</p> <p>Tier <input type="text" value="1"/> * Leaf Tier is usually 1 for leaf, but can be 2 for Border-Leaf in a 3-tier network</p> <p>Pod <input type="text"/></p> <p>Net <input type="text"/></p> <p>Enable BFD <input type="checkbox"/></p> <p>true if Bidirectional Forwarding Detection should be enabled</p> <p>General Manageability Fabric Backbone/BGP Multicast Gateway VPC-Peer Interfaces Auto-Config FCoE</p>									
<p>Loopback IP <input type="text" value="10.0.0.57"/> * Loopback IP address used by Backbone [example: 1.1.1.1.10-20]</p> <p>Secondary Loopback IP <input type="text" value="10.0.0.102"/> Must be set the same for VPC Peers [example: 11.11.11.10-20]</p> <p>Is this Device an RR? <input type="checkbox"/> TRUE if this device is also an RR</p> <p>BGP Neighbor Subnet <input type="text"/></p> <p>BGP AS # <input type="text" value="65000"/> * BGP Router Autonomous System Number</p> <p>Route Reflector IP <input type="text" value="10.0.0.51"/> IP Address of Route Reflector - NOT NEEDED if this device is an RR</p> <p>Second Route Reflector IP <input type="text"/> IP Address of Second Route Reflector - NOT NEEDED if this device is an RR</p>									
<pre>router ospf UNDERLAY router-id 10.0.0.57 interface loopback0 ip router ospf UNDERLAY area 0.0.0.0 router bgp 65000 router-id 10.0.0.57 neighbor 10.0.0.51 remote-as 65000 update-source loopback0 address-family l2vpn evpn send-community both interface hve1 no shutdown source-interface loopback0 host-reachability protocol bgp</pre>									

DCNM N5600 POAP

General Manageability Fabric Backbone/BGP Multicast Gateway VPC-Peer Interfaces Auto-Config FCoE

Which RP * Is this an RP and if so which one, none if not an RP

RP Array

VIEW ONLY - DO NOT MODIFY: Populated from Admin Fabric Encapsulation Settings

Add Remove

RP_ID	IP_FABRIC_RP	RP_MASK	FABRIC_GROUP
Rp1	1.1.2.1	32	227.2.2.0/24

This is not RP so chose NONE

Set RP manually here

Structure Parameter Values

RP Array

VIEW ONLY - DO NOT MODIFY: Populated from Admin Fabric Encapsulation Settings

Add Remove

RP_ID	IP_FABRIC_RP	RP_MASK	FABRIC_GROUP
Rp1	1.1.2.1	32	227.2.2.0/24

ip pim rp-address 1.1.2.1 group-list 227.2.2.0/24 bidir

fabric forwarding anycast-gateway-mac 2020.0000.00AA
 fabric forwarding switch-role leaf
 fabric database mobility-domain md0
 fabric database profile-map global
 dot1q default dynamic
 vni default dynamic
 fabric database timer recovery 30
 system fabric dynamic-vlans 2500-3500
 system fabric core-vlans 2500-2999

Inherited from Global Setting

System Dynamic VLANs 2500-3500
 System Fabric Dynamic VLANs (mandatory for Auto-Config)

Core Dynamic VLANs 2500-2999
 Core Fabric Dynamic VLANs [subset of System Fabric Dynamic VLANs] (mandatory for Auto-Config)

DCNM N5600 POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config	FCoE
<input checked="" type="checkbox"/> Enable VPC									
true if VPC should be configured									
VPC Domain ID	578								
Should be the same for both VPC Peers [example: 10.10.11.11,12.12]									
VPC Peer Channel	578								
VPC Peer link port channel number									
VPC Peer Interfaces	e1/3-4								
VPC Peer Link VLAN	99								
VPC Peer Link VLAN IP	10.1.99.57								
VPC Peer Link VLAN IP Prefix Len	24								
For use in IP Fabric only [example: 12.12.12.10-20]									
VPC Keep-alive Interface Option	management ▾								
Select keep-alive interface option: mgmt or L3									
Use MGMT IP as peer keepalive									
VPC Keep-alive Local IP Address	10.0.7.57								
management or L3 interface's IP address of the switch for keep-alive [ex local ip: 10.0.0.1, peer ip: 10.0.									
VPC Keep-alive Peer IP Address	10.0.7.58								
Management or L3 interface's IP address of the Peer switch for keep-alive [ex peer ip: 10.0.0.2, local ip:									
VPC Keep-alive IP Prefix	30								
Keep-alive IP Prefix									
VPC Keep-alive L3 Interface									
L3 interface, if keep-alive is using L3 interface [ex: e1/25]									
VPC Keep-alive VRF Name	management								
vrf name, if keep-alive is using L3 interface									
VPC Advertise Primary IP in BGP	<input type="checkbox"/>								
true if Primary IP should be advertised as next-hop in BGP									
If you don't want to use NVE 2 nd IP as BGP NH									

feature vpc

vpc domain 578

peer-keepalive destination **10.0.7.58** source **10.0.7.57**

delay restore 150

ipv6 nd synchronize

ip arp synchronize

interface Ethernet1/3

description "vpc-peer-link"

switchport mode trunk

channel-group 578 mode active

interface port-channel578

description "vpc-peer-link"

switchport mode trunk

spanning-tree port type network

vpc peer-link

```
interface Vlan99
no shutdown
no ip redirects
ip address 10.1.99.57/24
no ipv6 redirects
ip ospf network point-to-point
ip router ospf UNDERLAY area 0.0.0.0
ip pim sparse-mode
```

DCNM N5600 POAP

General **Manageability** **Fabric** **Backbone/BGP** **Multicast** **Gateway** **VPC-Peer** **Interfaces** **Auto-Config** **FCoE**

Int connecting to SPINE

P2P Fabric Interfaces

IP Unnumbered Fabric Interfaces []

Fabric Interfaces using IP unnumbered [example: e1/1-4,e2/2]

P2P Fabric Interfaces

P2P Fabric Interfaces with IP addresses and network mask lengths

FEX [{{e1/1,10.1.51.10},

Point-to-point Fabric Interfaces with IP addresses and network mask lengths

Slot # [e1x 10/1], Channel # [e1x 10/1], Interfaces to FEX [e1x e1/1], Model Type [e1x]

VPC Interfaces (switch/SF-Fex)

{}{{e1/1,10.1.51.10},

Select the VPC PortChannel ID as well as the Interface range

PortChannel Interfaces and AA Fex PortChannel Interfaces

{}{{e1/1,10.1.51.10},

Select the PortChannel ID as well as the interface range

Trunked Host Interfaces

{}{{e1/1,10.1.51.10},

Select Trunked Host Ports and Auto-config Triggers

FEX

Access Host Interfaces []

Unused Interfaces e1/1-2,e1/5-6,e1/8-

A list of unused ports [example: e1/5,e2/3-5]

P2P Fabric Interfaces

Point-to-point Fabric Interfaces with IP addresses and network mask lengths

Add **Remove**

IF_NAME	IP	PREFIX_LEN
e1/31	10.1.51.10	30

Slot # [e1x 10/1], Channel # [e1x 10/1], Interfaces to FEX [e1x e1/1], Model Type [e1x]

interface Ethernet1/31

no ip redirects
no switchport
mtu 9192

N6K set MTU 9192 automatically

interface Ethernet1/31

no ip redirects
no ipv6 redirects
ip ospf network **point-to-point**
ip router ospf **UNDERLAY** area **0.0.0.0**
ip pim sparse-mode

FEX

Slot # [e1x 10/1], Channel # [e1x 10/1], Interfaces to FEX [e1x e1/1], Model Type [e1x]

Add **Remove**

FEX_SLOT	CHANNEL_ID	FEX_MODEL	INTERFACES	AA_FEX
101	101	N2K-C223	e1/29	true
102	102	N2K-C223	e1/30	true

interface Ethernet1/29

description UPLINK FEX101

switchport mode **fx-fabric**

switchport trunk allowed vlan 1-98,100-4094

interface port-channel101

description UPLINK FEX101

switchport mode **fx-fabric**

switchport trunk allowed vlan 1-98,100-4094

fx associate 101

vpc 101

fex 101

pinning max-links 1

description "FEX0101 desc"

slot 101

provision model **N2K-C223TM-E-10GE**

channel-group 101

cisco

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

DCNM N5600 POAP

General **Manageability** **Fabric** **Backbone/BGP** **Multicast** **Gateway** **VPC-Peer** **Interfaces** **Auto-Config** **FCoE**

IP Unnumbered Fabric Interfaces

Fabric Interfaces using ip unnumbered [example: e1/1-4,e2/2]

P2P Fabric Interfaces {{e1/3;1;10.1.51.10;}}

Point-to-point Fabric Interfaces with IP addresses and network mask lengths

FEX {{101;101;N2K-C22;}}

Slot #[e1x 101], Channel #[e1x 101], Interface to FEX [ex e1/1], Model Type[ex]

VPC Interfaces (switch/ST-Fex) {{28,e1/7,active};{}}

Select the VPC/PortChannel ID as well as the Interface range

PortChannel Interfaces and AA Flex PortChannel Interfaces

Select the PortChannel ID as well as the Interface range

Trunked Host Interfaces {{Po28,dot1q,false,t}}

Select Trunked Host Ports and Auto-config Triggers

Access Host Interfaces

Select Access Host Ports and properties

Unused Interfaces e1/1-2,e1/5-6,e1/8-

A list of unused ports [example: e1/5,e2/3-5]

Interface Range CHANNEL_ID IF_NAMES PC_MODE
19 e1/7 active
29 e1/8 active

Auto-Configuration Trigger Type

IF_NAMES	TRIGGER	BPDUGUARD_EN	1G_PORT
Po28	dot1q	<input type="checkbox"/>	<input checked="" type="checkbox"/>
e101/1/24	vdp	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Interface Configuration

```

interface Ethernet1/1/7
switchport mode trunk
switchport trunk allowed vlan 1-98,100-4094
encapsulation dynamic dot1q
speed 1000
channel-group 28 mode active

interface port-channel28
switchport mode trunk
switchport trunk allowed vlan 1-98,100-4094
encapsulation dynamic dot1q
spanning-tree port type edge trunk
speed 1000
vpc 28
  
```

Interface Configuration (Continued)

```

interface Ethernet101/1/24
switchport trunk allowed vlan 1-98,100-4094
encapsulation dynamic vdp
speed 1000
spanning-tree port type edge trunk
  
```

Interface Configuration (Continued)

```

interface Ethernet1/1/24
switchport mode trunk
switchport trunk allowed vlan 1-98,100-4094
encapsulation dynamic vdp
speed 1000
  
```

Cisco Logo

DCNM N5600 POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config	FCoE
<p>Enable DFS/dot1Q Trigger <input checked="" type="checkbox"/></p> <p>true if Auto-configuration DFS/dot1q packet based triggers should be enabled</p> <p>Enable LLDP Trigger <input checked="" type="checkbox"/></p> <p>true if Auto-configuration LLDP triggers should be enabled</p> <p>Enable VMTracker Trigger <input checked="" type="checkbox"/></p> <p>Enable VMTracker Trigger</p> <p>vCenter Configuration <input type="text"/> </p> <p>Configure Multiple vCenters for VMTracker Trigger. Limit to a MAXIMUM of 4 connections.</p> <p>Enable EVB/VDP Trigger <input checked="" type="checkbox"/></p> <p>true if Auto-configuration EVB/VDP triggers should be enabled</p>	feature evb								

```
evb resource-wait-delay 25  
evb reinit-keep-alive 27  
evb batch-response disable  
evb mac 0180.C200.0000
```

DCNM N5600 POAP

POAP Definitions

- ✓ Generate or Upload Config
- ✓ Select Switches
- ✓ Template Parameters

Template: IIPFabric_N5600_Leaf_v02

Generate Parameter Values

General Config

Preview CLI

Serial Number	Switch Name	Administrativ...	Administrativ...	Management...	Management...	Default Gate...	Management...
FOC1902R0E3	N567-2-57	admin	cisco11	10.0.7.57	24	10.0.7.254	management

< Back Save Publish Cancel

Must publish for POAP to work
Saving is not enough

DCNM N5600 POAP

Return to Launchpad

Cisco Prime
Data Center
Network Manager

admin ▾ DCNM-SAN | DCNM-LAN | DM | CLI | About By Name

Dashboard ▾ Health ▾ Performance ▾ Inventory ▾ Reports ▾ Config ▾ Admin ▾

POAP Switch Definitions

Add	Edit	Delete	Publish	Write Erase and Reload	Change Image	Boot Log	Refresh Diff State
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Serial Number	Switch ID	Management IP	Switch Status	Publish Status	Bootscript Status	Diff State	Model
<input checked="" type="checkbox"/> FOCL902R0E3	N5672-57	10.0.7.57	<input checked="" type="checkbox"/> Ok	<input checked="" type="checkbox"/> Published	POAP script is finished	Diff Detected	N5600 IPFabric N5600-leaf v02
<input type="checkbox"/> FOCL836R2Q6	N5672-58	10.0.7.58	<input checked="" type="checkbox"/> Ok	<input checked="" type="checkbox"/> Published	POAP script is finished	Diff Detected	N5600 IPFabric N5600-leaf v02

If you click published, then you can see SWITCH configuration and fix any error

Selected 1 | Total 10

Generated Configuration

```

hardware ethernet store-and-forward-switching

hostname N5672-57
install feature-set fabric
feature-set fabric
feature fabric-forwarding
dfs eth distribute
nv overlay evpn
feature ospf
feature bgp
feature pim
feature interface-vlan
feature dhcp
feature ebgp
feature lldp
feature vpc
feature nv overlay
feature nvsegment-vlan-based
no feature retain
logging event link-status default
feature nxapi
ip pim ssm range 232.0.0.0/8
feature fex
feature lacp

```

Bootscript status and Diff State would be empty unless POAP is initiated from SWITCH

POAP Start from SW

N5672-57# write erase

Warning: This command will erase the startup-configuration.

Do you wish to proceed anyway? (y/n) [n] y

N5672-57# reload

The system is going down for reboot NOW!

2016 Feb 17 02:11:18 %S VDC-1 %S %PFMA-2-PS_FAIL: Power supply 1 failed or shutdown(Serial number PST1845H08A)

System is coming up ... Please wait ...

Abort Power On Auto Provisioning and continue with normal setup ?(yes/no)[n]:

2016 Feb 17 02:15:02 switch %S VDC-1 %S %VDC_MGR-2-VDC_ONLINE: vdc 1 has come online

Starting Power On Auto Provisioning..

2016 Feb 17 02:15:52 switch %S VDC-1 %S %POAP-2-POAP_INITED: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - POAP process initialized

2016 Feb 17 02:15:52 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - USB disk not detected

2016 Feb 17 02:15:52 switch %S VDC-1 %S %POAP-2-POAP_DISCOVER_START: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - POAP DHCP Discover phase started

2016 Feb 17 02:17:00 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Valid DHCP OFFER received from 10.0.7.23

2016 Feb 17 02:17:10 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Using DHCP, information received over mgmt0 from 10.0.7.23

2016 Feb 17 02:17:10 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Assigned IP address: 10.0.7.200

2016 Feb 17 02:17:10 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Netmask: 255.255.255.0

2016 Feb 17 02:17:10 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - DNS Server: 10.0.1.1

2016 Feb 17 02:17:10 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Default Gateway: 10.0.7.254

2016 Feb 17 02:17:10 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Script Server: 10.0.7.23

2016 Feb 17 02:17:10 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Script Name: poap_dcmn.py

2016 Feb 17 02:17:21 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - The POAP Script download has started

2016 Feb 17 02:17:21 switch %S VDC-1 %S %POAP-2-POAP_INFO: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - The POAP Script is being downloaded from [copy

tftp://10.0.7.23/poap_dcmn.py bootflash:scripts/script.sh vrf management]

2016 Feb 17 02:17:22 switch %S VDC-1 %S %POAP-2-POAP_SCRIPT_DOWNLOADED: S/N[FOC1902R0E3]-MAC[8C:60:4F:4D:F9:3C] - Successfully downloaded POAP script file

You go to SWITCH and erase configuration and Reboot or just connect new
SWITCH MGMT int into DFA FABRIC MGMT NETWORK and power up

POAP only starts after SYSTEM is up. Module failure can prevent POAP to start

After DHCP done, POAP Script is downloaded

DCNM POAP

```
2016 Feb 17 02:17:30 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : copy scp://root@10.0.7.23/var/lib/dcnm/n6000-uk9-kickstart.7.3.0.N1.0.254.bin n6000-uk9-kickstart.7.3.0.N1.0.254.bin.tmp vrf management- script.sh  
2016 Feb 17 02:17:38 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Completed Copy of Kickstart Image - script.sh  
2016 Feb 17 02:17:38 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : copy scp://root@10.0.7.23/var/lib/dcnm/n6000-uk9.7.3.0.N1.0.254.bin n6000-uk9.7.3.0.N1.0.254.bin.tmp vrf management- script.sh  
2016 Feb 17 02:18:11 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Completed Copy of System Image - script.sh  
2016 Feb 17 02:18:30 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : copy scp://root@10.0.7.23/var/lib/dcnm/cableplan.xml cableplan.xml.tmp vrf management - script.sh  
2016 Feb 17 02:18:32 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Completed Copy of Cable file - script.sh  
2016 Feb 17 02:18:32 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : copy scp://root@10.0.7.23/var/lib/dcnm/FOC1902R0E3/device-config device-config vrf management - script.sh  
2016 Feb 17 02:18:33 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Completed Copy of Config File - script.sh  
2016 Feb 17 02:18:33 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Setting the boot variables - script.sh  
2016 Feb 17 02:18:35 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : config terminal; boot system n6000-uk9.7.3.0.N1.0.254.bin - script.sh  
2016 Feb 17 02:18:47 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : copy running-config startup-config - script.sh  
2016 Feb 17 02:18:57 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : copy poap_1.cfg scheduled-config - script.sh  
2016 Feb 17 02:18:58 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : config terminal; boot system n6000-uk9.7.3.0.N1.0.254.bin - script.sh  
2016 Feb 17 02:18:58 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Copying the first scheduled cfg done - script.sh  
2016 Feb 17 02:18:58 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - CLI : copy poap_2.cfg scheduled-config - script.sh  
2016 Feb 17 02:18:59 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Copying the second scheduled cfg done - script.sh  
2016 Feb 17 02:18:59 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - INFO: Configuration successful - script.sh  
2016 Feb 17 02:18:59 switch %$ VDC-1 %$ %USER-1-SYSTEM_MSG: S/N[FOC1902R0E3] - FINISH: Clean up files - script.sh
```

Download OS Image

RELOAD

First Reboot

Box is online...

Applying scheduled configuration...
WARNING: This command will reboot the system

2nd Reboot to make sure reboot required
commands take effect

2016 Feb 17 02:40:37 N5672-57 %\$ VDC-1 %\$ %SYSLOG-2-SYSTEM_MSG: POAP Completed - LOGIN

User Access Verification
N5672-57 login: admin
Password:

DCNM POAP DEBUG

N5672-57# dir

```
1814 Mar 10 19:41:18 2016 20160310_193910_poap_5273_dhcp.pcap
430833 Mar 10 19:51:17 2016 20160310_193910_poap_5273_init.log
6040 Mar 10 20:10:46 2016 poap.log.19_41_31
```

N5672-57# show file 20160310_193910_poap_5273_init.log | grep -i DHCP

```
Thu Mar 10 19:39:10 2016:poap_debug_init : DHCP frames redirected to pcap(/bootflash/20160310_193910_poap_5273_dhcp.pcap) file
Thu Mar 10 19:40:00 2016:poap_dhcpdiscover_start: Start DHCP v4 session
Thu Mar 10 19:41:05 2016:Creating the DHCP DISCOVER timer
Mon Feb 22 23:54:39 2016:poap_handle_dhcp_offer: the given ip address is 10.0.7.201
Mon Feb 22 23:54:39 2016:poap_handle_dhcp_offer: tftp server name is not present in theheader
Mon Feb 22 23:54:39 2016:poap_handle_dhcp_offer: boot file name not present in theheader
Mon Feb 22 23:54:39 2016:poap_is_offer_valid: Valid DHCP OFFER received from 10.0.7.10
```

N5672-57# show file poap.log.19_41_31 | beg errors

```
##### POAP first(splitConfig) Configuration errors #####
##### POAP scheduled Configuration errors #####
```

If there was any configuration error from POAP script
You can see if some configuration commands didn't work

Also, you can ssh to DCNM, 172.16.166.53 and run tcpdump to see if DCNM is getting DHCP Discover from SWITCH

```
[root@sjrpdcnm1 ~]# ifconfig
eth0      Link encap:Ethernet HWaddr 00:50:56:A9:1A:0F
          inet addr:10.0.1.10  Bcast:10.0.1.255  Mask:255.255.255.0
                
```

Eth1 is DFA FABRIC MGMT INT

```
eth1      Link encap:Ethernet HWaddr 00:50:56:A9:6B:B0
          inet addr:10.0.7.10  Bcast:10.0.7.255  Mask:255.255.255.0
```

Use TCPDUMP on eth1 to get DHCP pkt from switch

```
[root@sjrpdcnm1 ~]# tcpdump -i eth1 port 67 or port 68
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth1, link-type EN10MB (Ethernet), capture size 65535 bytes
```

DCNM N7K ADMIN VDC POAP

POAP Definitions

Enter Switch Details

✓ Generate or Upload Config

Select Switches

Generate Config

Template Parameters

* Switches: TM13343273 Total 1

Note: The username/password is used by DCNM only to manage the switch and DCNM will not create the entered username/password in the switch.

* Switch Type: N7K

* Image Server: Default_SCp_Repository

* SystemImage: n7000-s2-dkg-7.3.0.D1.0.201.bin

* Kickstart Image: n7000-s2-kickstart-7.3.0.D1.0.201.b

* Config Server: Default_SCp_Repository

Choose N7K here

Import from CSV File...

Add Switches to Group:

Default_LAN

* Switch User Name:

* Switch Password:

Show password in clear text

< Back Next > Cancel

**Need to create ADMIN VDC using POAP
to create other VDCs by POAP**

DCNM N7K ADMIN VDC POAP

POAP Definitions

✓ Generate or Upload Config
✓ Select Switches

Template Parameters

Generate Config

Select Switch Config Template

Template	Base_N7K_AdminVDC_v5	<input type="button" value="View..."/>	<input type="button" value="Diff..."/>
Saved Settings	LEAF1.txt	<input type="button" value="Apply"/>	<input type="button" value="Manage..."/>

Template Parameters - 1 Switches : TBM13343273

	General	Manageability
Switch Name	N7K1	* The host name of the switch [example: switch[1-3]]
Is AdminVDC	<input checked="" type="checkbox"/>	true if this is an Admin VDC false for Default VDC
Nexus Series	N7000	* Select whether this device is a N7700 or N7000
Administrative Username	admin	* Administrative User which will be used to manage this device
Administrative Password	cisco111	* Plain text password of at least 8 characters
Management IP	10.0.7.50	* Management address for this device [example: 10.10.10.10-20]
Management Prefix	24	* Management Network Prefix
Default Gateway	10.0.7.254	* Default Gateway IP address
Console timeout	0	* Enter time-out in minutes, 0 to disable

Choose BASE ADMIN VDC Template

	General	Manageability
AAA Server IP	10.0.1.8	Radius, tacacs, or none if not using AAA
AAA Shared Secret	cisco	Name or IP Address of AAA Server if used
DNS Server IP	10.0.1.1	Shared secret if AAA Server is used
NTP Server IP	10.0.1.254	IP Address of DNS Server if used
SNMP IP Address		IP Address of NTP Server if used
Secondary SNMP IP Address		DCNM SNMP Server IP
Enable feature-set Fabric path	<input checked="" type="checkbox"/>	true if feature-set Fabric path should be installed
Enable feature-set mpls	<input checked="" type="checkbox"/>	true if feature-set mpls should be installed
Enable feature-set fabric	<input checked="" type="checkbox"/>	true if feature-set fabric should be installed
Enable feature-set fcoe	<input checked="" type="checkbox"/>	true if fcoe should be installed
Power Redundancy-mode	combined	Configure power supply redundancy-mode

Choose feature-set and power mode here

DCNM N7K VDC SPINE POAP

POAP Definitions

Generate or Upload Config

Select Switches

Template Parameters

Generate Config

Enter Switch Details

Enter comma separated Serial Numbers or Serial Number:VDC Name identifiers. Examples: 123456, 123457 or 456795:VDC1, 456795:VDC2

* Switches: TBM13342273:R_PSP1HE1

Total 1

SW TYPE IS VDC

Note: The username/password is used by DCNM only to manage the switch and DCNM will not create the entered username/password in the switch.

* Switch Type: **VDC**

* Image Server: Select...

* System Image: Select...

* Kickstart Image: Select...

* Config Server: Default_SCP_Repository

Import from CSV File...

< Back Next > Cancel

**Image portion is
graved out as this
is VDC**

DCNM N7K VDC SPINE POAP

Template Parameters - 1 Switches : IBM13343273:RP SPINE1

General		Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
Switch Name	RSPINE1	*	The host name of the switch [example: switch[1-3]]			
Administrative Username	admin	*	Administrative User which will be used to manage this device			
Administrative Password	cisco11	*	Plain text password of at least 8 characters			
Management IP	10.0.7.51	*	Mgmt0 address for this device [example: 10.10.10.10-20]			
Management Prefix	24	*	Management Network Prefix			
Default Gateway	10.0.7.254	*	Default Gateway IP address			
Console timeout	0	*	Enter timeout in minutes, 0 to disable			

Similar to N6K

Template Parameters - 1 Switches : IBM13343273:RP SPINE1

General		Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
SNMP IP Address			DCNM SNMP Server IP			
SNMP2 IP Address			Secondary SNMP Server IP			
AAA Server Type	tacacs	▼	radius, tacacs, or none if not using AAA			
AAA Server IP	10.0.1.8		Name or IP Address of AAA Server if used			
AAA Shared Secret	cisco		Shared secret if AAA Server is used			
DNS Server IP	10.0.1.1		IP Address of DNS Server if used			

**NO LDAP CONFIG as SPINE
doesn't need AUTO-
CONFIGURATION**

DCNM N7K VDC SPINE POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
<p>Switch Role <input type="text" value="spine"/> * The role of the switch. e.g. leaf, spine</p> <p>Enable BFD <input type="checkbox"/></p> <p>Link-State Routing Protocol <input checked="" type="radio" value="ospf"/> * Select link-state routing protocol</p>	<p>Net <input type="text"/></p> <p>Network Entity Title for IS-IS [example:49.0001.#####.####.00], NET will be auto-generated if not provided</p>	<p>General</p> <p>Enable Flood and Learn <input type="checkbox"/></p> <p>Click this if user activate Flood and Learn</p>	<p>Loopback IP <input type="text" value="10.0.0.51"/> * [example: 11.11.11.10-20]</p> <p>Loopback Prefix <input type="text" value="32"/> * Loopback Prefix</p> <p>BGP AS # <input type="text" value="65000"/> * BGP Router Autonomous System Number</p>	<p>router ospf UNDERLAY router-id 10.0.0.51</p> <p>interface loopback0 ip address 10.0.0.51/32 ip router ospf UNDERLAY area 0.0.0.0 ip pim sparse-mode</p>	<p>feature ospf</p>
<p>Is this Device an RR? <input checked="" type="checkbox"/></p> <p>BGP Neighbor Subnet <input type="text" value="10.0.0.48/28"/></p>	<p>TRUE if this device is also an BGP Router Reflector</p> <p>BGP Neighbor Subnet if this device is an BGP Router Reflector [example: 11.11.11.0/24]</p>	<p>We summarize RR Clients' update source IP here</p>			

DCNM N7K VDC SPINE POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
RP Mode bidir *					
Select Multicast RP Mode					
Which RP RP1 ▾					
This is only for bidir					
PIM BIDIR					
RPS_BIDIR [RP1,1.1.2.1,32,22]					
Populated from Admin Fabric Encapsulation Settings in BIDIR Mode					
Anycast RP Address <input type="text"/>					
RP Group <input type="text"/>					
RPS <input type="text"/>					

ip pim rp-address 1.1.2.1 group-list 227.2.2.0/24 bidir

interface loopback2

description primary loopback for PIM bidir for group 227.2.2.0/24

ip address 1.1.2.2/30

ip ospf network point-to-point

ip router ospf UNDERLAY area 0.0.0.0

ip pim sparse-mode

**Use lo0 for routing
lo1 for PIM RP**

**Get it from Fabric Encap
Setting**

Anycast RP Address for PIM Multicast
 Populated from Admin Fabric Encapsulation Settings in ASM Mode

RP Group
 RP Group max 256 addresses [example: 239.239.239.0/24]

RPS
 RPS Populated from Admin Fabric Encapsulation Settings in ASM Mode

DCNM N7K SPINE VDC POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
VDC Module Types f3	*	A list of module types for this VDC			
VDC Allocated Interfaces e4/1-8	*	A list of interfaces to allocate to this VDC [example: e1/1-4,e2/2]			
Breakout Cables		Breakout pairs of Slot and port range [example 1, 1-3]			
Fabric IP Prefix 30	*	IP Prefix to use for all Fabric interfaces			
Unnumbered Fabric Interfaces		A list of ip-unnumbered fabric ports [example: e1/1-4,e2/2]			
P2P Fabric Interfaces {e4/1,10.1.51.1,91}		Point-to-point Fabric interfaces and their IP addresses			
Port Channel Mode active		channel mode options are on, active, and passive			
P2P Fabric Port-channel		If P2P interfaces as port-channel			
Unused Interfaces		A list of unused ports [example: e1/5,e2/3-5]			

P2P Fabric Interfaces			
Point-to-point Fabric interfaces and their IP addressess			
Add	Remove	All LEAF connecting P2P interfaces	Make sure set MTU to 9192 if there is N6K
<input type="checkbox"/> e4/1	10.1.51.1	9192	
<input type="checkbox"/> e4/2	10.1.51.5	9192	
<input type="checkbox"/> e4/3	10.1.51.9	9192	
<input type="checkbox"/> e4/4	10.1.51.13	9192	
<input type="checkbox"/> e4/5	10.1.51.17	9192	

N7K1sh run vdc

vdc RPSPINE1 id 4

limit-resource **module-type f3**

- allow feature-set fabricpath
- allow feature-set mpls
- allow feature-set fabric
- allow feature-set fex
- allocate interface **Ethernet4/1-8**

Setting up VDC related config here

interface Ethernet4/1

mtu 9192

ip address 10.1.51.1/30

ip ospf network **point-to-point**

ip router ospf **UNDERLAY** area 0.0.0.0

ip pim sparse-mode

no shutdown

DCNM N7K LEAF VDC POAP

Template Parameters - 1 Switches : IBM13343273:LEAF1

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
VDC Name	LEAF1	*						
Host Name	LEAF1	*						
Administrative Username	admin	*						
Administrative Password	cisco11	*						
Management IP	10.0.7.53	*						
Management Prefix	24	*						
Default Gateway	10.0.7.254	*						
Console timeout	0	*						
Console Speed	9600	*						

DCNM N7K LEAF VDC POAP

Template Parameters - 1 Switches : TBM13343273:LEAF1

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
<input type="checkbox"/> Enable Secure LDAP	true if secure LDAP will be used							
LDAP IP Address 10.0.1.10 *	IP Address of the Auto-config LDAP Server							
LDAP Hostname sjrpdcnm1.sjrp.local *	Hostname of the Auto-config LDAP Server [input format must match the server certificate CN]							
LDAP2 IP Address	IP Address of the Secondary Auto-config LDAP Server							
LDAP2 Hostname	Hostname of the Secondary Auto-config LDAP Server [input format must match the server certificate CN]							
LDAP Username reader	LDAP query user (needed for secure LDAP)							
LDAP Password fabric *	Password to access LDAP for Auto-Config Queries (needed for secure LDAP)							
XMP Server IP 10.0.1.10	IP Address of the XMP Server							
XMP Server Name sjrpdcnm1.sjrp.local	FQDN of the XMP Server [name entered for DCNM OVA deployment]							
XMP Groups dcnm:fabric	Space Separated XMP Leaf Group Names							
XMPF Password	Password for device in XMP Server							
AAA Server Type tacacs *	radius, tacacs, or none if not using AAA							
AAA Server IP 10.0.1.8	Name or IP Address of AAA Server if used							
AAA Shared Secret cisco	Shared secret if AAA Server is used							
DNS Server IP 10.0.1.1	IP Address of DNS Server if used							

DCNM N7K LEAF POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
Switch Role <input type="text" value="leaf"/> *	The role of the switch, e.g. leaf, spine							
Link-State Routing Protocol <input checked="" type="radio" value="ospf"/> * <input type="radio" value="isis"/>	Select link-state routing protocol							
Net <input type="text"/>	Network Entity Title for IS-IS [example:49.0001#####.00], NET will be auto-generated if not provided							
Enable BFD <input type="checkbox"/>	true if BFD should be enabled							
General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
<p><input type="checkbox"/> Click this if user activate Flood and Learn</p> <p>Enable Flood and Learn <input type="checkbox"/></p> <p>Loopback IP <input type="text" value="10.0.0.53"/> * [example: 1.1.1.1.10-20]</p> <p>Second Loopback IP <input type="text" value="10.0.0.101"/> * [example: 1.1.1.1.10-20]</p> <p>Loopback Prefix <input type="text" value="32"/> * Loopback Prefix</p> <p>BGP AS # <input type="text" value="65000"/> * BGP Router Autonomous System Number</p> <p>Is this Device an BGP Router Reflector? <input type="checkbox"/></p> <p>BGP RR Client Subnet: <input type="text"/></p> <p>RR prefix peers [example: 1.1.1.1.0/24]</p> <p>Route Reflector IP <input type="text" value="10.0.0.51"/> * IP Address of BGP Route Reflector</p> <p>Second Route Reflector IP <input type="text"/></p> <p>IP Address of Second BGP Route Reflector</p> <p>Advertise ip <input type="checkbox"/></p> <p>To advertise primary ip for L2vpn next hop</p>								

feature ospf

router ospf UNDERLAY
router-id 10.0.0.53

interface loopback0
ip address 10.0.0.53/32

```
ip router ospf UNDERLAY area 0.0.0.0
ip pim sparse-mode
```

router bgp 65000

```
router-id 10.0.0.53
address-family ipv4 evpn
neighbor 10.0.0.51 remote-as 65000
update-source loopback0
address-family ipv4 unicast
send-community both
address-family l2vpn evpn
send-community both
```

VPC leaf so 2nd IP

DCNM N7K LEAF POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
RP Mode bidir *	Select Multicast RP Mode							
RP Array <code>{ {RP1,1.1.2.1,32,22}</code>	Populated from Admin Fabric Encapsulation Settings							
Anycast RP Address <code>1.1.2.1</code>	Anycast RP Addresses for PIM Multicast							
RP Group <code>227.2.2.0/24</code>	RP Group max 256 addresses [example: 239.239.239.0/24]							
General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
Anycast-Gateway-MAC <code>2020.0000.00AA</code>	Shared MAC address for all leaves							
System Bridge Domain <code>2500-3500</code>	System Fabric Dynamic VLANs							
Fabric (core) Bridge Domain <code>2500-2999</code>	Core Fabric Dynamic VLANs [subset of System Fabric Dynamic VLANs]							

```

ip pim rp-address 1.1.2.1 group-list 227.2.2.0/24 bidir
fabric forwarding anycast-gateway-mac 2020.0000.00AA
fabric database profile-map global
dot1q default dynamic
vni default dynamic
fabric database timer recovery 30
system bridge-domain 2500-3500
system fabric bridge-domain 2500-2999

```

DCNM N7K LEAF POAP

General	Managability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <div> <p>Enable VPC <input checked="" type="checkbox"/></p> <p>true if VPC should be configured</p> </div> <div> <p>VPC Domain ID <input type="text" value="534"/></p> <p>Should be the same for both VPC Peers [example: 10.10.11.1,12,12]</p> </div> <div> <p>VPC Peer Channel <input type="text" value="534"/></p> <p>VPC Peer link port channel number</p> </div> </div> </div>								
<div> <p>VPC Peer Interfaces <input type="text" value="e4/23"/></p> <p>VLAN outside the dynamic range for use in IP Fabric only</p> <p>VPC Peer Link VLAN <input type="text" value="99"/></p> <p>For use in IP Fabric only</p> <p>VPC Peer Link VLAN IP Prefix Len <input type="text" value="24"/></p> <p>For use in IP Fabric only</p> <p>VPC Keep-alive Interface Option <input type="text" value="L3"/></p> <p>Select keep-alive interface option: mgmt or L3</p> <p>VPC Keep-alive Local IP Address <input type="text" value="1.1.1.1"/></p> <p>management or L3 interface's IP address of the switch for keep-alive [ex local ip: 10.0.0.1, peer ip: 10.0.0.2, local ip: 1.1.1.2]</p> <p>Peer Keep-alive IP <input type="text" value="1.1.1.2"/></p> <p>Management or L3 interface's IP address of the Peer switch for keep-alive [ex peer ip: 10.0.0.2, local ip: 1.1.1.2]</p> <p>VPC Keep-alive IP Prefix <input type="text" value="24"/></p> <p>Keep-alive IP Prefix</p> <p>VPC Keep-alive L3 Interface <input type="text" value="e4/24"/></p> <p>L3 interface, if keep-alive is using L3 interface [ex: e1/25]</p> <p>VPC keep-alive VRF Name <input type="text" value="vpck"/></p> <p>vrf name, if keep-alive is using L3 interface</p> <p>VPC Interfaces (switch/ST-Fex) <input type="text" value="{{55,e4/21,true},{3}}"/></p> <p>Select the VPC IP on Channel ID as well as the Interface range</p> </div>								

VPC Interfaces (switch/ST-Fex)		
Select the VPC/PortChannel ID as well as the Interface range		
CHANNEL_ID	IF_NAME	PORT_TRACKING
<input type="checkbox"/> 55	e4/21	<input checked="" type="checkbox"/>
<input type="checkbox"/> 34	e4/22	<input checked="" type="checkbox"/>

DCNM N7K LEAF POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config
P2P Fabric Interfaces								
Point-to-point Fabric interfaces and their IP addresses								
VDC Module Types		f3 *						
		A list of module types for this VDC use only f3 for DFA functionality						
VDC Allocated Interfaces		e4/17-24 *						
		A list of interfaces to allocate to the DFA Leaf VDC [example : e1/1-4,e2/2]						
IP Unnumbered Fabric Interfaces								
Fabric interfaces using ip unnumbered [example: e1/1-4,e2/2]								
Breakout Cables								
Breakout pairs of Slot and port range [example 1, 1-3]								
FEX								
Slot#[ex 101], Channel#[ex 101], Model Type[ex N2K-C2248PO], Interfaces to FEX								
Fex Bringup time								
300 *								
P2P Fabric Interfaces								
{ {#4/17,10.1.51.2.9} } *								
Point-to-point Fabric interfaces and their IP addressess								
P2P Fabric Interface Prefix Len								
30 *								
IP Prefix to use for all Fabric interfaces								
Fabric Port Channel Mode								
active *								
Fabric Port Interface Prefix Len								
30 *								
IP Prefix to use for all Fabric interfaces								
P2P Fabric Port-channel								
If P2P interfaces as port-channel								
Port Channel Mode								
active *								
PORT_CHANNEL_HOSTS								
{ {po34,Trunks,mdu0} } *								
Select Mobility domain per interface range (this overlaps above ports)								
Host Interfaces								
Unused Interfaces								
e4/18-20 *								
A list of unused ports [example: e1/5,e2/3-5]								

P2P Fabric Interfaces

interface Ethernet4/17

mtu 9192

ip address 10.1.51.2/30

ip ospf network point-to-point

ip router ospf UNDERLAY area 0.0.0.0

ip pim sparse-mode

no shutdown

Point-to-point Fabric interfaces and their IP addresses

interface Ethernet4/17

mtu 9192

ip address 10.1.51.2/30

ip ospf network point-to-point

ip router ospf UNDERLAY area 0.0.0.0

ip pim sparse-mode

no shutdown

IP Unnumbered Fabric Interfaces

Fabric interfaces using ip unnumbered [example: e1/1-4,e2/2]

Host Interfaces

Select Mobility domain per interface range (this overlaps above ports)

Different from N6K.
Need to specify which MD that N7K will use

Add		Remove	
Interfaces	Type	mobility_domain	autoconfig_trigger_type
<input type="checkbox"/> po34	Trunks	▼ mdu0	▼ frame-snoop
<input type="checkbox"/> po55	Trunks	▼ mdu0	▼ VDP

interface port-channel34

vpc 34

service instance vni default

encapsulation dynamic frame-snoo profile mdu0

interface port-channel55

vpc 55

service instance vni default

encapsulation dynamic vdp profile mdu0

interface Ethernet4/21

channel-group 55 mode active

no shutdown

interface Ethernet4/22

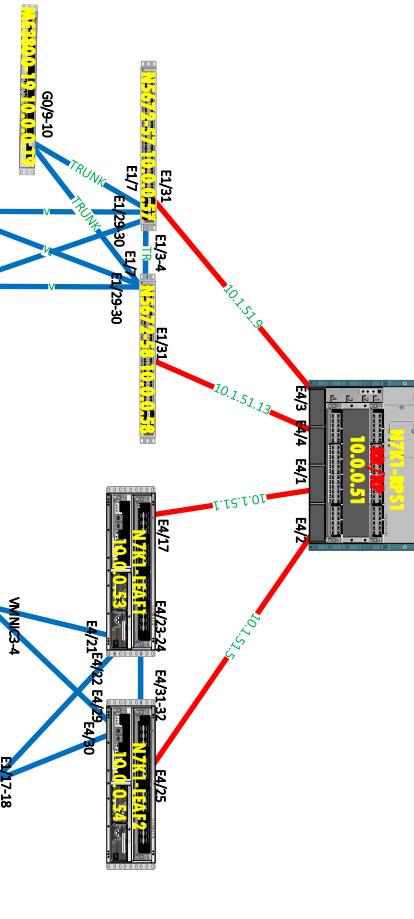
channel-group 34 mode active

no shutdown

DCNM N7K LEAF SPINE POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	Auto-Config										
<input checked="" type="checkbox"/> Enable DFS/dot1Q Trigger								true if Auto-configuration DFS/dot1q packet based triggers should be enabled										
	<input checked="" type="checkbox"/> Enable LLDP Trigger							true if Auto-configuration LLDP triggers should be enabled										
		<input checked="" type="checkbox"/> Enable VMTracker Trigger																
			<input checked="" type="checkbox"/> Enable VMTracker Configuration	{{sjrpvc1.10.0.1.2.s}}				Configure Multiple vCenters for VMTracker Trigger. Limit to a MAXIMUM of 4 connections.										
				<input checked="" type="checkbox"/> Enable VDP				Enable Autoconfig vdp feature										
					EVB MAC	0180.c200.0000		Shared EVB MAC address										
<p>vCenter Configuration</p> <p>Configure Multiple vCenters for VMTracker Trigger. Limit to a MAXIMUM of 4 connections.</p> <table border="1"> <thead> <tr> <th> Add</th> <th> Remove</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> </tbody></table>									 Add	 Remove								
 Add	 Remove																	

AFTER POAP FINISHED – FABRIC is Ready



```
N5672-57# sh bgp 12 evpn sum
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.0.0.57, local AS number 65000
BGP table version is 98, L2VPN EVPN config peers 1, capable peers 1
36 network entries and 55 paths using 4544 bytes of memory
BGP attribute entries [17/2448], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [3/12]
```

```
Neighbor          V   AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd
10.0.0.51        4 65000 4065    4015      98     0     0 2d17h 21
```

nve1 is up

Hardware: NVE

BW 0 Kbit,

```
N7K1-LEAF1# sh bgp 12 evpn sum
```

```
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.0.0.53, local AS number 65000
BGP table version is 91, L2VPN EVPN config peers 1, capable peers 1
36 network entries and 55 paths using 4544 bytes of memory
BGP attribute entries [18/2592], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [3/12]
```

```
Neighbor          V   AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd
10.0.0.51        4 65000 5974    5912      91     0     0 4d02h 0
```

```
N7K1-LEAF1# sh int nv 1
```

nve1 is up

admin state is up Hardware: NVE

Encapsulation VXLAN

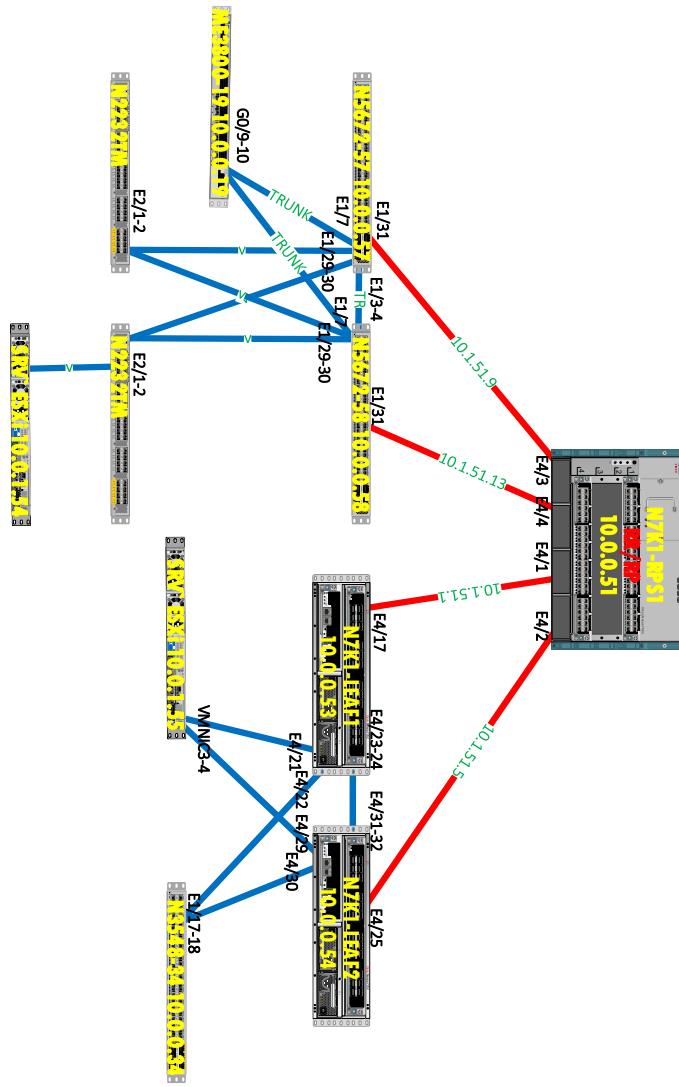
Auto-MDIX is turned off

RX

TX ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes

ucast: 0 pkts, 0 bytes - mcast: 39 pkts, 6318 bytes

AFTER POAP FINISHED – FABRIC is Ready



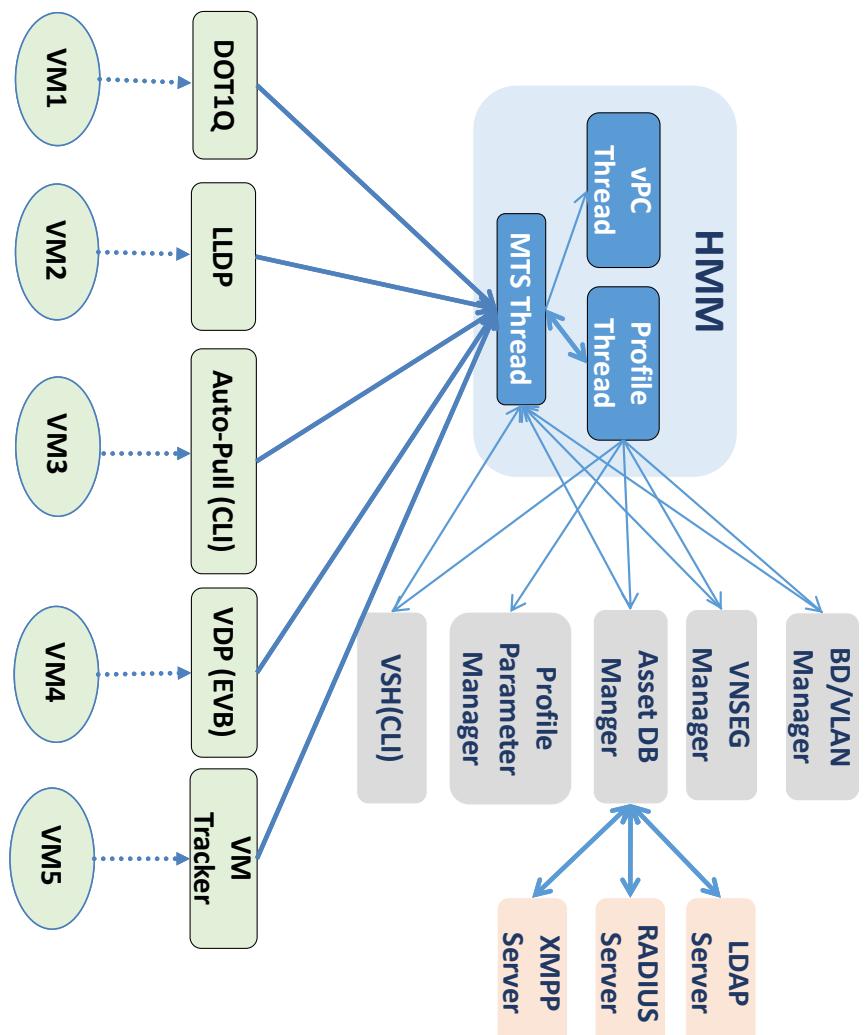
SO FBRIC is ready then WHAT?

There is no VRF, L2VNI, L3 VNI, EVPN, VRF, SVI configured yet???

Auto-configuration

- POAP populates configuration automatically for FABRIC side and Auto-Configuration configures ACCESS side.
- In a multi-tenancy network, consumers/tenants will be dynamically added/removed from the network
- Auto-configuration provides the ability to dynamically allocate/deallocate resources for every tenant on every leaf
 - For example, creating the VRF for the tenant, allocating the SW resources (such as a bridge-domain, vlan, vni)
- Auto-configuration can be triggered using various events generated on host add/delete
 - **Auto-Pull** - Use CLI to trigger dot1q or vni based configuration retrieval to test or pre-provision the network configuration.
- **VNI Discovery Protocol (VDP)** – A standard protocol that is implemented by Hypervisors to configure and unconfigure VMs created/deleted behind them
- **Dynamic Frame Snooping (DFS)** – First vlan tagged packet sent by the host triggers allocation of the resources. Last host removed from the network segment triggers deallocation of resources
- **VM Tracker** : Auto-Configuration based on VMWare VCenter events
- **LLDP** : Bare Metal Auto-Configuration based on LLDP protocol information from connected hosts

Software Components Overview



- HMM Threads
- MTS Thread – Handles MTS to/from other processes.
- Profile Writer Thread – Downloads profiles from ADBM, interacts with PD processes, applies profile via PPM.
- vPC Thread – Syncs profile to vPC peer.

1. Host Triggers send MTS to HMM with VLAN info
2. HMM notifies PPM to apply config profiles
3. PPM contacts ADBM if no local profile exist
4. ADBM query external DB, LDAP, Radius, XMPP, and etc to get config profiles
5. PPM notifies CLI to apply configuration

Auto-Configuration – Configuration Profile

- Configuration profile in DFA
 - Collection of commands used to instantiate a specific day-1 tenant-related configurations on a DFA leaf on need basis.
 - Commands are entered using variables for certain parameters, instead of entering the actual value
 - HMM has the mappings from the trigger to identify an unique profile to the profile name.
 - This profile can be downloaded from the remote server if it is not available on the switch
- Profile Types
 - Static Profile
 - Totally static or Parameterized - No parameters are downloaded from the remote server
 - Dynamic Profile
 - Here the parameters are downloaded from the remote server. Some of the parameters can also be automatically filled in by HMM. That would be the VLAN, VN-Segment (VNI) and also BGP asn.
- Universal Profile – from 7.1(0)N1(1)/7.2(0)D1(1)/7.1(1)
 - Enhanced configuration profiles with the ability to support optional parameters and non-disruptive profile refresh
 - Flexible and no need to specify VRF profiles in POAP any more
 - Updating and/or removing profile parameters (arguments or variables) without disrupting the traffic

Auto-Configuration Profiles

Cisco Prime Data Center Network Manager

Dashboard ▾ Health ▾ Performance ▾ Inventory ▾ Reports ▾ Config ▾ Admin ▾

Return to Launchpad

POAP Steps

- ✓ Generate or Upload Config
- ✓ Select Switch Config Template
- ✓ Save Settings
- ✓ LEAF-1
- ✓ IPFabric
- ✓ Copy
- ✓ Jobs
- ✓ Power-on Auto Provisioning (POAP)
- ✓ Template Parameters - 1 SW
- ✓ General
- ✓ Manageability
- ✓ Enable DFSDetect Trigger
- ✓ Provision
- ✓ SME
- ✓ Install Smartcard Drivers
- ✓ End Host
- ✓ Border Leaf / BorderPE Device Pairing
- ✓ Extended Partitions

DHCP Scope

Profiles (106 rows)

Profile Name	Type	Sub Type	Description	Forwarding Mode	Editable	Last Modified Time
serviceNetworkUniversalDynamicRoutingBPI	FPBD	network:universal	Profile for an enabled service network segment in the non-dc	none	yes	Thu Feb 18 02:31:02 PST 2016
serviceNetworkUniversalEfl3VpathServiceCdc	FPBD	network:universal	Profile for an enabled vPath L3 mode service network segme	proxy-gateway	yes	Thu Feb 18 02:31:02 PST 2016
serviceNetworkUniversalESChainStaticLBESProfile	FPBD	network:universal	Profile for an enabled service network segment in the non-dc	none	yes	Thu Feb 18 02:31:02 PST 2016
serviceNetworkUniversalESChainStaticLBESProfile	FPBD	network:universal	Profile for an enabled service network segment in the non-dc	none	yes	Thu Feb 18 02:31:02 PST 2016
serviceNetworkUniversalTfI3VpathServiceNc	FPBD	network:universal	Profile for an enabled vPath L3 mode service network segme	anycast-gateway	yes	Thu Feb 18 02:31:02 PST 2016
serviceNetworkUniversalTrStaticRoutingProfil	FPBD	partition:universal	Universal profile for a DFA partition.	none	yes	Thu Feb 18 02:31:02 PST 2016
vrf-common-universal-bl	FPBD	partition:universal	Universal profile for Border Leaf in a Two box solution. This w	none	yes	Thu Feb 18 02:31:02 PST 2016
vrf-common-universal-bl-dc-edge	FPBD	partition:universal	This profile is used on N7K box that acts as both Border Leaf	none	yes	Thu Feb 18 02:31:02 PST 2016
vrf-common-universal-dc-edge	FPBD	partition:universal	Universal profile for DC Edge router in a Two box solution. Th	none	yes	Thu Feb 18 02:31:02 PST 2016

Profile Instances (0 rows)

Organization Name	Partition Name	VRF Name	Segment ID	VLAN ID	Mobility Domain ID	Network Name
-------------------	----------------	----------	------------	---------	--------------------	--------------

You can see all pre-defined Profiles

Auto-Configuration

Profile is a collection of commands used to instantiate a specific configuration

Get a value from GUI then populate predefined configuration

```
config-profile vrf-common-evpn
vrf context $vrfName
vni $include_vrfSegmentId
rd auto
ip route 0.0.0.0/0 $include_serviceNodeIpAddress
address-family ipv4 unicast
route-target both auto
route-target both auto evpn
address-family ipv6 unicast
route-target both auto
route-target both auto evpn
router bgp $asn
vrf $vrfName
address-family ipv4 unicast
advertise l2vpn evpn
redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
maximum-paths ibgp 2
address-family ipv6 unicast
advertise l2vpn evpn
redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
maximum-paths ibgp 2
interface nve $nveld
member vni $include_vrfSegmentId associate-vrf
$xxx is input variable from GUI
```

```
config-profile defaultNetworkEvpnProfile
vlan $vlanId
vn-segment $segmentId
interface vlan $vlanId
vrf member $vrfName
ip address $gatewayIpAddress/$netMaskLength tag 12345
ip dhcp relay address $dhcpServerAddr1 use-vrf $vrfDhcp
ipv6 address $gatewayIpv6Address/$prefixLength tag 12345
fabric forwarding mode anycast-gateway
mtu $mtuValue
no shutdown
```

```
interface nve $nveld
member vni $segmentId
suppress-arp
mcast-group $mcastGroupIpAddress
evpn
vni $segmentId l2
rd auto
route-target import auto
route-target export auto
include profile any
```

```
L2 VNI related profile
- Config vlan/vn-segment
- Config SVI and connect to VRF
- Config Vni under NVE
- Configure VNI under EVPN
```

ADBM(Asset Database Manager)

- ADBM is responsible for looking up the configurations from external database based on the information passed from HMM, and send them back to HMM
- Supports RADIUS, LDAP and XMPP
- You can have multiple servers specified for a given type, but only one is active at any time
- Two types of query
 - Query for the profile name and the values of it's parameters if the profile is dynamically decided
- Query for the profile content when the profile is not defined in local switch
- Two types of dB types, corresponding to two types of queries
 - CLI - “fabric database type **network**”
 - To retrieve the profile name and values of the parameters associated with it
 - CLI - “fabric database type **profile**”
 - To retrieve the config-profile definition

```
fabric database type network
server protocol ldap host sjrpdcmn1.sjrp.local vrf management
db-table ou=networks,dc=cisco,dc=com key-type 1
db-security user cn=reader,dc=cisco,dc=com password 7 iwf1c
fabric database type profile
server protocol ldap host sjrpdcmn1.sjrp local vrf management
db-table ou=profilesIPFabric,dc=cisco,dc=com
db-security user cn=reader,dc=cisco,dc=com password 7 iwf1c
```

ADBM(Asset Database Manager)

- Show command
 - Show fabric database statistics

```
N5672-57# sh fabric database statistics
DB-Type          Requests   Dispatched  Not dispatched  Re-dispatched
-----          -----
network          0           0           0           0
profile          0           0           0           0
partition        0           0           0           0
bl-dci           0           0           0           0
host             0           0           0           0
-----          -----
TOTAL            0           0           0           0
Per Database stats:
  T Prot Server/DB
-----          ReqS    OK     NoRes   Err   TmOut  Pend
-----          -----
*ne ldap sjrpdcn1.sjrp.local      0       0       0       0       0       0
*pr ldap sjrpdcn1.sjrp.local      0       0       0       0       0       0
*pa ldap sjrpdcn1.sjrp.local      0       0       0       0       0       0
*ho ldap sjrpdcn1.sjrp.local      0       0       0       0       0       0
```

- Debug commands
 - debug adb trace | warning | error...
 - test fabric database client-request type network
 - This test CLI simulates the whole query and lookup process by sending ADBM a query with particular information such as segmentId, vlan and mobilityDomain etc.
 - If the lookup succeeds, the result will be printed out on console

Auto-Configuration GUI

Cisco Prime Data Center Network Manager

Dashboard ▾ Health ▾ Performance ▾ Inventory ▾ Reports ▾ Config ▾ Admin ▾

POAP Steps

- Return to Launchpad
- POAP Definitions**
 - Generate or Upload Config
 - Select Switches
- Template Parameters - 1 Step
- General
- Manageability
- DHCP Scope

Select Switch Config Tier

Template

Saved Settings

LEAF-1, LEAF-2

Template Parameters - 1 Step

General

Manageability

Enable DCFxIO Trigger

Archive

View

Compare

Copy

Delivery

Templates

Jobs

Power-On Auto Provisioning (POAP)

Fabric

Auto-configuration

Profiles

SMF

Provision

Install Smartcard Drivers

Border Leaf / BorderPE Device Pairing

Extended Partitions

Organizations (1 rows)

Organization/Partition Name	Description	Partition ID	Orchestration Engine	DCI
kk				

Networks (0 rows)

Network Name	Partition Name	Segment ID	VLAN ID	Mobility Domain	Profile Name	Subnet	Gateway	IP Range
--------------	----------------	------------	---------	-----------------	--------------	--------	---------	----------

Click Auto-Configuration

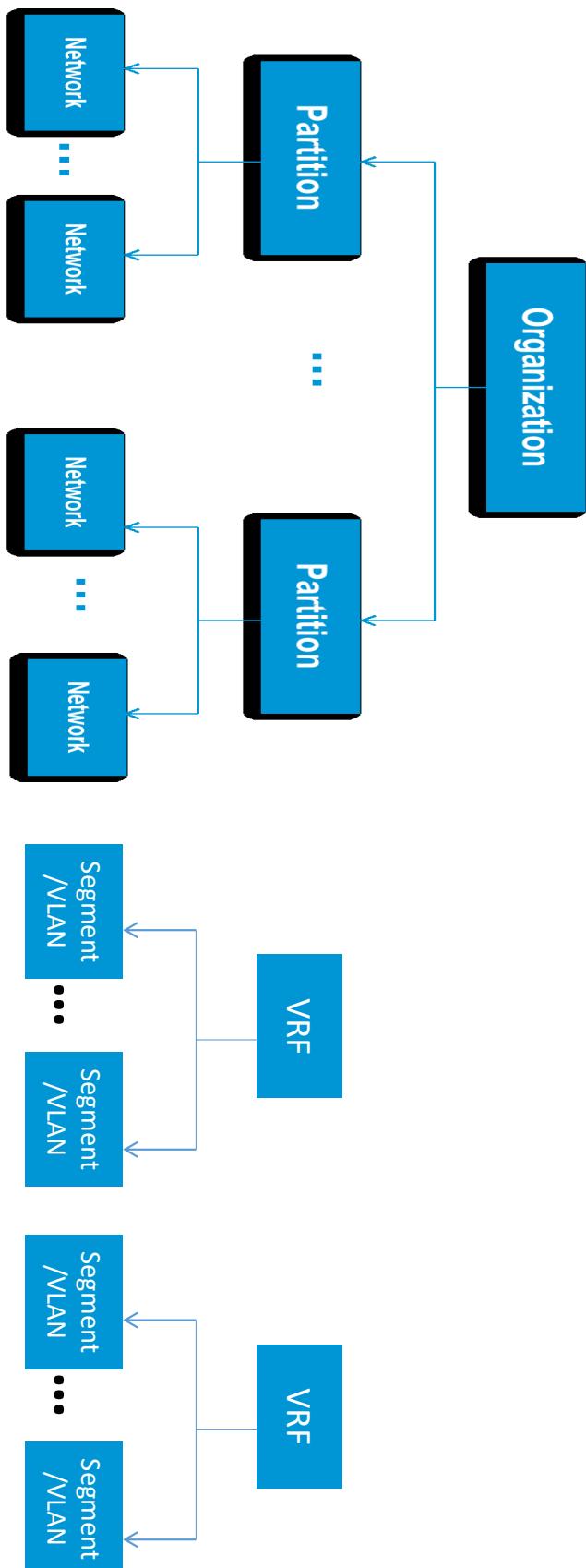
Click Add and two options, ORG and Partition

This two makes VRF, ORG:PARTITION is VRF name

cisco

Auto-Configuration Terminology

DCNM **vs.** **Networking**



Auto-Configuration – ORG:PART = VRF

Edit Organization

Name: kk
Description:
Orchestrator Name: kk

Edit Partition

Organization: kk

Name: a

VRF Name: kk:a

Partition ID: 50000

Description: kk's vrf a

PartID is from
General setting

OK Cancel

x

Let's create VRF

* Profile Name:

Choose proper profile

Profile Parameters

Required 'vrfName' format is 'organizationName:partitionName'

vrfName:

include_vrfSegmentId:

include_serviceNodeIpAddress:

asn:

nveId:

N7K supports 4 NVF interfaces

x

Auto-Configuration - NETWORK

Organizations (1 rows)

Organization/Partition Name	Description	Partition ID	Orchestration Engine	DCI ID	Extension Status	Profile	Service Node
kk a	kk's vrf a	50000	kk		1	Not Extended	vrf-common-evpn

Select proper ORG:PART first then click ADD

Add Network

* Organization: kk
 * Partition: a
 * VRF Name: kk:a
 * Network Name: bdi200
 * Multicast Group Address: 227.2.2.3
 * Network Role: Host Network
 * Gateway IPv4 Address: 10.120.0.254
 * Netmask Length: 24
 * Gateway IPv6 Address:
 * Prefix Length:
 * Network ID:

Profile Parameters

Required 'vrfName' format is 'organizationName;partitionName'

* Profile Name: defaultNetworkEvpnProfile

vlanId: 1200
 segmentId: 30002
 vrfName: kk:a
 gatewayIpAddress: 10.120.0.254
 netMaskLength: 24
 dhcpServerAddr1:
 vrfDhcp:

Select Seg ID only for VDP trigger
Use MD for any dot1q trigger

Segment ID Only
 Mobility Domain and VLAN

Segment ID: 30002
 Generate Seg ID
 VLAN ID: 1200
 Mobility Domain ID: mdo

Service Configuration Parameters

VM Manager IP:
 Static IP Start:
 Static IP End:
 vSwitch Controller Network Id:
 Distributed Virtual Switch Id:
 Secondary Gateway IPv4 Address:

OK **Cancel**

Auto-Configuration – LDAP Content on DCNM

We can use LDAP browser to access DCNM and see how LDAP entries are created from GUI

The screenshot shows two separate tables of LDAP configuration data, likely from a DCNM system, displayed in a JXplorer interface.

Table 1 (Top):

attribute type	value
orgName	kk
partitionName	a
partitionSegmentId	50000
segmentId	1200
include_vrfsegmentId=	
include_serviceNodeIpAddress=	
\$asn=	
\$nvId=1	
1	
kk's vrf a	
partition	kk:a
vrfName	vrf-common-avon
vrflProfileName	
dnsServerIpAddress	
mcastGroupIpAddress	
secondaryDnsServerIpAddress	
serviceChanneleDeploymentProfileName	
serviceNodeIpAddress	

Table 2 (Bottom):

attribute type	value
mobilityDomainId	md0
defaultNetworkEponProfile	
segmentId	30000
vlanId	1200
segmentId=1200	
segmentId=30000	
\$vrfName=kk:a	
\$gateWayIp=10.120.0.254	
\$netMaskLength=24	
\$dhcpServerAddr1=	
\$vrfIcp=	
\$dhcpServerAddr2=	
\$gateWayIp6Address=2001:120::254	
\$prefixLength=64	
\$mtuValue=	
\$nvId=1	
\$mcastGroupIpAddress=227.2.2.1	
\$include_vrfsegmentId=50000	
10.120.0.254	
2001:120::254	
227.2.2.1	
24	
bd1:200	
Host Network	
network	
objectClass	
orgName	kk
partitionName	a
prefixLength	64
vrfName	kk:a
description	
DvSId	
secondaryGatewayIpAddress	
staticIpEnd	
staticIpStart	
vSwitchControllerId	
vSwitchControllerNetworkId	

The screenshot shows a single table of LDAP configuration data, likely from a DCNM system, displayed in a JXplorer interface. A callout box highlights the 'orgName' and 'partitionName' fields.

attribute type	value
orgName	kk
partitionName	a
secondaryGatewayIpAddress	
staticIpEnd	
staticIpStart	
vSwitchControllerId	
vSwitchControllerNetworkId	

Callout Box Content:

These information will be downloaded by ABDM

Auto-Pull

- Typically, auto-configuration of network and tenant configuration is initiated via dynamic triggers.
- User may have a need to test or pre-provision the network configuration. The auto-pull CLI allows for this by retrieving configuration based on desired dot1q or vni.
- In general, Auto-configuration profiles remain applied until all hosts under a profile are removed. These hosts may be removed due to aging or because control protocol removed them.
- When a profile is auto-pulled, it remains applied until user explicitly clears the profile using the “clear fabric database host” CLI.
- Dynamic host triggers (e.g. from data frame snooping, VM Tracker, LLDP) are added alongside auto-pull host.
- If auto-pull is done first, instance name will be start with “instance_auto”.
- If dynamic trigger applies profile first and then auto-pull is issued for same dot1q, an auto-pull host will be added. Instance name will not change. See “Reload” section for implications of this ordering.

Auto-Pull with Dot1q

```
fabric database auto-pull dot1q 1800 interface po56
show fabric database host detail
Active Host Entries
Flags: L - Locally inserted, V - vPC+ inserted, R - Recovered, X - xlated Vlan
VLAN          VNI          STATE          FLAGS PROFILE(INSTANCE)
1800          11700        Profile Active  LV    defaultNetworkEvpnProfile(instance_auto_1800_1_369098801_1800)
Displaying Auto-Pull Ports
Interface      Encap          Flags          State          Profile
1800           LV            LV             Profile Active
po50
```

- Auto-pull with dot1q option will download and apply profile, and add a host entry for specified dot1q and interface.
- Profile and host entry will be flagged as LV.
- Instance name begins with “**instance_auto**”.
- While not strictly required, just as previous recommendation, auto-pull should be issued on both vPC peers.
- Issuing same CLI multiple times is a no-op.

Auto-Pull with VNI

```
fabric database auto-pull vni 11700 interface po50
show fabric database host detail
Active Host Entries
flags: L - Locally inserted, V - vPC+ inserted, R - Recovered, X - xlated Vlan
VNI          VLAN STATE           FLAGS PROFILE(INSTANCE)
11700        3000  Profile Active   LV    defaultNetworkEvpnProfile(instance_auto_11700_1_369098801_0)
Got no hosts for this profile
```

- Auto-pull with vni option will download and apply profile, but no host entry will be added.
- Profile will be flagged as LV.
- Instance name begins with “instance_auto”
- While not strictly required, just as previous recommendation, auto-pull should be issued on both vPC peers.
- Issuing same CLI multiple times is a no-op.

Default Profile Related Configuration after POAP

Running configuration related to Profile after POAP

```

ip host sjrpdcn1.sjrp.local 10.0.1.10
fabric database type network
server protocol ldap host sjrpdcn1.sjrp.local vrf management
db-table ou=networks,dc=cisco,dc=com key-type 1
db-security user cn=reader,dc=cisco,dc=com password fabr1c
fabric database type profile
server protocol ldap host sjrpdcn1.sjrp.local vrf management
db-table ou=profilesIPFabric,dc=cisco,dc=com
db-security user cn=reader,dc=cisco,dc=com password fabr1c
fabric database type partition
server protocol ldap host sjrpdcn1.sjrp.local vrf management
db-table ou=partitions,dc=cisco,dc=com
db-security user cn=reader,dc=cisco,dc=com password fabr1c
fabric database type host
server protocol ldap host sjrpdcn1.sjrp.local vrf management
db-table ou=hosts,dc=cisco,dc=com
db-security user cn=reader,dc=cisco,dc=com password fabr1c
fabric database mobility-domain mdo
fabric database profile-map global
vni default dynamic
dot1q default dynamic

configure profile vrf-tenant-profile
vlan $vrflVlanId
vn-segment $vrflSegmentId
interface vlan $vrflVlanId
vrf member $vrflName
    ip forward
    ipv6 forward
    no ip redirects
    no ipv6 redirects
    no shutdown

```

Auto-Configuration – Manual Deployment from GUI

Networks (2 rows)						
	Network Name	Partition Name	Segment ID	Mobility Domain	Profile Name	DHCP Scope
Add	Edit	Delete	Deploy Configuration	Undeploy Configuration	Logs	Print
<input type="checkbox"/>	bd1200	a	30000 30001	1200 1201	md0 md0	defaultNetworkEponProfile defaultNetworkEponProfile
<input type="checkbox"/>	bd1201	a				

Deploy Configuration

Deploy CLI config to selected switches. This will configure the auto-config process on those switches.

Select Device Type

Device Type	Device Name	Interface Name
<input checked="" type="checkbox"/> NSK-C5672UP	N5672-57 N5672-58	p028 p028
<input type="checkbox"/> NSK-C9372PX	N9372-42	
<input type="checkbox"/> NSK-C9372TX	N9372TX-43	
<input type="checkbox"/>	N9372TX-44	

Configuration is successfully pushed to device

OK

Before Auto-Configuration deployment, NO VLAN 1200, no VNI, no VRF exist

```
N5672-57# sh vrf
VRF-Name          VRF-ID State   Reason
default           1      Up      --
management        2      Up      --
```

```
N5672-58# sh vlan id 1200-1201 vn-segment
VLAN 1200-1201 not found in current VLAN database
```

```
N5672-58# sh nve vni
N5672-58#
```

```
N5672-57# sh vrf
VRF-Name          VRF-ID State   Reason
default           1      Up      --
kk:a management    5      Up      --
                           2      Up      --
```

```
N5672-57# sh nve vni
Interface VNI   Multicast-group  State Mode Type [BD/VRF] Flags
nve1       30000     227.2.2.1    Up   CP   L2 [1200]
nve1       50000     n/a          Up   CP   L3 [kk:a]      SA
```

Auto-Configuration – What really happens

```
N5672-57# debug adbml trace detail  
N5672-57# sh deb logfile kkabdm1.log  
2016 Mar 10 22:21:51.360631 hmm: Auto-pull triggered for dot1q 1200, vni 0 on interface port-channel19  
2016 Mar 10 22:21:51.361331 hmm: Match the global profile-map (default/UNI/2)  
2016 Mar 10 22:21:51.361347 hmm: Match the global map for UNI trigger 2:1200  
2016 Mar 10 22:21:51.361434 hmm: Adding active host on interface port-channel19, vlan 1200 (encap 1200), vni 0, mac -  
2016 Mar 10 22:21:51.361674 hmm: Updated auto_config vlan 1200 seg_id 0 in PS  
2016 Mar 10 22:21:51.362072 hmm: 1200: Sent to ADBM to get profile name  
2016 Mar 10 22:21:51.363353 hmm: Add ADBM attribute: mobility domain: mdo  
2016 Mar 10 22:21:51.363552 hmm: Add ADBM attribute: vlan_id: 1200  
2016 Mar 10 22:21:51.364082 hmm: called adbml_send_msg and got result 0  
2016 Mar 10 22:21:51.364118 hmm: Adding Entry 1 to PT  
2016 Mar 10 22:21:51.364124 adbml: adbml_get_query_key_type() return key_type 2 for database network  
2016 Mar 10 22:21:51.364137 hmm: [1200] ADBM Sent request for database type 1 and got req_id 1  
2016 Mar 10 22:21:51.364177 hmm: [1200] Updating Parent state to Profile Wait  
2016 Mar 10 22:21:51.364180 adbml: got TLV for mobility domain: mdo  
2016 Mar 10 22:21:51.364191 hmm: Get instance - Could not find profile with instance 0  
2016 Mar 10 22:21:51.364192 adbml: Got TLV for Vlan: 1200  
2016 Mar 10 22:21:51.364234 adbml: adbml_queue_enqueue() enqueued item 0x9fe375c and total is 1  
2016 Mar 10 22:21:51.364258 hmm: [1200] Updating Host state change [Err: No profile] => [Profile Wait] (HOST-VLAN port-channel19)  
2016 Mar 10 22:21:51.364272 adbml: adbml_handle_client_request() sent query to thread for db type network (1)  
2016 Mar 10 22:21:51.364285 adbml: LDAP query thread id 0x65162b90 got query.  
2016 Mar 10 22:21:51.364382 hmm: Increment outstanding ADBM request (0/50) -> (1/50)  
2016 Mar 10 22:21:51.365090 adbml: adbml_ldap_bind: enable_ssl:NO  
2016 Mar 10 22:21:51.366540 adbml: Attempting bind to sjrpdcnml.sjrp.local:0, u:cn=reader,dc=cisco,dc=com, p:iwfw1c  
2016 Mar 10 22:21:51.373215 adbml: ldap_bind succeeded  
2016 Mar 10 22:21:51.373256 adbml: adbml_bind_unbind_syslog: LDAP connect for sjrpdcnml.sjrp.local  
2016 Mar 10 22:21:51.373410 adbml: Allocated Ldap search filter (&(vlanId=1200)(mobilityDomainId=md0)) for key 2  
2016 Mar 10 22:21:51.373711 adbml: ldap_search sent, msgid 2  
2016 Mar 10 22:21:51.376221 adbml: Getting next reply for fd 20.  
2016 Mar 10 22:21:51.377869 adbml: Got LDAP_RES_SEARCH_RESULT (0x65)  
2016 Mar 10 22:21:51.377939 adbml: The number of entries returned was 1  
2016 Mar 10 22:21:51.377968 adbml: adbml_alloc_rep_msg(): setting dB type network  
2016 Mar 10 22:21:51.378011 adbml: adbml_ldap_process_one_reply(): allocate mem for rsp msg 0x9fcfcf124 for size 12  
2016 Mar 10 22:21:51.378052 adbml: NETWORK table
```

GUI triggered auto-pull on dot1 1200

HMM check profile name for vlan 1200 with ADBM

ADBML has no profile for vlan 1200

HMM ask ADBM to send out query

ADBML does LDAP Query on vlan1200/md0

Auto-Configuration – What really happens

```
2016 Mar 10 22:21:51.378142 adbm: Got 1 values for 'profileName'.  
2016 Mar 10 22:21:51.378224 adbm: Inserting ADBM_TLV_PROFILE_NAME defaultNetworkEvpnProfile
```

```
2016 Mar 10 22:21:51.378271 adbm: Got 1 values for 'gatewayIpAddress'.
```

```
2016 Mar 10 22:21:51.378346 adbm: Inserting ADBM_TLV_PROFILE_ARGV $gatewayIpAddress=10.120.0.254
```

```
2016 Mar 10 22:21:51.378374 adbm: Inserting ADBM_TLV_PROFILE_ARGV $netMaskLength=24
```

```
2016 Mar 10 22:21:51.378422 adbm: Got 14 values for 'configArg'.
```

```
2016 Mar 10 22:21:51.378449 adbm: Inserting ADBM_TLV_PROFILE_ARGV $vlanId=1200
```

```
2016 Mar 10 22:21:51.378474 adbm: Inserting ADBM_TLV_PROFILE_ARGV $segmentId=30000
```

```
2016 Mar 10 22:21:51.378529 adbm: Inserting ADBM_TLV_PROFILE_ARGV $vrfName=kk:a
```

```
2016 Mar 10 22:21:51.378558 adbm: Inserting ADBM_TLV_PROFILE_ARGV $gatewayIpAddress=10.120.0.254
```

```
2016 Mar 10 22:21:51.378589 adbm: Inserting ADBM_TLV_PROFILE_ARGV $netMaskLength=24
```

```
2016 Mar 10 22:21:51.378659 adbm: Inserting ADBM_TLV_PROFILE_ARGV $dhcpServerAddr1=
```

```
2016 Mar 10 22:21:51.378713 adbm: Inserting ADBM_TLV_PROFILE_ARGV $vrfDhcp=
```

```
2016 Mar 10 22:21:51.378775 adbm: Inserting ADBM_TLV_PROFILE_ARGV $gatewayIpv6Address=2001:1201::254
```

```
2016 Mar 10 22:21:51.378803 adbm: Inserting ADBM_TLV_PROFILE_ARGV $prefixLength=64
```

```
2016 Mar 10 22:21:51.378828 adbm: Inserting ADBM_TLV_PROFILE_ARGV $mtuValue=
```

```
2016 Mar 10 22:21:51.378853 adbm: Inserting ADBM_TLV_PROFILE_ARGV $nvneId=1
```

```
2016 Mar 10 22:21:51.378878 adbm: Inserting ADBM_TLV_PROFILE_ARGV $mcastGroupIpAddress=227.2.2.1
```

```
2016 Mar 10 22:21:51.378903 adbm: Inserting ADBM_TLV_PROFILE_ARGV $include_vrfSegmentId=50000
```

```
2016 Mar 10 22:21:51.378945 adbm: Got 1 values for 'vrfName'
```

```
2016 Mar 10 22:21:51.378971 adbm: Inserting ADBM_TLV_VRF kk:a
```

```
2016 Mar 10 22:21:51.379005 adbm: Got 1 values for 'mcastGroupIpAddress'.
```

```
2016 Mar 10 22:21:51.379031 adbm: Inserting ADBM_TLV_MCAST_GROUP 227.2.2.1
```

```
2016 Mar 10 22:21:51.379057 adbm: Inserting ADBM_TLV_PROFILE_DELIMITER
```

```
2016 Mar 10 22:21:51.379082 adbm: Done processed Ldap entry #1
```

```
2016 Mar 10 22:21:51.379112 adbm: Processed Ldap Entries 1 out of total 1
```

```
2016 Mar 10 22:21:51.379152 adbm: Clearing LDAP error
```

```
2016 Mar 10 22:21:51.379179 adbm: Send reply msg #1 with RET_SUCCESS token/req_id 1
```

```
2016 Mar 10 22:21:51.379267 hmm: Got callback from ADBM
```

```
2016 Mar 10 22:21:51.379325 hmm: Decrement outstanding ADBM request (1/50) -> (0/50)
```

ABDM got LDAP Query result back from DCNM

Auto-Configuration – What really happens

```
2016 Mar 10 22:21:51.379503 hmm: [1200] ADBM Mark query done [000000010] : network
2016 Mar 10 22:21:51.379578 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Name: defaultNetworkEvpnProfile
2016 Mar 10 22:21:51.379602 hmm: [1200] ADBM Mark to-do query flags [010000010] : profile content
2016 Mar 10 22:21:51.379634 hmm: Adding profile defaultNetworkEvpnProfile
2016 Mar 10 22:21:51.379651 hmm: Adding Instance Entry 1
2016 Mar 10 22:21:51.379666 hmm: Adding Entry 1 to PT
2016 Mar 10 22:21:51.379682 hmm: Added active profile defaultNetworkEvpnProfile with instance num 1 and instance name instance_def_1200_1
2016 Mar 10 22:21:51.379700 hmm: hmm_auto_config_syslog - msg_id 1, parent_entry 0x68b040e4 parent_entry type 1 vlan_id 0, cmd_data_vlan_id 0
key_bd_id 1200 seg_id 0 cmd_data_vni 0 key_seg_id 0
2016 Mar 10 22:21:51.379802 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $gatewayIpAddress=10.120.0.254
2016 Mar 10 22:21:51.379865 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $netMaskLength=24
2016 Mar 10 22:21:51.379892 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $vlanId=1200
2016 Mar 10 22:21:51.379936 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $segmentId=30000
2016 Mar 10 22:21:51.380016 hmm: Retrieve VNI: 3000 from ADBM
2016 Mar 10 22:21:51.380088 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $vrfName=kk:a
2016 Mar 10 22:21:51.380151 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $gatewayIpAddress=10.120.0.254
2016 Mar 10 22:21:51.380205 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $netMaskLength=24
2016 Mar 10 22:21:51.380244 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $dhcpServerAddr1=
2016 Mar 10 22:21:51.380286 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $vr-fDhcp=
2016 Mar 10 22:21:51.380355 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $gatewayIpv6Address=2001:1201::254
2016 Mar 10 22:21:51.380423 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $prefixLength=64
2016 Mar 10 22:21:51.380457 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $mtuValue=
2016 Mar 10 22:21:51.380492 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $mtuValue=
2016 Mar 10 22:21:51.380539 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $icastGroupIpAddress=227.2.2.1
2016 Mar 10 22:21:51.380591 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, Profile Parameter: $include_vrfSegmentId=50000
2016 Mar 10 22:21:51.380646 hmm: [1200] ADBM Receive [ADD] Entry: PARENT, DB: NETWORK, VRF TLV: kk:a
2016 Mar 10 22:21:51.380695 hmm: vrfname not found in profile
```

HMM found there is no VRF profile ☺

2016 Mar 10 22:21:51.382612 hmm: **Add ADBM attribute: vrf_name: kk:a**

2016 Mar 10 22:21:51.382761 adb: adb_handle_client_request() sent query to thread for db type partition (4)

2016 Mar 10 22:21:51.382794 **adb: Allocated Ldap search filter (vrfName=kk:a) for key 4**

2016 Mar 10 22:21:51.383171 hmm: Increment outstanding ADBM request (0/50) -> (1/50)

2016 Mar 10 22:21:51.383322 adb: ldap_search sent, msgid 3

2016 Mar 10 22:21:51.382676 adb: adb_get_query_key_type() return key type 4 for database type 4 and got req_id 2

2016 Mar 10 22:21:51.382687 hmm: [1200] ADBM Sent request for database type 4 and got req_id 2

2016 Mar 10 22:21:51.382700 adb: got TLV for vrf: kk:a

Auto-Configuration – What really happens

```
2016 Mar 10 22:21:51.384333 adbm: Getting next reply for fd 20...
2016 Mar 10 22:21:51.384582 adbm: Got LDAP_RES_SEARCH_RESULT (0x65)
2016 Mar 10 22:21:51.384616 adbm: The number of entries returned was 1

2016 Mar 10 22:21:51.384582 adbm: Got LDAP_RES_SEARCH_RESULT (0x65)
2016 Mar 10 22:21:51.384616 adbm: The number of entries returned was 1
2016 Mar 10 22:21:51.384644 adbm: adbm_alloc_rep_msg(): setting dB type partition
2016 Mar 10 22:21:51.384670 adbm: adbm_ldap_process_one_reply(): allocate mem for rsp msg 0x9fc46dc for size 12
2016 Mar 10 22:21:51.384701 adbm: **** Partition ****
2016 Mar 10 22:21:51.384732 adbm: Got 1 values for 'vrfProfileName'.
2016 Mar 10 22:21:51.384758 adbm: Inserting ADBM_TLV_PROFILE_NAME vrf-common-evpn
2016 Mar 10 22:21:51.384787 adbm: Got 1 values for 'vrfName'
2016 Mar 10 22:21:51.384812 adbm: Inserting ADBM_TLV_VRF_Kt:a
2016 Mar 10 22:21:51.384851 adbm: Got 5 values for 'configuring'.
2016 Mar 10 22:21:51.384887 adbm: Inserting ADBM_TLV_PROFILE_ARGV $vrfName=kk:a
2016 Mar 10 22:21:51.384903 adbm: Inserting ADBM_TLV_PROFILE_ARGV $include_vrfSegmentId=
2016 Mar 10 22:21:51.384928 adbm: Inserting ADBM_TLV_PROFILE_ARGV $include_serviceNodeIdIpAddress=
2016 Mar 10 22:21:51.384953 adbm: Inserting ADBM_TLV_PROFILE_ARGV $asn=
2016 Mar 10 22:21:51.384978 adbm: Inserting ADBM_TLV_PROFILE_ARGV $rnvId=1
2016 Mar 10 22:21:51.385009 adbm: Got 1 values for 'dcId'
2016 Mar 10 22:21:51.385036 adbm: Inserting ADBM_TLV_DCI_ID 1
2016 Mar 10 22:21:51.385071 adbm: Got 1 values for 'partitionSegmentId'.
2016 Mar 10 22:21:51.385097 adbm: Inserting ADBM_TLV_PARTITION_SEG_ID 50000
2016 Mar 10 22:21:51.385135 adbm: Inserting ADBM_TLV_PROFILE_DELIMITER
2016 Mar 10 22:21:51.385161 adbm: Done processed Ldap entry #1
2016 Mar 10 22:21:51.385185 adbm: Processed Ldap Entries 1 out of total 1
2016 Mar 10 22:21:51.385212 adbm: Clearing LDAP error
2016 Mar 10 22:21:51.385237 adbm: Send reply msg #1 with RET_SUCCESS token/req_id 2
2016 Mar 10 22:21:51.385306 himn: Got callback from ADBM
```

ADBM got VRF name from LDAP

```
2016 Mar 10 22:21:51.388071 himn: [1200] ADBM Run profile content check for 'vrf-common-evpn'
2016 Mar 10 22:21:51.388366 himn: [1200] ADBM Sent profile content query : [1200]
```

Auto-Configuration – What really happens

```
2016 Mar 10 22:21:51.388318 hnm: Check: 'vrf-common-evpn' Profile id exists (X) In profile content avl tree (V)
2016 Mar 10 22:21:51.388280 admn: LDAP query thread id 0x6a162b90 got query.
2016 Mar 10 22:21:51.388338 admn: Allocated Ldap search filter (profileName=vrf-common-evpn) for key 6

2016 Mar 10 22:21:51.390065 admn: Inserting ADBM_TLV_PROFILE_CONTENT:
2016 Mar 10 22:21:51.390109 admn: vrf context $vrfName
2016 Mar 10 22:21:51.390155 admn: vni $include_vrfSegmentId
2016 Mar 10 22:21:51.390180 admn: rd auto
2016 Mar 10 22:21:51.390205 admn: ip route 0.0.0.0/0 $include_serviceNodeIpAddress
2016 Mar 10 22:21:51.390230 admn: address-family ipv4 unicast
2016 Mar 10 22:21:51.390261 admn: route-target both auto
2016 Mar 10 22:21:51.390286 admn: route-target both auto evpn
2016 Mar 10 22:21:51.390311 admn: address-family ipv6 unicast
2016 Mar 10 22:21:51.390335 admn: route-target both auto
2016 Mar 10 22:21:51.390360 admn: route-target both auto evpn
2016 Mar 10 22:21:51.390384 admn: router bgp $asn
2016 Mar 10 22:21:51.390408 admn: vrf $vrfName
2016 Mar 10 22:21:51.390433 admn: address-family ipv4 unicast
2016 Mar 10 22:21:51.390457 admn: advertise l2vpn evpn
2016 Mar 10 22:21:51.390482 admn: redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
2016 Mar 10 22:21:51.390511 admn: maximum-paths ibgp 2
2016 Mar 10 22:21:51.390536 admn: address-family ipv6 unicast
2016 Mar 10 22:21:51.390560 admn: advertise l2vpn evpn
2016 Mar 10 22:21:51.390585 admn: redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
2016 Mar 10 22:21:51.390609 admn: maximum-paths ibgp 2
2016 Mar 10 22:21:51.390633 admn: interface nve $nvneId
2016 Mar 10 22:21:51.390658 admn: member vni $include_vrfSegmentId associate-vrf !
2016 Mar 10 22:21:51.390682 admn: ! done processed Ldap entry #1
2016 Mar 10 22:21:51.390708 admn: Done processed Ldap entry #1
2016 Mar 10 22:21:51.390733 admn: Processed Ldap Entries 1 out of total 1
```

ADBm sends out LDAP Query on VRF
PROFILE EVPN COMMON

Auto-Configuration – What really happens

```
2016 Mar 10 22:21:52.594545 adm:      admn_get_query_key_type() return key_type 6 for database profile
2016 Mar 10 22:21:52.594683 adm: Allocated Ldap search filter (profileName=defaultNetworkEvpnProfile) for key 6
2016 Mar 10 22:21:52.596928 adm: Got 1 values for 'configCommands'.
2016 Mar 10 22:21:52.596956 adm: Inserting ADBM_TLV_PROFILE_CONTENT:
2016 Mar 10 22:21:52.597007 adm:   vlan $vlanId^I
2016 Mar 10 22:21:52.597032 adm:     interface vlan $vlanId
2016 Mar 10 22:21:52.597056 adm:       vrf member $vrflName
2016 Mar 10 22:21:52.597081 adm:         ip address $gatewayIpAddress/$netMaskLength tag 12345
2016 Mar 10 22:21:52.597110 adm:           ip dhcp relay address $dhcpServerAddr1 use-vrf $vrflDhcp
2016 Mar 10 22:21:52.597136 adm:             ip dhcp relay address $dhcpServerAddr2 use-vrf $vrflDhcp
2016 Mar 10 22:21:52.597160 adm:               ipv6 address $gatewayIpv6Address/$prefixLength tag 12345
2016 Mar 10 22:21:52.597185 adm:                 fabric forwarding mode anycast-gateway
2016 Mar 10 22:21:52.597236 adm:                   mtu $mtuValue
2016 Mar 10 22:21:52.597261 adm:                     no shutdown
2016 Mar 10 22:21:52.597286 adm:                     interface nve $nvneId
2016 Mar 10 22:21:52.597310 adm:                       member vni $segmentId
2016 Mar 10 22:21:52.597335 adm:                         suppress-arp
2016 Mar 10 22:21:52.597359 adm:                           mcast-group $mcastGroupIpAddress
2016 Mar 10 22:21:52.597388 adm:                             evpn
2016 Mar 10 22:21:52.597413 adm:                               vni $segmentId 12
2016 Mar 10 22:21:52.597438 adm:                                 rd auto
2016 Mar 10 22:21:52.597462 adm:                                   route-target import auto
2016 Mar 10 22:21:52.597486 adm:                                     route-target export auto
2016 Mar 10 22:21:52.597511 adm:                                       include profile any
2016 Mar 10 22:21:52.597535 adm:                                         !
2016 Mar 10 22:21:52.597561 adm: Done processed Ldap entry #1
2016 Mar 10 22:21:52.597587 adm: Processed Ldap Entries 1 out of total 1
2016 Mar 10 22:21:52.597618 adm: Clearing LDAP error
2016 Mar 10 22:21:52.597643 adm: Send reply msg #1 with RET_SUCCESS token/req_id 4
2016 Mar 10 22:21:52.597723 hmm: Got callback from ADBM
```

ADBm sends LDAP Query on Network EVPN Profile

Auto-Configuration – What really happens

```
2016 Mar 10 22:21:53.803706 hmm: PPM CONFIG PROFILE REQ =====
2016 Mar 10 22:21:53..803735 hmm: OPER: 1
2016 Mar 10 22:21:53..803764 hmm: SRC -----
2016 Mar 10 22:21:53..803793 hmm: FLAGS: 1
2016 Mar 10 22:21:53..803822 hmm: PROFILE_NAME:
2016 Mar 10 22:21:53..803855 hmm: PROFILE_INSTANCE_NAME:
2016 Mar 10 22:21:53..803885 hmm: PROFILE_PARAM_LIST_NAME:
2016 Mar 10 22:21:53..803914 hmm: PROFILE_PARAM_INFO_Q:
2016 Mar 10 22:21:53..803947 hmm: gatewayIpAddress
2016 Mar 10 22:21:53..804015 hmm: netMaskLength
2016 Mar 10 22:21:53..804045 hmm: vIlanId
2016 Mar 10 22:21:53..804076 hmm: segmentId
2016 Mar 10 22:21:53..804105 hmm: vrfName
2016 Mar 10 22:21:53..804135 hmm: gatewayIpv6Address
2016 Mar 10 22:21:53..804165 hmm: prefixLength
2016 Mar 10 22:21:53..804199 hmm: nveId
2016 Mar 10 22:21:53..804229 hmm: macstGroupIpAddress
2016 Mar 10 22:21:53..804259 hmm: asn
2016 Mar 10 22:21:53..804288 hmm: DEST -----
2016 Mar 10 22:21:53..804317 hmm: DEST_PROFILE_NAME:
2016 Mar 10 22:21:53..804346 hmm: DEST_PROFILE_INSTANCE_NAME:
2016 Mar 10 22:21:53..804375 hmm: DEST_PROFILE_LIST_NAME:
2016 Mar 10 22:21:53..804403 hmm: DEST_PROFILE_PARAM_INFO_Q:
2016 Mar 10 22:21:53..804432 hmm: INCLUDE -----
2016 Mar 10 22:21:53..804461 hmm: INCLUDE_PROFILE_NAME:
2016 Mar 10 22:21:53..804494 hmm: INCLUDE_INSTANCE_NAME:
2016 Mar 10 22:21:53..804523 hmm: INCLUDE_LIST_NAME:
2016 Mar 10 22:21:53..804403 hmm: INCLUDE_PARAM_INFO_Q:
2016 Mar 10 22:21:53..804554 hmm: include_vrfSegmentId
2016 Mar 10 22:21:53..804584 hmm: vrfName
2016 Mar 10 22:21:53..804614 hmm: nveId
2016 Mar 10 22:21:53..804643 hmm: include_dciId
2016 Mar 10 22:21:53..804673 hmm: include_vrfSegmentId
2016 Mar 10 22:21:53..804702 hmm: vrfName
2016 Mar 10 22:21:53..804731 hmm: vrfName
2016 Mar 10 22:21:53..804764 hmm: asn
2016 Mar 10 22:21:53..804793 hmm: nveId
2016 Mar 10 22:21:53..804917 hmm: [1200] ppm_oper_req_send (PPM ADD) was successful
```

After all Q are done, HMM processes PPM
request for real configuration

Auto-Configuration – Manual Deployment

After ABDM/HMM/PPM/VSH pulled out profiles from DCNM, we can see downloaded profile on LEAF

```
configure profile vrf-common-evpn
vrf context $vr fName
vni $include_vrfSegmentId
rd auto
ip route 0.0.0.0/0 $include_serviceNodeIpAddress
address-family ipv4 unicast
route-target both auto
route-target both auto evpn
address-family ipv6 unicast
route-target both auto
route-target both auto evpn
router bgp $asn
vr f $vr fName
address-family ipv4 unicast
advertise 12vpn evpn
redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
maximum-paths ibgp 2
address-family ipv6 unicast
advertise 12vpn evpn
redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
maximum-paths ibgp 2
interface nve $nv eId
member vni $segmentId associate-vrf
vni $include_vrfSegmentId
```

```
configure profile defaultNetworkEvpnProfile
vlan $vlanId
vn -segment $segmentId
interface vlan $vlanId
vrf member $vr fName
ip address $gatewayIpAddress/$netMaskLength tag 12345
ip dhcp relay address $dhcpServerAddr1 use-vrf $vr fDhcp
ip dhcp relay address $dhcpServerAddr2 use-vrf $vr fDhcp
ipv6 address $gatewayIpv6Address/$prefixLength tag 12345
fabric forwarding mode anycast-gateway
mtu $mtuValue
no shutdown
interface nve $nv eId
member vni $segmentId
suppress-arp
mcast-group $mcastGroupIpAddress
evpn
vni $segmentId 12
rd auto
route-target import auto
route-target export auto
include profile any
```

Auto-Configuration – Manual Deployment

Also, we can see parameter-list that works with profile to create actual configuration commands

```
param-list include_param_list_name_kk:a_1
define asn string asn
define include_dciId string include_dciId
define include_vrfSegmentId string include_vrfSegmentId
define nvfId string nvfId
define vrfName string vrfName
instance kk:a_1
set asn 65000
set include_dciId 1
set include_vrfSegmentId 50000
set nvfId 1
set vrfName kk:a
param-list param_list
define bridgeDomainId integer bridge domain id
define vrfBdName string bd name
define vrfName string vrf name
define vrfSegmentId integer segment id
define vrfVlanId integer vlan id
instance param_inst_50000
set bridgeDomainId 0
set vrfBdName bd50000
set vrfName kk:a
set vrfSegmentId 50000
set vrfVlanId 2500
param-list param_list_name_instance_auto_1200_1_369098770_1200
define asn string asn
define gatewayIpAddress string gatewayIpAddress
define gatewayIpv6Address string gatewayIpv6Address
define mcastGroupIpAddress string mcastGroupIpAddress
define netMaskLength string netMaskLength
define nvfId string nvfId
define prefixLength string prefixLength
define segmentId string segmentId
define vianId string vianId
define vrfName string vrfName
instance instance_auto_1200_1_369098770_1200
set asn 65000
set gatewayIpAddress 10.120.0.254
set gatewayIpv6Address 2001:1201::254
set mcastGroupIpAddress 227.2.2.1
set netMaskLength 24
set nvfId 1
set prefixLength 64
set segmentId 30000
set vianId 1200
set vrfName kk:a
```

Auto-Configuration – Manual Deployment

```
N5672-57# sh run vlan 1200, 2500 expand-port-profile
vlan 1200
vn-segment 30000
vlan 2500
vn-segment 50000
interface Vlan1200
no shutdown
vrf member kk:a
ip address 10.120.0.254/24 tag 12345
ipv6 address 2001:1201::254/64 tag 12345
fabric forwarding mode anycast-gateway

interface Vlan2500
no shutdown
vrf member kk:a
no ip redirects
ip forward
ipv6 forward
no ipv6 redirects

vrf context kk:a
vni 50000
rd auto
address-family ipv4 unicast
route-target both auto
route-target both auto evpn
address-family ipv6 unicast
route-target both auto
route-target both auto evpn

router bgp 65000
vrf kk:a
address-family ipv4 unicast
advertise 12vpn evpn
redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
maximum-paths ibgp 2
address-family ipv6 unicast
advertise 12vpn evpn
redistribute direct route-map FABRIC-RMAP-REDIST-SUBNET
maximum-paths ibgp 2
route-map FABRIC-RMAP-REDIST-SUBNET permit 10
match tag 12345

evpn
vni 30000 12
rd auto
route-target import auto
route-target export auto

interface nve1
no shutdown
source-interface loopback0
host-reachability protocol bgp
member vni 30000
suppress-arp
cast-group 227.2.2.1
member vni 50000 associate-vrf
```

Why redistribute SVI subnet?
So, external network can reach any subnet in FABRIC

1200 is manually assigned by GUI
2500 is assigned from "system fabric core-vlans 2500-2999"

cisco

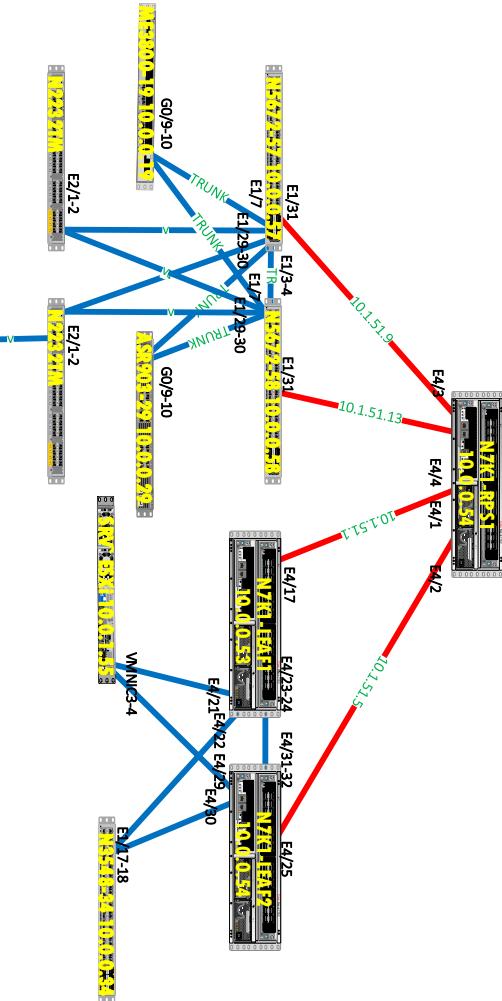
Auto-Configuration – Manual Deployment

HMM/PPM uses VSH to push config to switch
Always check accounting log for any failure

```
N5672-57# sh fabric database host det
Active Host Entries
  flags: L - Locally inserted, V - vPC+ inserted, R - Recovered, X - xlated Vlan
  VLAN VNI STATE
  1200 30000 Profile Active LV
    defaultNetworkEvpnProfile(instance_auto_1200_1_369098770_1200)
      Displaying Auto-Pull Ports
        Interface Encap Flags State
        Po19       1200   LV   Profile Active

N5672-57# sh fabric database host detail dot1q 1200
instance_index 1
Got Local and vPC Peer originated client trigger at 22:21:56
Number of associated interfaces: 1
Received Parameters from Database Manager at 22:21:51
Displaying parameters for profile defaultNetworkEvpnProfile and instance
instance_auto_1200_1_369098770_1200
parameter 0: $gatewayIpAddress=10.120.0.254
parameter 1: $netMaskLength=24
parameter 2: $vlanId=1200
parameter 3: $segmentId=30000
parameter 4: $vrName=kk:a
parameter 5: $dhcpServerAddr1=
parameter 6: $vrfdhcp=
parameter 7: $dhcpServerAddr2=
parameter 8: $gatewayIpv6Address=2001:1201::254
parameter 9: $prefixLength=64
parameter 10: $mtuValue=
parameter 11: $nvid=1
parameter 12: $mcastGroupIpAddress=227.2.2.1
parameter 13: $include_vrfSegmentId=50000
parameter 14: $asn=65000
Sent Apply to Configuration Manager at 22:21:53
Completed executing all commands at 22:21:55
Sent to vPC peer at 22:21:55
Completed executing all commands on vPC peer at 22:21:56
Displaying Auto-Pull Ports
  Interface Encap Flags State
  Po19       1200   LV   Profile Active
```

Auto-Configuration-Dot1Q Trigger on N6K



For vlan 1201/vni50000, let's trigger it by bring up svi1201 on ME3800-19
 ME3800-19 generates G-ARP in vlan 1201 when bring up SVI1201 and this triggers Auto-
 Configuration download

ME3800-19(config)#int vl 1201

ME3800-19(config-if)#no sh

```
N5672-57# sh fabric forwarding interface even auto-config | i 1201
2016 Mar 11 18:58:00.861737 hmm [5442]: [5450]: Adding with keys: seg-id 0,bd
1201
2016 Mar 11 18:58:00.861714 hmm [5442]: [5450]: Adding active host on interface
port-channel19, vlan 1201 (encap 1201), vni 0, mac 5897.1e9a.cc40
2016 Mar 11 18:58:00.861636 hmm [5442]: [5450]: [5450]: 1201:5897.1e9a.cc40 Got Dynamic
Profile Name (null)
2016 Mar 11 18:58:00.861472 hmm [5442]: [5450]: 1201:5897.1e9a.cc40 Got ARP
request on interface port-channel19
2016 Mar 10 22:21:53.804134 hmm [5442]: [5460]: gatewayIpv6Address
= 2001:1201::254
2016 Mar 10 22:21:53.803024 hmm [5442]: [5460]: [1200] Profile
defaultNetworkEvnProfile(1): Parameter $gatewayIpv6Address=2001:1201::254 and
token1.gatewayIpv6Address and token2.gatewayIpv6Address=2001:1201::254
2016 Mar 10 22:21:53.801201 hmm [5442]: [5460]: Inserting internal parameter
$azn=65000
2016 Mar 10 22:21:51.380388 hmm [5442]: [5450]: [1200] ADBM Receive [ADD] Entry:
PARENT, DB: NETWORK, Profile Parameter: $gatewayIpv6Address=2001:1201::254
```

HOST19 can ping def GW on N6K

ME3800-19#ping vrf host1201 10.120.1.254

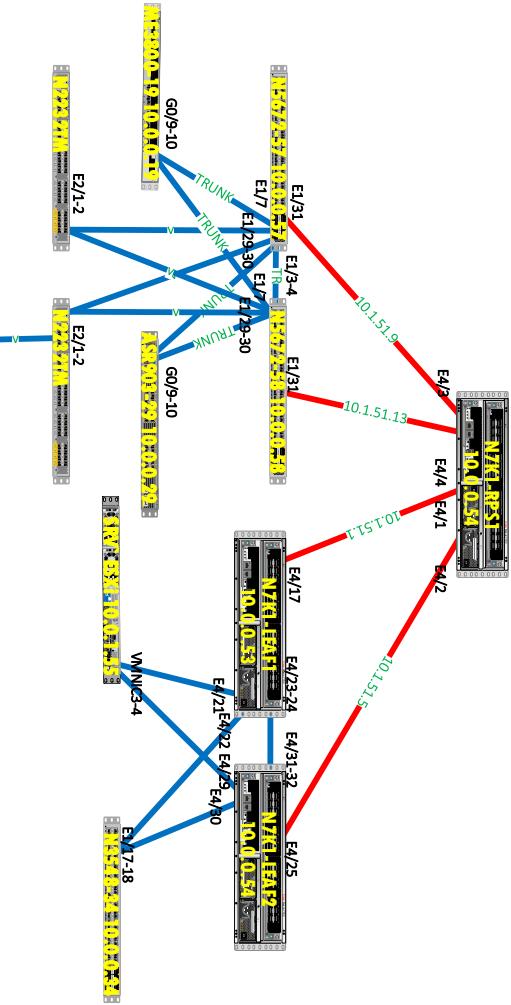
Type escape sequence to abort.

Sending 5, 100-byte ICMP Echoes to 10.120.1.254, timeout is 2 seconds:

!!!!

Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/4 ms

Auto-Configuration-Dot1Q Trigger on N6K



```
N5672-57# sh fabric database host detail
flags: L - Locally inserted, V - vPC+ inserted, R - Recovered, X - xlated VLAN
VLAN VNI STATE FLAGS PROFILE(INSTANCE)
1201 30001 Profile Active L defaultNetworkEvpnProfile(instance_def_1201_2)
Displaying Data Snooping Ports
Interface Encap Flags State
P019 1201 L Profile Active
```

leaf57 received PKT from HOST19, so flag is locally inserted

```
N5672-58# sh fabric database host detail
VLAN VNI STATE FLAGS PROFILE(INSTANCE)
1201 30001 Profile Active V defaultNetworkEvpnProfile(instance_def_1201_2)
Displaying Data Snooping Ports
Interface Encap Flags State
P019 1201 V Profile Active
```

leaf58 got synced from Leaf57 so flag is VPC inserted

```
N5672-57# sh mac add vlan 1201
Ports/SWID.SSID.LID
```

*	1201	2020.0000.00aa	static	0	F	F	sup-eth2		
*	1201	5897.1e9a.cc40	dynamic	220	F	F	Po19		
*	1201	8c60.4f2c.1481	static	0	F	F	nve1/10.0.0.58		

```
N5672-57# sh l2route evpn mac-ip evi 1201
```

Mac Address	Prod Host IP	Next Hop (s)
5897.1e9a.cc40	HMM 10.120.1.19	N/A

```
N5672-57# sh bgp 12vpn evpn vni-id 30001
Route Distinguisher: 10.0.0.57.33968 (L2VNI 30001)
* i[2]:[0]:[0]:[48]:[5897.1e9a.cc40]:[0]:[0.0.0.0]/216
* >1 10.0.0.102 100 0 i
* i[2]:[0]:[0]:[48]:[5897.1e9a.cc40]:[32]:[10.120.1.19]/272
* 10.0.0.102 100 0 i
* >1 10.0.0.102 100 32768 i
```

L2RB and EVPN routes are created

Auto-Configuration-Dot1Q Trigger on N7K

On N7K1-LEAF1 and LEAF2, bring up P034 on N3548-34 triggers auto-configuration

```
N7K1-LEAF1# sh fabric database host detail
Active Host Entries
flags: L - Locally inserted, V - vPC+ inserted, R - Recovered, X - xlated Vlan
VNI      BD ID STATE          FLAGS PROFILE(INSTANCE)
30000   3000  Profile Active  L
defaultNetworkEvpnProfile(instance_vni_30000_1)
Displaying hosts
```

Interface	Encap	Flags	State	Profile	Active	Client	Host
Po34	1200	L	Profile Active	PM			21000016B00040000

```
30001 3001 Profile Active L
defaultNetworkEvnProfile(instance_vni_30001_2)
```

Displaying hosts	Interface	Encap	Flags	State	Profile	Active	Client	Host
	Po34	1201	L	Profile Active	PM			21000016B1040000

N7K1-LEAF1# sh nve vni

Interface	VNI	Multicast-group	State	Mode	Type [BD/VRF]	Flags
nve1	30000	227.2.2.1	Up	CP	L2 [3000]	SA
nve1	30001	227.2.2.2	Up	CP	L2 [3001]	SA
nve1	50000	n/a	Up	CP	L3 [kk:a]	

N7K1-LEAF1# sh ip int b vrf kk:a

IP Interface Status for VRF "kk:a" (4)	IP Address	Interface Status
Interface		
Bd12500	forward-enabled	protocol-up/link-up/admin-up
Bd13000	10.120.0.254	protocol-up/link-up/admin-up
Bd13001	10.120.1.254	protocol-up/link-up/admin-up

N7K1-LEAF1# sh vni

VNI	Status	BD	VSI
			VSI???

	30000	30001	30002
ip address	Up	Up	Up
vrf member host1201			
ip address 10.120.1.34/24			

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

cisco

Auto-Configuration-Dot1Q Trigger on N7K

L3 VNI/VRF/NVE/EVPN configuration on N7K is same as n6K/N9K but VLAN/VNI is different.

```
feature nv overlay
feature vni
nv overlay vni
vni 30000-30001, 50000
system fabric bridge-domain 2500-2999

bridge-domain 3000-3001
member vni 30000-30001

interface Bdi2500
no shutdown
mtu 9192
vrf member kk:a
no ip redirects
ip forward
ipv6 address use-link-local-only
ipv6 forward
no ipv6 redirects

interface Bdi3000
no shutdown
vrf member kk:a
no ip redirects
ip address 10.120.0.254/24 tag 12345
ipv6 address 2001:1201:2:254/64 tag 12345
no ipv6 redirects
fabric forwarding mode anycast-gateway

interface nve1
no shutdown
source-interface loopback0
host-reachability protocol bgp
member vni 30000
suppress-arp
mcast-group 227.2.2.1
member vni 50000 associate-vrf

cisco
```

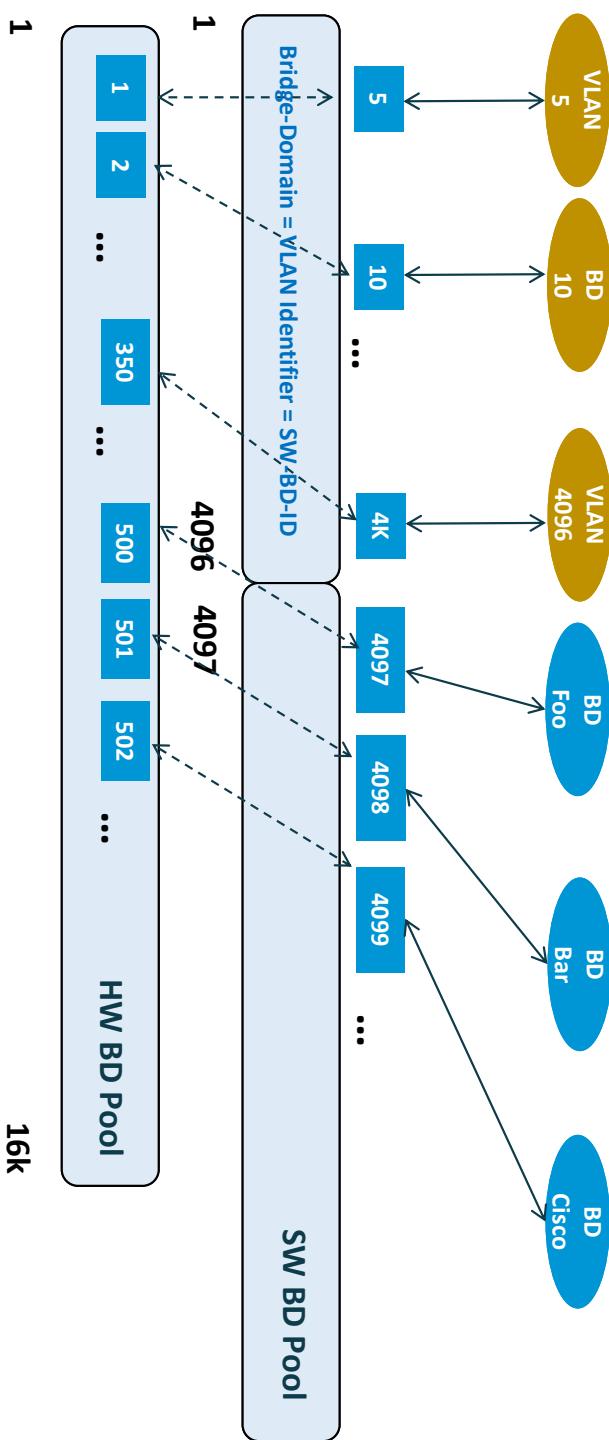
New Bridging Technology on N7K – Bridge Domain

- Introduced with N7K, we covered BD in MPLS FWD training on N7K briefly
- Overcome VLAN range limitation and solve VLAN address space collision problem
- EARL8 hardware provides hardware support for “*bridge domain (BD)*”
- In EARL8 all frames are associated with a logical interface or LIF
- During packet processing L2LU will associate src_port_index, vlan_id pair and map it to a BD and with LIF for bridging
- This allows port, vlan_id pair to extend from current 4K vlan_id to 16K BDs
- Bridge Domain adopted by all N7K Forwarding Engines (Eureka, Orion & Clipper)
- The L2 Forwarding Table tracks entries against Bridge Domain, this can map to the VLAN Id for a given VDC

VLAN	Bridge-Domain
Network Wide Entity	Switch local Entity
Represents actual 802.1q tag on the wire	A generic s/w entity representing L2 broadcast domain (VLAN or VN Segment can be mapped under BD)
Only numbered (1 to 4096)	Both Named and Numbered

SW BD's

- VLAN only services like PVLAN, VTP, FCoE etc. continue to operate on legacy VLAN and do not change
- The services that are required to be aware of > 4k broadcast domains or 24-bit segment-id are required to completely migrate to a **new SW BD based model**
- SW BD ID – numeric representation of named bridge-domain that is replacing VLAN ID.



N7K VNI, Bridge-domain and VSI

- VN-Segment Bridge has 1:1 mapping between VNI and Broadcast Domain
- VN-Segment has global significance in the L2 network while Bridge-domains remain local to the switch
- Bridge-domains would have VNIs as members
- **VNI Service Instance (VSI)**
 - Tenant traffic received as “Dot1Q” tag need to be classified to the VN-Segment assigned to those tenants.
 - Existing legacy IEEE 802.1Q switches and End-host/Servers should be able to connect to VN-Segment network. This capability is provided by VNI Service Instance(VSI).
 - VNI Service Instance ports on the VN-Segment capable switch allows to map the Dot1Q tagged frames received on that port uniquely to a VN-Segment
- **Numbered VSI**
 - Under any physical port or port-channel, numbered VSIs can be configured
 - VSI ids ranging from 1-4094
- **Default VSI**
 - Under any physical port or port-channel, default VSI can be configured
 - VSI id is **4095**, Always admin up

N7K VNI Configuration

```
N7K1-LEAF1(config)# feature vni  
Feature vni requires F3 or newer linecards  
system bridge-domain 2500-3500  
system fabric bridge-domain 2500-2999  
vni 30000-30001, 50000  
bridge-domain 30000-3001  
member vni 30000-30001  
  
encapsulation profile vni kk  
  dot1q 20 vni 20000  
  untagged vni 30000  
  
encapsulation profile vni md0  
  encapsulation dynamic frame-snoop profile md0  
  encapsulation dynamic vdp profile md0  
  
Interface port-channel11  
  service instance 10 vni  
  encapsulation profile KK default  
  
interface port-channel134  
  vpc 34  
  service instance vni default  
  encapsulation dynamic frame-snoop profile md0
```

1. Enable feature VNI
Feature vni' will be enabled only when the vdc has been limited to F3

2. Carved out BD range form VLAN range
VLAN should not be created for this BD range
Fabric Bridge-domain range will be used only for L3 VNI

2. create a VNI

3. Add VNI to BD.
Assigned BD range to VNI range. BD 3000 → VNI 3000, BD 3001 → VNI 30001

4. Create an Encapsulation Profile template for Auto-Configuration
Map vlan 20 → vni 2000

5. Create numbered VSI under port and adding ENCAP Profile
Create Default VSI under port and adding dynamic ENACP profile for auto-configuration
cannot create a default VSI and a numbered VSI under the same parent port

N7K EVPN Verification

```
N7K1-LEAF1# sh vlan brief
VLAN Name          Status    Ports
1 default         active   Po534, Eth4/18, Eth4/19
                                Eth4/20, Eth4/23
99 VLAN099        active   Po534, Eth4/23

N7K1-LEAF1# sh vlan internal bd-info bd-to-vlan all-bd
VDC Id BD Id Vlan Id
----- -----
2      25      1
2      27      4043
2      28      4042
2      29      4044
2      30      4040
2      31      4041
2      32      4045
2      33      99
2      34      3000
2      35      2500
2      36      3001
                                2500 is for L3VNI BD
                                3000-3001 is L2 VNI BD

N7K1-LEAF1# sh vlan internal info vlan-db vlan 3000
Entry No: 0
-----
vlan id 3000
vdc 2 bd_id = 34, bd_id_orig = 34, vlan_type = USER_VLAN(1)
sdb_vlan_type USER_VLAN(1, err 0x0), oper = up
vlan_state = active, vlan_oper = up, sdb_oper_state up(1, err 0x0)
oper_up technologies: CE VXLAN oper down technologies: configured
technologies: CE VXLAN
vlan_shut_state = no-shutdown, vlan_modes = (1)fabricpath
sdb_vlan_mode 1(fabricpath, err 0x0), n_access_ports 0, n_native_ports 0
mtu = 1500, internal_state = 0, cfg_flags 128
vlan_name = VLAN3000 segment_id = 0(err 0x0, sdb 0)
```

```
N7K1-LEAF1# sh spanning-tree bridge-domain 3000
BD3000
Spanning tree enabled protocol rstp
Root ID Priority 35768
Address 0023.04ee.c016
This bridge is the root
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Bridge ID Priority 35768 (priority 32768 sys-id-ext 3000)
Address 0023.04ee.c016
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Interface Role Sts Cost Prio.Nbr Type
----- -----
VSI-Po34.4095 Desg FWD 1 128.4129 (vPC) P2p
```

```
module-4# sh hardware internal mac port 21 table cbl vlan
```

Color-Based Logic (CBL) used for STP

FWD State in HW

		INGRESS
Disabled State	2-1199, 1202-4031, 4036-4095	
Forwarding State	0,1200-1201, 4032-4035	
Blocked State	1	
Learning State		

		EGRESS
Disabled State	2-2999, 3002-4031, 4036-4095	
Forwarding State	0,3000-3001, 4032-4035	
Blocked State	1	
Learning State		

Ingress shows vlan no incoming packets are tagged with dot1Q

Egress shows BD no

N7K EVPN Verification

```

N7K1-LEAF1# sh vni
VNI          Status   BD          VSI
30000        Up       3000        VSI-port-channel134. 4095
30001        Up       3001        VSI-port-channel134. 4095
50000        Up       2500        No VSI for L3 VNI.
                                         only for routing

N7K1-LEAF1# sh bridge-domain
Bridge-domain 3000 (2 ports in all)
Name:: Bridge-Domain3000
Administrative State: UP
Operational State: UP
vni30000
nve1

VSI-Po34..4095  Default VNI on PO34 is added to BD 3000
Bridge-domain 3001 (2 ports in all)
Name:: Bridge-Domain3001
Administrative State: UP
Operational State: UP
vni30001
nve1

VSI-Po34..4095

N7K1-LEAF1# sh service instance vni
VSI          Admin Status    Oper Status #BD
VSI-port-channel134.4095 Up           Up      2
VSI-port-channel155.4095 Up           Up      0
                                         SW BD IDs to be brought down:
                                         Admin Config Information:
                                         admin config PSS key(<7, 0x651134ffff0210>
                                         port_name(), admin_link_logging_enable(1)
                                         state(up), cfg-status(valid))
                                         Operational (Runtime) Information:
                                         runtime PSS key(<7, 0x331234ffff0210>
                                         state(up), Local ID(0x95)
                                         Allowed SW BD IDs: (3000-3001)  BD 3000 and 3001 is only allowed here
                                         Operational SW BD IDs: 3000-3001
                                         Errored SW BD IDs:
                                         Last bringup_ok(TRUE); VNSEG sdb state(TRUE)
                                         reason(None)
                                         iod(149), bringdown_reason(0)
                                         Encap tags()
                                         Information from SDB Query (IM call)
                                         GLDB PSS key(NULL)
                                         admin state(up), runtime state(up), mtu(1500),
                                         delay(1), bandwidth(10000000), Encap tags()
                                         VSI interface(net_val(0), status(0x40290000)) from parent port(0x16000021)/Vlan(0)
                                         VSI interface FSM state(3)

No errors on VSI interface
cisco
  
```

N7K EVPN Verification

VLAN/BD	MAC Address	Type	age	Secure	NTFY	Ports/SWID	SSID/LID
G 3000	0026.51c9.5042	static	-	F	F	VPC Peer-Link(R)	
G 3000	0026.51c9.5043	static	-	F	F	Sup-eth1(R)	
*	3000	30f7.0d9c.39bc	dynamic	~~~	F	VSI-Po34.4095	
E 3000	5897.1e9a.cc40	dynamic	-	F	F	nve1/10.0.0.102	
E 3000	64f6.9de3.2520	dynamic	-	F	F	nve1/10.0.0.102	

N7K1-LEAF1# sh mac address-table bridge-domain 3000

N7K1-LEAF1# sh l2route evpn mac-ip evi 3000

Mac Address	Prod Host IP	Next Hop (s)
30f7.0d9c.39bc	HMM	10.120.0.34
5897.1e9a.cc40	BGP	10.120.0.19
		N/A
		10.0.0.102

N7K1-LEAF1# sh arp vrf kk:a

Interface	Peer-IP	State	LearnType	Uptime	Router-Mac
nve1	10.0.0.102	Up	CP	04:02:03	0200.0a00.0066

N7K1-LEAF1# sh nve interface

Interface	nve1	State	encapsulation:	VXLAN
		Up		

Interface: nve1, State: Up, encapsulation: VXLAN

VPC Capability: VPC-VIP-Only [notified]

Local Router MAC: 0026.51c9.5043

Host Learning Mode: Control-Plane

Source-Interface: Loopback0 (primary: 10.0.0.53, secondary: 10.0.0.101)

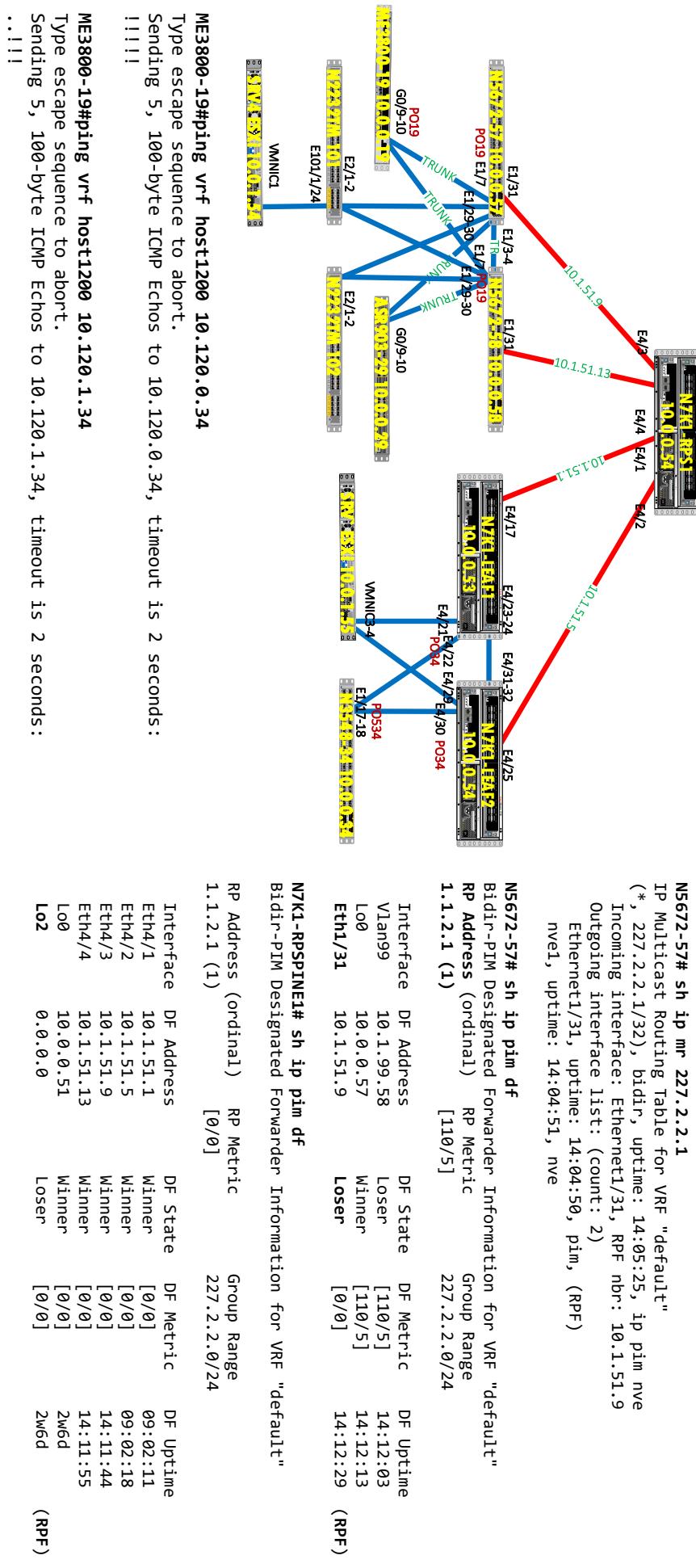
N7K1-LEAF1# sh nve vni

Interface	VNI	Multicast-group	State	Mode	Type [BD/VRF]	Flags
nve1	30000	227.2.2.1	Up	CP	L2 [3000]	SA
nve1	30001	227.2.2.2	Up	CP	L2 [3001]	SA
nve1	50000	n/a	Up	CP	L3 [kk:a]	

N7K1-LEAF1# sh bgp 12 evpn vni-id 30000

Network	Next Hop	Metric	LocPrf	Weight	Path
Route Distiguisher: 10.0.0.53:35767 (L2VNI 30000)					
*>1[2]:[0]:[0]:[48]:[30f7.0d9c.39bc]:[0]:[0.0.0.0]/216	10.0.0.101	100	32768 i	0	i
* i	10.0.0.101	100	32768 i	0	i
*>i[2]:[0]:[0]:[48]:[5897.1e9a.cc40]:[0]:[0.0.0.0]/216	10.0.0.102	100	0 i	0	i
* i	10.0.0.102	100	0 i	0	i
*>i[2]:[0]:[0]:[48]:[64f6.9de3.2520]:[0]:[0.0.0.0]/216	10.0.0.102	100	0 i	0	i
* i	10.0.0.102	100	0 i	0	i
*>1[2]:[0]:[0]:[48]:[30f7.0d9c.39bc]:[32]:[10.120.0.34]/272	10.0.0.101	100	32768 i	0	i
* i	10.0.0.101	100	0 i	0	i
*>i[2]:[0]:[0]:[48]:[5897.1e9a.cc40]:[32]:[10.120.0.19]/272	10.0.0.102	100	0 i	0	i
* i	10.0.0.102	100	0 i	0	i
*>i[2]:[0]:[0]:[48]:[64f6.9de3.2520]:[32]:[10.120.0.29]/272	10.0.0.102	100	0 i	0	i
* i	10.0.0.102	100	0 i	0	i

BIG Picture



member vni 30000

卷之三

mcast-group 22/.2.2.1

```
N5672-57# sh ip mrt 227.2.2.1
IP Multicast Routing Table for VRF "default"
(*, 227.2.2.1/32), bidir, uptime: 14:05:25, ip pim nve
Incoming interface: Ethernet1/31, RPF nbr: 10.1.51.9
Outgoing interface list: (count: 2)
    Ethernet1/31, uptime: 14:04:50, pim, (RPF)
nve1, uptime: 14:04:51, nve
```

```
N5672-57# sh ip pim df
Bidir-PIM Designated Forwarder Information for VRF "default"
RP Address (ordinal)          RP Metric      Group Range
1.1.2.1 (1)                   [11@5]           227.2.2.0/24
```

Interface	DF	Address	DF	State	DF	Metric	DF	Uptime
Vlan99	10.1.99.58		Loser	[110/5]			14:12:03	
Lo0	10.0.0.57		Winner	[110/5]			14:12:13	
Eth1/31	10.1.51.9		Loser	[0/0]			14:12:29	(RPF)

N7K1-RspINE1# sh ip pim df
Bidir-PIM Designated Forwarder Information for VRF "default"

KP Address (ordinal)	KP Metric	Group Range
1.1.2.1 (1)	[0/0]	227.2.2.0/24

ME3800-19#ping vrf host1200 10.120.0.34
Type escape sequence to abort.

Sending 5, 100-byte ICMP Echos to 10.120.0.34, timeout is 2 seconds:

- - - -

ME3800-19#ping vrf host1200 10.120.1.34
Type escape sequence to abort.

Cisco

VDP (VSI Discovery Protocol) Overview

- VDP is defined as part of IEEE standard 802.1Qbg (Edge Virtual Bridging)
- VDP Exchanges Virtual End Host information with the nearest VDP capable bridge via various TLVs
- VDP Uses Edge Control Protocol (ECP) as the transport protocol for its PDUs. For reliability and sequencing.
- VDP TLVs flow in both directions with VM Interface's network information
 - TLVs sent by Station (N1kv) to the Leaf are called VDP Requests
 - TLVs sent by Leaf to the Station are called VDP Responses.
- Whenever a VM is provisioned and powered on, station VDP sends a request with the network information to the Leaf
- The VDP on the Leaf will process this and provision any needs of the VM on the Leaf by HMM/PPM/ABDM and send a response back to Station

Network Information

- TLV is filled with Filter Formats indicating various parameters of the Network info.
- For Vinci first phase We support the segmentation based TLVs.

Number of entities (2 octets)	GroupID (4 octets)	MAC address (6 octets)	PS (1 bit)	PCP (3 bits)	VID (12 bits)

- MAC Address : Mac Address of the VM's interface Itself.
- PS,PCP, VID : 802.1q encapsulation information associated with the VM's interface
- Group ID : This is the segment ID that is associated with the VM.

Additional Network Information

- IP Address : N1kv can pass the IP Address of a VM in its VDP request.
In a separate Filter Format as follows.

Format	Value
VID/IP	0xF1
MAC/VID/IP	0xF2
GroupID/VID/IP	0xF3
GroupID/MAC/VID/IP	0xF4

IP field

AFI (2 bytes)	LEN (1 byte)	ADDR (N bytes)

- VM Name : N1KV can pass Name of the VM to the Leaf. This is passed in a proprietary Cisco OUI TLV.



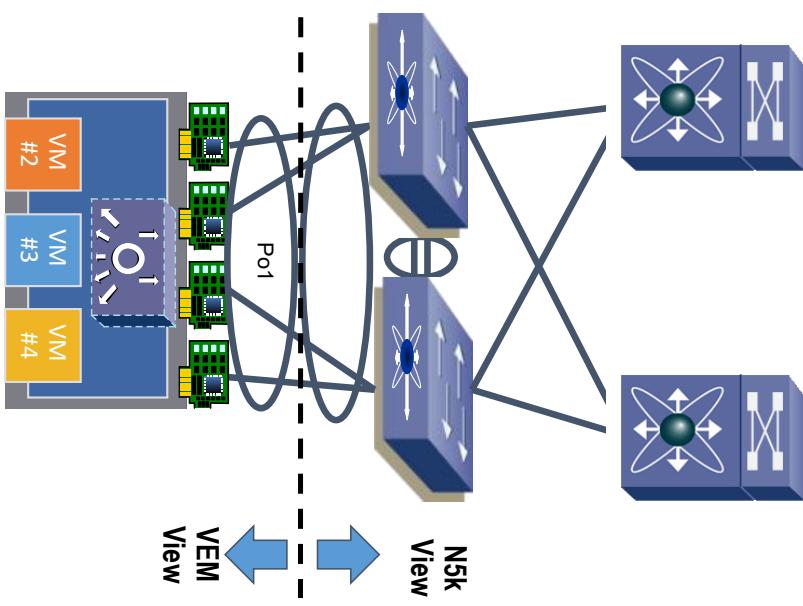
VDP Sequence



Nexus 1kv Directly Connected to Nexus 5000

- Recommended Design: LACP port-channel on the Nexus 1000V, LACP vPC configured on the Nexus 5000
- This is the recommended design for any “clustered” upstream switches (vPC, VSS, VBS)

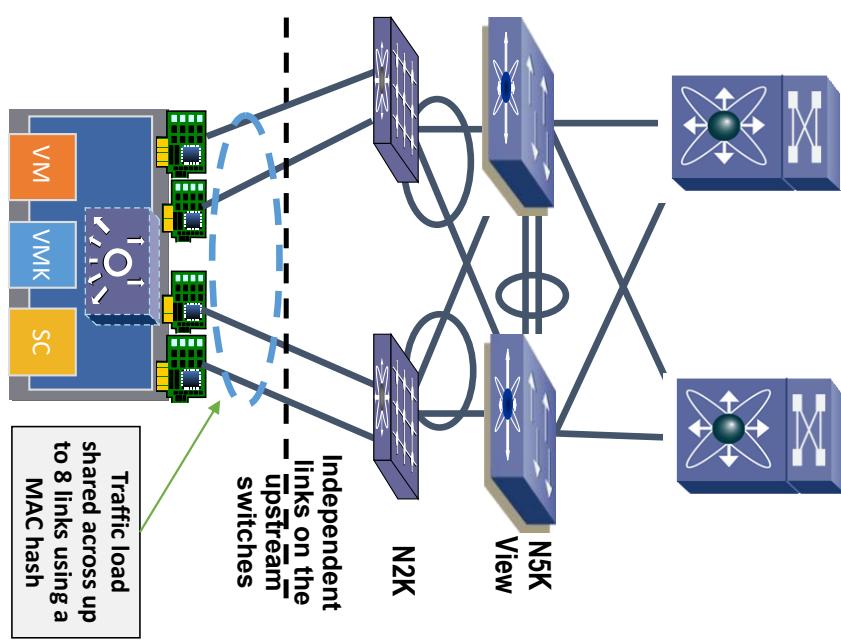
Supported



Connecting Nexus 1000V to Nexus 2k (2148)

- Recommended Design: vPC-HM on the Nexus 1kv
- No port-channel on host facing ports for N2148.
- If a failover occurs, all the traffic pinned to an interface will be migrated to the other interfaces.

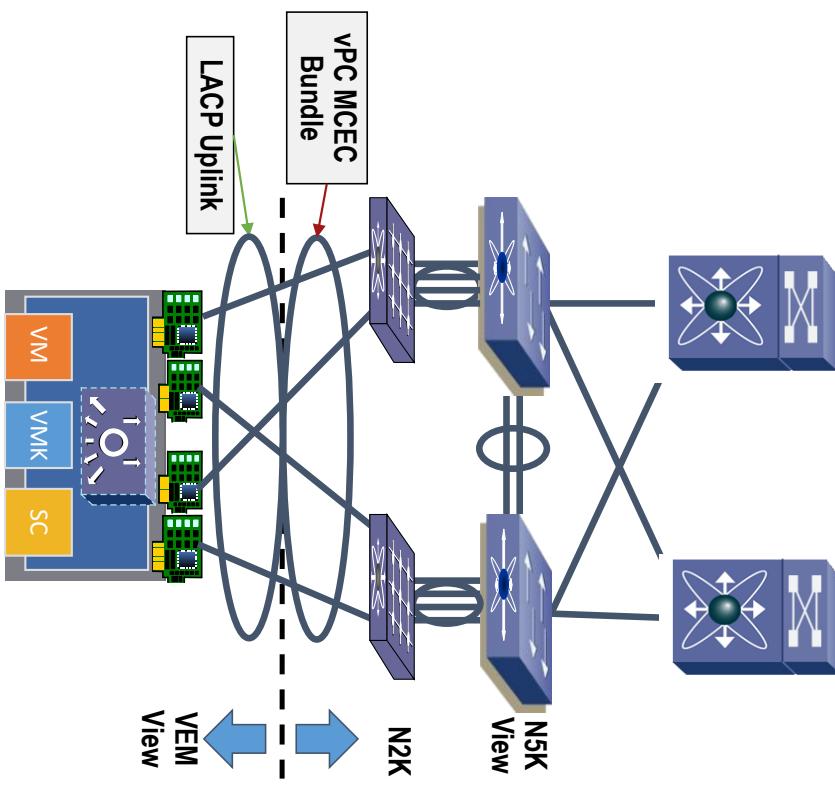
Supported



Connecting Nexus 1000V to Nexus 2248/2232

- vPC between a pair of N5K/N2K allows symmetrical Etherchannels
- N2248/N2232 allows host facing port-channel.

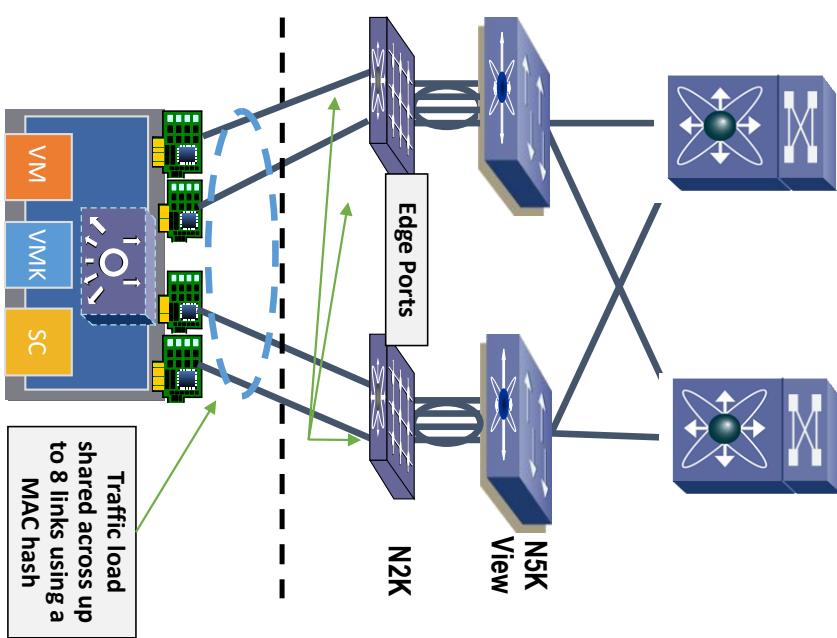
Supported



Connecting Nexus 1000V to Nexus 2k (2148)

- Recommended Design: vPC-HM on the Nexus 1kv
- No port-channel on host facing ports for N2148.
- No vPC on N5K
- If a failover occurs, all the traffic pinned to an interface will be migrated to the other interfaces.

Not Supported



DCNM Auto-Configuration – VDP Network

Add Network

* Organization:	kk
* Partition:	a
VRF Name:	kka
Network Name:	bd1203
Multicast Group Address:	227.2.2.8
* Network Role:	Host Network
Gateway IPv4 Address:	10.120.3.254
Netmask Length:	24
Gateway IPv6 Address:	
Prefix Length:	

Profile Parameters

Required 'vrflName' format is 'organizationName:partitionName'

* Profile Name:	defaultNetworkEvpnProfile
vlanId:	
segmentId:	30003
vrflName:	kka
gatewayIpAddress:	10.120.3.254
netmaskLength:	24
dhcpServerAddr1:	
vrfdhcp1:	

Network ID

Segment ID Only

Segment ID: 30003

Service Configuration Parameters

VM Manager IP:

Static IP Start:

Static IP End:

vSwitch Controller Network Id:

Distributed Virtual Switch Id:

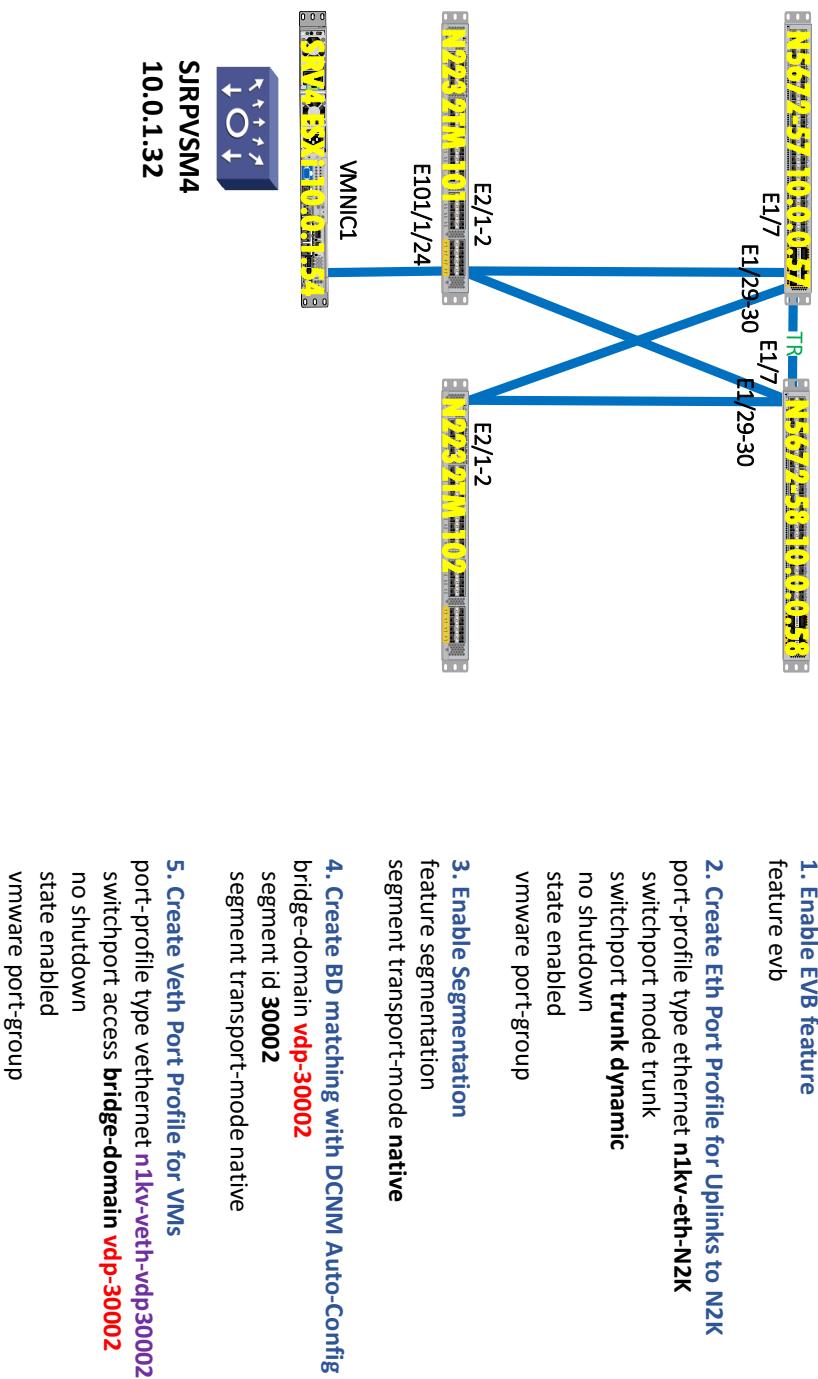
Secondary Gateway IPv4 Address:

Buttons

OK Cancel

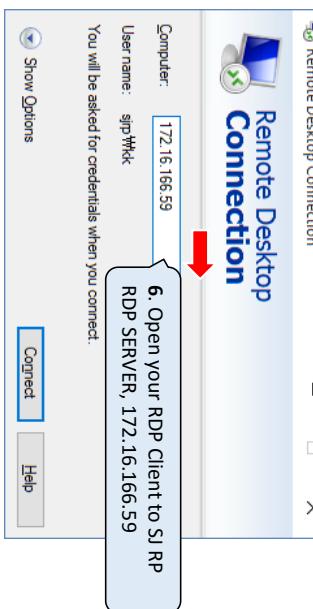
N1KV VDP CONFIG – VSM4 to N6K

Telnet to 10.0.1.32 from rp1 or rp2.cisco.com

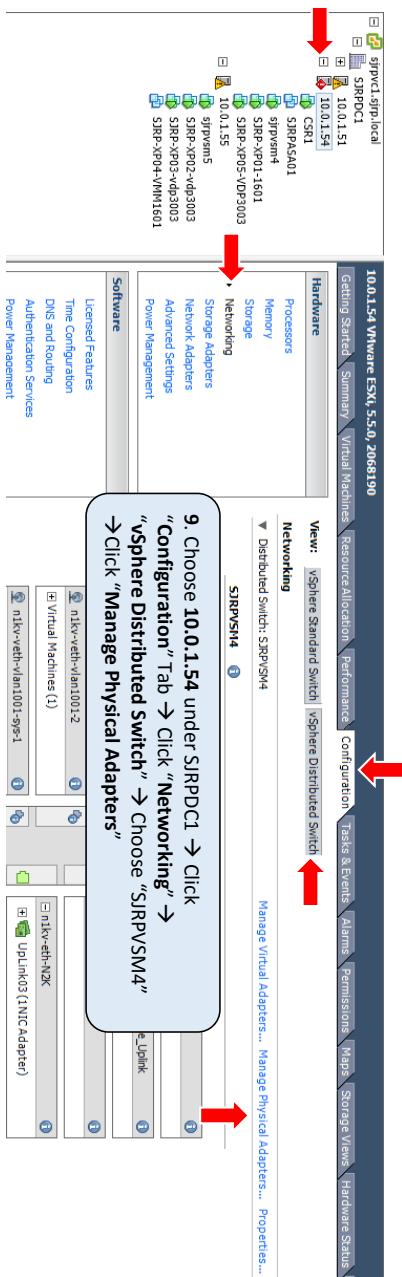
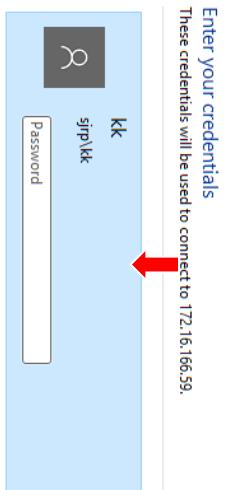


N1KV VDP Configuration – VMWARE

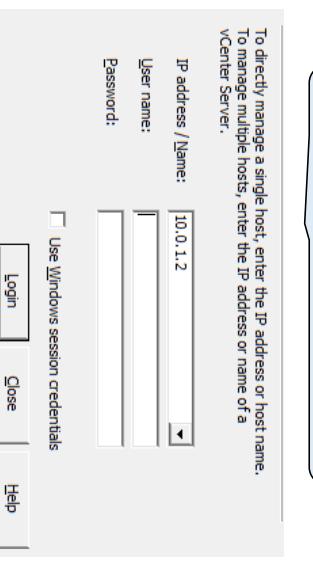
Let's add VMWARE SRPESX4's VMNIC1 into N1K ETH profile we create before



7. Use your SJ RP UID/PW.
Make sure you type **sjp\uid** as username
These credentials will be used to connect to 172.16.166.59.



9. Choose 10.0.1.54 under SRPD1 → Click “Configuration” Tab → Click “Networking” → “vSphere Distributed Switch” → Choose “SRPvSM4” → Click “Manage Physical Adapters”



8. Make sure you chose “n1kv-eth-N2K” then “Click to Add NIC” and select proper VMNIC.

This is only required when you have to add new VMNIC to newly created N1K ETH Profile
In our lab, No need to do this as all VMNIC is already assigned to N1K ETH Profile

N1KV VDP Configuration – VMWARE

Let's create a XP VM from Tempalte

11. Home → Inventory → "VMs and Templates"

12. Right-Click "SJRPP-XP-TEMP" and click "Deploy VM from this temp"

**13. Type VM name.
Please follow naming rule.**

14. Choose proper ESX host where NIK VDP is configured

**15. Choose SJRP-SAN as storage
This will allow VM Move.
SJRPP-SAN is iSCSI Storage**

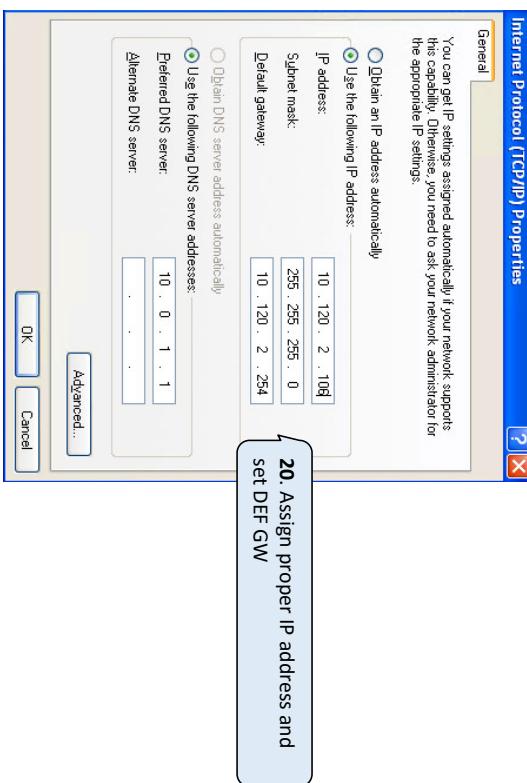
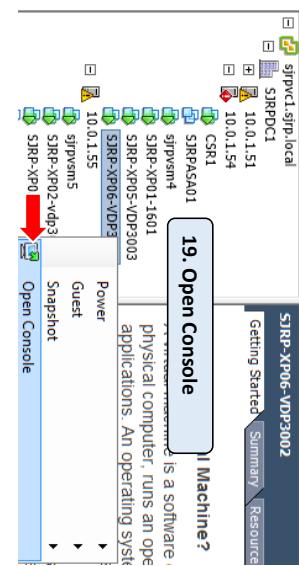
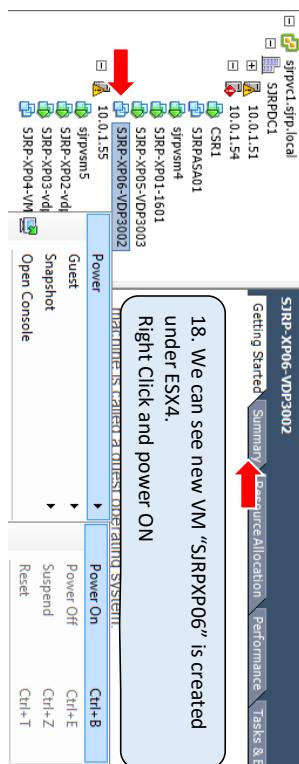
16. Select "Edit ..." in next screen

17. choose NIC and select proper NIK veth PORT you created

18. Creation of the virtual machine (VM) does not include automatic in system. Install a guest OS on the VM after creating the VM.

N1KV VDP Configuration – VMWARE

Start XP VM and change IP address, then PING Default GW



N1K VDP Verification

Check EVB is enabled

```
sJRPVSM4# show evb
Role : VDP Station
VDP Mac Address : 0180.C200.0000
VDP Resource Wait Delay : 20(17 secs)
VDP Reinit Keep Alive : 20(10 secs)
```

Check ECP status and counters

```
sJRPVSM4# show ecp
ECP Max Retries : 3
ECP Retransmission Timer Exp : 14(163840 micro seconds)
ECP Max Retries : 3
ECP Retransmission Timer Exp : 14(163840 micro seconds)
TX Sequence No : 1
Retry Count : 0
TX Count : 0
TX Count Errors : 0
In TX Queue : 0
RX Count : 0
RX Sequence : 0
```

Check SIRP-XP02 EVB VSI info

```
sJRPVSM4# sh evb vsi interface ve 6
LTL : 55 [module: 2]
Segment : 30003
MAC IP : 0050.56A9.6F9E
VSI State : 3
State Machine State : 7
Rwd Expiry Count : 115920
Last CMD Time : 8
Last RSP Time : 6
```

SW BD must be assigned at VEM level

Check VEM Module BD information

```
sJRPVSM4# module vem 3 execute vemcmd show segment 30003
BD 407, vdc 1, segment id 30003, segment group IP 0.0.0.0, encap NATIVE,
vff_mode Anycast, swid 4096, VLAN 3000, 2 ports, "vdp-30003",
Segment Mode: Unicast
VTEP DSN: 0 , MAC DSN: 0, BGP VTEP DSN: 0
Portlist:
19 vnic1
55 SIRP-XP05-VDP3003.eth0
```

Encap must be NATIVE
 Check if vlan is assigned
 If VDP failed, then no vlan assigned

Check if SIRP-XP belong to which Veth

```
SJRPVSM4# sh int status | i XP
Veth6      SIRP-XP05-VDP3003, connected 30003
Veth11     SIRP-XP06-VDP3002, connected 30002
```

Must both UPLINK and VM int are showed up here

N6K VDP Verification

N5672-58# sh ecn detail

ECP (Edge Control Protocol)

Retrans timer init	:	14
Max retries	:	3
Mode	:	Native (Point-to-Point)
No. rx packet	:	2866
No. tx packet	:	2866

1 Plugin(s):

ULP ID	Description	Status
1	VDP Plugin	Enabled

E1/7 E1/29-30 E1/7 E1/29-30

E2/1-2 E2/1-2

E1/101/1/24 VMNIC1

N5672-57# sh ecn detail

ECP (Edge Control Protocol)

Retrans timer init

:

14

Max retries

:

3

Mode

:

Native (Point-to-Point)

No. rx packet

:

2866

No. tx packet

:

2866

1 Session(s):

Interface	S-Vlan	Peer MAC	Sess ID	RxSeq	TxSeq
Eth101/1/24	1	-	1	31605	2458

[1] Interface: Ethernet101/1/24
Index: 0x1f6405c0

We can see LEAF58 received
ECP/EVB packets from N1kv

S-Vlan: 1

Sess ID: 1

Rx seq: 31605

Tx seq: 2458

Rx	615 packets	1 duplicate	0 drop
Tx	1228 packets	1022 retry	159 error

Seeing pkt errors

1 Plugin(s):
ULP ID Description Status

1 VPP Plugin Enabled

0 Session(s):
Interface S-Vlan Peer MAC Sess ID RxSeq TxSeq

Leaf57 shows nothing as ECP/EVB
packets are reached to LEAF58 by
l2 vpc hash

N6K VDP Verification

NON-WORKING EXAMPLE.		Interface	Profile ID	M-state	E-State	RWD/RKA
Mgr ID	VSI ID					
00000000 7533005056A94784	Po55			INIT	-	-
Role	: vdp bridge					
VDP MAC address	: 0180.c200.0000 (Nearest Bridge)					
No. of associated:	2					
E-state:	P - Pre-associated					
PR - Pre-associated with reservation						
A - Associated						
D - De-associated						
S - Standby						
Mgr ID	VSI ID	Interface	Profile ID	M-state	E-State	RWD/RKA
00000000 7532005056A95F2C	Eth101/1/24	00007532	WAIT_CMD	A	1332	
00000000 7533005056A95F9E	Eth101/1/24	00007533	WAIT_CMD	A	1332	
MGR:	00000000000000000000000000000000					
VST ID:	00000000000000000000000000000000					
Host name:	SJRP-XP03-vdp3003					
Interface:	0050.5653.894f					
S-Channel:	[02] VDP_BDG_INIT					
Station:						
Machine State:	-					
Entry State:	-					
Reason:	Could not apply profile >>>					
Recycle in:	330 sec					
VSI entry table						
No. of VSI entries:	2					
No. of associated:	2					
E-state:	P - Pre-associated					
PR - Pre-associated with reservation						
A - Associated						
D - De-associated						
S - Standby						
Mgr ID	VSI ID	Interface	Profile ID	M-state	E-State	RWD/RKA
00000000 7532005056A95F2C	Eth101/1/24	00007532	WAIT_CMD	A	1332	
00000000 7533005056A95F9E	Eth101/1/24	00007533	WAIT_CMD	A	1332	
MGR:	00000000000000000000000000000000					
VST ID:	00000000000000000000000000000000					
Host name:	SJRP-XP03-VDP3003					
Interface:	Ethernet101/1/24 (vPC 102423)					
S-Channel:	1					
Station:	0050.565d.cfb9					
Machine State:	[08] VDP_BDG_WAIT_STATION_CMD					
Entry State:	[03] VSI_ASSOCIATE					
Keep-Alive:	1332 sec					
Profile ID:	00007533					
Filter[01]:	group=30003 vlan=3000 mac=0050.56a9.6f9e ip=10.120.3.105					

N6K VDP Verification

N5672-58# show evb hosts detail

EVB Host table
No. of Hosts: 2
No. of VSIs: 2
Flags: + - Multiple addresses
> - Cisco OUI L3 address

Host Name	VNI	Vlan BD	Mac-Address	IP-Address	Interface
SJRP-XP05-VDP3003	30003	3000	0050.56a9.6f9e	10.120.3.105	Eth101/1/24
	30002	3001	Profile Active	L	defaultNetworkEvpnProfile(instance_vni_30002_4)
					FLAGs PROFILE(INSTANCE)
					VNI STATE
					Displaying VDP hosts
					Interface
					Eth101/1/24
					3001
					L
					Profile Active
					000000000000007532005056A95F2C
					VSI-ID

Host Name: SJRP-XP05-VDP3003

Host UUID: 3530032392033372038362031382039

VSI ID: 00000000000000000000007533005056A95F9E

Interface: Ethernet101/1/24 (vpc 102423)

Station: 0050.565d.cfb9

VNI: 30003

VLAN: 3000

BD: 3000

MAC: 0050.56a9.6f9e

IP: 10.120.3.105

Must LEAF and NIKV's assigned
vlan for BD are matched here

```
SJRPVSM4# module vcm 3 execute vcmcmd show segment 30003
BD 407, vdc 1, segment id 30003, segment group IP 0.0.0.0, encaps NATIVE,
vff_mode Anycast, swbd 4096, VLAN 3000, 2 ports, "vdp-30003"
Segment Mode: Unicast
VTEP DSN: 0 , MAC DSN: 0, BGP VTEP DSN: 0
Portlist:
19 vminic1
55 SJRP-XP05-VDP3003.eth0
```



VMTRACKER Overview

- VDP take cares of VMs behind N1Kv but what about VMs behind VMWARE DV(S)(distributed Virtual Switch)?
- Virtual Machine Tracker (VM Tracker) enables you to do the following:
 - Identify the Cisco Nexus device port that is used for each VM
 - Identify the VLAN requirements of each VM
 - Track the movement of VMs from one host (ESXi) to another
 - Track VM configuration changes such as additions, deletions, or modifications of VLANs, and configure VLANs on Cisco Nexus device ports accordingly
 - Track the additions or deletions of VMs and hosts, and configure VLANs on Cisco Nexus device ports accordingly
 - Track whether the VMs are in the power on or power off state, and configure VLANs on local Cisco Nexus device ports for a VM only if the VM is in the power on state
- VM Tracker uses VMware's vSphere SDK to communicate with vCenter.
 - Support up to 4 VCENTER connections
 - VM Tracker module establishes a secure connection (HTTPS) with vCenter. On successful connection, VM tracker first fetches all the host and VM related information and creates a local database of following items.
- Available on N3K/N5K/N6K/N7K/N9K without DFA

VM Tracker Configuration

```
feature vmtracker
```

```
vmtracker fabric auto-config
```

```
vmtracker connection 1
remote ip address 10.0.0.1.2 port 80 vrf management
username sjrp\kyukim password 5 ng
connect
```

Enable feature VMTRACKER

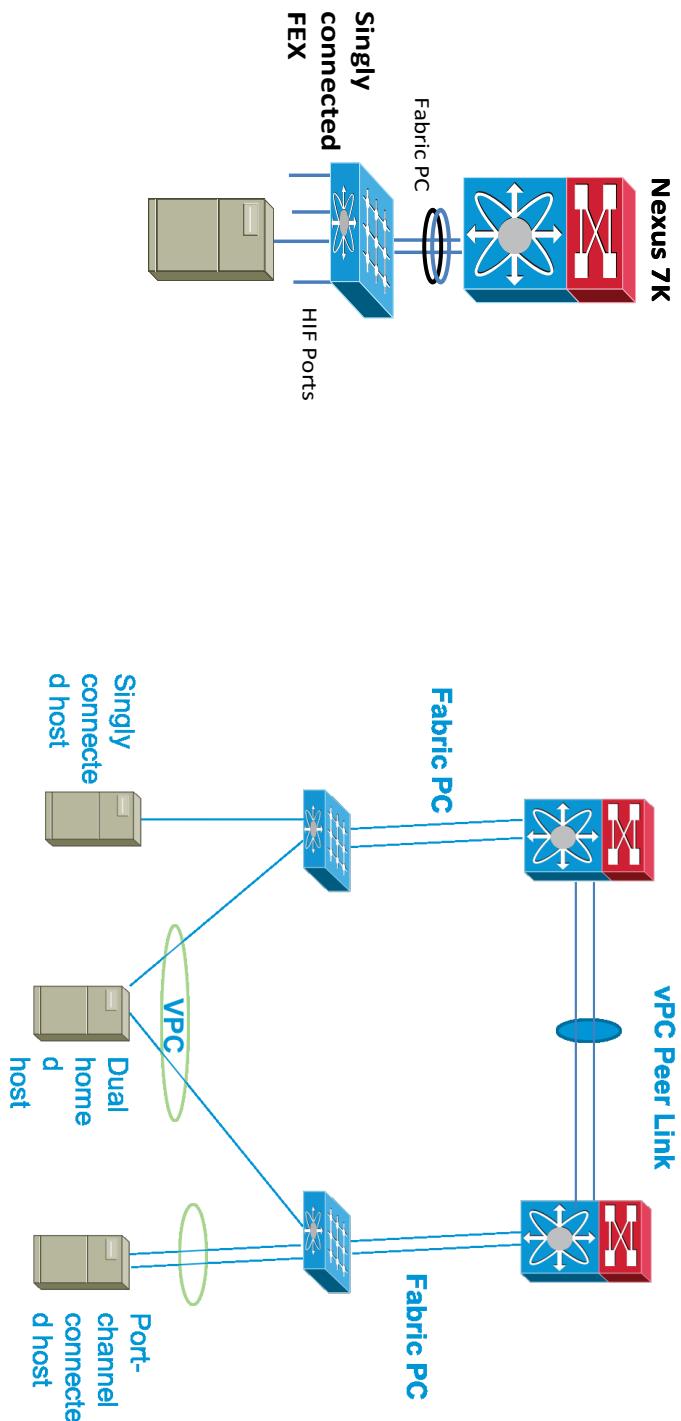
CLI is disruptive. User is recommended to disconnect all VMTracker connections before this configuration/ unconfiguration
When unconfigure this CLI, auto-configuration configured configuration on the switch will be removed and VMTracker fallback to classical VMTracker behavior (VLAN created by VMT)

Connect to VCENTER. Make sure provide proper vcenter ADMIN ID/PW

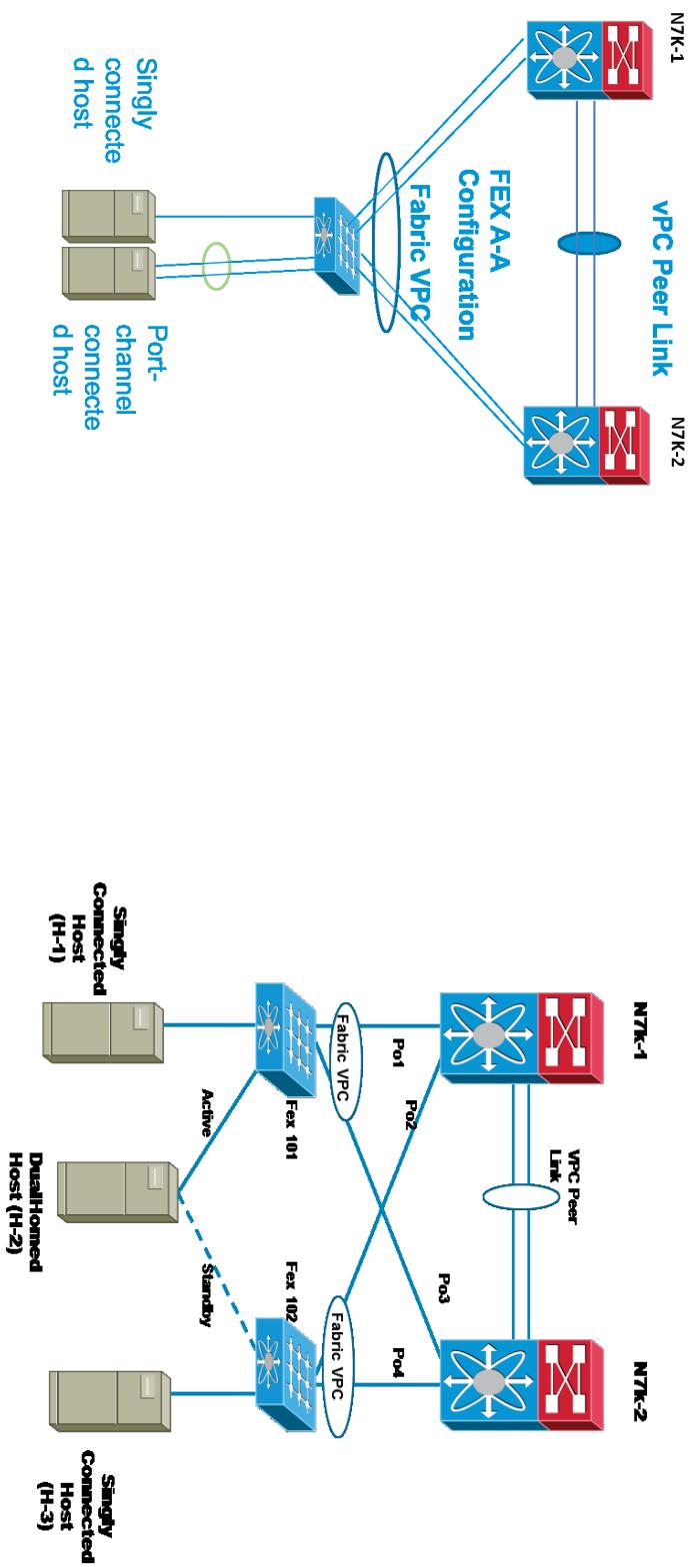
Same as other Trigger, Apply VMTRACKER on Interface level

```
interface port-channel15
vpc 55
service instance vni default
encapsulation dynamic vmtracker profile mdo
```

VM Tracker Supported Design



VM Tracker Unsupported Design



VMTRACKER LAB

Let's take look at SJRPESX5, 10.0.1.55

Device	Speed	Connection	Switch	MAC Address	Observed IP ranges	Wake on LAN Supported
Intel Corporation 82576 Gigabit Network Connection vmnic1	1000 Full	Negotiate	SJRPDVSSA sjrpvsm5	00:26:cb:c0:76:47 00:26:cb:c0:76:46	0.0.0.1-255.255.255.254 (V... None	Yes
Cisco Systems Inc Cisco VNIC Ethernet NIC vmnic2	10000 Full	Negotiate				No
Cisco Systems Inc Cisco VNIC Ethernet NIC vmnic3	10000 Full	Negotiate				No
Cisco Systems Inc Cisco VNIC Ethernet NIC vmnic5	10000 Full	Negotiate	SJRPDVSSB sjrpvsm5	a4:4c:11:13:89:4e a4:4c:11:13:89:4e	None	No

Network Adapters

Hardware

- Processors
- Memory
- Storage
- Networking
- Storage Adapters
- Network Adapters
- Advanced Settings

Getting Started | **Summary** | **Virtual Machines** | **Resource Allocation** | **Performance** | **Configuration** | **Tasks & Events** | **Alarms** | **Permissions** | **Maps** | **Storage Views** | **Hardware Status** | **Update Manager**

View: **vSphere Standard Switch** | **vSphere Distributed Switch**

Networking

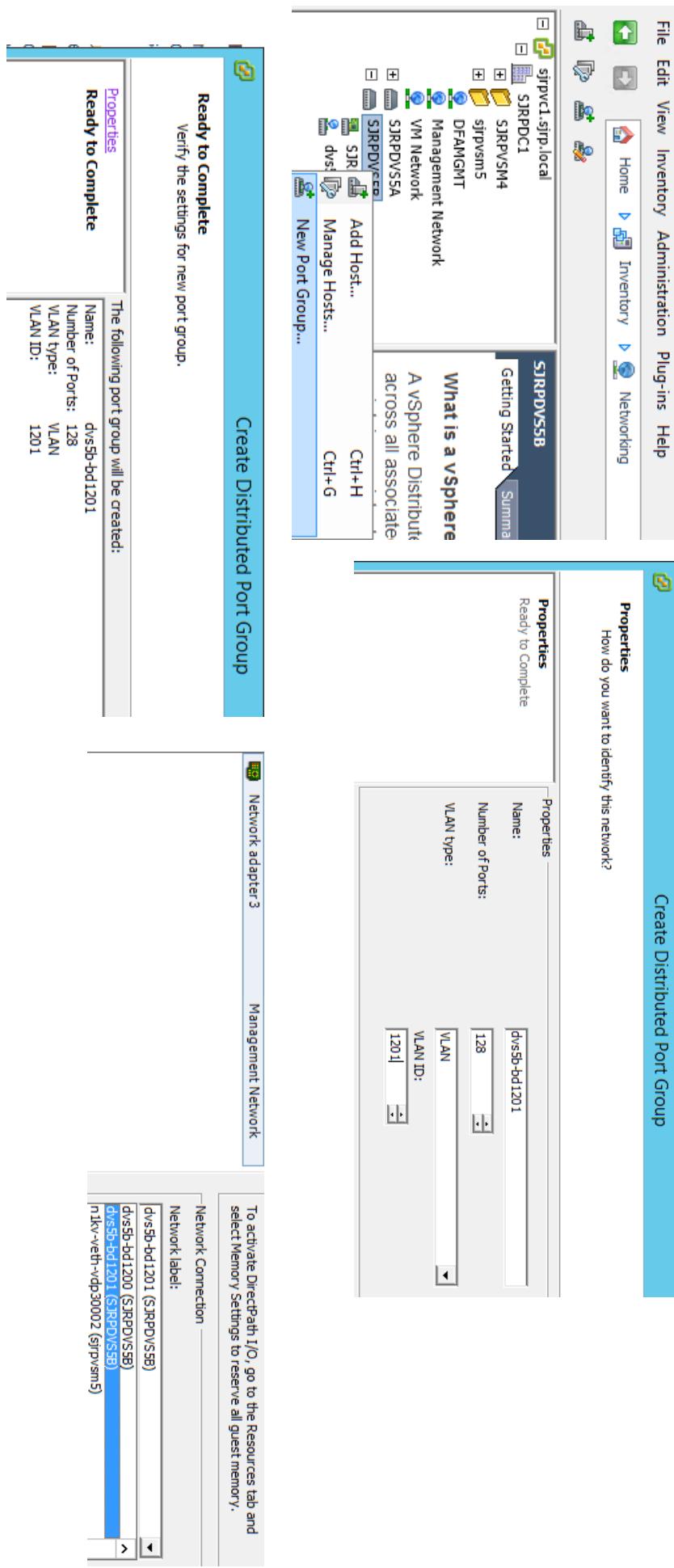
- Processors
- Memory
- Storage
- Networking
- Storage Adapters
- Network Adapters
- Advanced Settings
- Power Management
- Distributed Switch: sjrpvsm5
- Distributed Switch: SJRPDVSSA
- Manage Virtual Adapters... | Manage Physical Adapters... | Properties...
- Manage Virtual Adapters... | Manage Physical Adapters... | Properties...
- Manage Virtual Adapters... | Manage Physical Adapters... | Properties...
- Manage Virtual Adapters... | Manage Physical Adapters... | Properties...

Software

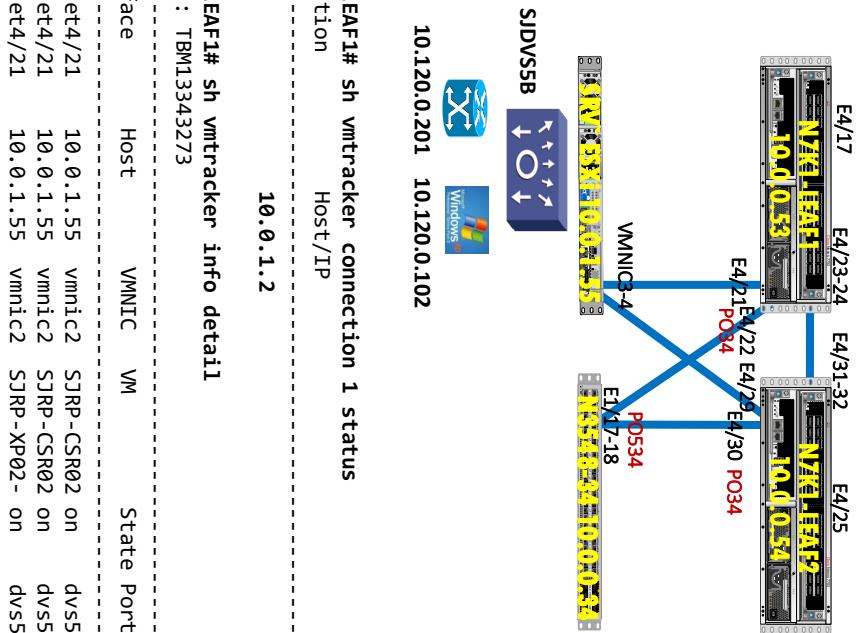
There are 4 VMNICs on ESX5 and 3 DVs (distributed virtual switch)

- **VMNIC0** -----3750-15 G1/0/29, **SJRPVSM5(N1KV)**, use for mgmt and OOB traffic
- **VMNIC1**-----N9372-42 E1/25, **SJRPDVSSA(VMWARE DVS)**, use for N9K vmtracker test
- **VMNIC2**-----N7K1-LEAF1 E4/21, **SJRPDVSSB(VMWARE DVS)**, use for N7K vmtracker
- **VMNIC3**-----N7K1-LEAF2 E4/29, **SJRPDVSS5B(VMWARE DVS)**,

VMTRACKER-Creating DVS Port Group



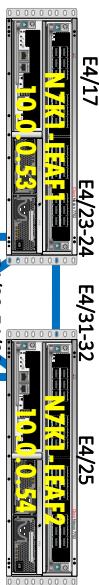
VMTRACKER Debug



```
N7K1-LEAF1# sh vmtracker event-history
-----
Event History (Conn:1 NumEv:4 IP:10.0.1.2)
-----
EventId Time Event Msg
-----
8652 Mar 16 2016 20:25:12:388912 event.DVPortgroupCreatedEvent.fullFormat
8649 Mar 16 2016 20:17:28:812848 event.VmReconfiguredEvent.fullFormat
8643 Mar 16 2016 20:15:17:698331 event.DVPortgroupCreatedEvent.fullFormat
8621 Mar 16 2016 18:32:38:813909 event.esx.clear.net.dvport.connectivity.y.restored.fullFormat (esx.clear.net.dvport.connectivity.y.restored)

N7K1-LEAF1# sh system internal vmtracker info ?
<CR>
>> Redirect it to a file
>> Redirect it to a file in append mode
all Print vmtracker all info
counters Print counter information
dvs-info Print DVS Info
event-history Print event history info
host-cdp Print host CDP info
host-dvs-portgroup Print host DVS portgroup info
host-dvs-switch Print host DVS Switch info
host-lldp Print host LLDP info
host-portgroup Print host portgroup info
host-unconnected Print host which are not connected
host-vn Print host Virtual Machine info
host-vswitch Print host vSwitch info
switch Print switch info
switch-device-id Print all switch device ids
time-info Print info related to duration of various events
vpc-sync-config Print vPC related connection info
| Pipe command output to filter
```

VMTRACKER Debug



```
N7K1-LEAF1# sh mac address-table bridge-domain 3000
VLAN/BD MAC Address Type age Secure NTFY Ports/SWID.SSID.LID
-----+-----+-----+-----+-----+
G 3000 0026.51c9.5042 static - F F sup-eth1(R)
G 3000 0025.51c9.5043 static - F F VPC Peer-Link(R)
* 3000 0050.56a9.5899 dynamic ~~~ F F VSI-Eth4/21.4095
E 3000 5897.1e9a.cc40 dynamic - F F nve1/10.0.0.102
E 3000 64f6.9de3.2520 dynamic - F F nve1/10.0.0.102
```

```
N7K1-LEAF1# sh service instance vni interface e4/21 det
```

VSI: VSI-Ethernet4/21.4095(default)
If-index: 0x35194fff

Admin Status: Up
Oper Status: Up

Auto-configuration Mode: VMTracker
encapsulation profile vni md0

dot1q vni BD

Dot1q VNI BD

1200	30000	3000
1201	30001	3001

10.120.0.201 10.120.0.102

```
N7K1-LEAF1# sh bridge-domain
```

Bridge-domain 3000 (3 ports in all)

Name:: Bridge-Domain3000
Administrative State: UP

vni30000
Operational State: UP

VSI-Eth4/21.4095

nve1

VSI-Po34.4095

```
Bridge-domain 3001 (3 ports in all)
```

Name:: Bridge-Domain3001
Administrative State: UP

vni30001
Operational State: UP

VSI-Eth4/21.4095

nve1

VSI-Po34.4095

POAP on VDC

POAP Switch Definitions

Selected 2 | Total 11



	Add	Edit	Delete	Publish	Write Erase and Reload	Change Image	Boot Log	Refresh Diff State			
	Serial Number	Switch ID	Management IP	Status	Switch Status	Publish Status	Bootscript Status	Diff State	Model	Template/Config File Name	Bootscript Last Updated
<input type="checkbox"/>	FOCL902R0E3	N5672-57	10.0.7.57	<input checked="" type="checkbox"/> ok	Published	Published	POAP script is finisher	Diff Detected	IPFabric N5600	Leaf v02	Thu Mar 10 04:38:46 GM
<input type="checkbox"/>	FOCL836R2Q6	N5672-58	10.0.7.58	<input checked="" type="checkbox"/> ok	Published	Published	POAP script is finisher	Diff Detected	IPFabric N5600	Leaf v02	Thu Mar 10 04:38:48 GM
<input type="checkbox"/>	TBM13343273	N7K1	10.0.7.50	<input checked="" type="checkbox"/> ok	Published	Published	POAP script is finisher	Diff Detected	N7K	Base N7K AdminVDC v5	Thu Feb 18 18:30:14 GM
<input type="checkbox"/>	SAL18495M25	N9372-42	10.0.7.42	<input checked="" type="checkbox"/> ok	Published	Published	POAP script is finisher	Diff Detected	N9K	IPFabric N9300 Leaf v02 kk	Thu Mar 3 10:49:46 GM
<input type="checkbox"/>	SAL19017CEU	N9372-43	10.0.7.43	<input checked="" type="checkbox"/> ok	Published	Published	Downloaded image:d	Diff Detected	N9K	IPFabric N9300 Leaf v02 kk	Thu Mar 3 10:49:31 GM
<input type="checkbox"/>	SAL18516Q80	N9372-44	10.0.7.44	<input checked="" type="checkbox"/> ok	Published	Published	POAP script is finisher	Diff Detected	N9K	IPFabric N9300 Leaf v02 kk	Thu Mar 3 03:04:38 GM
<input type="checkbox"/>	FOX18446P9Z	N9504-41	10.0.7.41	<input checked="" type="checkbox"/> ok	Published	Published	POAP script is finisher	Diff Detected	N9K	IPFabric N9500 Spine v02 k	Thu Mar 3 03:38:14 GM
<input checked="" type="checkbox"/>	TBM13343273:LEAF1	LEAF1	10.0.7.53	<input checked="" type="checkbox"/> ok	Published	Published	VOAP script is finisher	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 02:44:51 GM
<input type="checkbox"/>	TBM13343273:LEAF2	LEAF2	10.0.7.54	<input checked="" type="checkbox"/> ok	Published	Published	VOAP script is finisher	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 02:36:21 GM
<input type="checkbox"/>	TBM13343273:RPSPI	RPSPI	10.0.7.51	<input checked="" type="checkbox"/> ok	Published	Published	VOAP script is finisher	Diff Detected	VDC	IPFabric VDC Spine v02	Mon Feb 22 06:40:20 GM
<input type="checkbox"/>	TBM13343273:PE1	PE1	10.0.7.1	<input checked="" type="checkbox"/> ok	Published	Published	VOAP script is finisher	Diff Detected	VDC	N7K1-PE1-running-config	Tue Mar 15 04:49:32 GM
<input type="checkbox"/>	TBM13343273:LEAF1	LEAF1	10.0.7.53	<input checked="" type="checkbox"/> ok	Published	Published	Downloading startup config:downloading /var/lib/dcmny/	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 03:57:52
<input type="checkbox"/>	TBM13343273:LEAF2	LEAF2	10.0.7.54	<input checked="" type="checkbox"/> ok	Published	Published	Downloading startup config:downloading /var/lib/dcmny/	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 03:57:52
<input type="checkbox"/>	TBM13343273:LEAF1	LEAF1	10.0.7.53	<input checked="" type="checkbox"/> ok	Published	Published	Delete VDC:LEAF1	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 03:58:11
<input type="checkbox"/>	TBM13343273:LEAF2	LEAF2	10.0.7.54	<input checked="" type="checkbox"/> ok	Published	Published	Delete VDC:LEAF2	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 03:58:11
<input type="checkbox"/>	TBM13343273:LEAF1	LEAF1	10.0.7.53	<input checked="" type="checkbox"/> ok	Published	Published	Create VDC:LEAF1	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 04:00:52
<input type="checkbox"/>	TBM13343273:LEAF2	LEAF2	10.0.7.54	<input checked="" type="checkbox"/> ok	Published	Published	Create VDC:LEAF2	Diff Detected	VDC	IPFabric VDC Leaf v02 bufl	Wed Mar 16 04:00:52

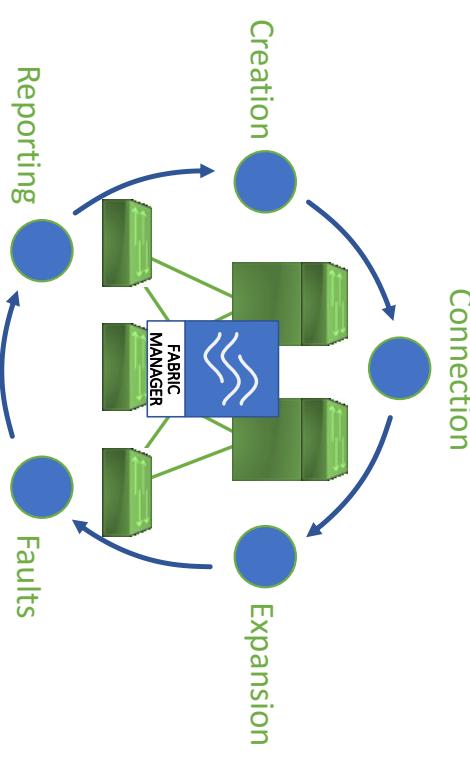
Nexus Fabric Manager

Cisco Nexus Fabric Manager

Intelligent Fabric Lifecycle Management

- *Fabric-wide focus* – auto-configuration and management of fabric running stand-alone NX-OS mode
- Automation based on knowledge of underlying fabric architecture
- Designed to simplify fabric management through its various lifecycle phases
- *Delivered via VXLAN-based architecture*

Fabric Management Lifecycle

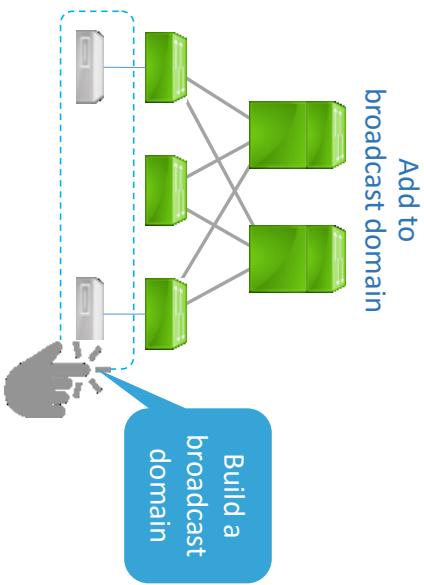


Fabric Management Workflows

- Optimized for *fabric management workflows*
 - Help network ops quickly support business needs
- Switch features managed based on workflows

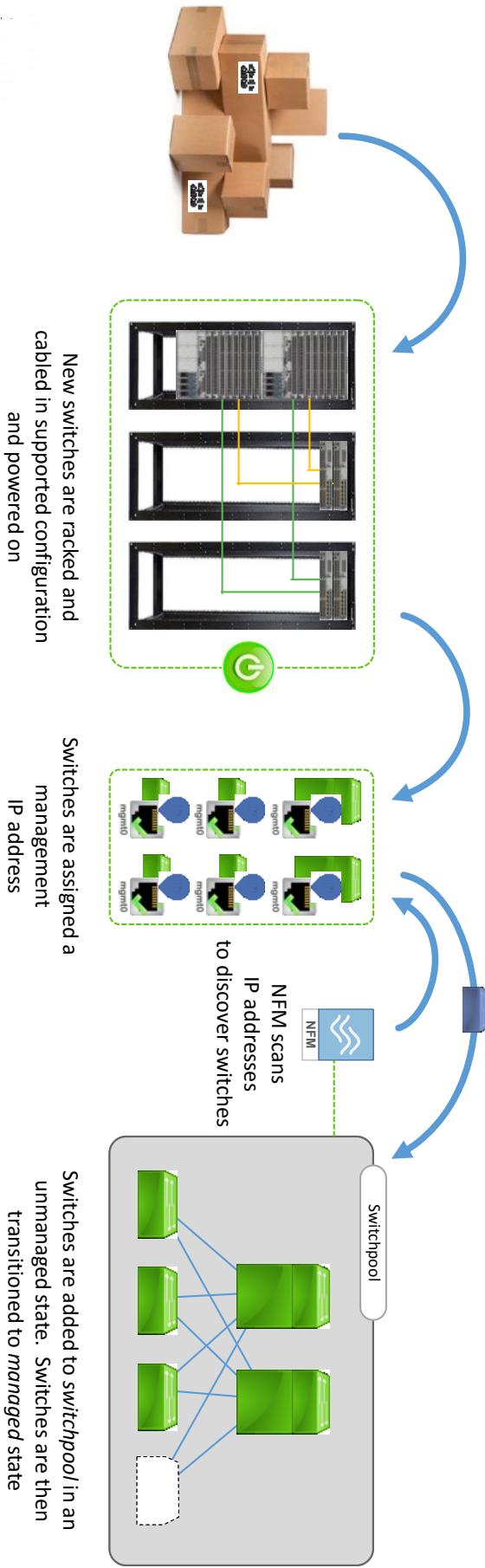
Sample fabric management workflows

1. **Create a fabric**
 - NFM creates and manages HA-enabled fabric
 - NFM configures Fabric Underlay automatically
2. **Add a new switch to the fabric**
 - NFM discovers, adds, and configures new switch
3. **Create a broadcast domain**
 - NFM creates and manages VLANs and VXLAN topology



Building a New Fabric

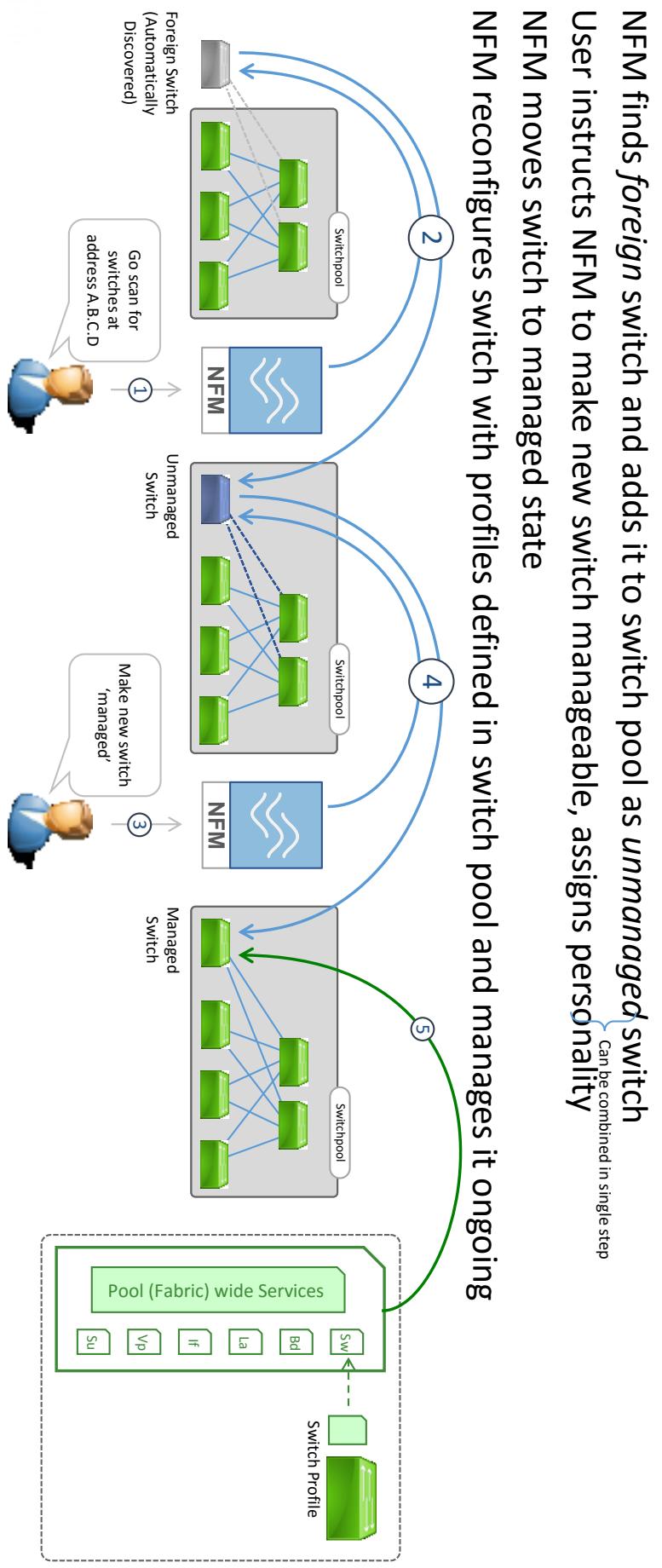
- NFM manages switches as part of a supported topology
- Switches are discovered by mgmt0 addresses or as neighbor via LLDP and optionally added to switchpool
 - Use POAP server and configure new switches with mgmt ip, mgmt gw, admin id/pw
 - Each switch is given a *personality* within fabric (eg. leaf, spine)
 - NFM configures all Fabric underlay configuration automatically
 - User configures Host connectivity manually from GUI or REST API



Adding a New Leaf Switch

User wishes to add new leaf switch to managed switch pool

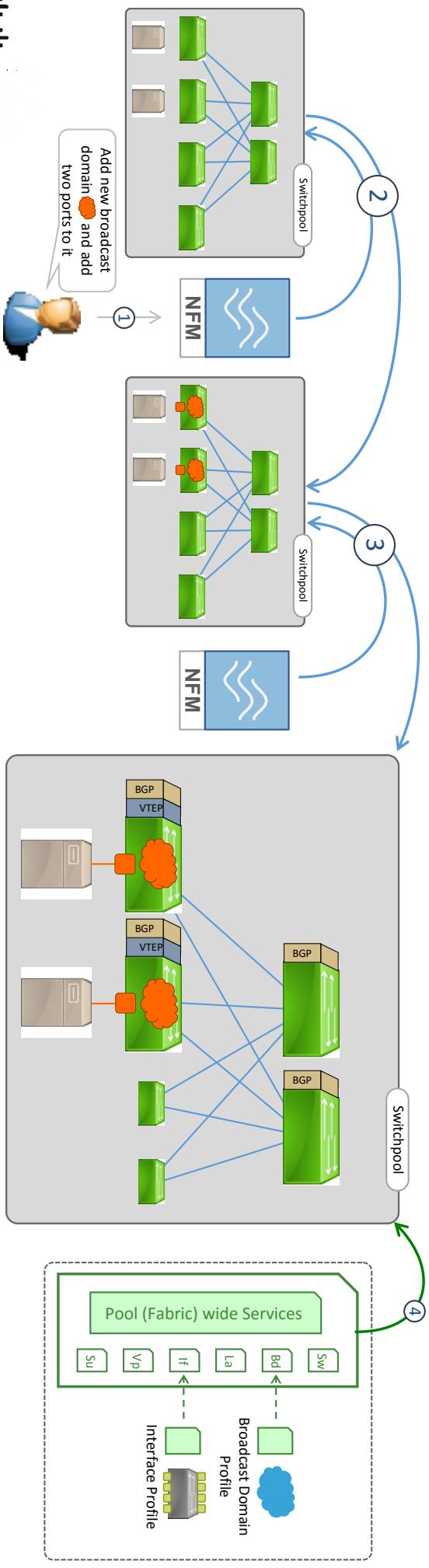
1. User instructs NFM to scan address/range/subnet for switch(es) to add to switch pool – or selects discovered neighbor
2. NFM finds *foreign* switch and adds it to switch pool as *unmanaged switch*
3. User instructs NFM to make new switch manageable, assigns **personality**
Can be combined in single step
4. NFM moves switch to managed state
5. NFM reconfigures switch with profiles defined in switch pool and manages it ongoing



Creating a Broadcast Domain

User wishes to create new broadcast domain and attach two devices

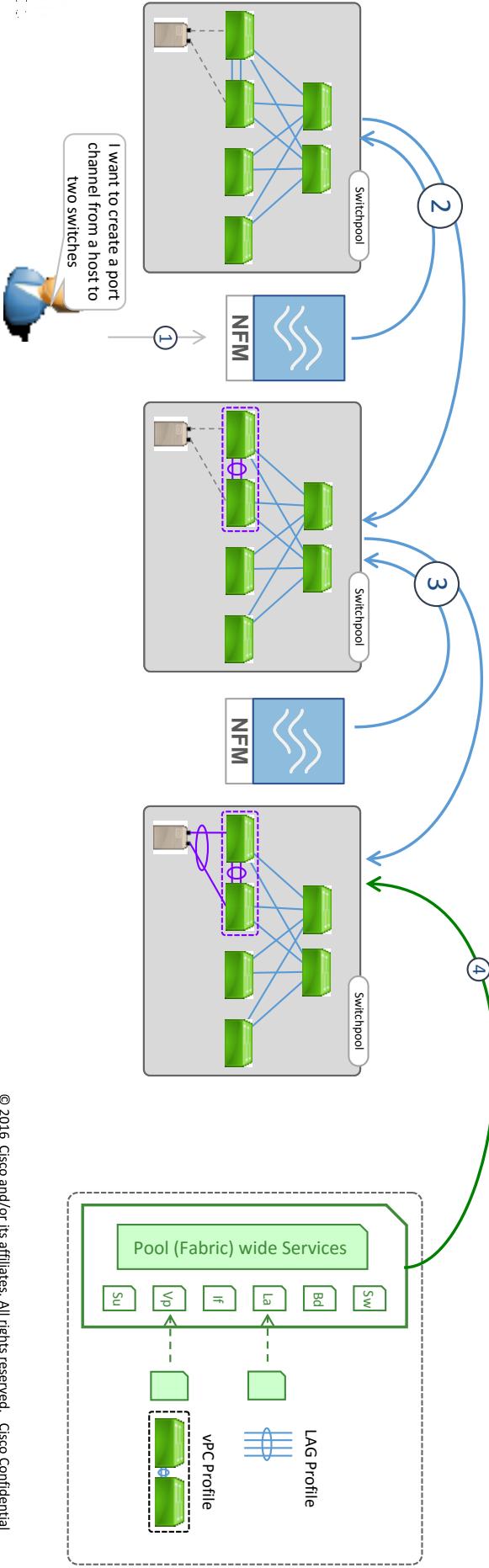
1. User configures BD, GW, VRF from GUI or use REST API to push
2. User instructs NFM to add new broadcast domain for two servers on two switches to communicate
3. NFM finds two switches and creates broadcast domain (VLAN) and puts ports in new VLAN
4. NFM builds two VTEPs, assigns new VNID, configures MP-BGP/EVPN on leafs and spines
5. NFM configures broadcast domain and interfaces with profiles defined in switch pool and manages them ongoing



Adding a Dual-Connected Host

User wishes to add a dual-connected server to leaf nodes

1. User instructs NFM to connect a dual-attached host via a port channel to two different switches
2. NFM recognizes that links exist between two desired switches and a vPC can be built
3. NFM builds full vPC configuration including peer link port channels. NFM assigns all required addressing and IDs
4. NFM builds two separate port channels to the host and ties together all the configuration for vPC



Upgrading the Fabric

- Switches are added into 'upgrade groups'
 - Image and upgrade policy are applied to groups
 - Images are stored in Nexus Fabric Manager
 - Edit groups to change image to upgrade to next release

The screenshot shows the 'Images' page in the Cisco Nexus Fabric Manager interface. The top navigation bar includes the Cisco logo, the title 'Cisco Nexus Fabric Manager', and user information ('admin'). The left sidebar has links for 'SWITCHPOOL', 'Home Profiles', 'Racks', 'VRFs', 'Faults', 'MANAGEMENT', 'Upgrades', and 'Images'. The main content area has filters for 'Filter by status' (dropdown), 'Filter expression' (text input), 'Sort by' (dropdown set to 'Name'), and a 'Details' button. A table lists two images:

Name	Description	Version	Release date	Status
7.0(3)I2(2)(nxos.7.0.3.I2.2.bin)	-	7.0(3)I2(2)	97 days ago (Sun, Nov 8, 11:29PM)	Active
7.0(3)I2(2a)(nxos.7.0.3.I2.2a.bin)	-	7.0(3)I2(2a)	61 days ago (Mon, Dec 14, 12:23PM)	Active

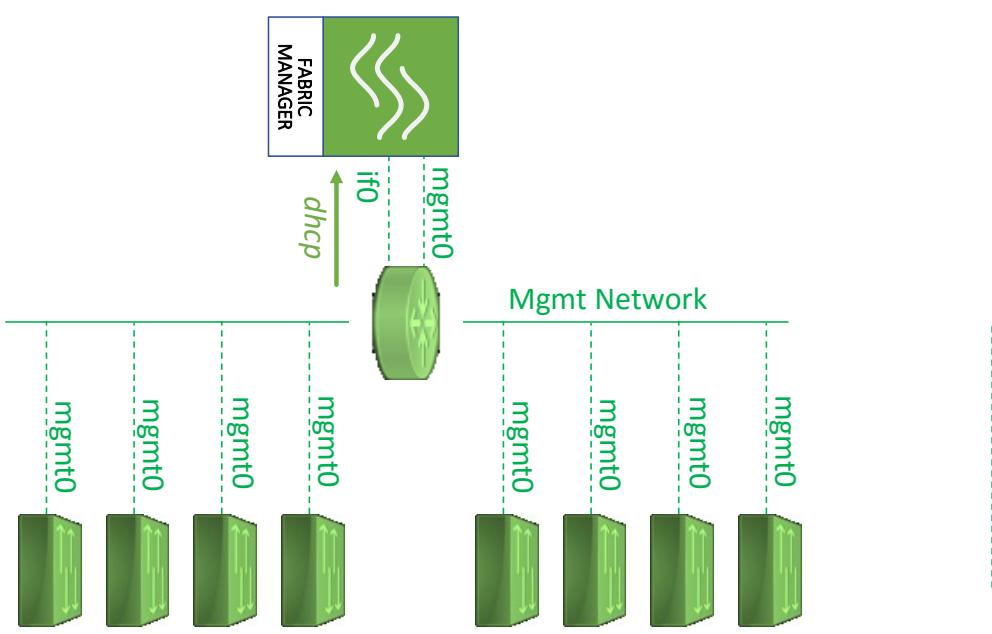
At the bottom, there are navigation icons for first, previous, next, last, and a page size selector ('1 / 1'). A footer note says '1 - 2 of 2 items'.

ADMINISTRATION
Users
System
Account Settings

cisco

Virtual Appliance Requirements

- VMware ESXi v5.5 (no current KVM support)
- 32 GB of RAM
- 8 vCPUs
- 100 GB disk
- Two vNICs
 - Mgmt0 – OOB access to NFM
 - if0 – Auto Fabric Provision, same as POAP, connects to Leaf/Spine's mgmt0 network



NFM OVA Deployment

Deploy OVF Template...

Source
Select the source location.

Source	
OVF Template Details	Name and Location
<input checked="" type="checkbox"/> Host / Cluster	Host / Cluster
<input checked="" type="checkbox"/> Resource Pool	Resource Pool
<input checked="" type="checkbox"/> Disk Format	Disk Format
<input checked="" type="checkbox"/> Ready to Complete	Ready to Complete
Deploy from a file or URL	
<input type="text" value="E:\DOWNLOADED-1.1.258.ovf"/>	
<input type="button" value="Browse..."/>	
Enter a URL to download and install the OVF package from the Internet, or specify a location accessible from your computer, such as a local hard drive, a network share, or a CD/DVD drive.	

Host / Cluster
On which host or cluster do you want to run the deployed template?

Host / Cluster	
OVF Template Details	Name and Location
<input checked="" type="checkbox"/> SRRPDC1	SRRPDC1
<input checked="" type="checkbox"/> 10.0.1.51	10.0.1.51
<input checked="" type="checkbox"/> 10.0.1.54	10.0.1.54
<input checked="" type="checkbox"/> 10.0.1.55	10.0.1.55
<input checked="" type="checkbox"/> Ready to Complete	Ready to Complete

Storage
Where do you want to store the virtual machine files?

Storage	
OVF Template Details	Select a destination storage for the virtual machine files:
Name and Location	<input type="button" value="VM Storage Profile: ..."/>
Host / Cluster	<input type="button" value="A"/>
Storage	<input type="button" value="VMFS5"/>
<input checked="" type="checkbox"/> datastore1	Name: datastore1
<input checked="" type="checkbox"/> Non-SSD	Drive Type: Non-SSD
<input checked="" type="checkbox"/> 1.0TB	Capacity: 1.0TB
<input checked="" type="checkbox"/> 835.53GB	Provisioned: 835.53GB
<input checked="" type="checkbox"/> 741.07GB	Free: 741.07GB
<input checked="" type="checkbox"/> VMFS5	Type: VMFS5
<input checked="" type="checkbox"/> Support	Thin Provisioned: Support
<input checked="" type="checkbox"/> srrp-sani	Name: srrp-sani
<input checked="" type="checkbox"/> Non-SSD	Drive Type: Non-SSD
<input checked="" type="checkbox"/> 506.75GB	Capacity: 506.75GB
<input checked="" type="checkbox"/> 63.85GB	Provisioned: 63.85GB
<input checked="" type="checkbox"/> 451.80GB	Free: 451.80GB
<input checked="" type="checkbox"/> VMFS5	Type: VMFS5
<input checked="" type="checkbox"/> Support	Thin Provisioned: Support
<input checked="" type="checkbox"/> Ready to Complete	Ready to Complete

OVF Template Details
Verify OVF template details.

NFM OVA Deployment

Disk Format

In which format do you want to store the virtual disks?

Source	Datastore:	datastore1.1
OVF Template Details	Available space (GB):	741.1
Name and Location	Host / Cluster:	Storage
Host / Cluster	Disk Format:	Network Mapping
Storage		Ready to Complete
	<input checked="" type="radio"/> Thick Provision Lazy Zeroed	
	<input type="radio"/> Thick Provision Eager Zeroed	
	<input type="radio"/> Thin Provision	

Network Mapping

What networks should the deployed template use?

Source	Map the networks used in this OVF template to networks in your inventory
OVF Template Details	
Name and Location	
Host / Cluster	
Storage	
Disk Format	
Network Mapping	
Ready to Complete	

Source Networks	Destination Networks
bridged	VM Network
SF-AMGT	VM Network

Ready to Complete

Are these the options you want to use?

Source	When you click Finish, the deployment task will be started.
OVF Template Details	Deployment settings:
Name and Location	OVF file: C:\DOWN\Nfm.1.0.1.ova
Host / Cluster	Download size: 292.6 MB
Storage	Size on disk: 100.0 GB
Disk Format	Name: SRRPNFM1
Network Mapping	Folder: SRRPDC1
Ready to Complete	Host/Cluster: 10.0.151
	Datastore: datastore1
	Disk provisioning: Thick Provision Lazy Zeroed
	Network Mapping: "bridged" to "VM Network"

Power on after deployment

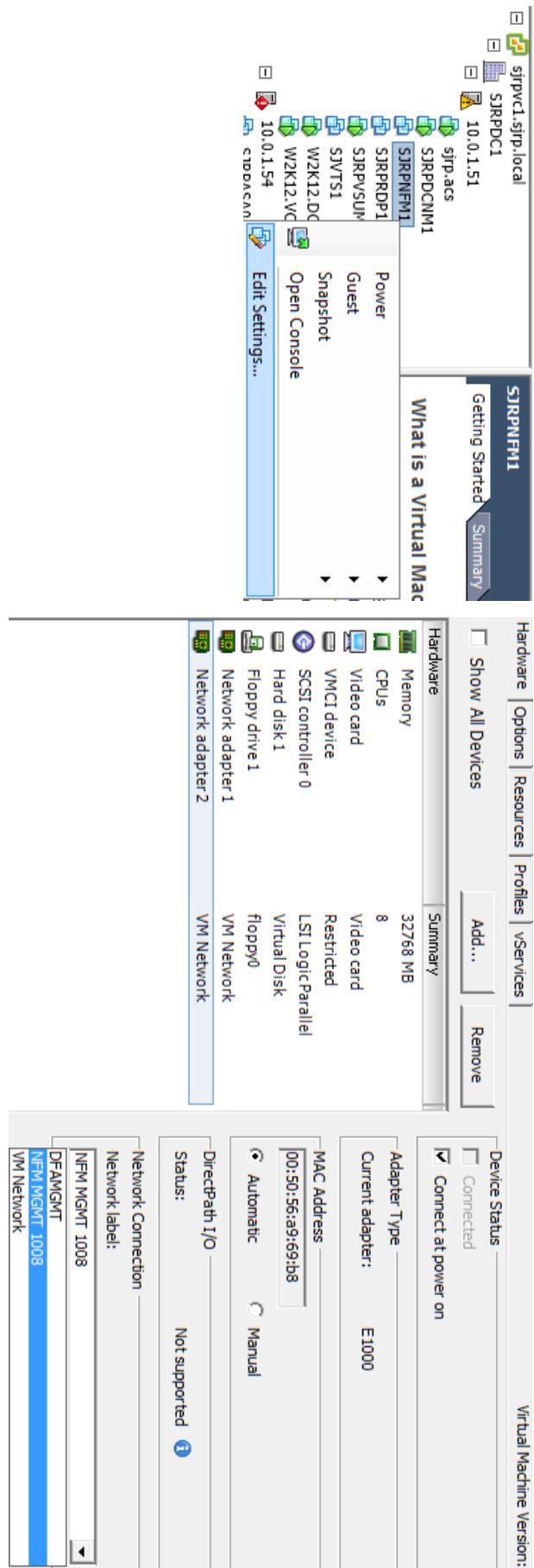
Help

≤ Back

Finish

NFM OVA Deployment for AFP

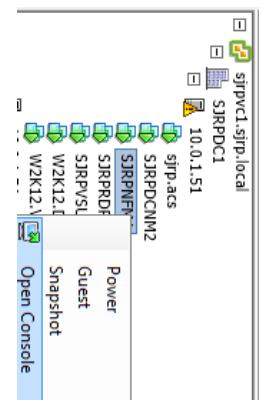
If you want to use AFP, then must edit VM network adapter setting



Change 2nd NIC to NFM MGMT 1008 which is connected to SRRP BB vlan 1008 for N9K mgmt0 interfaces

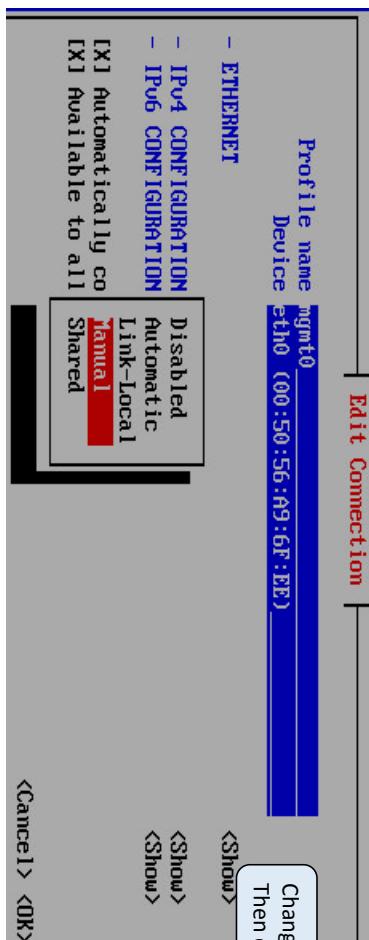
NFM Initial Setup

Need to configure NFM Network after NFM boots

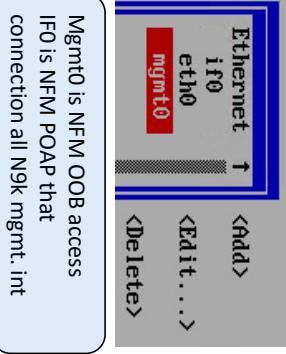


Change pw for root as cisco123
Change pw for admin as admin123 → NFM GUI Access

Choose "edit connection"
<Cancel> <OK>

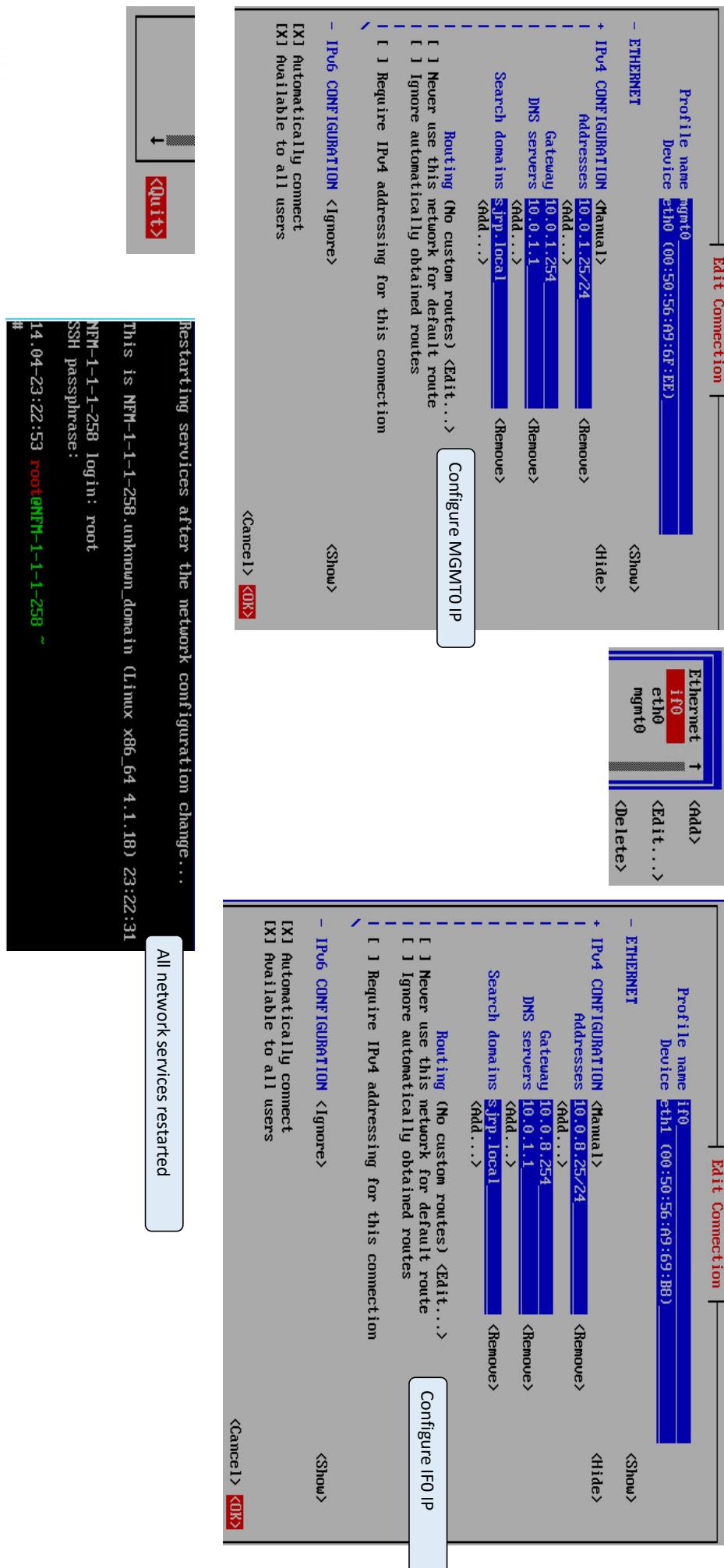


Change mgmt0 int from disabled to Manual
Then click Show



Mgmt0 is NFM OOB access
If0 is NFM POAP that
connection all N9k mgmt. int

NFM Initial Setup



cisco

NFM Initial Check-up

```
Ssh to 172.16.166.56 as root/cisco123  
14.04-23-24:27 root@NFM-1-1-1-258 ~  
#
```

```
* service esm status  
* Wrapper process is running with PID=13592  
* Elastic Services Manager process is running with PID=13614
```

ESM is Apache service + NFM processes
You can restart ESM with "service esm restart"

```
# netstat -l  
Active Internet connections (only servers)  
Proto Recv-Q Send-Q Local Address          Foreign Address        State  
tcp     0      0 localhost:4200           0.0.0.0:*
```

Proto	Recv-Q	Send-Q	Local Address	Foreign Address	State
tcp	0	0	localhost:4200	0.0.0.0:*	LISTEN
tcp	0	0	0.0.0.0:sunrpc	0.0.0.0:*	LISTEN
tcp	0	0	0.0.0.0:38352	0.0.0.0:*	LISTEN
tcp	0	0	0.0.0.0:ftp	0.0.0.0:*	LISTEN
tcp	0	0	0.0.0.0:ssh	0.0.0.0:*	LISTEN
tcp	0	0	localhost:postgresql	0.0.0.0:*	LISTEN
tcp6	0	0	localhost:8005	[::]:*	LISTEN
tcp6	0	0	[::]:sunrpc	[::]:*	LISTEN
tcp6	0	0	localhost:http-alt	[::]:*	LISTEN
tcp6	0	0	localhost:http-alt	[::]:*	LISTEN
tcp6	0	0	[::]:http	[::]:*	LISTEN
tcp6	0	0	[::]:8083	[::]:*	LISTEN
tcp6	0	0	[::]:ssh	[::]:*	LISTEN
tcp6	0	0	[::]:8086	[::]:*	LISTEN
tcp6	0	0	localhost:postgresql	[::]:*	LISTEN
tcp6	0	0	[::]:8088	[::]:*	LISTEN
tcp6	0	0	[::]:15003	[::]:*	LISTEN
tcp6	0	0	[::]:https	[::]:*	LISTEN
tcp6	0	0	[::]:45918	[::]:*	LISTEN

```
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500  
      inet 10.0.1.25 netmask 255.255.255.0 broadcast 10.0.1.255  
            inet6 fe80::250:56ff:fe9:6fee prefixlen 64 scopeid 0x20<link>  
              ether 00:50:56:a9:6f:ee txqueuelen 1000 (Ethernet)  
                RX packets 2030 bytes 149955 (146.4 kB)  
                TX packets 2986 bytes 3627701 (3.4 MiB)  
                  TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0  
  
eth1: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500  
      inet 10.0.8.25 netmask 255.255.255.0 broadcast 10.0.8.255  
            inet6 fe80::250:56ff:fe9:69b8 prefixlen 64 scopeid 0x20<link>  
              ether 00:50:56:a9:6e:b8 txqueuelen 1000 (Ethernet)  
                RX packets 0 bytes 0 (0.0 B)  
                TX packets 0 dropped 0 overruns 0 frame 0  
                  TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0  
  
lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536  
      inet 127.0.0.1 netmask 255.0.0.0  
            inet6 ::1 prefixlen 128 scopeid 0x10<host>  
              loop txqueuelen 0 (Local Loopback)  
                RX packets 8420 bytes 1747568 (1.6 MiB)  
                TX packets 8420 bytes 1747568 (1.6 MiB)  
                  TX errors 0 dropped 0 overruns 0 frame 0  
                    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
```

NFM SJ RP Static LAB



SJ RP Static LAB is open to all Cisco Engineers
Telnet to rp1 or rp2.cisco.com
Then, telnet 10.0.8.4X for each leaf or spine access

SPINE 41
Leaf 42
vPC Leaf 43/44

NFM GUI

Log on to 172.16.166.56 with admin/admin123

The screenshot shows two browser windows side-by-side. Both windows are titled 'Cisco Nexus Fabric Manager' and have the URL 'https://172.16.166.56/login'. The left window displays the 'Log in' screen with fields for 'Username' (containing 'kyuhwan') and 'Password' (containing 'cisco'). A 'Log in' button is visible at the bottom. The right window shows the 'Switchpool' interface, which is currently empty. The top navigation bar includes links for 'OVERVIEW', 'INTERFACES', 'SWITCHES', 'NEIGHBORS', 'BROADCAST DOMAINS', 'HISTORY', 'FAULTS', and 'SETTINGS'. Below the navigation bar, there are buttons for 'Physical view', 'Info', and 'Not in'. A message box in the center of the page says 'First time you will see empty SWITCH POOL'.

NFM SW Pool Setting

Set important SW pool settings

The screenshot shows the 'Switchpool' settings page. The top navigation bar includes links for OVERVIEW, INTERFACES, SWITCHES, NEIGHBORS, BROADCAST DOMAINS, HISTORY, FAULTS, and SETTINGS. The 'GENERAL' tab is selected. A red arrow points to the 'Last modified' field, which shows '23 hours ago (Yesterday @ 3:25PM) by admin'. Below this, the 'Available count' is listed as 101. The bottom section shows two tables: 'Unallocated VLANs' (1800-1900) and 'Available VNIDs' (65523).

SWITCHPOOL SETTINGS

The screenshot shows the 'Switchpool' settings page with the 'GENERAL' tab selected. A callout box highlights the 'Auto-import switches' section, which contains the following configuration:

- Auto-import switches**: default OFF, discovered switch CDP/LLDP tables to auto import connected switches with matching credentials
- Enable/disable the automated creation of port channels on switch facing interfaces**:
- default ON, Best to set before any fabric build.**
- Cannot change once some port channels auto-created**

The 'Switchpool wide VLANs' field is set to 'Vlan that will be used for NFM EVPN'.

vPC cluster VLAN

3966

Overlay BACKUP Routing SVI VLAN for vPC

Define the VLAN ID used to program SVI interfaces for vPC L3 peer link. The NFM automatically leverages both L2 and L3 peer links for vPC. This VLAN ID must be excluded from the switchpool wide VLANs configured above.

NFM SW Pool Setting

Default profiles
Select profiles to be associated by default to each switch or interface role

Leaf switch
sjrp-SW

Spine switch
sjrp-SW

Host facing interface
None

Switch facing interface
None

Uplink Interface
None

SW profiles used for AUTO-IMPORT

SWITCHPOOL SETTINGS

GENERAL MORE UNDERLAY

Reserved VLANs
3968-4095

Set the pool of VLANs the NFM must not use on the switches. Use this pool to match any reserved VLANs on each switch in the switchpool. VLAN 1 is always reserved and does not need to be specified.

VNID allocation offset
20000

Set the numeric offset between VLAN ID and VNID used by NFM during VXLAN configuration

Switchpool default VRF
Configure the recirculation VLAN and VNID for the default tenant VRF. This VRF is the default for all the gateways associated to broadcast domains.

VLAN ID
3967

By default, all BD/SVIs are belong to "Switchpool-Default" VRF
VNIID
1677215

NFM SW Pool Setting

SWITCHPOOL SETTINGS

X
GENERAL MORE UNDERLAY X

Gateway MAC address

CA:BB:D3:59:9B:8E

Set the virtual MAC address used for all L3 gateways defined through the NFM

BGP router AS number

65535

OSPF router instance tag

100

Set the dedicated OSPF router instance tag used by the NFM for the underlay. This instance tag must not conflict with other OSPF router instances on the switches.

Underlay pool

IP address pool

10.1.0.0/16

Used for underlay automatic IP addressing

Set the IP address pool used for underlay spine-leaf links and loopback interfaces, including those used for vPC configuration and BGP routers

NFM SW Profile

To automatically import switches, need to create switch profiles with admin id/pw

The screenshot shows the NFM interface with the following details:

- Header:** Cisco Nexus Fabric Manager, admin, alert icon.
- Left Sidebar:** SWITCHPOOL, Home, Profiles, Racks, VRFs, Faults.
- Main Content:** A table titled "CREATE SWITCH PROFILE" with tabs for GENERAL, MORE, SYSLOG, SNMP, CDP, LLDP, and NTP. It shows a single row being edited:
 - Name:** sjp-sw-profil
 - Username:** admin
 - Description:** Set the username for management access authentication
 - Note:** "username/pw must be matched on all switches to be imported into NFM automatically"
 - Password:** (redacted)
 - Message of the day:** (redacted)
- Bottom Buttons:** SAVE CHANGES, CANCEL.
- Bottom Note:** "Also you can set other common configuration, like NTP, SYSLOG, etc..." (in a blue box).

NFM SW Profile

Message of the day

Welcome to SJRP NFM Fabric

Set the banner message shown to users logging into the CLI ⓘ

Beacon LED

Turn on/off the beacon LED

Load balancing method

Source and Destination MAC address

Select the load balancing method to be used by all port channels on the switch

Image

None

The image to run on the switch. This image will be applied when an upgrade is started on the switch or when the switch is brought up using Auto Fabric Provisioning. If an image is configured on the switch, it will override any images configured in upgrades and profiles.

Extra CLI commands

line con
exec-ti 0
line vty
exec-ti 0

Configure extra features that NFM UI doesn't support

Switch features not available through the UI can be managed by entering extra CLI commands.
Refresh the switch to apply new changes.

NFM Fabric Discovery – Auto-Import

The screenshot shows the Cisco Nexus Fabric Manager (NFM) interface. In the top navigation bar, the user is logged in as 'admin'. The main menu includes 'OVERVIEW', 'INTERFACES', 'SWITCHES' (which is selected), 'NEIGHBORS', 'BROADCAST DOMAINS', 'HISTORY', 'FAULTS', and 'SETTINGS'. Below the menu, there are filters for 'Filter by role' and 'Filter expression', and sorting options for 'Sort by ID' and 'Select all'. A prominent blue button labeled 'Import new switch' is visible. The main content area displays a message 'No items found'.

NFM Auto-Import

- Add one seed switch manually
- Seed switch discovers connected SW by LLDP/CDP
- NFM uses SW Profile's username/pw and import all discovered switches
- Requirement
 - Switches must be configured with username/password, MGMT IP, MGMT DEF GW
 - Built-in AFP (Auto Fabric Provision) server can be used to prepare switches

ADD SWITCH

GENERAL MORE

Management address

Specify one switch mgmt. IP address

Enter the IP address of the switch to import. By default switches are imported in monitored mode and are automatically assigned a role. The role can be changed now under the 'More' tab or later before putting the switch into managed mode.

Serial number

optional

Management state

Monitored

The management state determines whether the NFM will actively manage this object or only monitor it in read-only mode. The switch serial number must be entered to enable this setting.

Username

The switch management username can be taken from a default profile configured on the switchpool or from the profile configured on the switch. If a username is configured on the switch, it will override any profile value.

You must make sure username/pw is matched with Switch's configured username/pw to avoid failure

The switch management password can be taken from a default profile configured on the switchpool or from the profile configured on the switch. If a password is configured on the switch, it will override any profile value.

TEST CREDENTIALS

You can test SSH connection to SWITCH.

SAVE CHANGES

NFM Fabric Discovery – Auto-Import

ADD SWITCH

Name: Use existing name on switch GENERAL MORE

Description:

Role:
By default, NFM Recognize 95xx as Spine and 93xx as Leaf
You can manually change
Select the role of the switch. Select 'Auto' to let the NFM decide.

Profile:
Specify which SW profile to auto import switches

A profile is a configuration template containing commonly used settings. These settings will be applied when no user provided overrides are in effect.

Image:
The image to run on the switch. This image will be applied when an upgrade is started on the switch or when the switch is brought up using Auto Fabric Provisioning. If an image is configured on the switch, it will override any images configured in upgrades and profiles.

SAVE CHANGES

NFM searches and install AGENT on SWITCH

Switchpool

OVERVIEW | INTERFACES | **SWITCHES** | NEIGHBORS | BROADCAST DOMAINS | HISTORY | FAULTS | SETTINGS

Filter by role: Filter expression: Sort by: ID Select all

ID	NAME	ROLE	IP ADDRESS	STATUS
1	SW1 (N9504-41)	Spine / NSR-C9504	10.0.8.41	SEARCHING... (N9504-41)
2	SW2 (N9504-41)	Spine / NSR-C9504	10.0.8.41	INSTALLING AGENT... (N9504-41)

TEST CREDENTIALS

NFM Fabric Discovery – Auto-Import

Switchpool

OVERVIEW	INTERFACES	SWITCHES	NEIGHBORS	BROADCAST DOMAINS	HISTORY	FAULTS	SETTINGS
Filter by role ▾	Filter expression		Sort by	ID ▾	<input type="checkbox"/> Select all		
<div style="border: 1px solid #ccc; padding: 5px;"> FSW N9372-42 1 Interface @10.0.7.42 N9K-C9372PK / N9372-42 </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> SW SEARCHING... (N9372-42) 10.0.7.42 </div>							

Switchpool

OVERVIEW	INTERFACES	SWITCHES	NEIGHBORS	BROADCAST DOMAINS	HISTORY	FAULTS	SETTINGS
Leaf, Spine ▾	Filter expression		Sort by	ID ▾	<input type="checkbox"/> Select all		
<div style="border: 1px solid #ccc; padding: 5px;"> SW SW1 (N9504-41) 52 Interfaces @10.0.8.41 Spine / N9K-C9504 </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> SW SW2 (N9372-43) 54 Interfaces @10.0.8.43 Leaf / N9K-C9372X </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> SW SW3 (N9372-42) 54 Interfaces @10.0.8.42 Leaf / N9K-C9372X </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> SW SW4 (SWITCH) 52 interfaces @10.0.8.44 Leaf / N9K-C9372X </div>							

FD Foreign Device

FSW Foreign Switch – N5K, N7K

- Discovered by CDP/LDP and put into NEI tab
- FD objects are persistent even if deleted
- N2K must be connected to N9K by CLI
- vSwitch discovered, VM is not

Error message if N9K version is not supported

fault1
on **SW4 (N9372-44)**
Switch software version 7.0(3)(1)(3) is not supported.
Recently (Today @ 12:24PM)

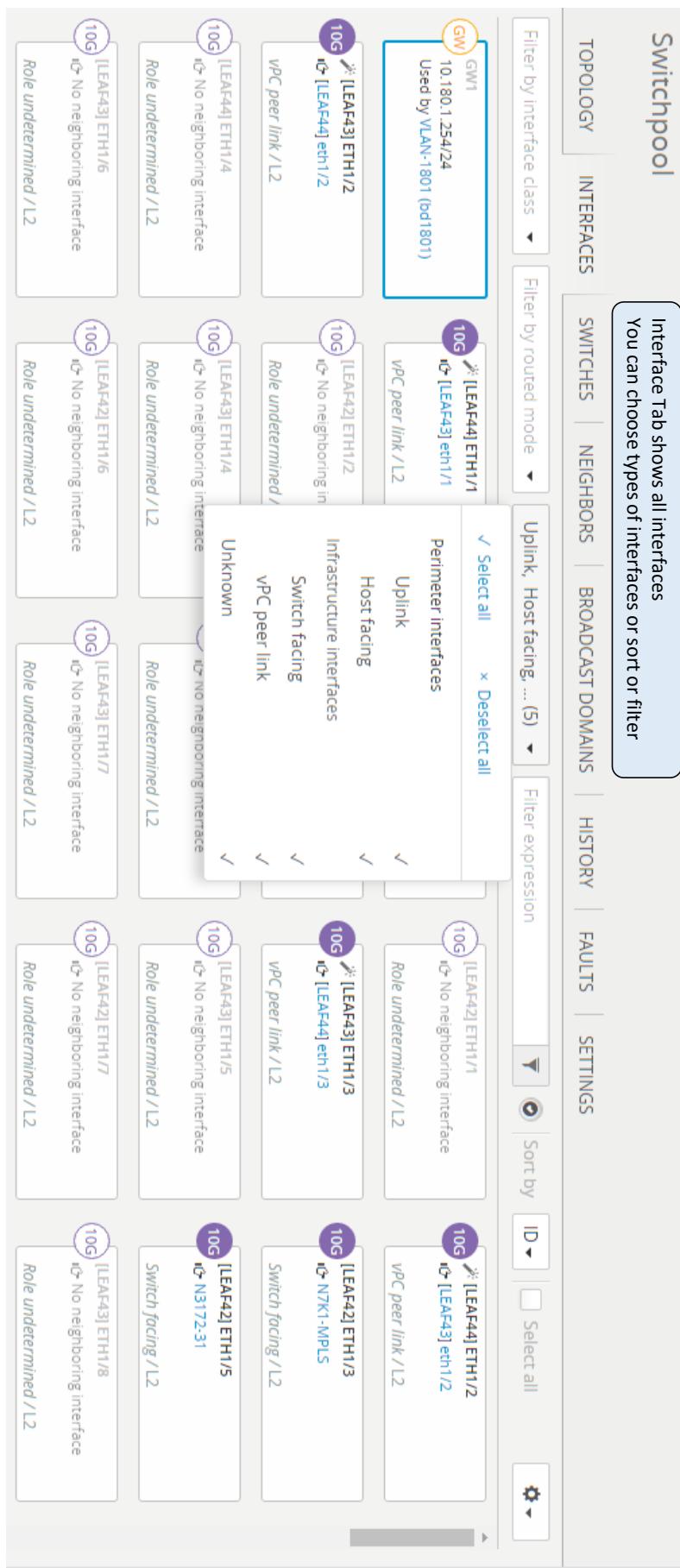
OVERVIEW

OVERVIEW	INTERFACES	SWITCHES	NEIGHBORS	BROADCAST DOMAINS	HISTORY	FAULTS	SETTINGS
Filter by type ▾	Filter expression		Sort by	Device ID ▾	<input type="checkbox"/> Select all		
<div style="border: 1px solid #ccc; padding: 5px;"> FD 3750-17 4 Interfaces @6.6.3 Cisco WS-C3750G-24PS </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> FD 3750-18 4 Interfaces @10.1.20.4 Cisco WS-C3750G-24PS </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> FD 3750X-28 1 Interface @10.160.1.28 Cisco WS-C3750X-12S </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> FD ASR1004-10 1 Interface @10.160.1.10 Cisco ASR1004 </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> FD N3064-33 1 Interface @172.16.156.221 SRPvSM1 </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> FD N3172-31 2 Interfaces N3K-C3064PQ-10GE / N3064-33 Nexus1000V / Sfpvsm1 </div>							

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

cisco

NFM Fabric Discovery – Auto-Import



NFM 1.1 shows [SwitchName] Interface X/Y
Also you can see the character of interface

Also, you can see the character of interface

NFM Fabric Discovery – Auto Fabric Provision

NFM AFP

- Built-in POAP functions with DHCP server, tftp Server
- Add all switches manually by specifying serial no, mgmt ip, username/pw for each switches

Interface Tab shows all interfaces
You can choose types of interfaces or sort or filter

The screenshot shows the 'ADD SWITCH' screen in the NFM AFP interface. It includes fields for Management address (10.0.8.41), Name (N9504-41), Description, Role (Auto), Profile (SjNFM-SW), and Image (admin). A note indicates that the management state is set to 'Monitored'. Below the main form, a note states: 'If choose "Managed", Auto-Create will start immediately'. Another note says: 'You can specify username/pw or use SW profile in "MORE" tab'. A third note states: 'The switch management password can be taken from a default profile configured on the switch, if a password is configured on the switch, it will override any profile value.' At the bottom, there are 'TEST CREDENTIALS' and 'SAVE CHANGES' buttons.

Management address
10.0.8.41

Name
N9504-41

Description

Role
Auto

Profile
SjNFM-SW

A profile is a configuration template containing commonly used settings. These settings will be applied when no user provided overrides are in effect.

Image
admin

The switch management username can be taken from a default profile configured on the switch or from the profile configured on the switch. If a username is configured on the switch, it will override any profile value.

You can specify username/pw or use SW profile in "MORE" tab

.....

The switch management password can be taken from a default profile configured on the switch or from the profile configured on the switch. If a password is configured on the switch, it will override any profile value.

TEST CREDENTIALS

SAVE CHANGES

TEST CREDENTIALS

SAVE CHANGES

NFM Fabric Discovery – Auto Fabric Provision

NFM AFP

- Built-in POAP functions with DHCP server, tftp Server
- Add all switches manually by specifying serial no, mgmt ip, username/pw for each switches

Interface Tab shows all interfaces
You can choose types of interfaces or sort or filter

The screenshot displays two separate windows for adding switches via Auto Fabric Provisioning (AFP).
Top Window (Import new switch):

- Management address:** 10.0.8.41 (highlighted in yellow).

Enter the IP address of the switch to import. By default switches are imported in monitored mode and are automatically assigned a role. The role can be changed now under the 'More' tab or later before putting the switch into managed mode.
- Serial number:** FOX1844GP9Z
- Role:** Auto
- Description:** N9504-41

Bottom Window (ADD SWITCH):

- Management address:** 10.0.8.41 (highlighted in yellow).
- Serial number:** FOX1844GP9Z
- Role:** Auto
- Description:** N9504-41

Annotations:

- If choose "Managed", Auto-Create will start immediately**: A callout points to the 'Role' dropdown in the top window.
- You can specify username/pw or use SW profile in "MORE" tab**: A callout points to the 'Role' dropdown in the bottom window.
- Specify which version to be loaded with POAP**: A callout points to the 'Profile' dropdown in the bottom window.
- TEST CREDENTIALS**: Buttons located at the bottom of both windows.
- SAVE CHANGES**: Buttons located at the bottom right of both windows.

NFM Fabric Discovery – Auto Fabric Provision



All 4 switches are manually imported
Search Failed as switches are not online yet.

Start POAP by wr erase and reload switch

N9372-42# write erase

Warning: This command will erase the startup-configuration.

Do you wish to proceed anyway? (y/n) [n] y

N9372-42# reload

This command will reboot the system. (y/n)? [n] y

2016 Apr 15 05:43:11 N9372-42 %\\$ VDC-1 %\\$ %PLATFORM-2-PFM_SYSTEM_RESET: Manual system restart from Command Line Interface

2016 Apr 15 05:45:54 switch %\\$ VDC-1 %\\$ %ASCII-CFG-2-CONF_CONTROL: System ready
Starting Auto Provisioning ...

2016 Apr 15 05:45:59 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INITED: POAP process initialized
2016 Apr 15 05:46:20 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_DHCP_DISCOVER_START: POAP DHCP Discover phase started
2016 Apr 15 05:46:49 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: Using DHCP, information received over mgmt0 from 10.0.8.27
2016 Apr 15 05:46:49 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: Assigned IP address: 10.0.8.42
2016 Apr 15 05:46:49 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: Netmask: 255.255.255.0
2016 Apr 15 05:46:49 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: DNS Server: 1.1.1
2016 Apr 15 05:46:49 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: Default Gateway: 10.0.8.254
2016 Apr 15 05:46:49 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: Script Server: 10.0.1.27
2016 Apr 15 05:46:49 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: Script Name: poap.py
2016 Apr 15 05:48:23 switch %\\$ VDC-1 %\\$ %POAP-2-POAP_INFO: The POAP Script download has started

NFM Fabric Discovery – Auto Fabric Provision

Check from NFM server

ssh to 172.16.166.53 as root

```
14.04-22:31:25 root@NFM-1-1-1-258~  
# tf /var/log/dhcpd.log  
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: DHCPDISCOVER from 88.1d:fc:ca:3f:48 via eth1  
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: DHCPOFFER on 10.0.8.42 to 88.1d:fc:ca:3f:48 via eth1  
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: Found client uid SA18495M25 len 12  
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: Cisco vendor callout using /opt/esm/server/bin/ciscoDhcpPlugin.py SAL18495M25 true  
Apr 14 22:32:40 NFM-1-1-1-258 dhcpd[14329]: DHCPREQUEST for 10.0.8.42 (10.0.8.27) from 88.1d:fc:ca:3f:48 via eth1  
Apr 14 22:32:40 NFM-1-1-1-258 dhcpd[14329]: DHCPACK on 10.0.8.42 to 88.1d:fc:ca:3f:48 via eth1
```

Use TF to see real time DHCP debug on NFM

vi /opt/esm/data/ztp-tfp/poap.py

NFM customized POAP python script location

```
# cd ./opt/esm/data.local/ztp-images/  
# ls  
Each Switch Provisioned CONFIG for AFP  
9300.txt  
9300.txt.md5  
9500.txt  
9500.txt.md5  
FOX1844GP9Z_provConfig.txt  
FOX1844GP9Z_provConfig.txt.md5  
SAL18495M25_provConfig.txt  
SAL18495M25_provConfig.txt.md5  
SAL18516QBO_provConfig.txt  
SAL18516QBO_provConfig.txt.md5  
SAL19017CEU_provConfig.txt  
SAL19017CEU_provConfig.txt.md5  
nxos.7.0.3.12.2b.bin  
nxos.7.0.3.12.2b.bin.md5  
nxos.7.0.3.12.2c.bin  
nxos.7.0.3.12.2c.bin.md5
```

cat SAL18495M25_provConfig.txt

```
switchname N9372-42  
interface mgmt0  
vrf member management  
ip address 10.0.8.42/24  
vrf context management  
ip route 0.0.0.0/0 10.0.8.254  
username kk password 5  
$5$Kz4n0CdYd1$ig5aK.7IDJhsDfib.guFoBNl8pjw5hHzcITRxoO.C role  
network-admin  
guestshell enable  
feature ssp-server  
feature nxapi  
nxapi https port 443  
nxapi use-vrf management  
copp profile strict  
boot nxos bootflash:/nxos.7.0.3.12.2c.bin  
system vlan 3968 reserve
```

cat SAL18495M25_provConfig.txt

LEAF42 specific provisioned Config

NFM Fabric Auto-Create

Switches are in Monitored mode and need to be changed to Managed mode for auto-create

Switchpool

The screenshot shows the NFM Switchpool interface. There are four switches listed: SW1 (N9504-41), SW2 (N9372-43), SW3 (N9372-42), and SW4 (SWTC). Each switch has a status of "Monitored". A context menu is open over SW4, with the "Edit selected" option highlighted. Other options in the menu include "Import new switch" and "Import peer switch".

SWITCH SETTINGS (MULTIPLE)

Select all 4 SW and click "Edit Selected"

The screenshot shows the NFM Switch Settings (Multiple) interface for four switches. The "MORE" tab is selected. A callout box points to the "Edit Selected" button in the bottom right corner. The interface includes tabs for GENERAL, MORE, SYSLOG, SNMP, CDP, LLDP, and NTP.

Profile

sfp-sw

Select the profile to apply. A profile is a configuration template containing commonly used settings.

Description

Moving switches to managed mode starts fabric configuration
Intra-fabric port channels (can be disabled)
Underlay IP addressing
Underlay interior gateway protocol (OSPF)
Bi-directional Forward Detection (BFD)

Management state

Monitored

Change to Managed

Managed

Monitored

NFM Auto-Create

```
2016 Apr 1 01:57:25 N9372-42 %$ VDC-1 %$ clis[13425]: !!!!! WARNING:  
'LAN_ENTERPRISE_SERVICES_PKG' LICENSE NOT FOUND ON THE SYSTEM !!!!. You  
have tried to enable a licensed feature [ospf] without installing the  
'LAN_ENTERPRISE_SERVICES_PKG' license, which has not been found on the system.  
2016 Apr 1 01:57:25 N9372-42 %$ VDC-1 %$ clis[13425]: System supports honor  
based license.Feature will be enabled and fully functional.License usage for feature  
will appear after licence intall and a reload.
```

As soon as you change SW to managed, getting license warning as features are automatically configured

```
N9372-42# sh run | i feature  
feature nxapi  
feature ospf  
feature bgp  
feature interface-vlan  
feature vn-segment-vlan-based  
feature lacp  
feature lldp  
feature bfd  
feature nv overlay  
  
Fri Apr 1 01:57:30 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; feature  
vn-segment-vlan-based(SUCCESS)  
Fri Apr 1 01:57:32 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; feature  
bfd (SUCCESS)  
Fri Apr 1 01:57:34 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; feature  
lacp (SUCCESS)  
Fri Apr 1 01:57:35 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; feature  
lldp (SUCCESS)  
Fri Apr 1 01:57:37 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; nv  
overlay evpn (SUCCESS)  
Fri Apr 1 01:57:37 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; cdp  
enable (SUCCESS)  
Fri Apr 1 01:57:37 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; banner  
motd _welcome to SJ RP EVPN Fabric_ (SUCCESS)  
Fri Apr 1 01:57:37 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; lldp  
holdtime 120 (SUCCESS)  
Fri Apr 1 01:57:37 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; lldp timer  
2 (SUCCESS)  
Fri Apr 1 01:57:37 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; port-  
30 (SUCCESS)  
Fri Apr 1 01:57:37 2016:type=update:id=nginx:user=admin:cmd=configure terminal ; port-  
channel load-balance src-dst ip-l4port (SUCCESS)
```

From accounting log, we can
see NFM pushes
configurations by NX-API

NFM Auto-Create

You will see some Faults reported but all of them will be resolved when PO completed.
It takes 3-5 min

Switchpool

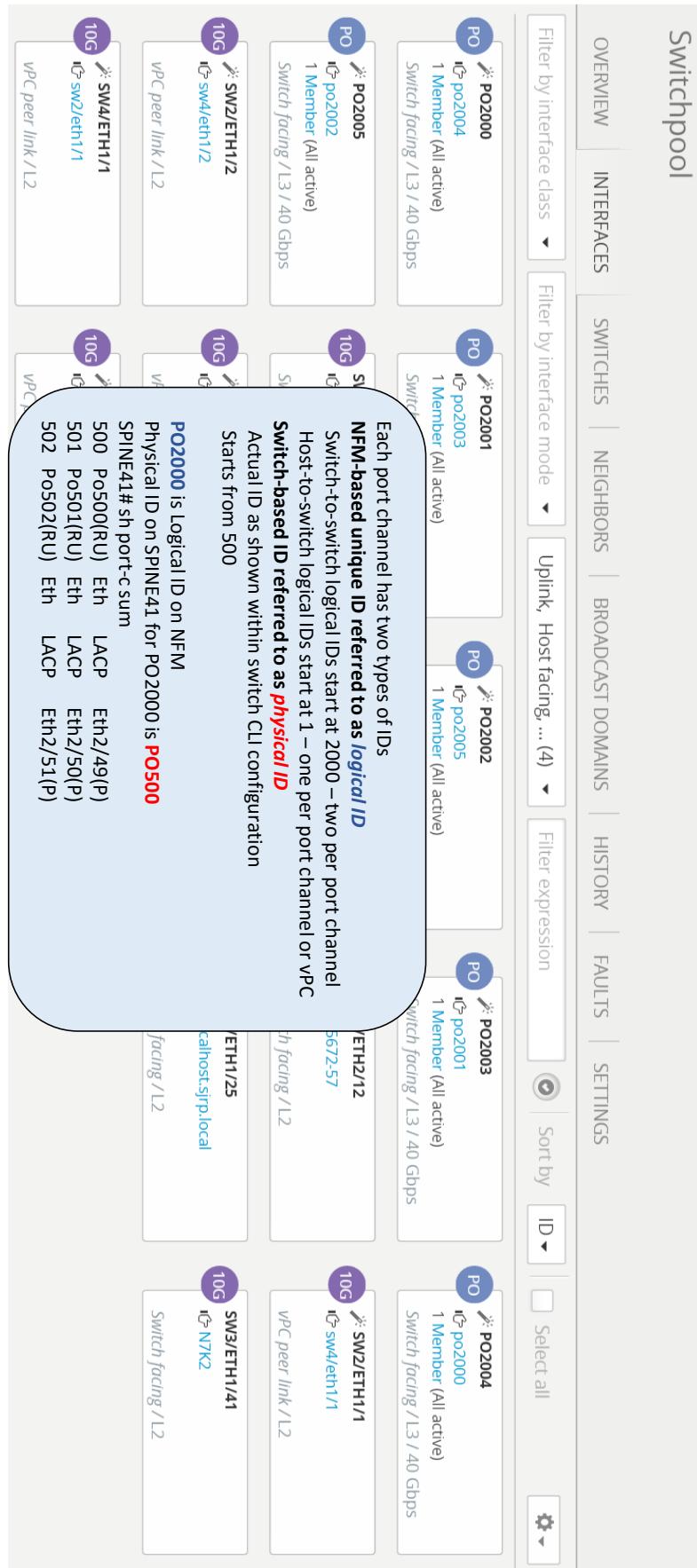
Cisco Nexus Fabric Manager

admin 6

TOPOLOGY	INTERFACES	SWITCHES	NEIGHBORS	BROADCAST DOMAINS	HISTORY	FAULTS	SETTINGS
Filter by role ▾	Filter expression		▼ Sort by ID ▾	<input type="checkbox"/> Select all			
FSW SJRPNTK2-PE2 3 interfaces @10.0.7.2 N7K-C7010 / SJRPNTK2-PE2		FSW N7K1-MPLS 1 Interface @10.0.7.3 N7K-C7010 / N7K1-MPLS	SW SW1 (N9504-41) 52 interfaces @10.0.8.41 Spine / N9K-C9504	SW SW3 (N9572-43) 54 interfaces @10.0.8.43 Leaf / N9K-C9572T		SW fault4 on po2000 None of the member interfaces are currently active Recently (Today @ 0:02AM)	
SW SW4 (N9372-44) 54 interfaces @10.0.8.44 Leaf / N9K-C9372T		SW SW1 (N9504-41) 52 interfaces @10.0.8.41 Spine / N9K-C9504	SW fault5 on po2001 None of the member interfaces are currently active Recently (Today @ 0:02AM)	SW SW3 (N9572-43) 54 interfaces @10.0.8.43 Leaf / N9K-C9572T	SW fault6 on po2002 None of the member interfaces are currently active Recently (Today @ 0:02AM)	SW fault7 on po2003 /po2002 None of the member interfaces are currently active Recently (Today @ 0:02AM)	SW fault8 on po2004 /po2001 None of the member interfaces are currently active Recently (Today @ 0:02AM)
SW fault9							

NFM Auto-Create

We can see all PO Channels are created and up



ON SPINE41

```
N9504-41# sh ip int b
Lo501          10.1.0.0   protocol-up/link-up/admin-up
Po500          10.1.0.11  protocol-up/link-up/admin-up
Po501          10.1.0.9   protocol-up/link-up/admin-up
Po502          10.1.0.13  protocol-up/link-up/admin-up
```

SPINE has 3 PO to 3 LEAF

```
N9504-41# sh port-csum
500 Po500(RU) Eth LACP   Eth2/49(P)
501 Po501(RU) Eth LACP   Eth2/50(P)
502 Po502(RU) Eth LACP   Eth2/51(P)
```

Spine has only one loopback 501

```
interface loopback501
description Used by NFM for EVPN routing
ip address 10.1.0.0/32
ip ospf network point-to-point
ip router ospf 100 area 0.0.0.0
ip ospf bfd
```

L2EVPN NEI to all 3 LEAFs

```
router ospf 100
bfd
router-id 10.1.0.0
redistribute static route-map local-into-ospf
```

```
route-map local-into-ospf permit 10
match route-type local
```

```
router bgp 65535
```

```
router-id 10.1.0.0
neighbor 10.1.0.1 remote-as 65535
bfd
remote-as 65535
```

```
update-source loopback501
address-family l2vpn evpn
```

```
send-community both
route-reflector-client
```

```
ip address 10.1.0.11/31
```

```
no ipv6 redirects
```

```
ip ospf network point-to-point
```

```
ip router ospf 100 area 0.0.0.0
```

```
ip ospf bfd
```

ON LEAF43

```
N9372-42# sh ip int b
Lo500      10.1.0.4    protocol-up/link-up/admin-up
Lo501      10.1.0.1    protocol-up/link-up/admin-up
Po500     10.1.0.8    protocol-up/link-up/admin-up
```

N9372-42# sh port-csum | i Po

Group	Port-	Type	Protocol	Member Ports
500	Po500(RU)	Eth	LACP	Eth1/51(P)

interface Ethernet1/51

```
no switchport
mtu 2048
channel-group 500 mode active
no shutdown
```

interface port-channel500

```
description This interface has been created by Nexus Fabric Manager at
10.0.1.27. This port-channel was auto-created between sw3 and sw1
```

```
no switchport
mtu 2048
bfd interval 50 min_rx 50 multiplier 3
no ip redirects
ip address 10.1.0.8/31
no ipv6 redirects
ip ospf network point-to-point
ip router ospf 100 area 0.0.0.0
ip ospf bfd
```

Leaf has two loopbacks, 500 for NVE Source and 501 for ROUTING

```
interface loopback500
description Used by NFM for VXLAN termination (source-interface of nve1)
ip address 10.1.0.4/32
ip ospf network point-to-point
ip router ospf 100 area 0.0.0.0
ip ospf bfd
```

```
interface loopback501
description Used by NFM for EVPN routing
ip address 10.1.0.1/32
ip ospf network point-to-point
ip router ospf 100 area 0.0.0.0
ip ospf bfd
```

```
N9372-42# sh vrf
VRF-Name          VRF-ID State Reason
default           1 Up    --   no shutdown
management        2 Up    --   vrf member underlay-l3
switchpool-default 3 Up    --   ip forward
underlay          4 Up    --   ipv6 address use-link-local-only
```

vrf context underlay

```
address-family ipv4 unicast
```

vrf context switchpool-default

```
vni 16777215
rd auto
address-family ipv4 unicast
route-target both auto
route-target both auto evpn
router bgp 65535
vrf underlay-l3
address-family ipv4 unicast
advertise l2vpn evpn
```

```
interface Vlan3967
no shutdown
vrf member underlay-l3
ip forward
ipv6 address use-link-local-only

interface nve1
no shutdown
description Used by NFM for VXLAN termination
source-interface loopback500
host-reachability protocol bgp
member vni 16777215 associate-vrf
```

```
N9372-42# sh vxlan
Vlan      VN-Segment
=====  =====
3967    16777215

cisco
```

NFM Port-Channel

Click Port-Channel NO will get to PO window

Logical port channel po2000

TX

RX

Members

Member	Status
sw1 [SPINE41]	1h
sw4 [LEAF44] eth1/5/1	1h

Events

Event ID	Description	Property	Original value	New value
task9410	Update auto-created port-channel po2000. Succeeded 1 minute ago (Today @ 12:14PM).			
task9404	Update auto-created port-channel po2000. Succeeded 1 minute ago (Today @ 12:14PM).			

NFM vPC Creation

NFM doesn't configure VPC until configure HOST interfaces that connects to two Leafs

Switchpool

OVERVIEW INTERFACES SWITCHES NEIGHBORS BROADCAST DOMAINS HISTORY FAULTS SETTINGS

Filter by interface class ▾ Filter by interface mode ▾ Uplink, Host facing, ... (5) ▾ 3750-17

Sort by ID ▾ Select all

1G SW2/ETH1/43 1G SW2/ETH1/44 1G SW4/ETH1/5 1G SW4/ETH1/6

iC 3750-17 iC 3750-17 iC 3750-17 iC 3750-17

Role undetermined / L2 Role undetermined / L2 Role undetermined / L2 Role undetermined / L2

Use filter for specific HOST or SW for vPC PO

Edit selected

INTERFACE SETTINGS(MULTIPLE)

X

GENERAL SETTINGS MORE SETTINGS

Profile None

Select the profile to apply. A profile is a configuration template containing commonly used settings.

Description //

Role Uplink

Select the role of the interface based on the attached device. Select 'Auto' to let NFM decide.

Changed ROLE from AUTO to "Uplink"
This is only required if you are creating vPC to L2 SW
Also, You can leave it as AUTO

Check all 4 Interfaces

cisco
State
Enable/disable the interface

NFM vPC Creation

OVERVIEW | **INTERFACES** | **SWITCHES** | **NEIGHBORS** | **BROADCAST DOMAINS** | **HISTORY** | **FAULTS** | **SETTINGS**

Filter by interface class ▾ Filter by interface mode ▾ Switch facing, vPC peer link, ... (3) ▾ 3750-17

Sort by ID ▾ Select all

1G SW2/ETH1/43 1G SW2/ETH1/44 1G SW3/ETH1/5

1G SW2/ETH1/43 1G SW2/ETH1/44 1G SW3/ETH1/5

Role undetermined / L2 ✓ Role undetermined / L2 ✓

CREATE NEW LOGICAL PORT CHANNEL

Name: vpc-17 **GENERAL** **MORE** **ROUTING**

Profile: None

Type NAME

LOGICAL PORT CHANNEL SETTINGS

Add to logical port channel

PO1 Type NAME

Port channel mode: LACP

Set LACP On

Select whether members use LACP or be static

Untagged broadcast domain: None

Select the untagged broadcast domain to use for an interface. The VLAN ID associated with a broadcast domain for an interface corresponds to the access VLAN ID for access interfaces, or the native VLAN ID for trunk interfaces.

Role: Uplink

Match ROLE to Uplink Only for L2 SW You can leave it as AUTO

State: Enabled(disable the interface)

Selected (4) Moving (0) Excluded (0)

Valid candidate members that will be added to the port channel:

- SW2/eth1/43
- SW2/eth1/44
- SW3/eth1/5
- SW3/eth1/6

NEIGHBORS | **BROADCAST DOMAINS** | **HISTORY** | **FAULTS** | **SETTINGS**

e mode ▾ Uplink, Switch facing ... (4) ▾ 3750-17

Edit selected

Add to new broadcast domain

Add to broadcast domain

PO PO1 (VPC-3750-17)

No neighboring interface

4 Members (All active)

L2 / 4 Gbps

Now, add to new PO

LOGICAL PORT CHANNEL SETTINGS

Add to logical port channel

VRF: None

Routed mode:

Select whether the interface will be static or dynamic mode

IP address: For L3 PO

Set the IP address and prefix of the interface, for example '192.168.1.1/24'

MAC address: None

Select the tenant VRF to associate with this interface

MTU: 1500

Set the administrative maximum transmission unit (MTU) size

SAVE CHANGES

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

CIS

NFM vPC Creation

N9372-43# sh port-c sum

500 Po500(RU)	Eth	LACP	Eth1/51(P)
501 Po501(SU)	Eth	LACP	Eth1/43(P) Eth1/44(P)
502 Po502(SU)	Eth	LACP	Eth1/1(P) Eth1/2(P) Eth1/3(P)
503 Po503(SU)	Eth	LACP	Eth1/45(P) Eth1/46(P)

Legend:

(*) - local vPC is down, forwarding via vPC peer-link

N9372-43# sh bgp l2vpn evpn sum

BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.1.0.1, local/AS number 65535
BGP table version is 3, L2VPN EVPN config peers 1, capable peers 1
0 network entries and 0 paths using 0 bytes of memory
BGP attribute entries [0/0], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [0/0]

vpc domain 1

peer-switch

peer-keepalive destination 10.0.8.44

peer-gateway

ipv6nd synchronize

ip arp synchronize

use MGMT IP as vpc

peer-keepalive

vpc 1.

interface port-channel501

vpc 1.

interface port-channel502

vpc peer-link

interface port-channel503

vpc 2.

vpc status

vPC Peer-link status

vPC Port Status Active Vlans

vPC id Port Status Active Vlans

1 Po502 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

interface loopback500

description Used by NFM for VXLAN termination (source-interface of nve1)

ip address 10.1.0.4/32

ip address 10.1.0.7/32 secondary

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

ip router ospf 100 area 0.0.0.0

ip ospf bfd

vPC status

id Port Status Consistency Reason Active Vlans

1 Po501 up success success -

2 Po503 up success success -

ip ospf network point-to-point

NFM VPC to SW18

The screenshot shows the NFM interface for creating a new logical port channel. The main window displays a list of interfaces: SW2/ETH1/45 (1G), SW2/ETH1/46 (1G), SW3/ETH1/17 (1G), and SW3/ETH1/18 (1G). The interface SW3/ETH1/18 is selected. A context menu is open over this interface, showing options: 'Edit selected', 'Add to new broadcast domain', and 'Add to broadcast domain'. The 'Add to new broadcast domain' option is highlighted.

CREATE NEW LOGICAL PORT CHANNEL

GENERAL SETTINGS MORE SETTINGS

Name vpc-SW18

Profile None

Select the profile to apply. A profile is a configuration template containing commonly used settings.

Description

Role

Uplink

Select the role of the logical port channel based on the attached device. Select 'Auto' to let NFM decide.

State

Enable/disable the interface

SELECTED (4) **MOVING (0)** **EXCLUDED (0)**

Valid candidate members that will be added to the port channel

IP address

LACP

Select whether member interfaces will use active mode LACP or be static

Port channel type

Switched (L2) or Routed (L3) mode

Untagged broadcast domain

None

Select the untagged broadcast domain to use for an interface. The VLAN ID associated with a broadcast domain for an interface corresponds to the access VLAN ID for access interfaces, or the native VLAN ID for trunk interfaces.

VRF

None

Select the tenant VRF to associate with this interface

SAVE CHANGES

NFM vPC to N1K VSM4

Uplink, Host facing... (5) ▾ **sjpsm** Sort by ID ▾ Select all

Role	Host facing	Interface	Description
1G	SW2/ETH1/24	1G sjpsm1	Role undetermined / L2 ✓
1G	SW4/ETH1/24	1G sjpsm1	Role undetermined / L2 ✓

UpLink, Host facing... (5) ▾ **sjpsm** Sort by ID ▾ Select all

Role	Host facing	Interface	Description
1G	SW2/ETH1/24	1G sjpsm1	Host facing / L2 ✓
1G	SW4/ETH1/24	1G sjpsm1	Host facing / L2 ✓

CREATE NEW LOGICAL PORT CHANNEL

Name **vpc-sjpsm14** GENERAL SETTINGS MORE SETTINGS

vPC Peer-link status

Role	Host Facing	Interface	Description
None	Select the profile to apply. A profile is a configuration template containing commonly used settings.		
Description			

interface Ethernet1/24

After changed to HOST FACING, BPDUGUARD Enabled

switchport mode trunk
switchport trunk allowed vlan none
spanning-tree bpduguard disable

interface Ethernet1/24

switchport mode trunk
switchport trunk allowed vlan none
spanning-tree bpduguard enable

vPC status

id	Port	Status	Active vLans	Reason	Active Group Port-Channel	Type	Protocol	Member Ports
1	Po502	up	1,3966-3967		vpc 2			

Host Facing

Select the role of the logical port channel based on the attached device. Select 'Auto' to let NFM decide.

State Enable/disable the interface

SELECTED (2) MOVING (0) EXCLUDED (0)

Valid candidate members that will be added to the port channel

SW4/eth1/24
SW2/eth1/24

SAVE CHANGES

NFM Leaf 42 to SW33

Uplink, Host facing, ... (5) ▾ 3064

Edit selected

Role	AUTO
Auto	
Host facing	
Switch facing	
Uplink	
vPC peer link	

Role assignment failed / 12 ✓

mode ▾ Uplink, Host facing, ... (5) ▾ 3064

Edit selected

Add to new broadcast domain
Add to broadcast domain
Add to logical port channel

LOGICAL PORT CHANNEL SETTINGS

GENERAL SETTINGS MORE SETTINGS

Name host33

Profile PO4

Description None

Select the profile to apply. A profile is a configuration template containing commonly used settings.

PO4 (HOST33)

No neighbouring interface
1 Member (all active)

Uplink / L2 / 10 Gbps

interface Ethernet1/17

switchport mode trunk
switchport trunk allowed vlan none
spanning-tree bpduguard disable
channel-group 501 mode active

Role

Uplink

Select the role of the logical port channel based on the attached device. Select 'Auto' to let NFM decide.

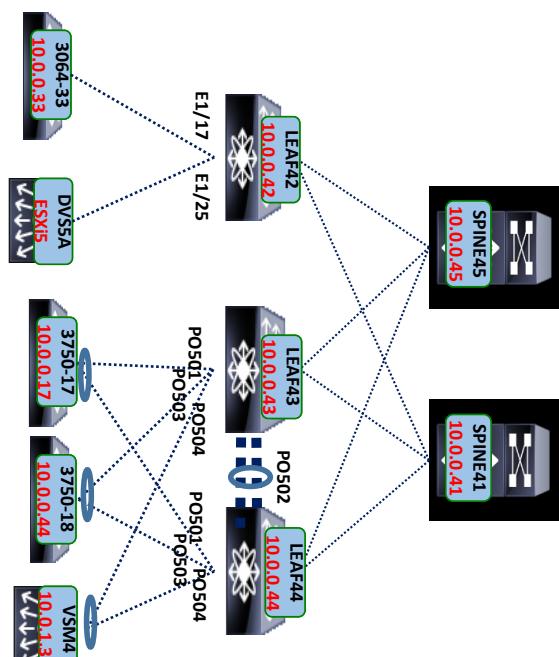
State

Enable/disable the interface ⓘ

SAVE CHANGES

Setting Role as UPLINK disables BPDU GUARD

NFM After Auto-Create



```
N9504-41# sh ip os nei
OSPF Process ID 100 VRF default
Total number of neighbors: 3
Neighbor ID          Pri State      Up Time Address           Interface
10.1.0.1              1 FULL/       23:46:57 10.1.0.8   Po501
10.1.0.2              1 FULL/       23:46:59 10.1.0.10  Po500
10.1.0.3              1 FULL/       23:46:58 10.1.0.12  Po502
```

```
N9372-42# sh bgp 12 evpn sum
Neighbor ID          V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down
10.1.0.0              4 65535    1431     1431      3     0     0 23:45:15 0
N9372-42# sh int nve 1
nve1 is up
admin state is up, Hardware: NVE
Description: Used by NFM for VXLAN termination
MTU 9216 bytes
Encapsulation VXLAN
Auto-mdix is turned off
```

```
RX
  ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
TX
  ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
```

VRF Creation

The screenshot shows the Cisco Nexus Fabric Manager interface. On the left, a navigation bar includes links for Switchpool, Profiles, Faults, Racks, VRFs, Images, Upgrades, Users, and Account settings. The main area displays a table of VRFs:

VRF	Description	State
(SWITCHPOOL DEFAULT)	No members	<input checked="" type="checkbox"/>
kk01		<input type="checkbox"/>

A tooltip for the 'default VRF' row states: "You can use this VRF for all traffic". A tooltip for the 'Create new VRF' button says: "You can use this VRF for all traffic".

Below the table, a modal window titled "ADD VRF" is open. It has tabs for "GENERAL" and "L3". The "GENERAL" tab contains fields for "Name" (set to "kk01") and "Description". A tooltip for the "Name" field states: "Only Name is required and NFM will automatically assign VLAN ID for L3 VNI when GW is assigned to VRF". The "L3" tab contains a "VNI" field set to "0". A tooltip for this field states: "L3 VNI for VRF is not created until GW is assigned to VRF".

On the left side of the screen, several terminal sessions are displayed:

- N9372-42# sh vrf | i kk**
- N9372-42# sh ip int b vrf kk2**
- N9372-43# sh vrf | i kk**
- N9372-44# sh vrf | i kk**

Output from the first session:

```

Vrf "KK1" is created on all LEAFs after VRF
creation on NFM without and GW
attachment to VRF

```

Output from the second session:

```

IP Interface Status for VRF "kk2"(6)
Interface          IP Address      Interface Status
kk01

```

Output from the third session:

```

N9372-43# sh vrf kk01 detail
VRF-Name: kk01, VRF-ID: 5, State: Up
VPNID: unknown
RD: 0:0
VNI: 0
Max Routes: 0 Mid-Threshold: 0
Table-ID: 0x80000005, AF: IPv6, Fwd-ID: 0x80000005, State: Up
Table-ID: 0x00000005, AF: IPv4, Fwd-ID: 0x00000005, State: Up

```

NFM BD Creation

BD = VLAN/VNI

The screenshot shows the NFM interface with the 'Switchpool' tab selected. A modal window titled 'CREATE NEW BROADCAST DOMAIN' is open, asking for a 'VLAN ID' which is set to 'bd1801'. The background shows other tabs like 'OVERVIEW', 'INTERFACES', etc.

BD is just created on NFM but not deployed to any LEAF yet

```
N9372-42# sh vxlan
Vlan      VN-Segment
=====
3967    16777215
=====
Only default L3 VNI for ""switchpool-default" exist"
```

The screenshot shows the 'CREATE NEW BROADCAST DOMAIN' dialog. It has fields for 'Name' (bd1801), 'VLAN ID' (bd1801), 'Description', 'State' (disabled), and 'Include all perimeter interfaces' (disabled). A note in the dialog states: 'NFM assigns lower VLAN ID from Switchpool wide VLANs or you can assign manually'.

Gateway
The MAC address for all gateways can be configured from the switchpool settings 'Underlay' tab.

IP address: From NFM 1.1, GW is included in BD creation UI
If don't configured IP/VRF, then BD is L2 VNI only
10.180.1.254/24
Set the IP address and prefix of the gateway, for example '192.168.1.1/24'

VRF: kk1
Select the tenant VRF to associate with this gateway

MTU: 9216
Set the administrative maximum transmission unit (MTU) size

cisco

BD Deployment

Method 1: Overview UI. Select Interface from UI and Right-Click

From Switchpool → Overview
Double-Clicking SW icon extends all interface connections

Select Interfaces that you want to deploy BD

Interface with "Add to New BD"

Add to broadcast domain

- Add to new broadcast domain
- Import new switch
- Edit selected
- Add to logical port channel
- Add to new upgrade
- Refresh topology
- Expand all
- Collapse all

After adding int into BD

Vlan	VN-Segment
3967	16777215
N9372-43# sh vxlan	Vlan
1801	VNI-Segment
3967	21801
3967	16777215

ADD TO EXISTING BROADCAST DOMAIN

Add to: VLAN-1801 (bd1801) ▾

VLAN-1801 (bd1801) TIED (1) EXCLUDED (0)

Candidate members that will be added to the broadcast domain

po1 (vpc-3750-17)

ADD

Port STP Forwarding

Port	STP Forwarding
Po501	1801
Po502	1,1801,3966-3967

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

cisco

BD Deployment

Method 2: Interface UI. Select multiple Interfaces from UI

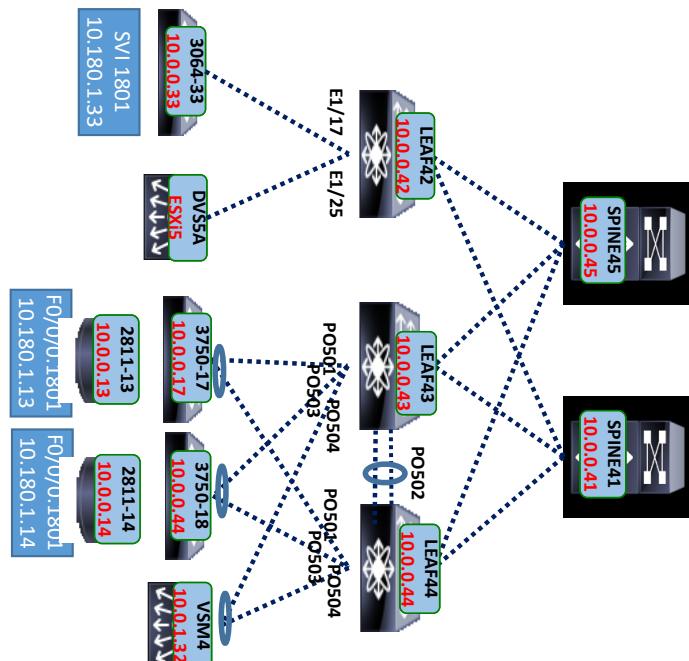
The screenshot shows the Cisco Interface UI with the following steps:

- Switchpool:** A list of logical ports (PO1-PO4) with their details. PO1 and PO2 are selected.
- Add to Existing Broadcast Domain:** A modal dialog where the selected ports (PO1 and PO2) are being added to the broadcast domain "VLAN-1801 (bd-1801)".
- Broadcast Domain VLAN-1801 (bd-1801):** A list of members for BD-1801. It shows two entries: VLAN-1801 (bd-1801) and VLAN-1802 (bd-1802). The first entry has 4 Members (All active), while the second has 0 Members.
- VLAN-1802 (bd-1802):** A list of members for BD-1802. It shows one entry: VLAN-1802 (bd-1802) with 0 Members.

Annotations highlight the following areas:

- Filter or Sort from Interface and select multiple interfaces:** A callout pointing to the interface list in the top-left.
- ADD TO EXISTING BROADCAST DOMAIN:** A callout pointing to the "Add to:" dropdown in the modal dialog.
- Click BD or members of BD opens BD info:** A callout pointing to the BD list in the middle-left.
- BD with 0 members appears as empty Circle:** A callout pointing to the empty circle icon for BD-1802.

BD Deployment Test



```
N3064-33# sh run int v1 1801
interface Vlan1801
no shutdown
mac-address 0000.1801.0033
vrf member host1801
ip address 10.180.1.33/24
```

```
N3064-33# ping 10.180.1.13 vrf host1801
PING 10.180.1.13 (10.180.1.13): 56 data bytes
36 bytes from 10.180.1.33: Destination Host Unreachable
Request 0 timed out
64 bytes from 10.180.1.13: icmp_seq=1 ttl=254 time=1.754 ms
64 bytes from 10.180.1.13: icmp_seq=2 ttl=254 time=1.47 ms
```

```
N9372-42# sh nve peers
Interface Peer-IP          State LearnType Uptime   Router-Mac
nve1      10.1.0.7          Up    CP        03:33:41  88f0.3187.1f3b
```

```
N9372-42# sh mac add v1 1801
```

*	1801	0000.1801.0033	dynamic	0	F	F	Po501
*	1801	0015.f905.8200	dynamic	0	F	F	nve1(10.1.0.7)
*	1801	001a.6c70.3697	dynamic	0	F	F	sup-eth1(R)
G	1801	881d.fcfa.3f4f	static	-	F	F	

```
N9372-43# sh mac address-table v1 1801
```

*	1801	0000.1801.0033	dynamic	0	F	F	nve1(10.1.0.6)
+	1801	0015.f905.8200	dynamic	0	F	F	Po503
+	1801	001a.6c70.3697	dynamic	0	F	F	Po501
G	1801	88f0.3187.1f3b	static	-	F	F	vPC Peer-Link(R)
G	1801	88f0.3187.8c4b	static	-	F	F	sup-eth1(R)

BD Deployment Test

```
N9372-42# sh vxlan
1801      21801
1802      21802
1900      16777214
3967      16777215
```

interface nve1

```
no shutdown
description Used by NFM for VXLAN termination
```

```
source-interface loopback500
host-reachability protocol bgp
```

```
member vni 21801
ingress-replication protocol bgp
```

```
member vni 21802
ingress-replication protocol bgp
```

```
member vni 16777214 associate-vrf
member vni 16777215 associate-vrf
```

NFM uses BGP Replication now

evpn

```
vni 21801:12
```

```
rd auto
```

```
route-target import auto
```

```
vni 21802:12
```

```
rd auto
```

```
route-target import auto
```

```
vni 16777214:12
```

```
rd auto
route-target import auto
```

```
route-target export auto
vni 16777215:12
```

```
rd auto
route-target import auto
route-target export auto
```

NFM created non-necessary EVPN for L3 VNI.
Known Cosmetic BUG. Will be fixed in 1.2.

cisco

NFM Switch SSH from GUI to Switch

Launch SSH session from GUI to Switch

Switchpool

TOPOLOGY	INTERFACES	SWITCHES	NEIGHBORS	BROADCAST DOMAINS	HISTORY	FAULTS	SETTINGS
Filter by role ▾	Filter expression	Sort by ID ▾	<input type="checkbox"/> Select all				
FSW N7K1-MPLS Interface @10.0.7.3	FSW SIRPNPK2-PF2 3 interfaces @10.0.7.2	SW SW1 (N9504-41) 52 interfaces @10.0.8.41 Spine /N9K-C9504					
NTK-CT010 / N7K1-MPLS	NTK-CT010 / SIRPNPK2-PF2	SW SW4 (N9372-44) 54 interfaces @10.0.8.44 Leaf /N9K-C9372PX					
SW SW2 (N9372-42) 10 Peer switch sw4 (N9372-44)	SW SW2 (N9372-42) 10 Peer switch sw3 (N9372-43)	SW SW2 (N9372-42) 10 Peer switch sw2 (N9372-42)					
Leaf /N9K-C9372PX	Leaf /N9K-C9372TX	Leaf /N9K-C9372TX					
1. Click SWITCH							
2. Click							
3. Opens new browser tab with SSH access							

Switchpool / Switch sw2 (N9372-42 @ 10.0.8.42)

OVERVIEW	INTERFACES	NEIGHBORS	HISTORY	FAULTS
SYSTEM	CPU LOAD	SW Switch	Reachable	
EDIT CLI REFRESH DELETE		sw2		
2. Click CLI				

The copyright rights to certain works contained in this software are owned by other third parties and used and distributed under their own licenses, such as open source. This software is provided "as is," and unless otherwise stated, there is no warranty, express or implied, including but not limited to warranties of merchantability and fitness for a particular purpose. Certain components of this software are licensed under the GNU General Public License (GPL) version 2.0 or the GNU Lesser General Public License (LGPL) Version 2.1 or Lesser General Public License (LGPL) Version 2.0. A copy of each such license is available at <http://www.gnu.org/licenses/gpl-2.0.php> and <http://www.opensource.org/licenses/gpl-3.0.html> and <http://www.gnu.org/licenses/old-licenses/lgpl-2.1.php> and <http://www.gnu.org/licenses/library.txt>.

NFM Switch Upgrade

Upload Image to NFM first

The screenshot shows the Cisco Nexus Fabric Manager interface. The top navigation bar includes 'Switchpool' (highlighted with a red box), 'Home', 'Profiles', 'Racks', 'VRFs', 'Faults', and 'Images'. The 'Images' tab is selected and highlighted with a blue box. The main content area displays a progress bar for uploading a new image:

- Step 1:** Shows the 'UPLOAD NEW IMAGE' dialog with a 'CHOOSE FILE' button (highlighted with a red box).
- Step 2:** Shows the 'UPLOAD NEW IMAGE' dialog with a progress bar at 35% (highlighted with a red box).
- Step 3:** Shows the 'UPLOAD NEW IMAGE' dialog with a progress bar at 35%.
- Step 4:** Shows the 'Open' file dialog with the file 'nxos.7.0.3.12.3.bin' selected (highlighted with a red box).
- Step 5:** Shows the 'UPLOAD NEW IMAGE' dialog with a progress bar at 35%.

Below the upload progress, the 'Images' table lists the uploaded image:

Name	Description	Version	Release date	Status
nxos.7.0.3.12.3.bin		7.0(3)I2(3)	23 days ago (Tue, Mar 22, 5:04AM)	Active

A callout box labeled '6. Image is listed after upload finished' points to the table.

NFM Switch Upgrade

Switchpool

TOPOLOGY	INTERFACES	SWITCHES	NEIGHBORS	BROADCAST DOMAINS	HISTORY	FAULTS	SETTINGS
Leaf	Filter expression	<input type="button" value="▼"/> <input type="radio"/> <input type="radio"/> Sort by	<input type="checkbox"/> ID <input type="checkbox"/> Selected				

SW1 (SPINE41)
52 Interfaces @10.0.8.41

SW2 (LEAF42)
54 Interfaces (1-1) @10.0.8.42
No peer switch
Leaf / N9K-C9372PX

SW3 (LEAF43)
54 Interfaces
Peer switch
Import new switch
Leaf / N9K-C9372PX

Add to new upgrade

CREATE UPGRADE

Name: UPG5 to 2.2C

GENERAL MORE

Description:

2. Type Name and Choose version to upgrade

Image: 7.0(3)J2(2c)S9 (nxos.7.0.3J2.2c.bin)

Select the image to apply to switches that are part of this upgrade. Images specified in switch profiles, or on a switch directly, will override this setting.

SELECTED (1) EXCLUDED (0)

3. Parallel is default

SAVE CHANGES

Created new Upgrade
Upgrade to 2.2C

The following switches will be added to the upgrade configuration

sw2 (LEAF42)

cisco

SAVE CHANGES

NFM Switch Upgrade

Upgrades

Filter expression Sort by Name ▾

Name ▾	Description	State	Strategy	Image	Members	Start time	End time
UPG to 2.2C	-	Inactive	Parallel	7.0(3)I2(2c).S9 (nxos 7.0.3.I2.2c...)	1	-	-

1. Click Upgrade Obj

Upgrade UPG to 2.2C

OVERVIEW MEMBERS FAULTS

Upgrade	Inactive
EDIT START STOP DELETE	

2. Click Start

Filter expression Sort by ID ▾ Select all

Upgrade UPG to 2.2C

OVERVIEW MEMBERS FAULTS

Leaf, Spine ▾

Leaf, Spine Sort by ID ▾ Select all

SW SW2 (LEAF42)
54 Interfaces (1:1) @10.0.8.42
No peer switch
Leaf / N9K-C9372PX

You can remove member

Switches

sw2 (LEAF42) using image 7.0(3)I2(2c).S9 (nxos 7.0.3.I2.2c.b..
(version 1f1fe8341-f57-c-334f-
b25c-3c70a26559720)

Members 1
Start time -
End time -
Switches

Upgrade started UPG to 2.2C

SW2 (LEAF42) Upgrade pending (0%)

cisco

NFM Switch Upgrade

Switchpool

- Home**
- Profiles**
- Racks**
- VRFs**
- Faults**

MANAGEMENT

- Upgrades**
- Images**

ADMINISTRATION

- Users**
- System**
- Account Settings**

GENERAL

Created	Last modified
1 minute ago (Today @ 3:07PM) by admin	1 minute ago (Today @ 3:07PM) by admin

SWITCHES

sw2 (LEAF42) using image 7.0(3)I2c S9 (nxos 7.0.3.I2c.b.. (version 1f1e8341-f57c-334fb25c-3c70a26559720)

Upgrades

Filter expression Sort by Name ▾

UPG to 2.2C

Name	Description	State	Strategy	Image	Members	Start time	End time
1. Click Upgrade OBJ		Inactive	Parallel	7.0(3)I2c S9 (nxos 7.0.3.I2c...	1	-	-

Upgrade UPG to 2.2C

OVERVIEW MEMBERS FAULTS

Upgrade Inactive

EDIT START STOP DELETE

2. Click Start

Filter expression Sort by ID ▾ Select all

Leaf, Spine

Upgrade UPG to 2.2C

OVERVIEW MEMBERS FAULTS

Leaf, Spine

Leaf / N9K-C9372PX

SW2 (LEAF42)

54 Interfaces (1:1) @10.0.8.42

No peer switch

Upgrade pending

You can remove member

Remove selected

Upgrades started UPG to 2.2C

NFM copies image to LEAF Then, initiates "install all nxos xxxx" on LEAF

SW2 (N9372-42) Upgrading... (25%)

3 % progress keep increasing until reaches 100%

cisco

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

NFM Fault Management

1. Red Alert box with no of Fault not resolved

2. Click then it shows current Faults

3. Click Fault No opens FAULT Window

Most of Faults are resolved automatically if problem is gone, like Port down but came up later

The screenshot shows the Cisco Nexus Fabric Manager interface. At the top, there's a navigation bar with 'Switchpool' selected. Below it, a 'Faults' section shows a single fault named 'fault107' on 'sw1 (SPINE41)'. A tooltip for this fault states: 'There was an unexpected failure when applying required configuration command 'copy running-config startup-config' with error message 'CLI execution error''. The 'GENERAL' tab of the fault details page is selected, showing 'Created' and 'Last modified' both as '1 day ago (Yesterday @ 9:21AM)'. The 'REMEDIAL ACTIONS ON FAULT107' section contains several options: 'TO RESOLVE THIS FAULT PLEASE CONSIDER THE FOLLOWING ACTIONS:', 'Resolve this fault and re-apply configuration', 'Manually override and resolve fault', 'Take action', and 'Take action'. A note says '4. You can ACK or take actions to resolve by click "Remedial Actions" tab'. Another note says '5. You can choose proper action'. The bottom of the page shows 'Severity' as 'Critical' and 'Related objects' as 'sw1 (SPINE41)'.

NFM EVPN KW and Collaboration Procedure

NFM KW

- **Technology: Data Center and Storage Networking**
 - **Sub-technology: VXLAN EVPN(Includes NFM/VTS support) → WW-RP**
 - Problem Code: Configuration Assistance
 - Problem Code: Error Messages, Logs, Debugs
 - Problem Code: Licensing → [WW-Licensing](#)
 - Problem Code: Interoperability
 - Problem Code: Product Feature/Function Question
 - Problem Code: Software Failure
 - **Sub-technology: NFM(Nexus Fabric Manager) Server → [WW-DC-Software](#)/[WW-NMS](#)/[AMM-NMS](#)**
 - Problem Code: Configuration Assistance
 - Problem Code: Data Corruption
 - Problem Code: Error Messages, Logs, Debugs
 - Problem Code: Licensing → [WW-Licensing](#)
 - Problem Code: Install, Uninstall, or Upgrade
 - Problem Code: Interoperability
 - Problem Code: Software Failure
- Provide a checklist for new KW to CIN team so they can route cases properly to avoid misQ

NFM Support Detail

- WW-DC-Software(NAM) handles
 - NFM Server Installation/Upgrade/Operation
 - Server failed to process request from GUI client or failed send change to VTEP
 - Connectivity issue between server and VTEP due to server issue
 - DC-Software verifies that NFM is not receiving traffic from TCPDUMP on NFM CLI
 - NFM Appliance Password Recovery
- RP handles
 - Initial setup with VTEP Discovery/Operation/EVPN Configuration from GUI
 - VTEP failed to process push change from GUI/Server
 - EVPN forwarding issue (host to host/server/external connectivity problem)
 - Connectivity issue between server and VTEP due to network issue
 - RP verifies network connectivity from VTEP to NFM server with ELAM, SPAN, TCPDUMP
 - VTEP Upgrade from NFM GUI
 - RP creates NFM Upgrade objects and start VTEP upgrade
 - RP checks if VTEP receive “copy scp: bootflash:” and “install all nxos ..” from NFM
 - If VTEP doesn’t receive these CLI, open collab with NFM Srv team
 - AFP (Auto Fabric Provisioning- POAP)
 - RP check DHCP process from VTEP console output/dhcp log and check DHCPD log on NFM
 - Open collab if NFM doesn’t create DHCP OFF/ACK received DHCP DISC/REQ

NFM GUI Basic Troubleshooting

1. **NFM GUI is not reachable**
2. RP checks if NFM appliance is pingable
3. **NFM is not pingable, RP checks basic network connectivity then open collab with DC-Software**
 - RP CSE can check NFM Appliance IP addresses are correct and interface is up
 - RP CSE can run nmtui to reconfigure ip address for NFM
 - RP CSE open collab with DC-Software for any linux network configuration issue
4. **NFM is pingable but NFM GUI is not reachable, RP CSE ssh to NFM and check ESM service status**

```
# service esm status
* Wrapper process is running with PID=13630
* Elastic Services Manager process is running with PID=13652
[ ok ]
```
5. **ESM is not running, try to restart ESM**

```
# service esm restart
* Stopping dhcpcd ...
* Stopping esm ...
* Starting esm ...
* Running: - [ ok ]
```
6. **Open collab with DC-Software if GUI is still not reachable after ESM restarted**

Auto Fabric Provision Troubleshooting

1. RP check console output regarding POAP process

N9372-42# reload

This command will reboot the system. (y/n)? [n] y

```
2016 Apr 15 05:43:11 N9372-42 %$ VDC-1 %$ %PLATFORM-2-PFIM_SYSTEM_RESET: Manual system restart from Command Line Interface
2016 Apr 15 05:45:54 switch %$ VDC-1 %$ %ASCII-CFG-2-CONF_CONTROL: System ready
Starting Auto Provisioning ...
2016 Apr 15 05:45:59 switch %$ VDC-1 %$ %POAP-2-POAP_INITED: POAP process initialized
2016 Apr 15 05:46:20 switch %$ VDC-1 %$ %POAP-2-POAP_DISCOVER_START: POAP DHCP Discover phase started
2016 Apr 15 05:46:49 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Using DHCP, information received over mgmt0 from 10.0.8.27
2016 Apr 15 05:46:49 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Assigned IP address: 10.0.8.42
2016 Apr 15 05:46:49 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Netmask: 255.255.255.0
2016 Apr 15 05:46:49 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Default Gateway: 10.0.8.254
2016 Apr 15 05:46:49 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Script Server: 10.0.1.27
2016 Apr 15 05:46:49 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Script Name: poap.py
2016 Apr 15 05:48:23 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: The POAP Script download has started
```

2. If POAP failed with DHCP failure, RP CSE SSH to NFM

```
2016 Apr 15 06:58:05 switch %$ VDC-1 %$ %POAP-2-POAP_DISCOVER_START: POAP DHCP Discover phase started
2016 Apr 15 06:58:13 switch %$ VDC-1 %$ %POAP-2-POAP_FAILURE: POAP DHCP discover phase failed
```

Check DHCPD log on NFM and see if NFM received DHCP DISC or not.

If DHCP DISC/ACK received but NFM doesn't create DHCP OFF or ACK, open collab with NFM

```
# tft /var/log/dhcpd.log
Apr 14 04:22:31:25 root@NFM-1-1-1-258 ~
# tft /var/log/dhcpd.log
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: DHCPDISCOVER from 88:1d:fc:ca:3f:48 via eth1
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: DHCPOFFER on 10.0.8.42 to 88:1d:fc:ca:3f:48 via eth1
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: Found client uid SAL18495M25 len 12
Apr 14 22:32:22 NFM-1-1-1-258 dhcpd[14329]: Cisco vendor callout using /opt/esm/server/bin/ciscoDhcpPlugin.py SAL18495M25 true
Apr 14 22:32:40 NFM-1-1-1-258 dhcpd[14329]: DHCPREQUEST for 10.0.8.42 (10.0.8.27) from 88:1d:fc:ca:3f:48 via eth1
Apr 14 22:32:40 NFM-1-1-1-258 dhcpd[14329]: DHCPCACK on 10.0.8.42 to 88:1d:fc:ca:3f:48 via eth1
```

NFM Fabric Upgrade Troubleshooting

1. RP upload images to NFM and create “upgrade obj” on NFM for switches

Upgrades							
Filter expression		Sort by		Name ▾			
Name ▾	Description	State	Strategy	Image	Members	Start time	End time
UPG to 2.2C	-	Inactive	Parallel	7.0(3)I2(2c)S9 (nxos.7.0(3)I2.2c....)	1	-	-

2. RP starts VTEP upgrade from NFM GUI
3. If NFM GUI shows upgrade is not started, then RP CSE check if VTEP’s “show accounting log” and “dir” to see if VTEP received “copy scp: bootflash:” and “install all nxos ..” from NFM
4. Open collab with DC-Software if VTEP didn’t receive these commands from NFM and Upgrade
5. If VTEP start Upgrade but failed or crashed, RP CSE will open collab with DCSW or CATS
6. If VTEP is upgraded and reloaded but NFM Upgrade OBJ is not updated correctly or hang, RP CSE open collab with DC-Software

Virtual Topology System

VTS Discovery

The screenshot shows a web browser window with the following details:

- Address Bar:** https://10.0.1.26:8443/VTS/resetpassword.htm
- Toolbar:** Includes back, forward, search, and other standard browser icons.
- Page Content:**
 - Most Visited:** VC, NFM, CCO, DW, Cisco Secure ACS Login, Cisco Virtual Topology...
 - VTS Discovery Logo:** Cisco Virtual Topology System
 - User Information:** Logged in as: admin
 - Action Buttons:** Cancel (outline), Change Password (outline).

Please change your password after first time login.

* Current Password	<input type="text"/>
* New Password	<input type="text"/> ?
* Confirm Password	<input type="text"/> ?

Cancel

Change Password

VTS Discovery

ual Topology System

Logged in as: admin

Welcome to Virtual Topology System

The Cisco Virtual Topology System is a standards-based, open software-overlay management and provisioning system. It automates data center network fabric provisioning for virtual and physical infrastructure.

The system offers a rich set of automation and programmability capabilities across the entire Cisco Nexus switching portfolio to enhance the deployment of cloud-based services.

Guide me to
setup

I will setup
myself

Get Started

VTS Discovery

Virtual
Topology
System



Logged in as: admin



System Settings

Cisco VTS enables you to define global system wide parameters for enabling flexible overlays on the data center fabric.

Domain ID



Domain ID specifies the BGP site-of-origin attribute

DHCP Server IP



DHCP Server IP can take a valid IP4 address. A valid IP4 octet contains no more than three digits. Each octet has to be in range 0-255. The only valid IP4 address starting with 0 is 0.0.0.0

* AnyCast GW Mac



AA:BB:CC:DD:EE:FF

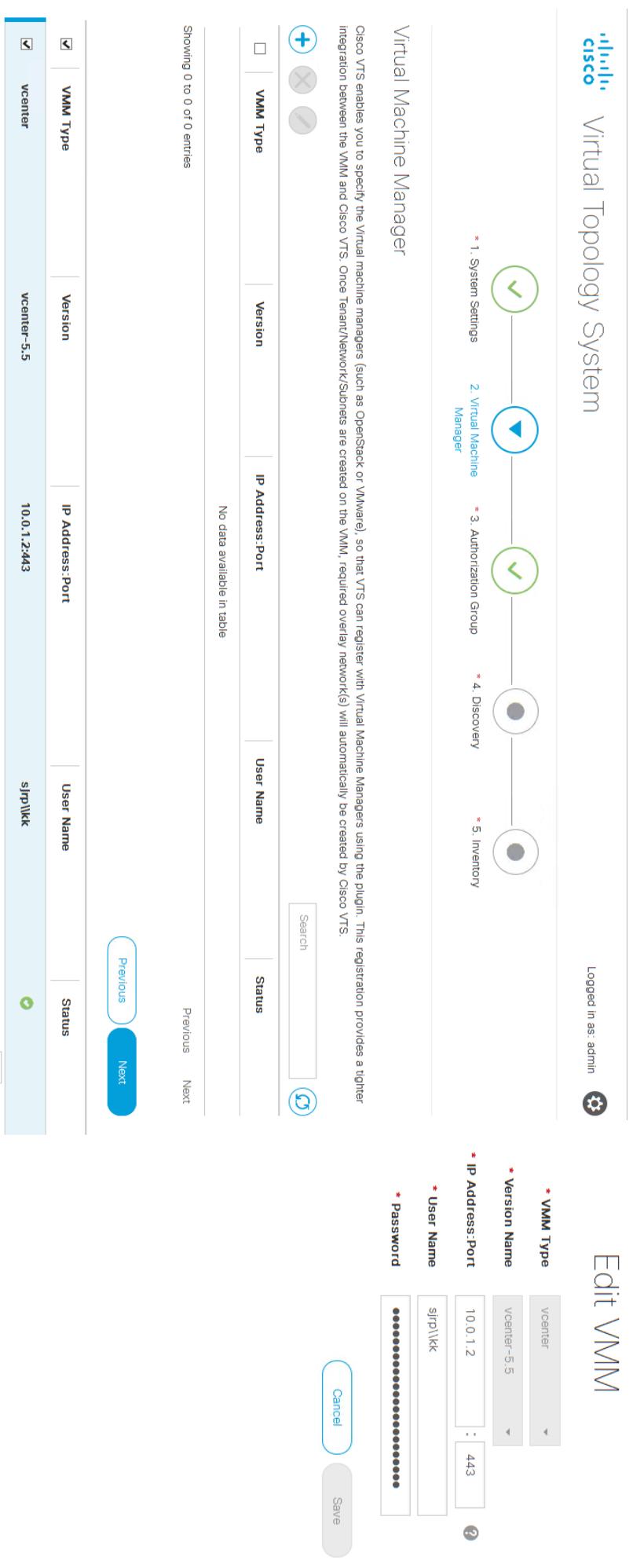
Cancel

Submit

Previous

Next

VTS Discovery



VTS GUI Initial Setup

Virtual Topology System

Logged in as: admin

1. System Settings 2. Virtual Machine Manager 3. Authorization Group 4. Discovery 5. Inventory

Edit User

* Auth Group Name: default

* Controller User Name: admin

* Device User Name: admin

* Password:

Authorization Group

Cisco VTS enables you to create authorization groups and assign devices you import into Cisco VTS, to these groups. Authorization groups are used to group devices with the same credentials (i.e. usernames and passwords). Once the authorization groups are created, all the devices under these groups may be accessed without specifying the credentials every time they are accessed. If the same credentials are used for accessing all devices, one authorization group can be used.

Auth Group Name	Controller User Name	Device User Name
<input type="checkbox"/> default		

Previous **Next**

VTS GUI Initial Setup

VTS Discovery

cisco

DCNM N9K POAP

Enter comma separated Serial Numbers or Serial Number/VDCA identifiers. Examples: 123456, 123457 or 456795\VDCA1, 456795\VDCA2

* Switches: Total 1

Import from CSV File...

Note: The username/password is used by DCNM only to manage the switch and DCNM will not create the entered username/password in the switch.

* Add Switches to Group:

* Switch User Name:

* Switch Password:
 Show password in clear text

Config Server:

< Back Next > Cancel

DCNM N9K POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces
Switch Role leaf *	The role of the switch, e.g. leaf, spine						
Link-State Routing Protocol ospf *	Select link-state routing protocol						
Net <input type="text"/>	Network Entity Title for I-S-I-S [example:49.0001.####.##.##.00], NET will be auto-generated if not provided						
Enable BFD <input type="checkbox"/>	TRUE if Bidirectional Forwarding Detection should be enabled						
General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces
Loopback0 IP 10.0.0.43 *	[example: 11.11.11.10.20]						
BGP AS # 65000 *	BGP Router Autonomous System Number						
Route Reflector IP 10.0.0.41 *	IP Address of Route Reflector						
Second Route Reflector IP <input type="text"/>	IP Address of Second Route Reflector						
General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces
Anycast RP Address 10.0.0.41 *	Anycast RP Address for PIM Multicast						
RP Group <input type="text"/> *	RP Group max 256 addresses [example: 239.239.239.0/24]						

DCNM N9K POAP

Template Parameters - 1 Switches : SAL19017CEU

	General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces	
Anycast-Gateway-MAC	<input type="text" value="2020.0000.00AA"/>	*							
Shared MAC address for all leaves									
Enable VMTracker	<input checked="" type="checkbox"/>	Enable Autoconfig vmtracker feature							
vCenter IP Address	<input type="text" value="10.0.1.2"/>	IP Address of the vCenter Server							
vCenter Port	<input type="text" value="80"/>	vCenter Server port							
vCenter Username	<input type="text" value="sjp\kk"/>	vCenter query user (needed for secure vmtracker)							
vCenter Password	<input type="text" value="kk"/>	Password to access vCenter for Auto-Config Queries (needed for vmtracker)							
Enable Auto-Pull	<input checked="" type="checkbox"/>	Enable Autoconfig auto-pull feature							
Global Mobility Domain	<input type="text" value="md0"/>	Fabric-wide mobility domain string							
System Dynamic VLANs	<input type="text" value="2500-3500"/>	*	System Fabric Dynamic VLANs						
Core Dynamic VLANs	<input type="text" value="2500-2999"/>	*	Core Fabric Dynamic VLANs [subset of System Fabric Dynamic VLANs]						

DCNM N9K POAP

General	Manageability	Fabric	Backbone/BGP	Multicast	Gateway	VPC-Peer	Interfaces
<input checked="" type="checkbox"/> Enable VPC	true if VPC should be configured						
VPC Domain ID 434	Should be the same for both VPC Peers [example: 10,10,11,11,12,12]						
VPC Peer IP 10.1.1.1	[example: 10.10.10.11,10.10.10.10]						
VPC Peer Channel 434	VPC Peer link port channel number						
VPC Peer Interfaces e1/1-2	A comma and dash separated list of interfaces, e.g. e6/6-12,e7/7 to be used for the VPC Peer Link						
VPC Peer Link VLAN 99	VLAN outside the dynamic range for use in IP Fabric only						
VPC Peer Link VLAN IP 10.1.0.99	For use in IP Fabric only						
VPC Peer Link VLAN IP Prefix 30	For use in IP Fabric only						
Loopback1 IP 10.0.0.97	*	[example: 12.11.11.10-20]					
Loopback1 secondary IP 10.0.0.100	Must be set the same for VPC Peers [example: 13.11.11.10-20]						

DCNM N9K POAP

cisco

DCNM N9K SPINE POAP

Template Parameters - 1 Switches : FOX1844GP9Z

General	Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
Switch Role <input type="text" value="spine"/> *					
The role of the switch. e.g. leaf, spine					
Link-State Routing Protocol <input checked="" type="radio" value="ospf"/> * <input type="radio" value="isis"/> *					
Select link-state routing protocol					
Net <input type="text"/>					
Network Entity Title for IS-IS [example:49.0001.####.####.###]					
TRUE if Bidirectional Forwarding Detection should be enabled					
Enable BFD <input type="checkbox"/>					

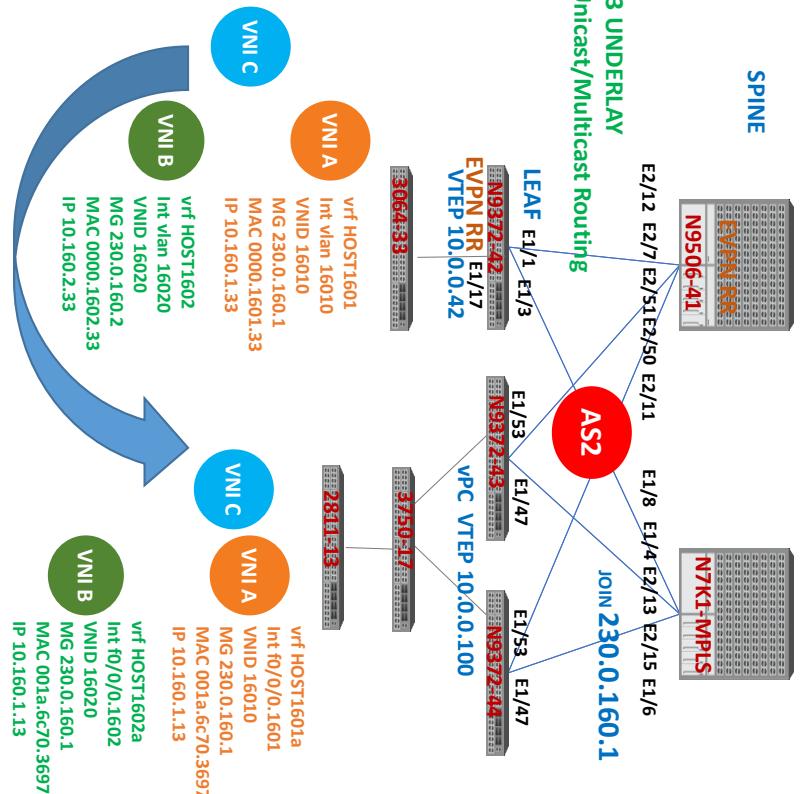
Template Parameters - 1 Switches : FOX1844GP9Z

General	Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
Loopback IP <input type="text" value="10.0.0.41"/> *					
[example: 11.11.11.10-20]					
Loopback Prefix <input type="text" value="32"/> *					
Loopback Prefix					
BGP Client Subnet <input type="text" value="10.41.0.0/16"/> *					
Backbone Subnet [example: 11.11.11.0/24]					
BGP AS # <input type="text" value="65000"/> *					
BGP Router Autonomous System Number					
Route Reflector IP <input type="text" value="10.0.0.41"/> *					
IP Address of Route Reflector					
Second Route Reflector IP <input type="text"/>					
IP Address of Second Route Reflector					

General	Manageability	Fabric	Backbone/BGP	Multicast	Interfaces
Anycast RP Address <input type="text" value="10.0.0.100"/> *					
Anycast RP Address for PIM Multicast					
RP Group <input type="text" value="227.2.2.0/24"/> *					
RP Group max 256 addresses [example: 239.239.239.0/24]					
RP Array <input type="text"/>					
Populated from Admin Fabric Encapsulation Settings					
Unused Interfaces <input type="text"/>					
A list of unused ports [example: e1/5,e2/3-5]					

VXLAN/EVPN Update

VXLAN EVPN – IBGP



```
ON SPINE41 - RR
router bgp 2
router-id 10.0.0.41
log-neighbor-changes
address-family ipv4 unicast
template peer vtep-peer
remote-as 1
update-source loopback0
address-family ipv4 unicast
send-community ext
route-reflector-client
address-family l2vpn evpn
send-community ext
route-reflector-client
neighbor 10.0.0.43
inherit peer vtep-peer
neighbor 10.0.0.44
inherit peer vtep-peer
```

Address-family ipv4 unicast for default routing IP service

ON LEAF43 - RR Client
router bgp 2
router-id 10.0.0.43
log-neighbor-changes
address-family ipv4 unicast
neighbor 10.0.0.41 remote-as 1
update-source loopback0
address-family ipv4 unicast
address-family l2vpn evpn
send-community extended
neighbor 10.0.0.42 remote-as 1
update-source loopback0
address-family ipv4 unicast
address-family l2vpn evpn
send-community extended
vrf evpn-tenant-kk1
address-family ipv4 unicast
advertise 12vpn evpn

All you really need on LEAF is L2VPN EVPN AF and send-community ext

Under address-family ipv4 unicast of each tenant
VRF instance, enable advertising EVPN routes

vrf HOST1601 int vlan 16010 VNIID 16010 MG 230.0.160.1 MAC 0000.1601.33 IP 10.160.1.33	vrf HOST1601a int f0/0/0.1601 VNIID 16010 MG 230.0.160.1 MAC 0001a.6c70.3697 IP 10.160.1.13	vrf HOST1602a int f0/0/1.1602 VNIID 16020 MG 230.0.160.2 MAC 0000.1602.33 IP 10.160.2.33
----------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------

VXLAN EVPN – eBGP Unique AS

All LEAFs has unique AS NO
ON SPINE41 - RR

```
route-map rm-nh-unchanged permit 10
  set ip next-hop unchanged
```



```
router bgp 2
  router-id 10.0.0.41
  address-family ipv4 unicast
    next-hop route-map rm-nh-unchanged
      retain route-target all
```

```
neighbor 10.0.0.43 remote-as 43
  update-source loopback0
  address-family ipv4 unicast
  send-community ext
  route-reflector-client
  address-family 12vpn evpn
  send-community ext
  route-reflector-client
```

route-map rm-nh-unchanged out

Set route policy filtering for next-hop not to be changed

Retain routes with RT when advertise to eBGP
Make ASBRs to import all route-targets not locally defined in Inter-AS

out

On LEAF43 – RR Client

```
router bgp 43
  router-id 10.0.0.43
  address-family ipv4 unicast
    neighbor 10.0.0.41 remote-as 1
    update-source loopback0
```

No difference in LEAF BGP configuration

Set NEI RM to keep next-hop unchanged

```
vrf HOST1601
  Int vlan 16010
  VNID 16010
  MG 230.0.160.1
  MAC 0000.1601.33
  IP 10.160.1.33

vrf HOST1602
  Int vlan 16020
  VNID 16020
  MG 230.0.160.2
  MAC 0000.1602.33
  IP 10.160.2.33

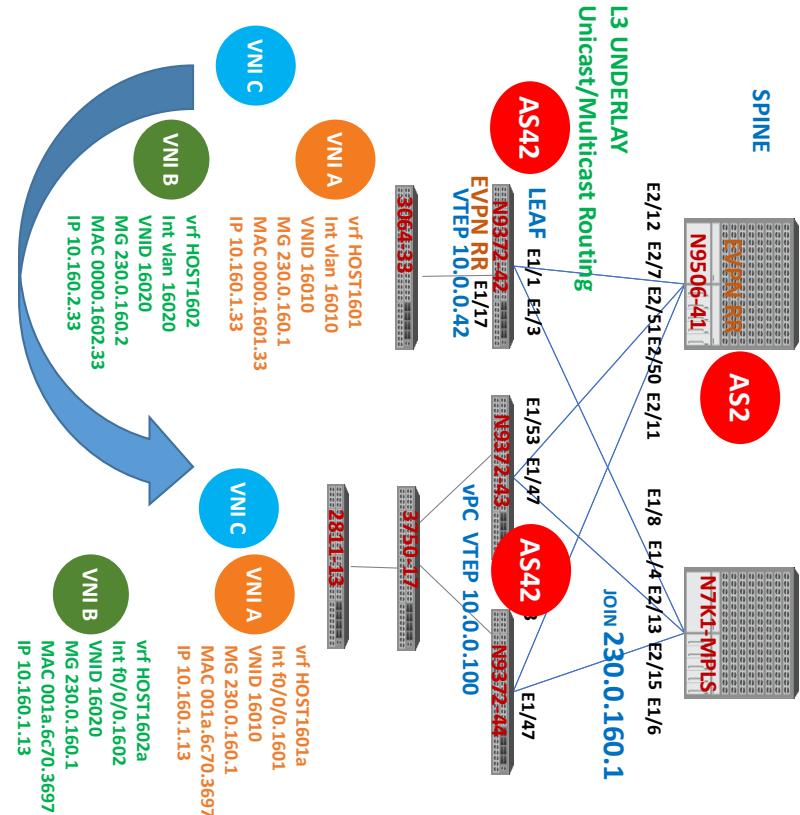
vrf HOST1602a
  Int fo 0/0.1602
  VNID 16020
  MG 230.0.160.1
  MAC 001a.6c70.3697
  IP 10.160.1.13
```

```
vrf evpn-rr
  address-family 12vpn evpn
    send-community extended
  vrf evpn-tenant-kk1
    address-family ipv4 unicast
    advertise 12vpn evpn
```

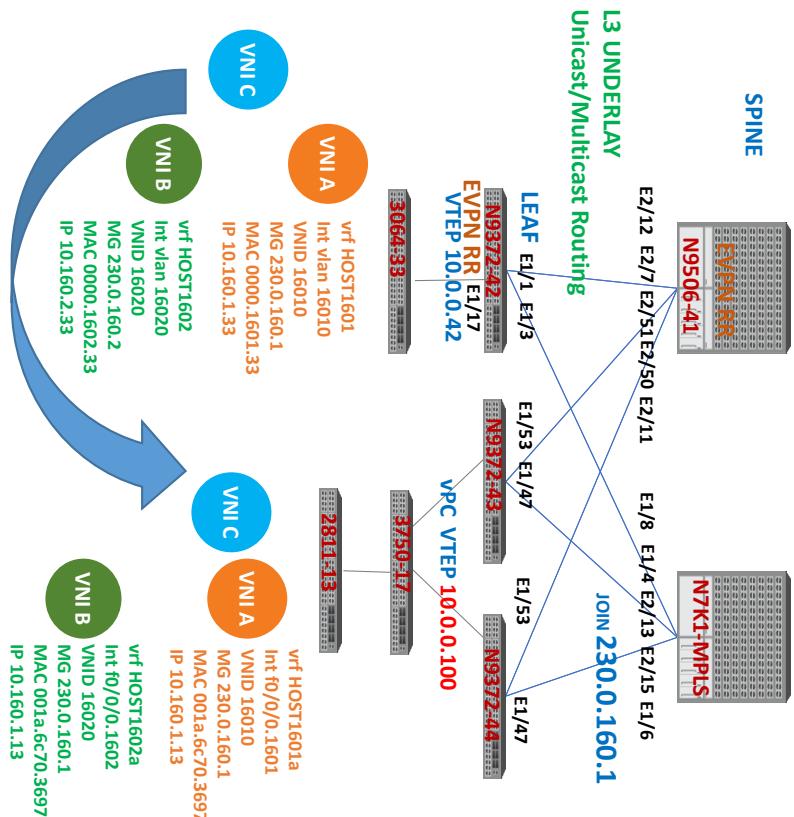
route policy filtering for next hops

- BGP next-hop filtering allows you to specify that when a next-hop address is checked with RIB
- If the route map rejects the route, the next-hop address is treated as unreachable.

VXLAN EVPN – eBGP Same LEAF AS



VXLAN EVPN – vPC VTEP

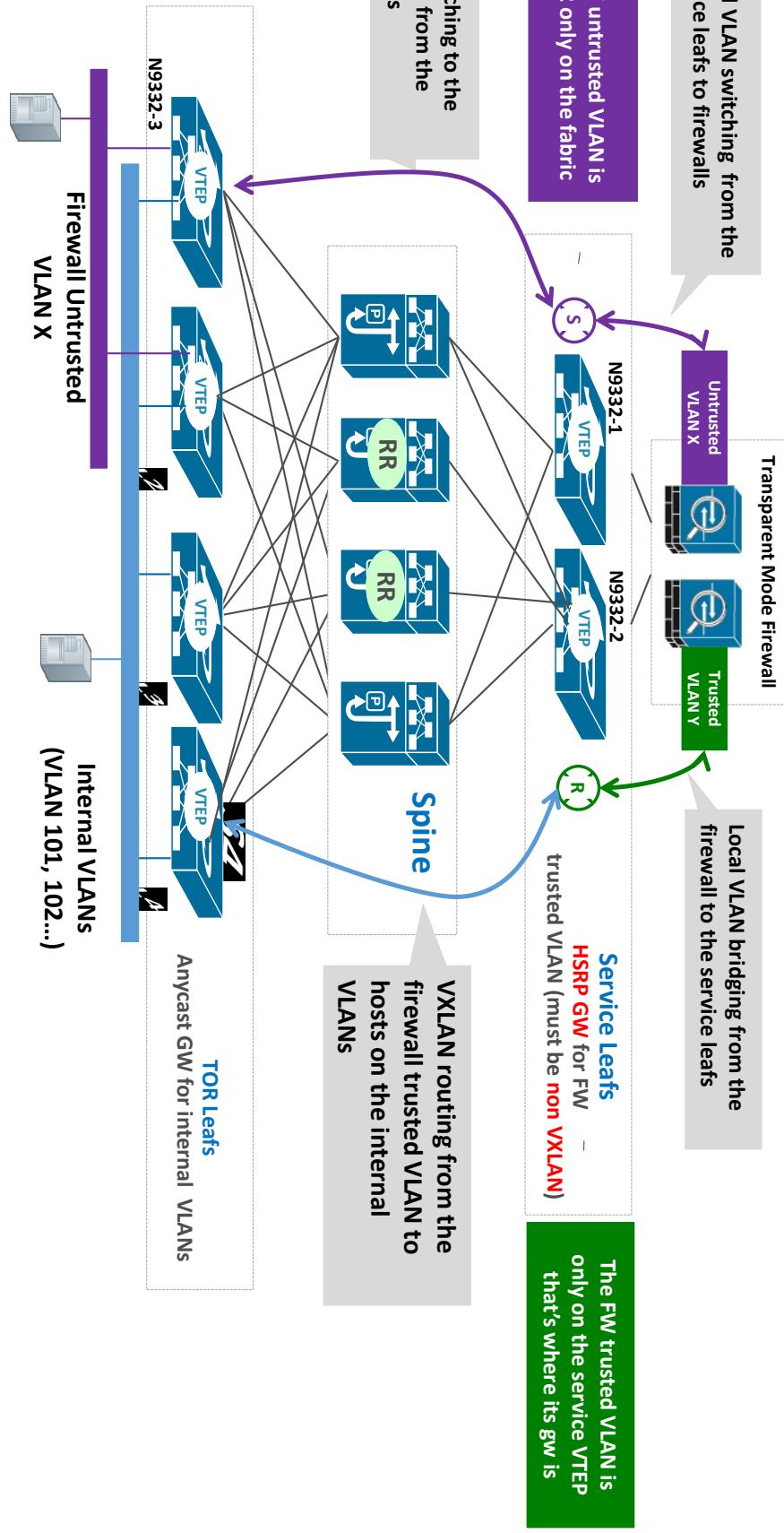


- Bind NVE to the loopback int that is not used for underlay routing protocols
 - VPC put NVE Source Loopback on VPC Secondary SW into Admin Shut when peer-link is shut
 - This will cause underlay routing issue if NVE uses underlay routing protocol loopback as source
 - VPC Peer must have
 - Consistent VLAN to VNI-seg mapping
 - Consistent NVE 1 binding to same loopback interface
 - Same secondary IP
 - Different Primary IP
 - Consistent VNI to multicast group mapping
 - vPC Recommendation
 - peer-gw, peer-sw, ip arp sync
 - Increase STP hello timer to 4 sec
 - Underlay Backup SVI with PIM enable
 - Overlay backup SVI with PIM enabled if DHCP relay is required
- On LEAF 43**
- ```
interface nve1
no shutdown
source-interface loopback1
```
- On Leaf44**
- ```
interface nve1
no shutdown
source-interface loopback1
```
- © 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

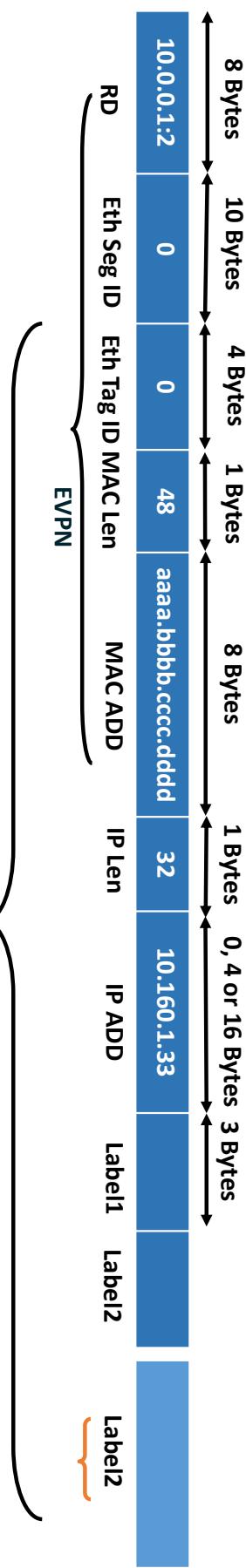
Default GW Support on VXLAN F&L and EVPN

- **VXLAN F&L**
 - Before 7.0(3)I2, VXLAN F&L can't have SVI on VTEP
 - Must use external router as default GW
 - After 7.0(3)I2, VXLAN F&L only support "Centralized GW"
 - Recommend SVI on a pair of vPC devices with FHRP
 - **Don't ever create SVI for same vlan/vxlan on any other VTEP**
 - Only one SVI can exist per VLAN on a pair of vPC VTEP or standalone VTEP
- **EVPN**
 - Recommended to use the anycast gateway feature on all VTEPs
 - Anycast gateway is configured for a specific VNI, all VTEPs must configure anycast GW for that VNI

Transparent Firewall Insertion to VXLAN EVPN Fabric



MPLS VPN Control Plane



MP-iBGP update with RD, RT, and label

- The Label (for the VPNv4 prefix) is assigned only by the PE whose address is the next-hop attribute
 - PE routers rewrite the next-hop with their own address (loopback)
 - “Next-hop-self” towards MP-iBGP neighbors by default
- PE addresses used as BGP next-hop must be uniquely known in the backbone IGP
 - **Do Not Summarize the PE Loopback Addresses in the Core**