The bridge to possible

Сетевой марафон Cisco:Классика LAN

День 5. Сетевые фабрики «сделай сам»: VXLAN EVPN

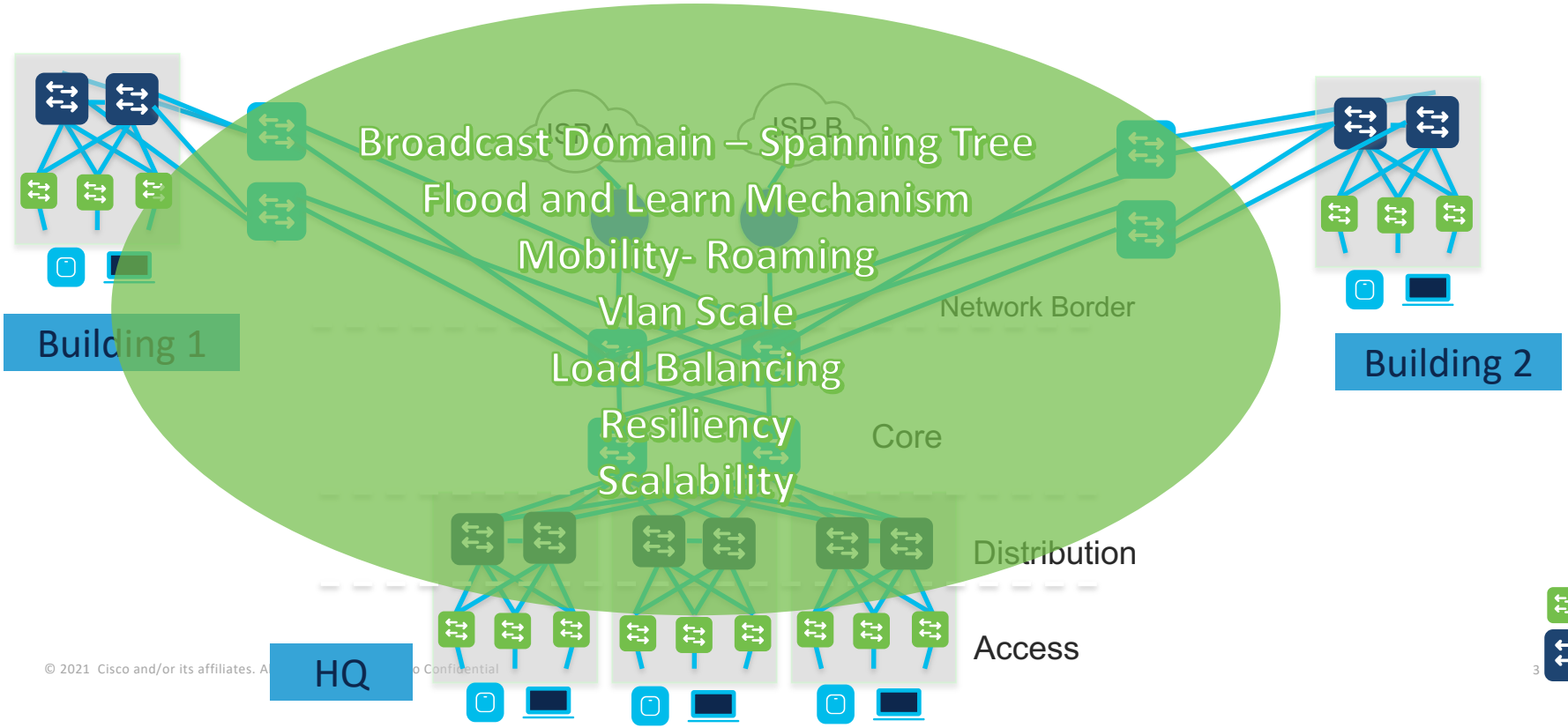
Михаил Окунев
Системный Архитектор
26 марта 2021

Agenda

- Introduction to Campus Fabrics
- VXLAN with BGP EVPN
  - Overview
  - Underlay
  - Control & Data Plane
  - Multi-Tenancy
  - StackWise Virtual Redundancy
  - *Tenant Routed Multicast (TRM)*
  - External Connectivity Options
  - *Services*
- *Underlay Designs and Configurations*

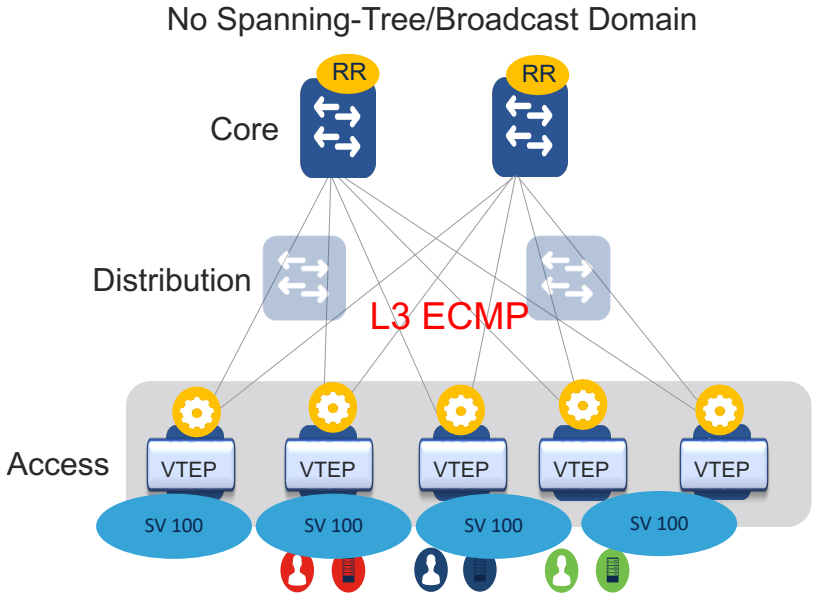
© 2021 Cisco and/or its affiliates. All rights reserved. Cisco Public

Typical Campus Network Challenges



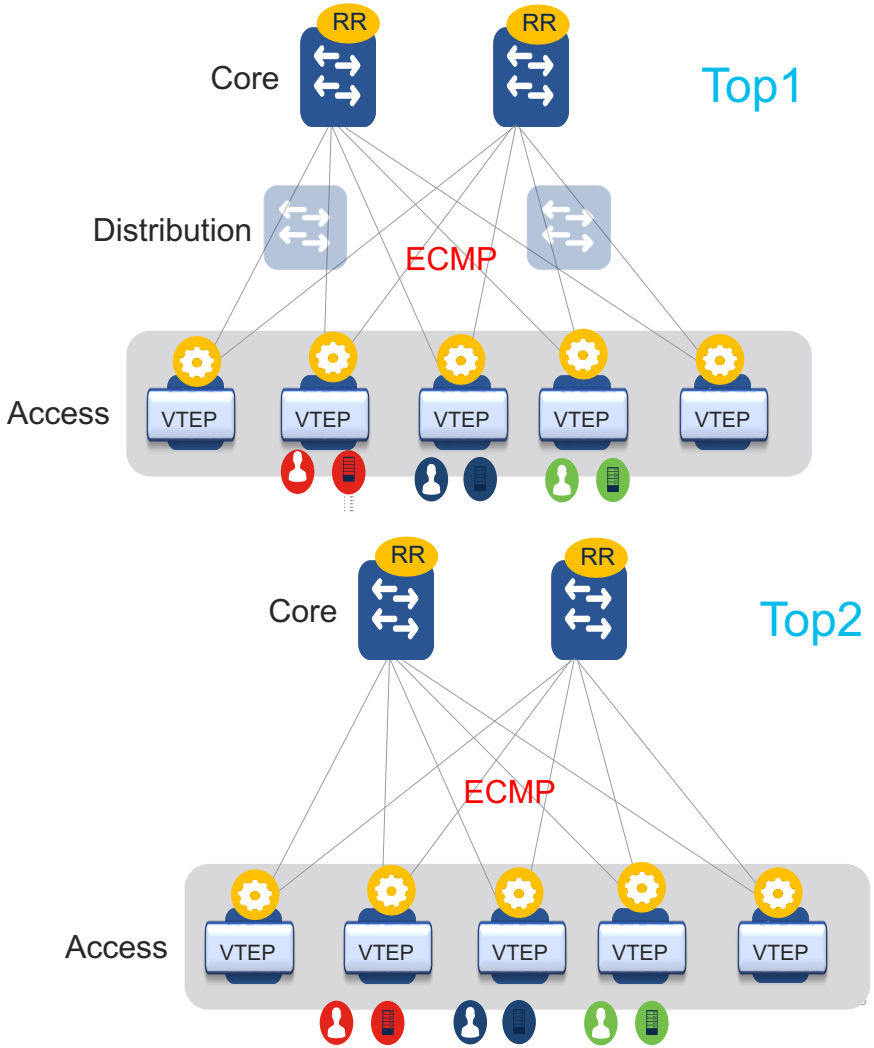
Why EVPN?

- Any subnet, anywhere, rapidly
- Layer 3 ECMP Links End-End
- No Flooding with BGP Control Plane
- Extensible Scale & Resiliency
- Distributed Gateway on all End Point Switches



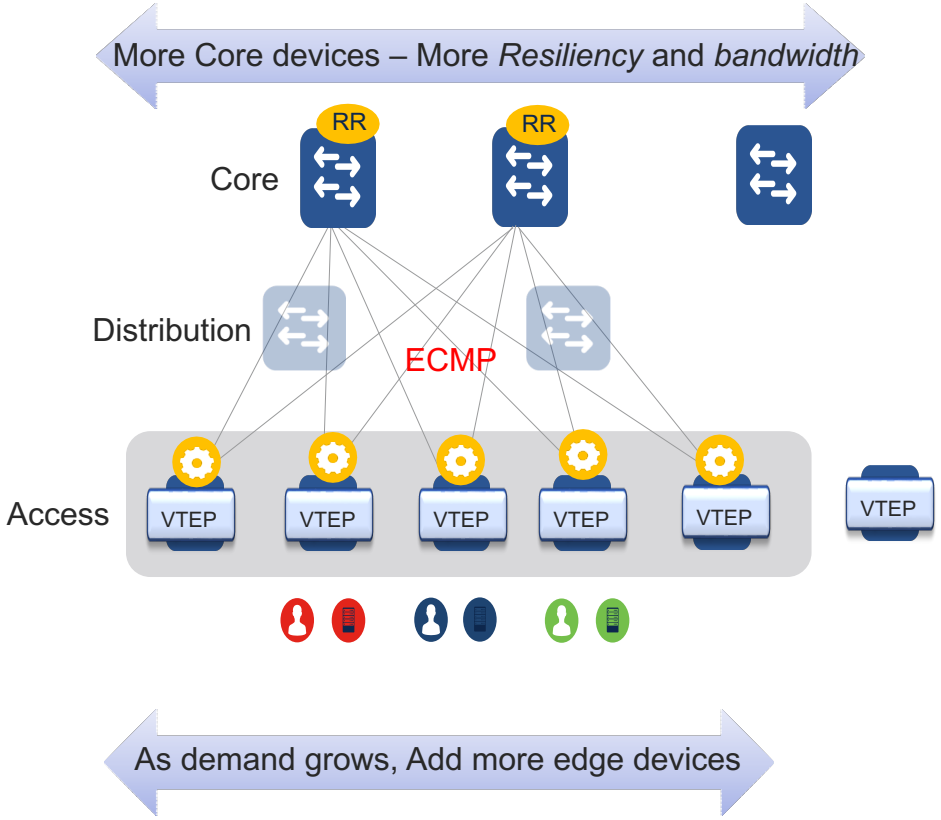
Flexible Topologies

- High Bi-Sectional Bandwidth
- Wide ECMP: Unicast or Multicast
- Uniform Reachability, Deterministic Latency
- High Redundancy: Node/Link Failure
- Line rate, low latency, for all traffic



Fabric Scalability

- Fabric size: Hundreds to 10s of Thousands of 1/10G ports
- Variety of Building Blocks:
  - Varying Size
  - Varying Capacity
  - Desired oversubscription
  - Modular and Fixed
- Scale Out Architecture
  - Add Edge and core devices, external connectivity as the demand grows



VXLAN with BGP EVPN:
Overview

Understanding Overlay Technologies

Overlay Services

- Layer 2
- Layer 3
- Layer 2 and Layer 3

Tunnel Encapsulation

Underlay Transport Network

Control Plane

- Peer Discovery mechanism
- Route Learning and Distribution
  - –Local Learning
  - –Remote Learning

Data Plane

- Overlay Layer 2/Layer 3 Unicast traffic
- Overlay Broadcast, Unknown Unicast, Multicast traffic (BUM traffic) forwarding
  - –Multicast

Why VXLAN?

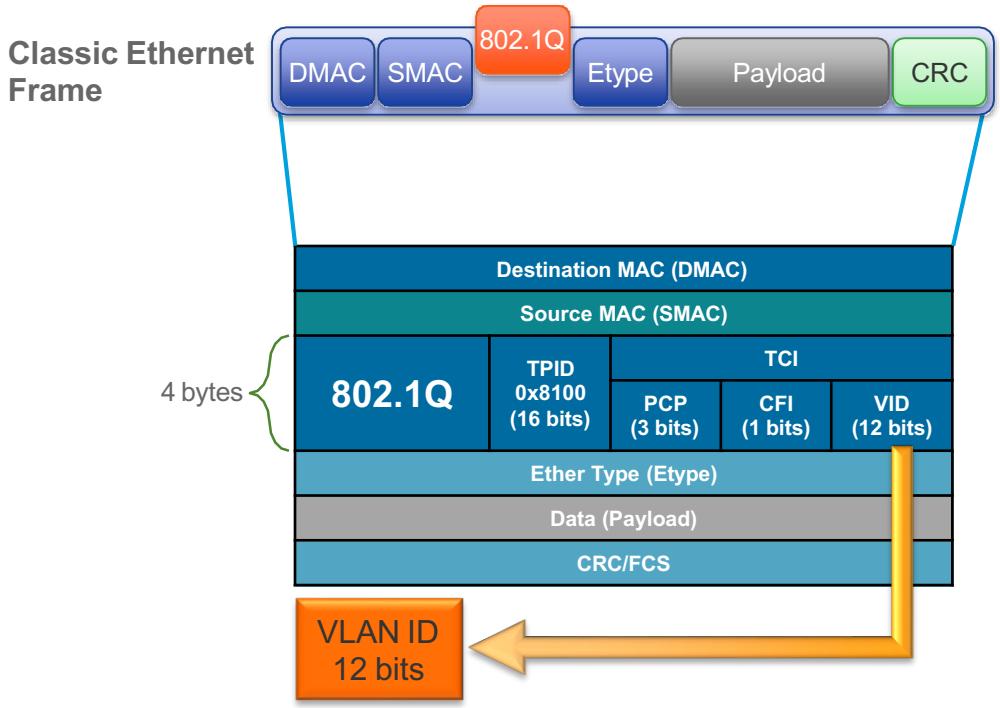
VXLAN provides a Network with Segmentation, IP Mobility, and Scale

- "Standards" based Overlay (RFC 7348)
- Leverages Layer-3 ECMP – all links forwarding
- Increased Name-Space to 16M identifier
- Integration of Physical and Virtual
- It's SDN ☺

Overview

Classic Ethernet IEEE 802.1Q Frame Format

- Traditionally VLAN is expressed over 12 bits (802.1Q tag)
  - Limits the maximum number of segments in a Campus to 4096 VLANs



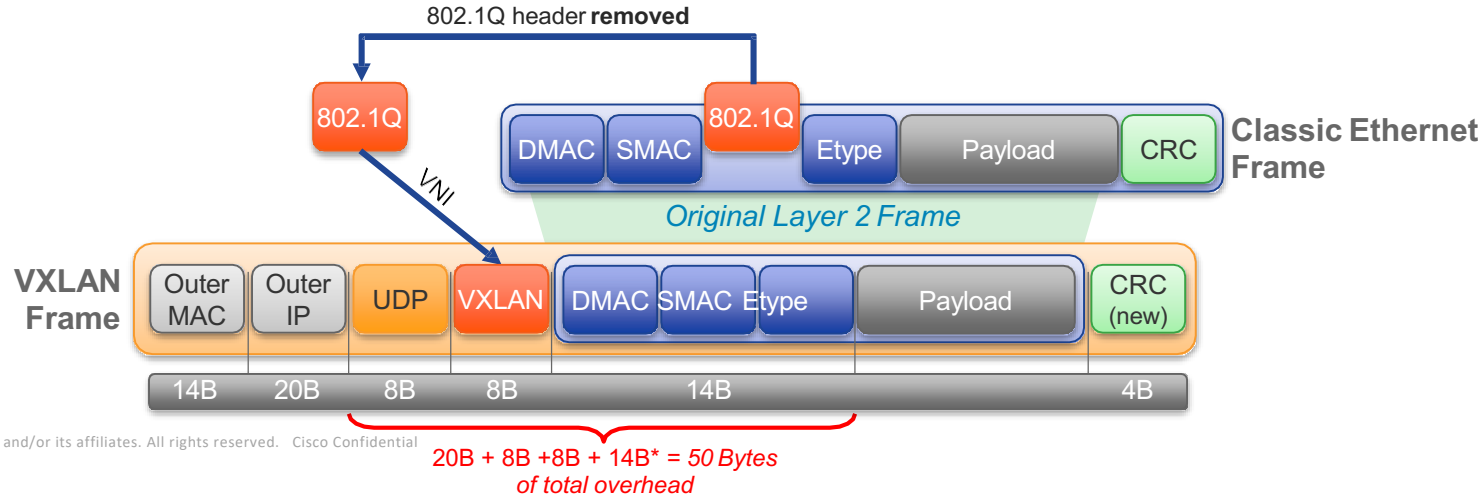
TPID = Tag Protocol Identifier, TCI = Tag Control Information,
PCP = Priority Code Point,
CFI = Canonical Format Indicator, VID = VLAN Identifier

Overview

Introducing VXLAN

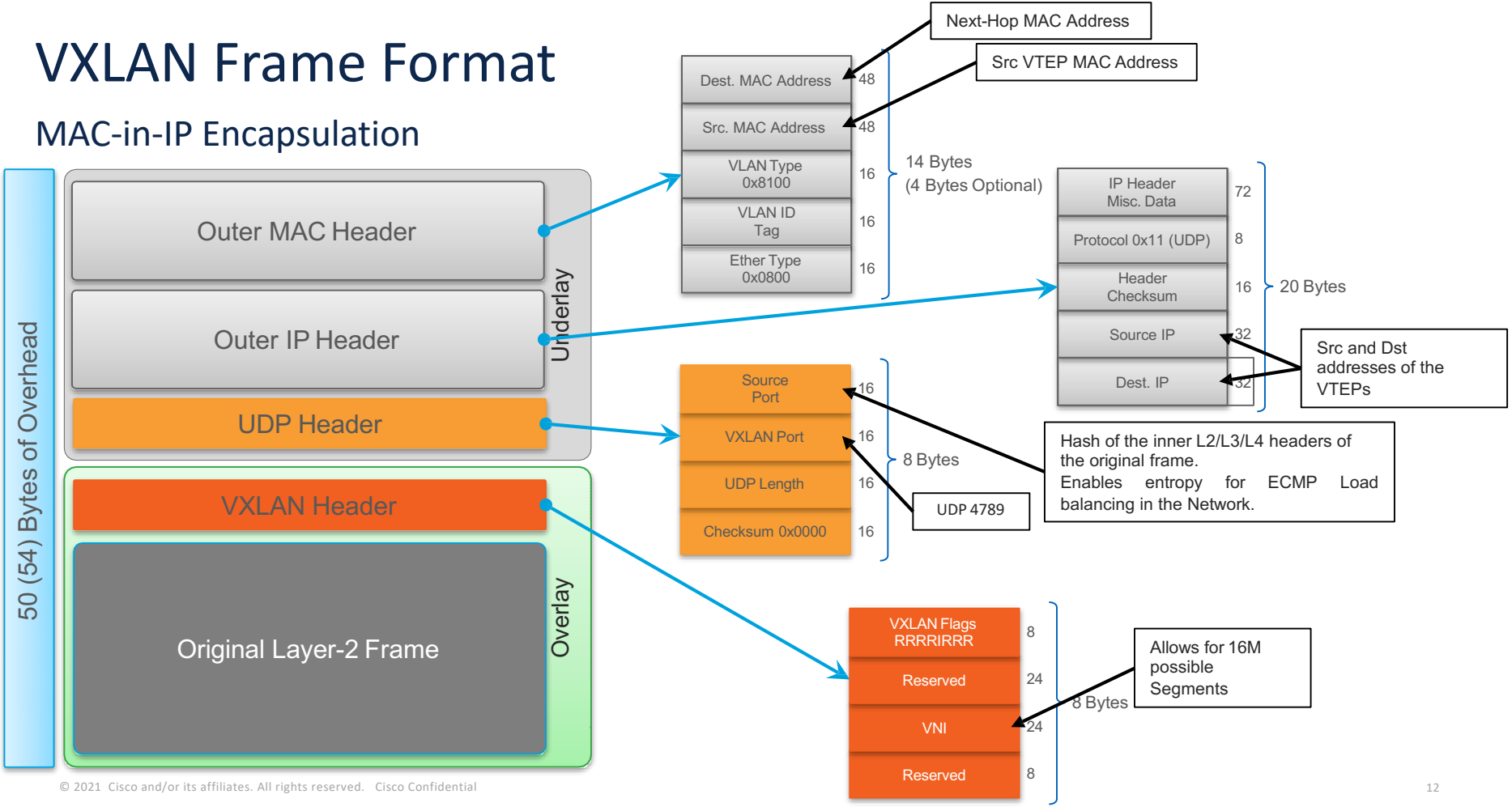
- Traditionally VLAN is expressed over 12 bits (802.1Q tag)
  - Limits the maximum number of segments in a Campus to 4096 VLANs

- VXLAN leverages the VNI field with a total address space of 24 bits
  - Support of ~16M segments
- The VXLAN Network Identifier (VNI/VNID) is part of the VXLAN Header

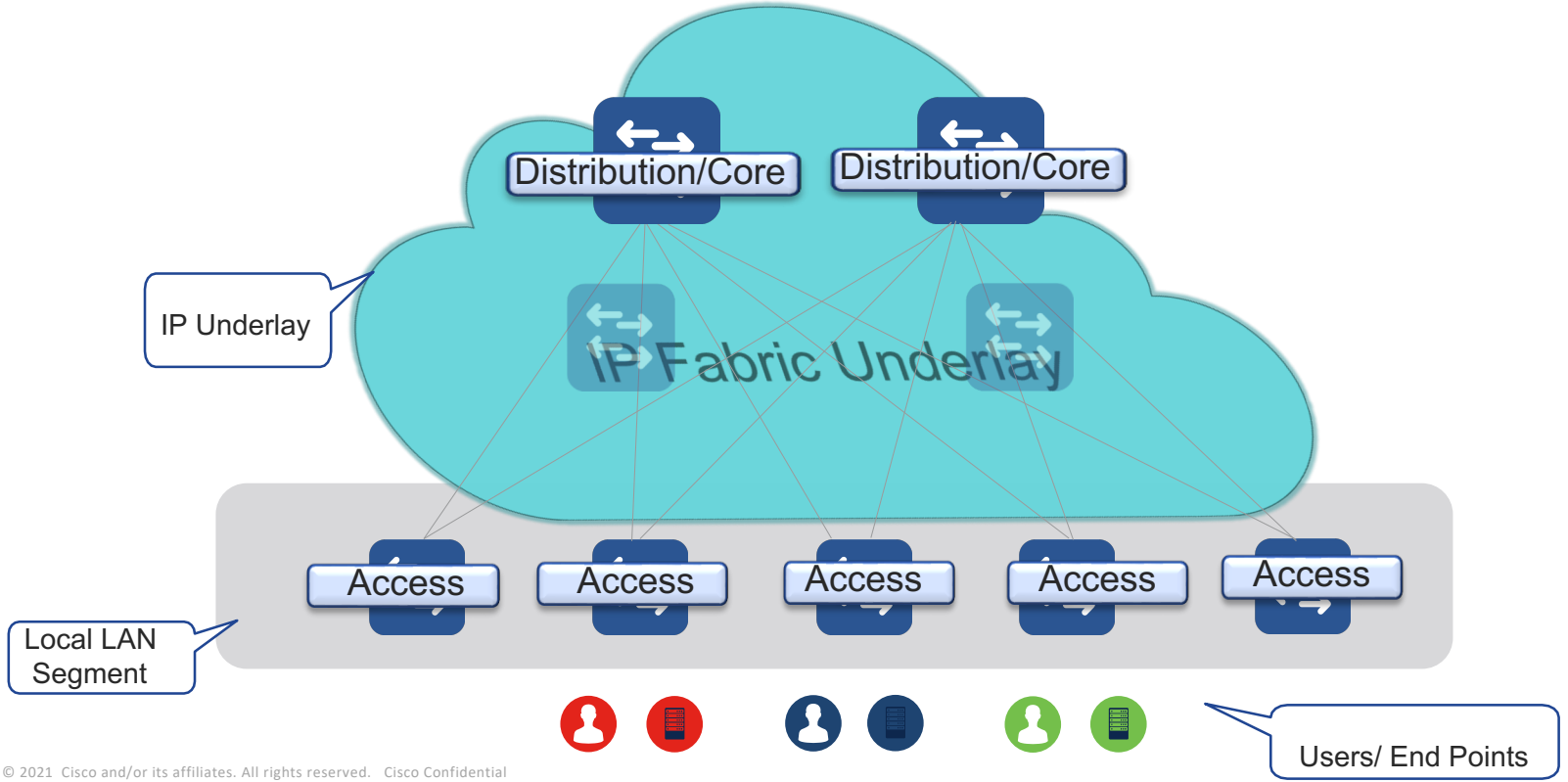


VXLAN Frame Format

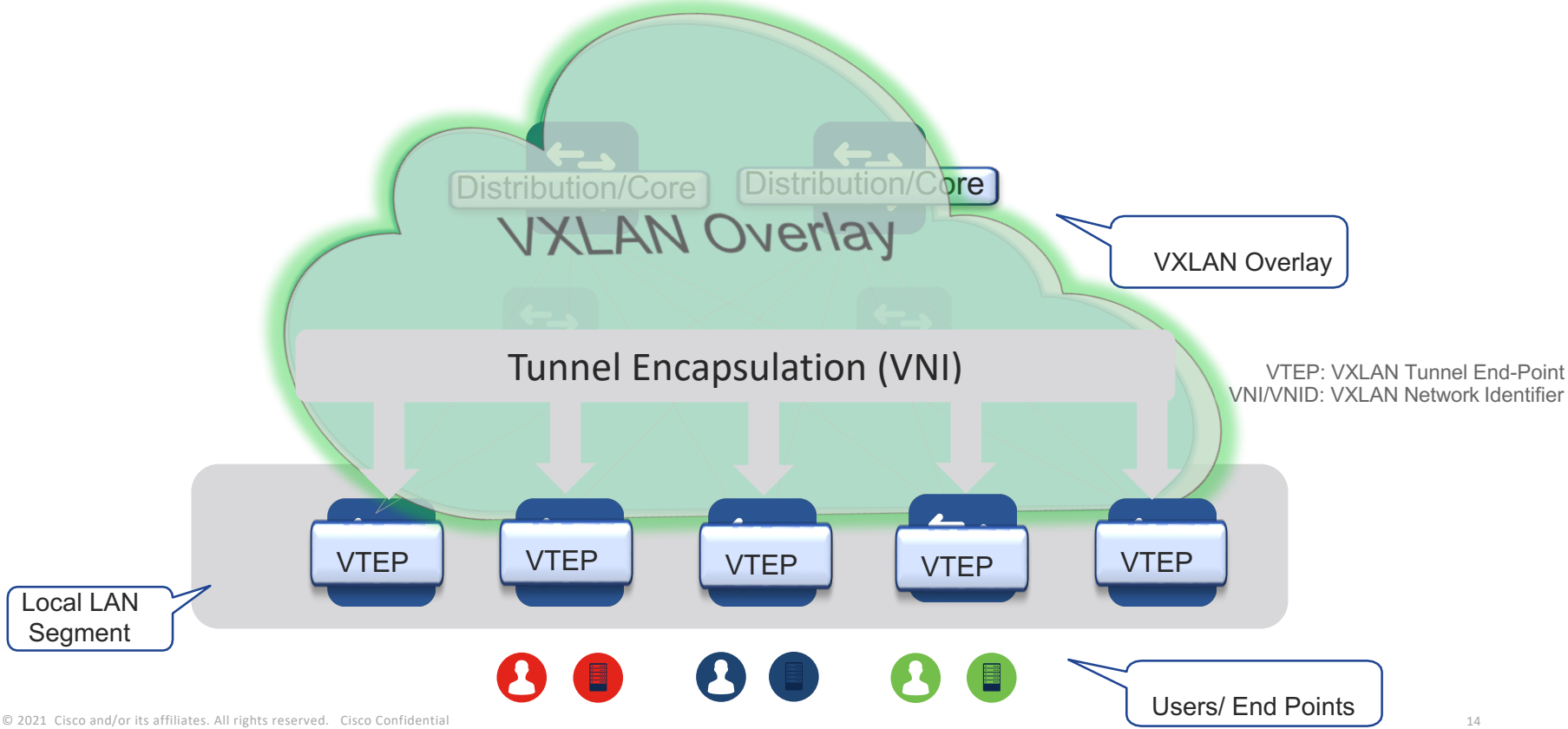
MAC-in-IP Encapsulation



VXLAN Taxonomy – Underlay Network



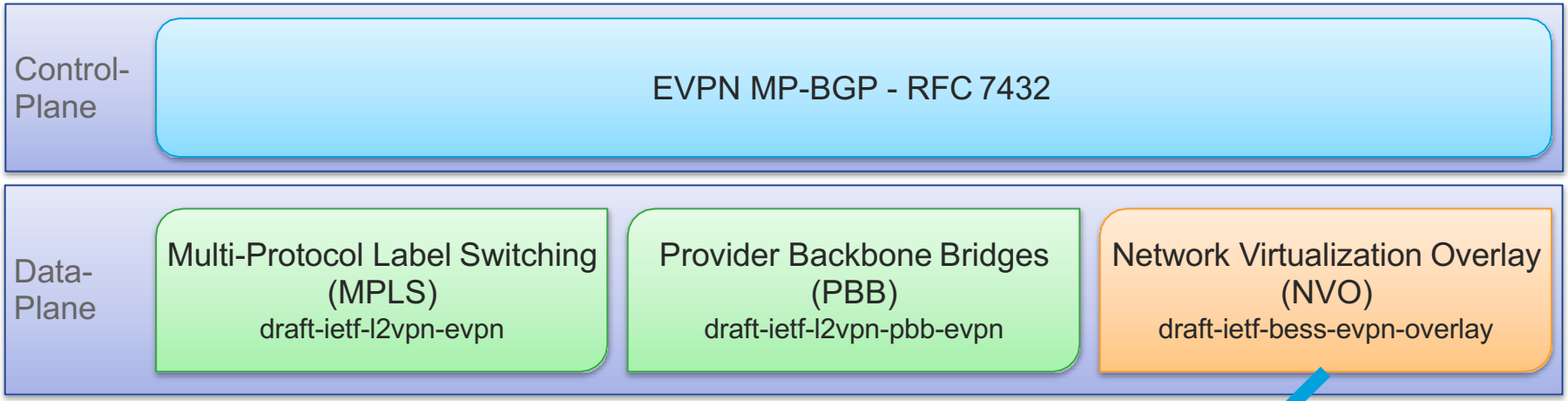
VXLAN Taxonomy – Overlay Network



What is VXLAN with BGP EVPN?

- Standards based Overlay (VXLAN) with Standards based Control-Plane (BGP)
- Layer-2 MAC and Layer-3 IP information distribution by Control-Plane (BGP)
- Forwarding decision based on Control-Plane (minimizes flooding)
- Integrated Routing/Bridging (IRB) for Optimized Forwarding in the Overlay
- Multi-Tenancy At Scale

EVPN – Ethernet VPN



- EVPN over NVO Tunnels (ie VXLAN)
- Provides Layer-2 and Layer-3 Overlays over simple IP Networks

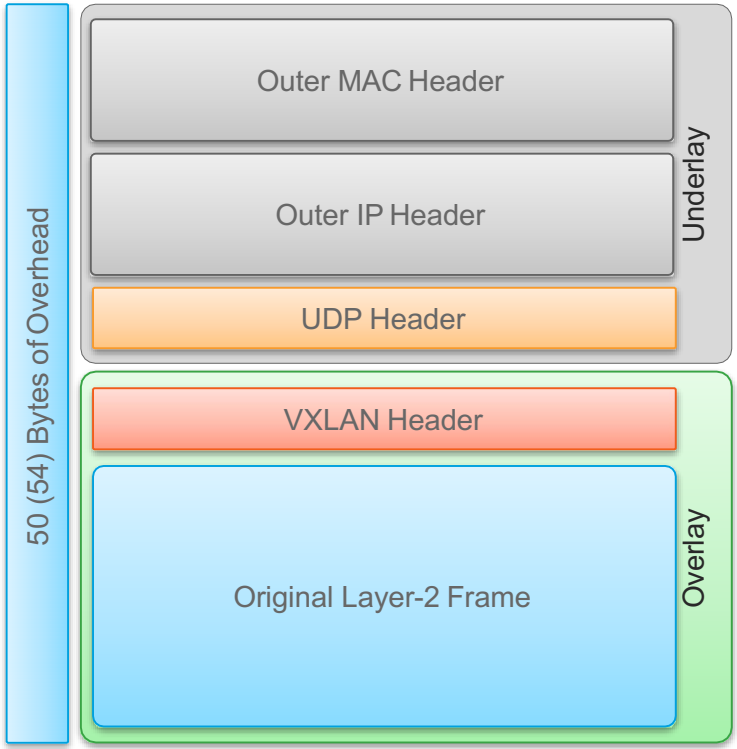
Only Supported in Cat9k

Cisco’s VXLAN related IETF RFCs & Drafts

ID	Title	Category
RFC 7348	Virtual eXtensible Local Area Network	Data Plane
RFC 7432	BGP MPLS based Ethernet VPNs	Control Plane
draft-ietf-bess-evpn-overlay	A Network Virtualization Overlay Solution using EVPN	Control Plane
draft-ietf-bess-evpn-inter-subnet-forwarding	Integrated Routing and Bridging in EVPN	Control Plane
draft-ietf-bess-l2vpn-evpn-prefix-advertisement	IP Prefix Advertisement in E-VPN	Control Plane
draft-tissa-nvo3-oam-fm	NVO3 Fault Management / OAM	Management Plane

VXLAN with BGP EVPN:
Underlay

MTU and VXLAN

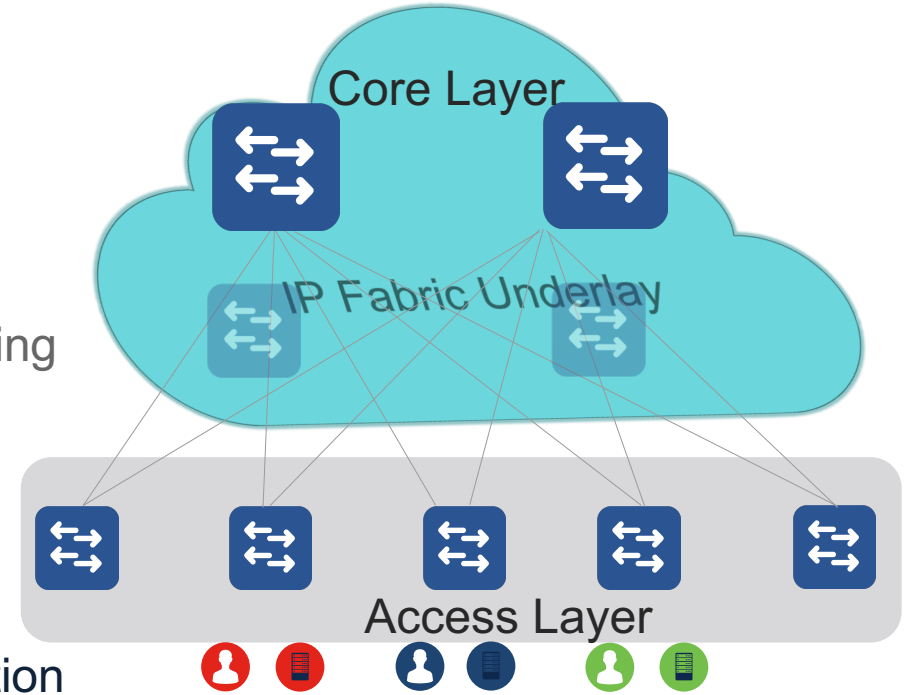


- VXLAN adds 50 Bytes (or 54 Bytes) to the Original Ethernet Frame
- Avoid Fragmentation by adjusting the IP Networks MTU
- Using a MTU of 9216* Bytes accommodates VXLAN Overhead plus other application MTU

*Cisco Catalyst 9k switches only support 9198 Byte for Layer-3 Traffic

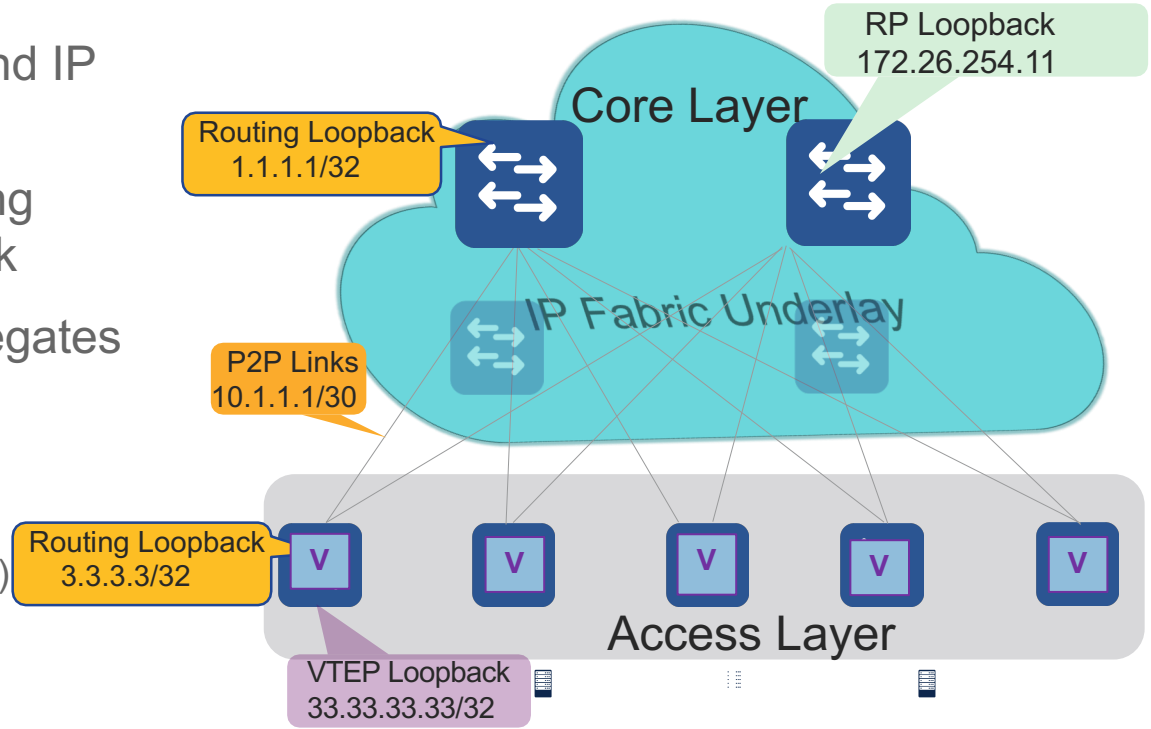
Deployment Considerations

- MTU and Overlays
  - Only 9198 Bytes supported on Cat 9k
- Unicast Routing Protocol and IP Addressing
  - ISIS, OSPF and BGP
- Multicast for BUM* Traffic Replication
  - PIM ASM only supported for BUM
- BUM Traffic Handling by Ingress Replication



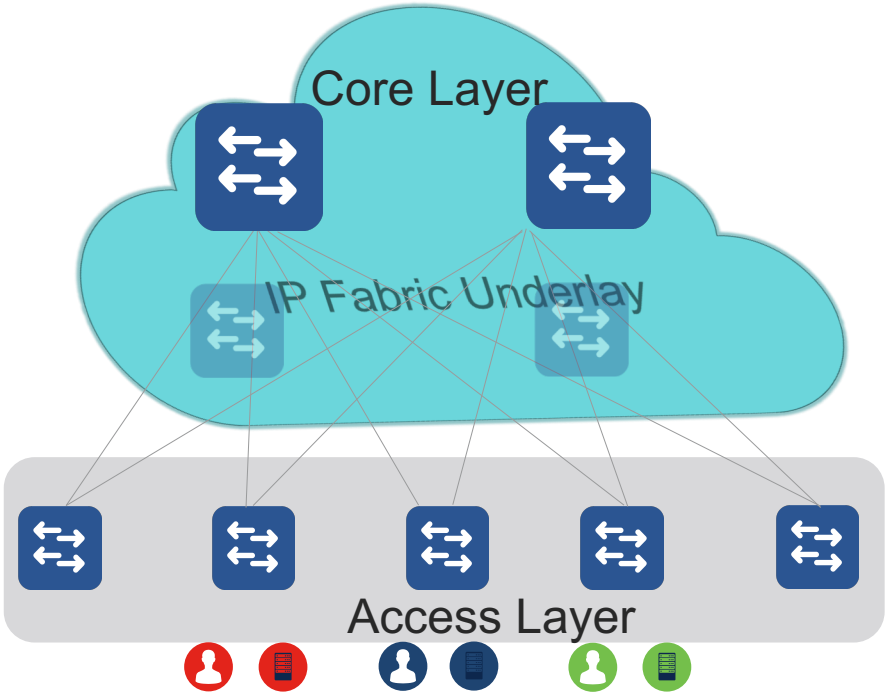
Building your IP Network – Interface Principles

- Know your IP addressing and IP scale requirements
- Separate VTEP from Routing Protocol from RP* Loopback
- Best to use individual Aggregates for the Underlay
  - Unicast Routing p2p** Links
  - Unicast Routing Loopbacks
  - VTEP (NVE) Loopback
  - Multicast Routing Loopback (RP)
- IPv4 only (today)



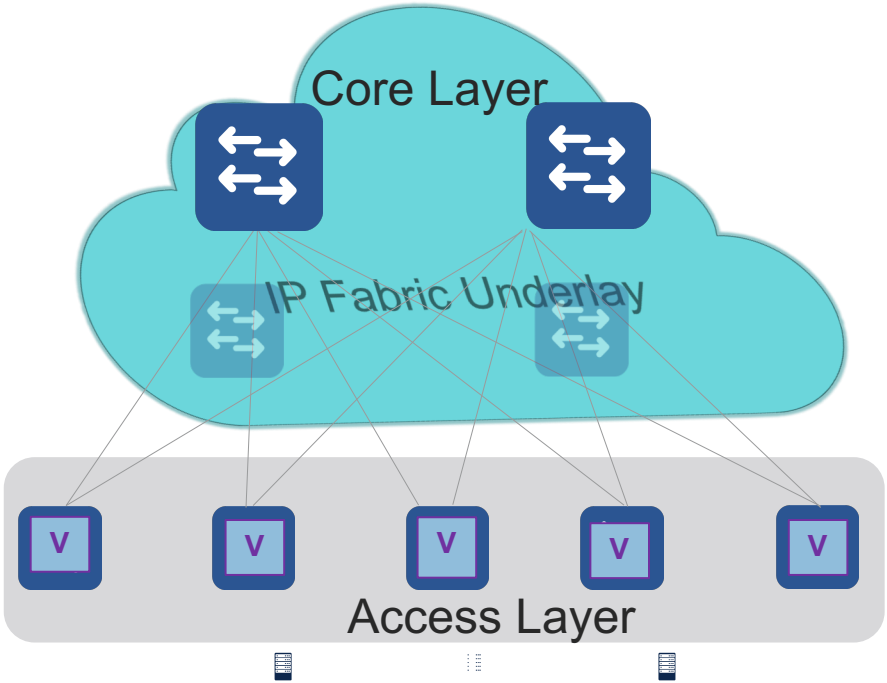
Building your IP Network

- Routed Ports/SVI's
  - Layer-3 Interfaces between Access and Core (no switchport) Or SVI's
  - For each Point-2-Point (P2P) connection, minimum /31 required
  - Alternative, use IP Unnumbered (/32)
- Use Loopback as Source-Interface for VTEP (NVE*)



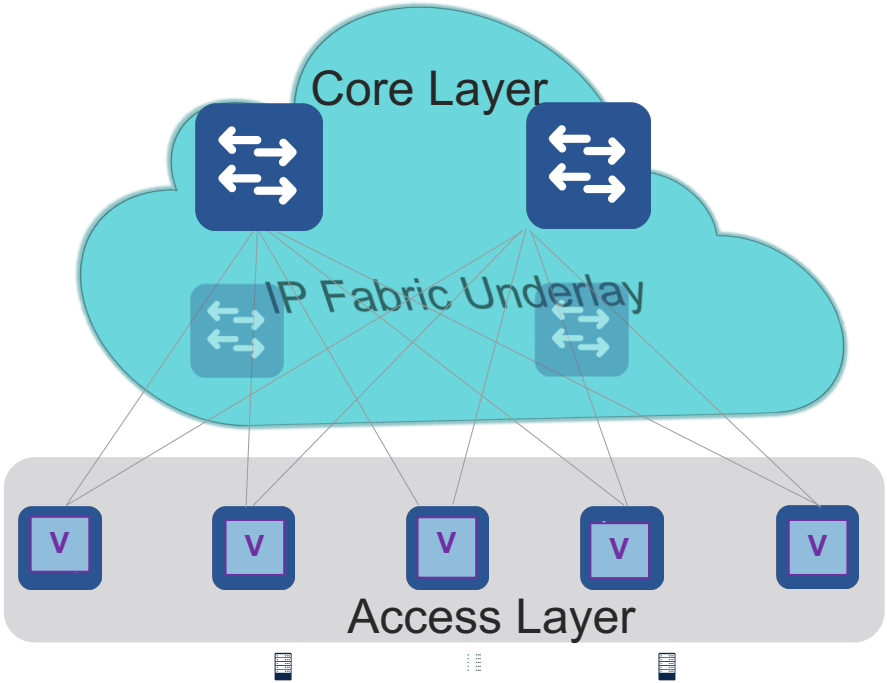
Building your IP Network – Routing Protocols; OSPF

- OSPF – watch your Network type!
  - Network Type Point-2-Point (P2P)
    - Preferred (only LSA type-1)
    - No DR/BDR election
    - Suits well for routed interfaces/ports (optimal from a LSA Database perspective)
    - Full SPF calculation on Link Change



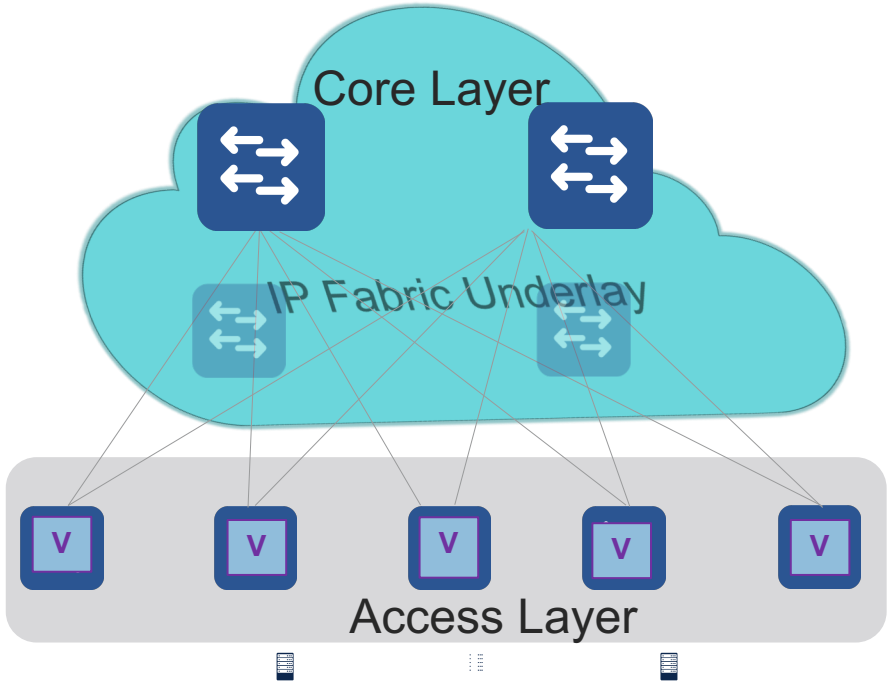
Building your IP Network – Routing Protocols; IS-IS

- IS-IS – what was this CLNS?
  - Independent of IP (CLNS)
  - Well suited for routed interfaces/ports
  - No SPF calculation on Link change; only if Topology changes
  - Fast Re-convergence
  - Not everyone is familiar with it



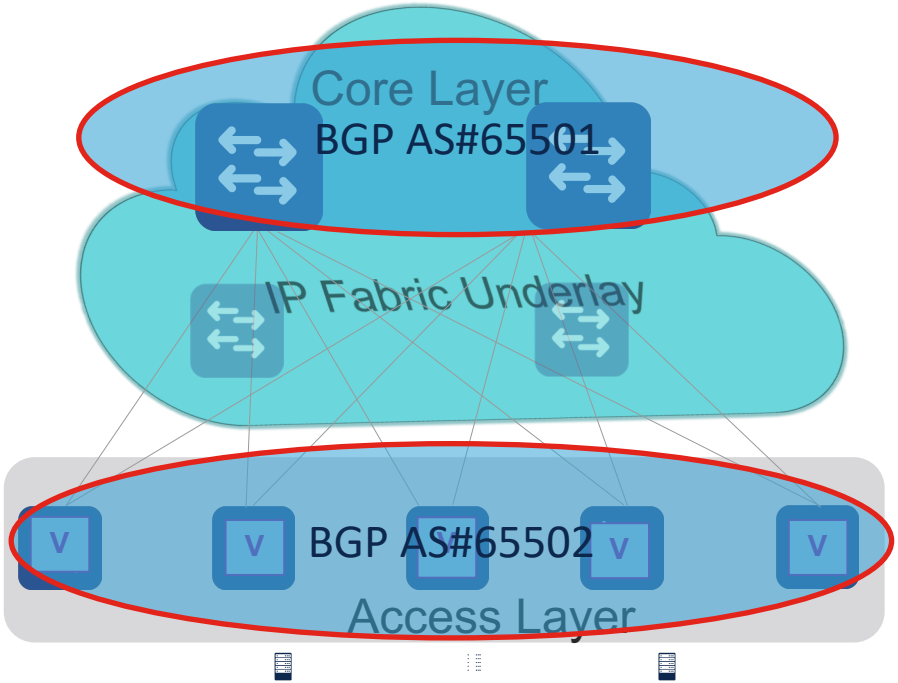
Building your IP Network – Routing Protocols; eBGP

- eBGP – Service Provider style
  - Two Different Models
    - Two-AS
    - Multi-AS
  - BGP is a Distance Vector Protocol (well, actually Path Vector)
    - AS* are used to calculate the Path (AS_Path)
  - If Underlay is eBGP, your Overlay becomes eBGP



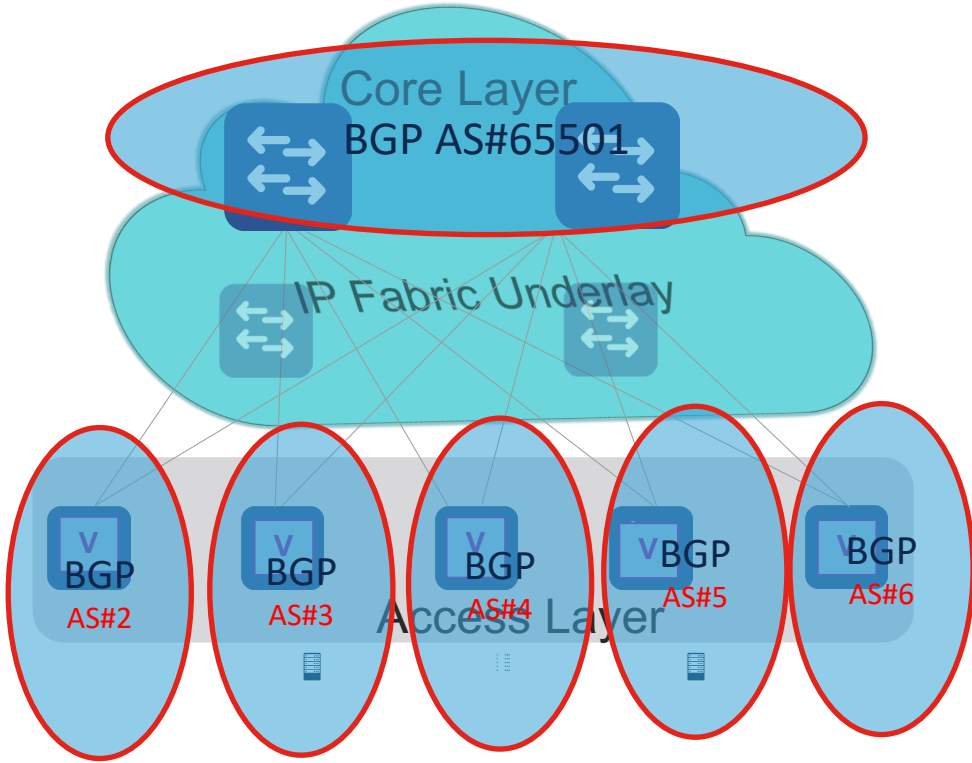
Building your IP Network – Routing Protocols; eBGP

- eBGP – TWO-AS, yes it works!
  - eBGP peering for Underlay-Routing based on physical interface
    - 2 Cores = 2 BGP Peering per Edge
    - Advertise all Infrastructure Loopbacks
  - eBGP peering for Overlay-Routing (EVPN)
    - Loopback to Loopback Peering
    - 2 Cores = 2 BGP Peering
  - Requires some BGP config knobs
    - Disable BGP AS-Path check
    - Next-Hop needs to be Unchanged
    - Retain all Routes on Core (not a RR)

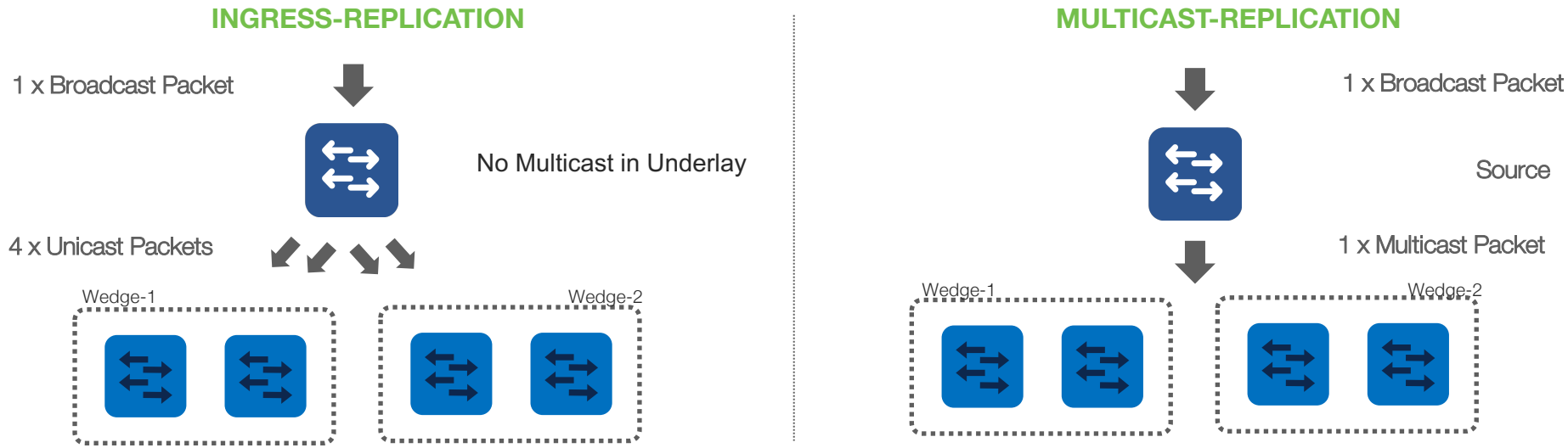


Building your IP Network – Routing Protocols; eBGP

- eBGP – Multi-AS
  - Total of 4 eBGP Peering (with 2 Cores)
  - eBGP peering for Underlay-Routing based on physical interface
    - 2 Cores = 2 BGP Peering per Edge
    - Advertise all Infrastructure Loopbacks
  - eBGP peering for Overlay-Routing (EVPN)
    - Loopback to Loopback Peering
    - 2 Cores = 2 BGP Peering
- Requires some BGP config knobs
  - Next-Hop needs to be Unchanged
  - Retain all Routes on Core (not a RR)



VxLAN BUM Replication Options



- 2 mechanics to handle **B**roadcast, **U**nknown Unicast and Link-Local **M**ulticast (BUM):
  - Ingress-Replication – Convert each BUM packet to multiple Unicast packets and transmit to each remote VTEP
  - Multicast-Replication – Convert each BUM packet to single Multicast packets and transmit in Underlay network
- Multicast replication offers significant system, network and end-user level performance benefits

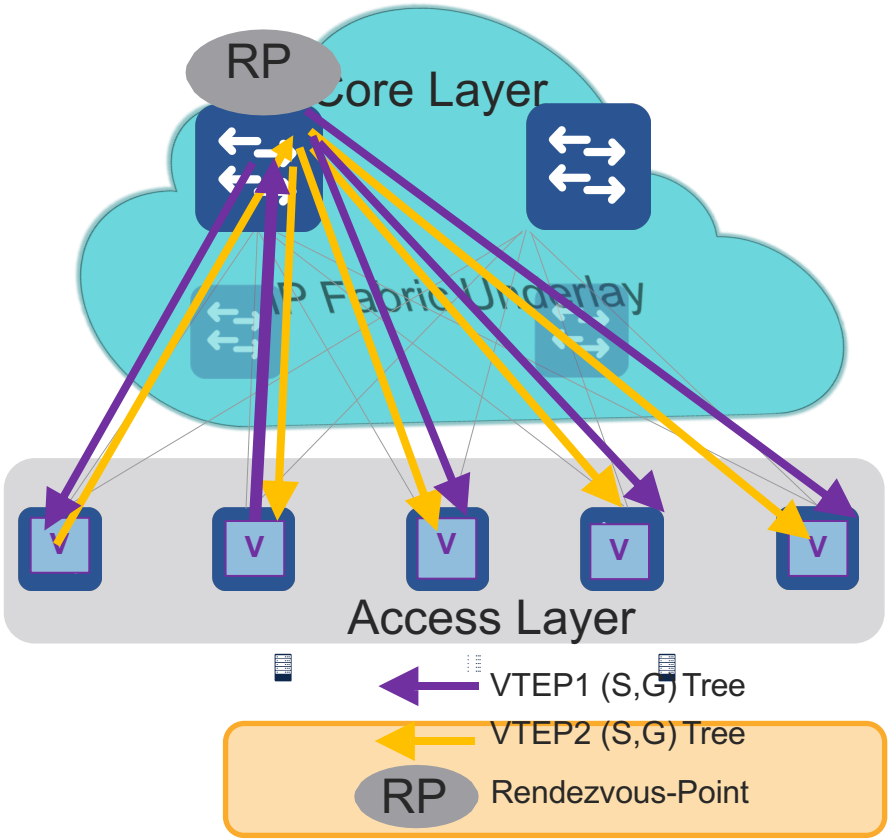
Multicast Enabled Underlay for BUM

Only PIM ASM is supported on Catalyst 9k

- Multi-Destination Traffic (Broadcast, Unknown Unicast, etc.) needs to be replicated to ALL VTEPs serving a given VNI
  - Each VTEP is Multicast Source & Receiver
- For a given VNI, all VTEPs act as a Sender and a Receiver
- Aggregation Switches make good Rendezvous-Point (RP) Locations in Topologies
- Reserve a range of Multicast Groups (Destination Groups/DGroups) to service the Overlay and optimize for diverse VNIs

Multicast Enabled Underlay – PIM ASM

- PIM Sparse-Mode (ASM)
- Redundant Rendezvous-Point using PIM Anycast-RP or External RP
- Source-Tree or Unidirectional Shared-Tree (Source-Tree shown)
  - Shared-Tree will always use RP for forwarding
- 1 Source-Tree per Multicast-Group per VTEP (each VTEP is Source & Receiver)

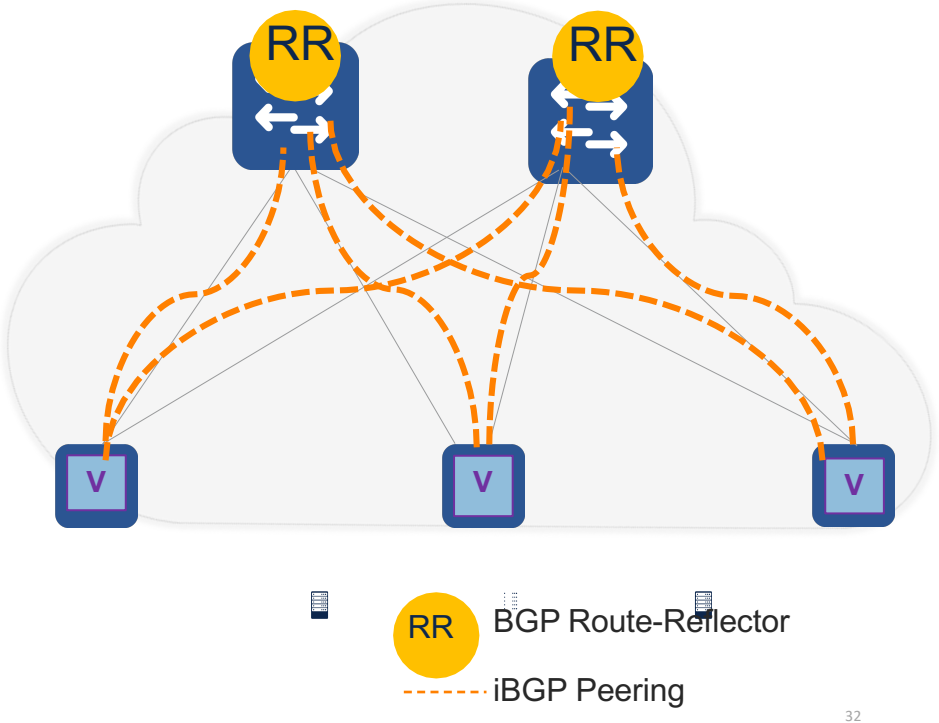


*ASM: Any-Source Multicast

VXLAN with BGP EVPN:
Control & Data Plane

Multiprotocol BGP (MP-BGP) Primer

- Multiprotocol BGP (MP-BGP)
- Extension to Border Gateway Protocol (BGP) - RFC 4760
- VPN Address-Family:
  - Allows different types of address families (e.g. VPNv4, VPNv6, L2VPN EVPN, MVPN)
  - Information transported across single BGP peering



MP-BGP EVPN Route Type(s)

- New BGP EVPN NLRI format is defined in RFC 7432

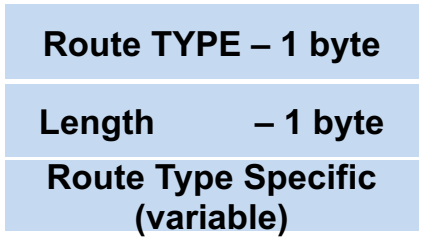
- The Route Type field defines the encoding of the rest of the EVPN NLRI (Route Type specific EVPN NLRI).

- RFC 7432 defines 4 different route types:
  - Route Type 1 - Ethernet Auto-Discovery (A-D) route
  - **Route Type 2 - MAC/IP advertisement route**
  - **Route Type 3 - Inclusive Multicast Route → EVPN Ingress Replication (IR) (unicast mode for BUM)**
  - Route Type 4 - Ethernet Segment Route

- Draft https://tools.ietf.org/html/draft-ietf-bess-evpn-prefix-advertisement-09 defines:
  Route Type 5 -IP Prefix Route → Layer-3 VNI Route

- Route-type 2 or MAC/IP Advertisement route is for host MAC or MAC-IP
- Route-type 5 or IP Prefix route will be used for the advertisement of IP prefixes only

EVPN NLRI

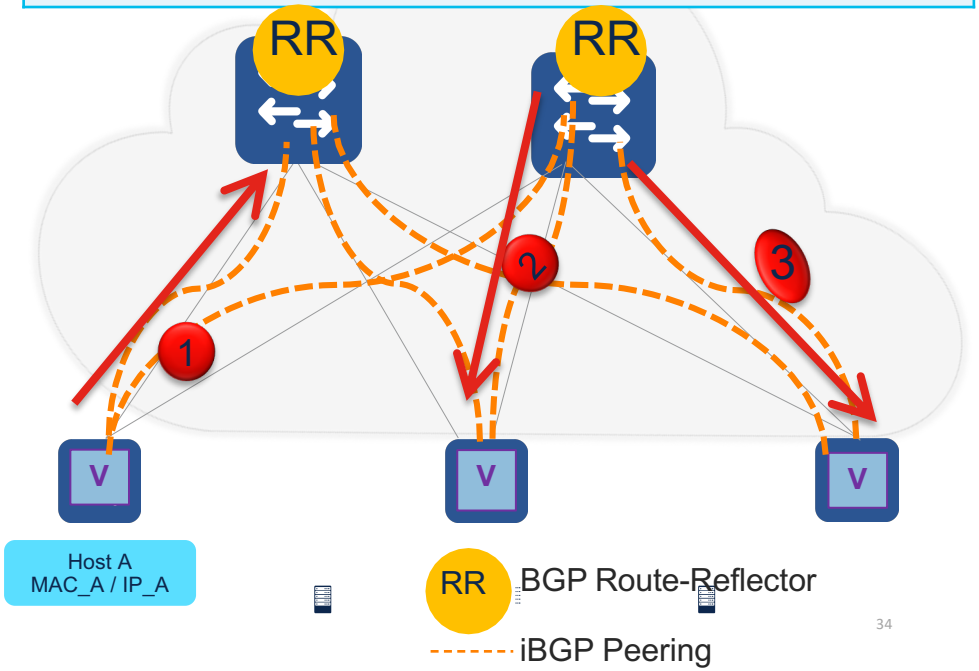


NLRI = Network Layer Reachability Information

MAC/IP Advertisement route
"MAC or MAC/IP host Advertisement (Route-Type 2)"

- Host "A" attaches to Edge Device (VTEP)
- VTEP V1 advertises Host "A" reachability information
  - MAC and L2VNI [mandatory]
  - IP and L3VNI [optional]
  - depending on ARP
- Additional Attributes advertised
  - MPLS Label 1 (Layer-2 VNI)
  - MPLS Label 2 (Layer-3 VNI)
  - Extended Communities

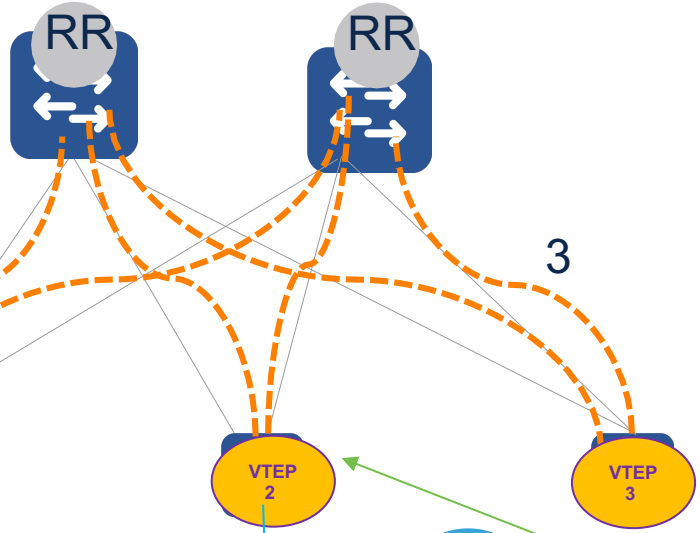
Route Type	MAC, IP	L2VNI	Layer-3 VNI ("VRF")	NH	Encap	Seq
2	MAC_A, IP_A	30001	50001	IP_V1	8:VXLAN	0



EVPN Control Plane --- ARP Suppression

Minimize Flood-&-learn behavior for host learning via ARP/ND Relay

Route Type	MAC, IP	VNI	NH	Encap	Seq
2	MAC_A, IP_A	30001	IP_V1	VXLAN	0



2

VTEP-1 receives and intercepts the ARP Request. Checked in its own host table.

. If it has an match for IP-B, It will change the ethernet header broadcast address to the destination unicast address and will forward to the destination

1

Host-A sends ARP request for IP-B

4

ARP response is received at the source and then traffic will be forwarded

3

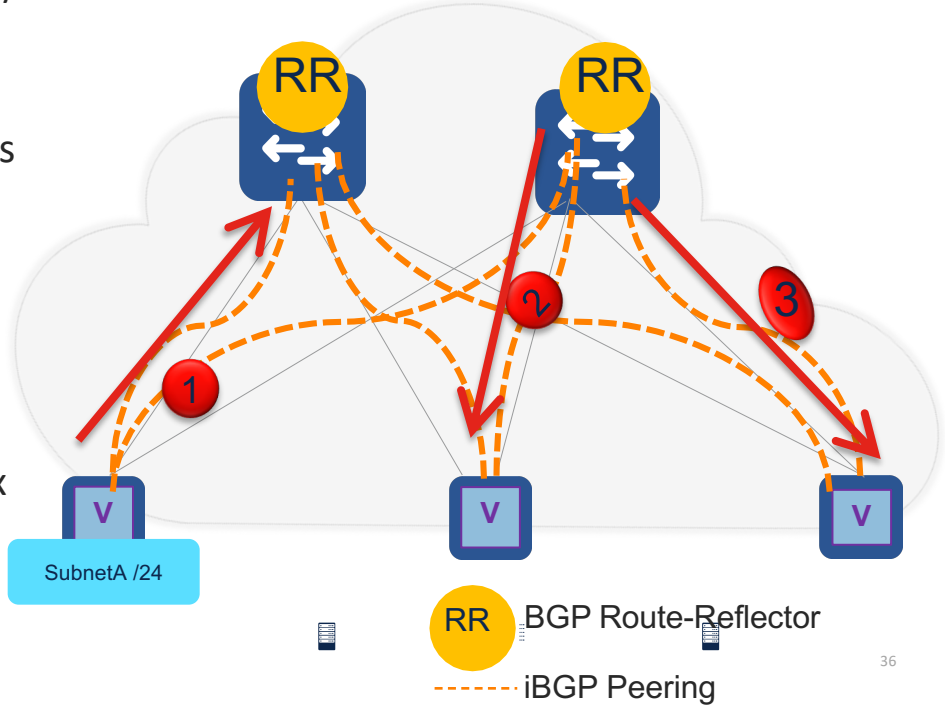
VTEP-2 receives and intercepts the ARP Request. Checked in its own host table.

. If it has an match for IP-B, It will respond with ARP reply by updating the mac-address in the ARP header

Protocol Learning & Distribution
"Subnet Route Advertisement (Route-Type 5)"

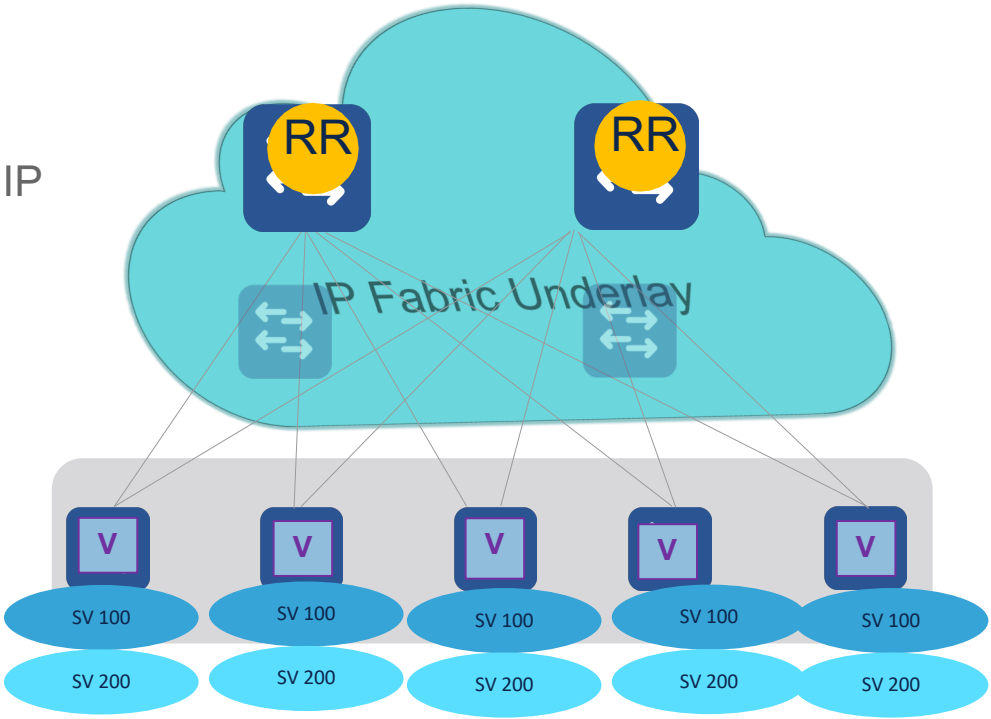
- IP Prefix Redistribution
  - From "Direct" (connected), Static or dynamically learned Routes
- VTEP V1 advertises local Subnet through redistribution of "Direct" (connected) routes
  - IP Prefix, IP Prefix Length, and Layer-3 VNI
- Additional route attributes advertised
  - MPLS Label (Layer-3 VNI)
  - Extended Communities
- Multiple VTEPs can announce same IP Prefix

Route Type	MAC, IP	Layer-3 VNI ("VRF")	NH	Encap
5	Subnet_A/24	50001	IP_V1	8:VXLAN



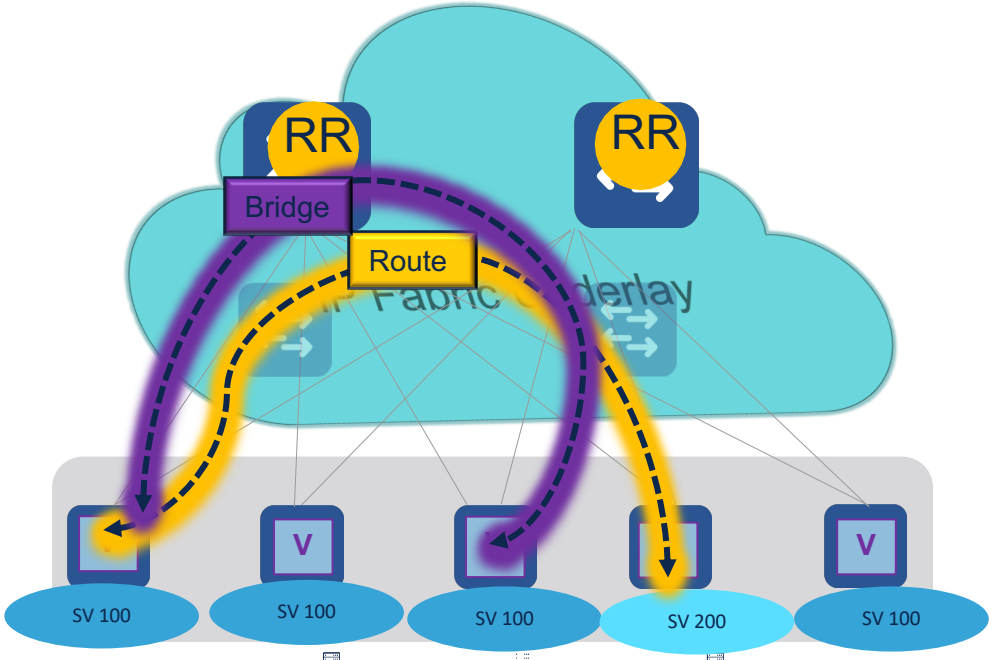
Distributed IP Anycast Gateway

- Distributed Inter-VXLAN Routing at Access Layer (Edge)
  - All edge switches share same gateway IP and MAC Address for a given Subnet
- Gateway is always active
  - no redundancy protocol, hello exchange etc.
- Distributed state - Smaller ARP Tables



SVI 100, Gateway IP: 192.168.1.1, Gateway MAC: AG:AG:AG:AG:AG:AG
SVI 200, Gateway IP: 10.10.10.1, Gateway MAC: AG:AG:AG:AG:AG:AG

Distributed IP Anycast Gateway

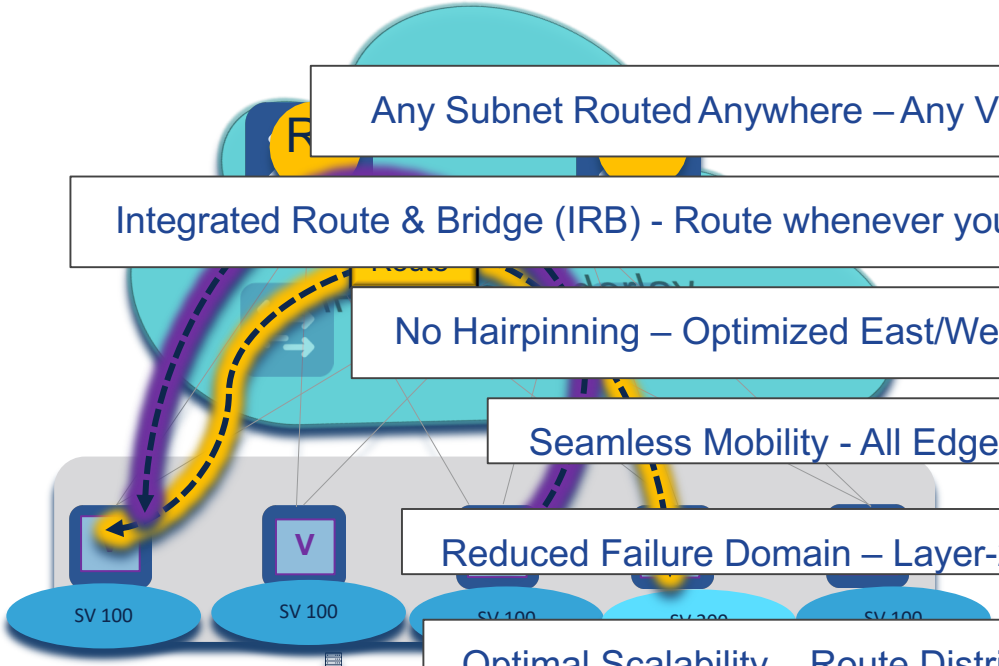


Host1
MAC: AA:AA:AA:AA:AA:AA
IP: 192.168.1.11
VLAN 100
VXLAN VNI 30001

Host3
MAC: CC:CC:CC:CC:CC:CC
IP: 192.168.1.33
VLAN 100
VXLAN VNI 30001

Host2
MAC: BB:BB:BB:BB:BB:BB
IP: 10.10.10.22
VLAN 200
VXLAN VNI 30002

Distributed IP Anycast Gateway



Any Subnet Routed Anywhere – Any VTEP can serve any Subnet

Integrated Route & Bridge (IRB) - Route whenever you can, Bridge when needed

No Hairpinning – Optimized East/West and North/South Routing

Seamless Mobility - All Edge's share same Gateway MAC

Reduced Failure Domain – Layer-2/Layer-3 Boundary at Edge

Optimal Scalability – Route Distributed & closest to the Host

Host1
MAC: AA:AA:AA:AA:AA:AA
IP: 192.168.1.11
VLAN 100
VXLAN VNI 30001

Host3
MAC: CC:CC:CC:CC:CC:CC
IP: 192.168.1.33
VLAN 100
VXLAN VNI 30001

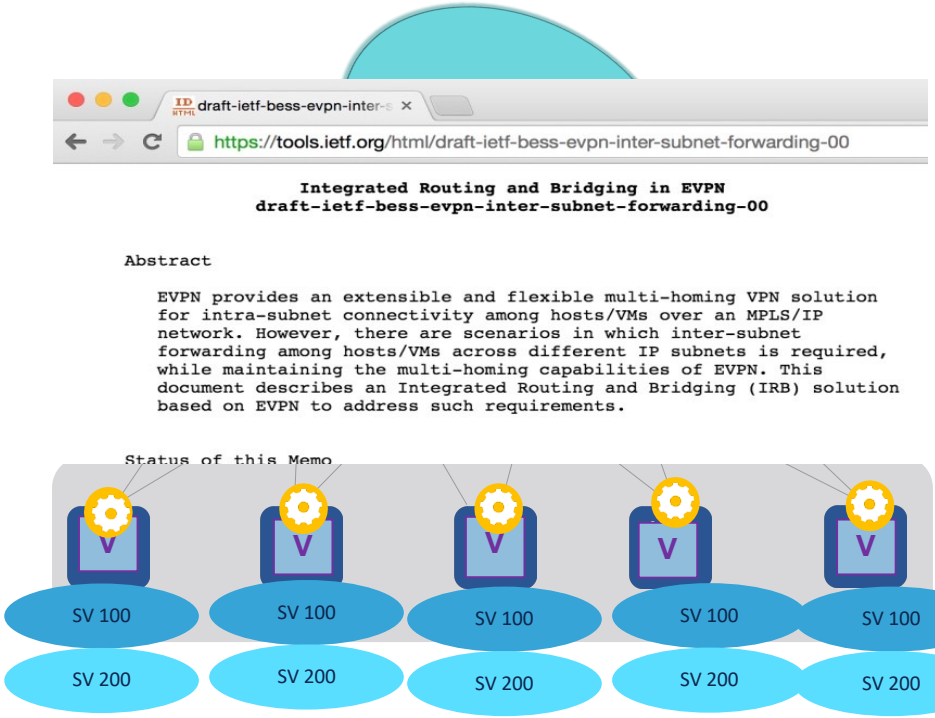
MAC: BB:BB:BB:BB:BB:BB
IP: 10.10.10.22
VLAN 200
VXLAN VNI 30002

Integrated Routing and Bridging (IRB)

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

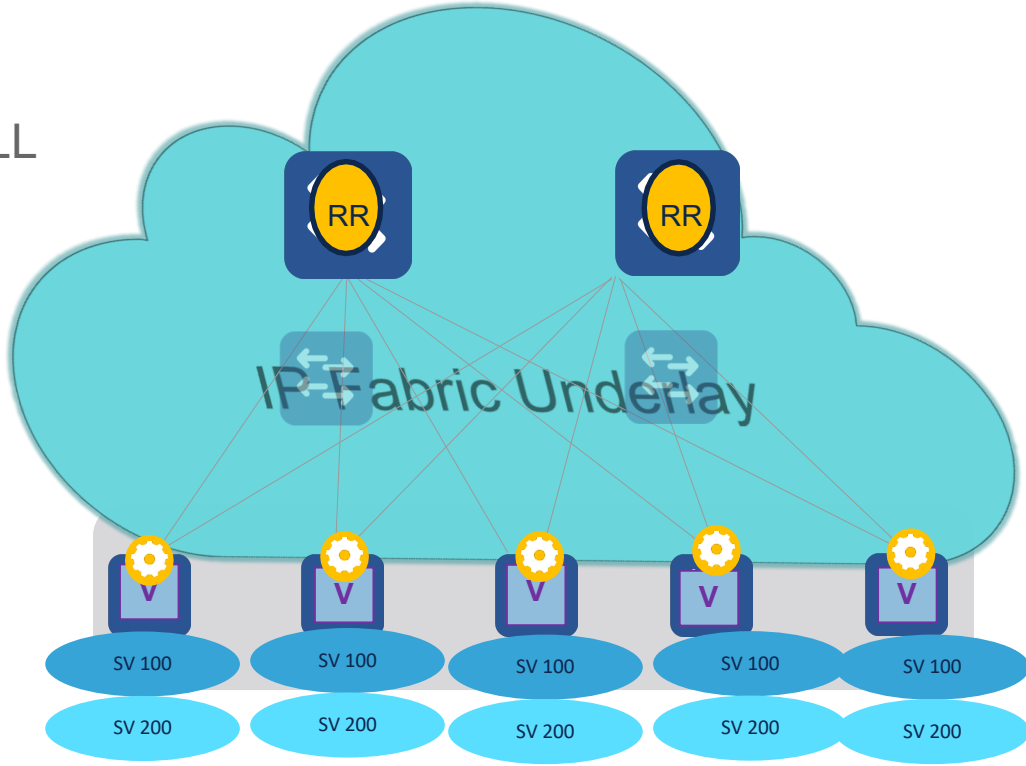
- Asymmetric
  - Uses an “asymmetric path” from the Host towards the egressing port of the VTEP vs. the way back
- Symmetric*
  - Uses an “symmetric path” from the Host towards the egressing port of the VTEP vs. the way back

*Implemented by Cisco’s VXLAN/EVPN



Consistent Configuration

- Logical Configuration (VLAN, VRF, VNI) consistently instantiated on ALL edge's
- Optimal for Consistency
  - Every VLAN/VNI Everywhere
- Sub-Optimal for Scale
  - Instantiates Resources (VLAN/VNI) even if no End-Point uses it

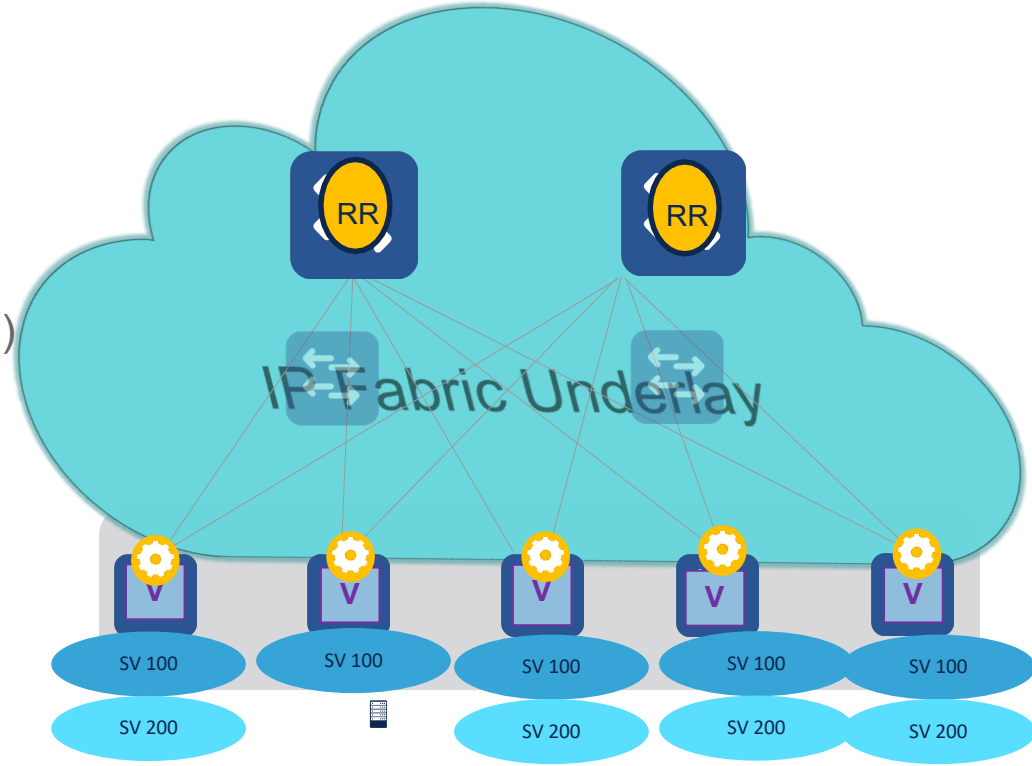


Scoped Configuration

- Logical Configuration (VLAN, VRF, VNI) scoped to edge's with respective connected End-Points

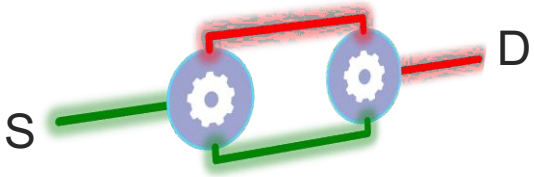
- Optimal for Scale
  - Instantiates Resources (VLAN/VNI) where End-Points are connected

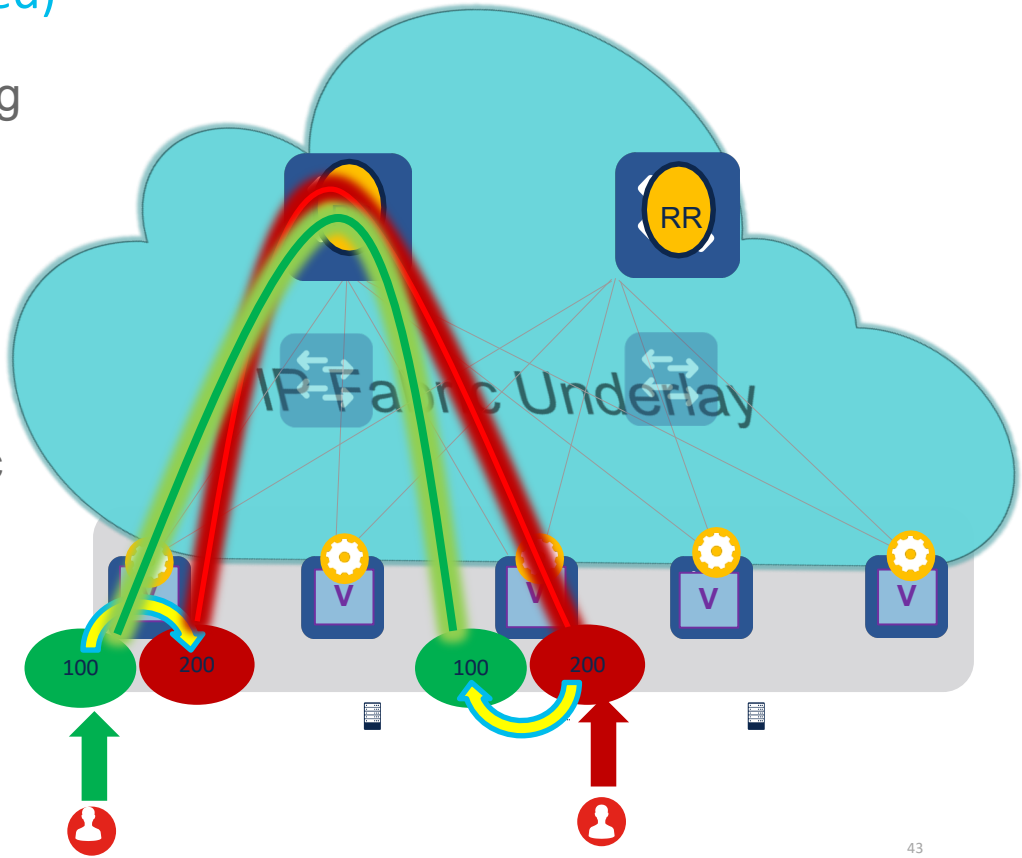
- Consistency with End-Points
  - Configuration Consistency depends on End-Points



Asymmetric IRB (Not Supported)

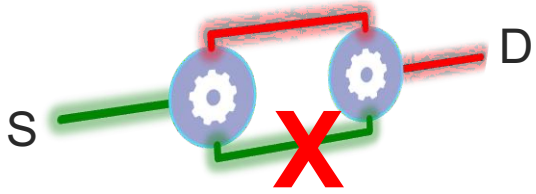
- Similar to todays Inter-VLAN routing
- Requires to follow a consistent configuration of VLAN and L2VNI across all Switches
- Post routed traffic will leverage destination Layer 2 Segment (L2VNI), same as for bridged traffic

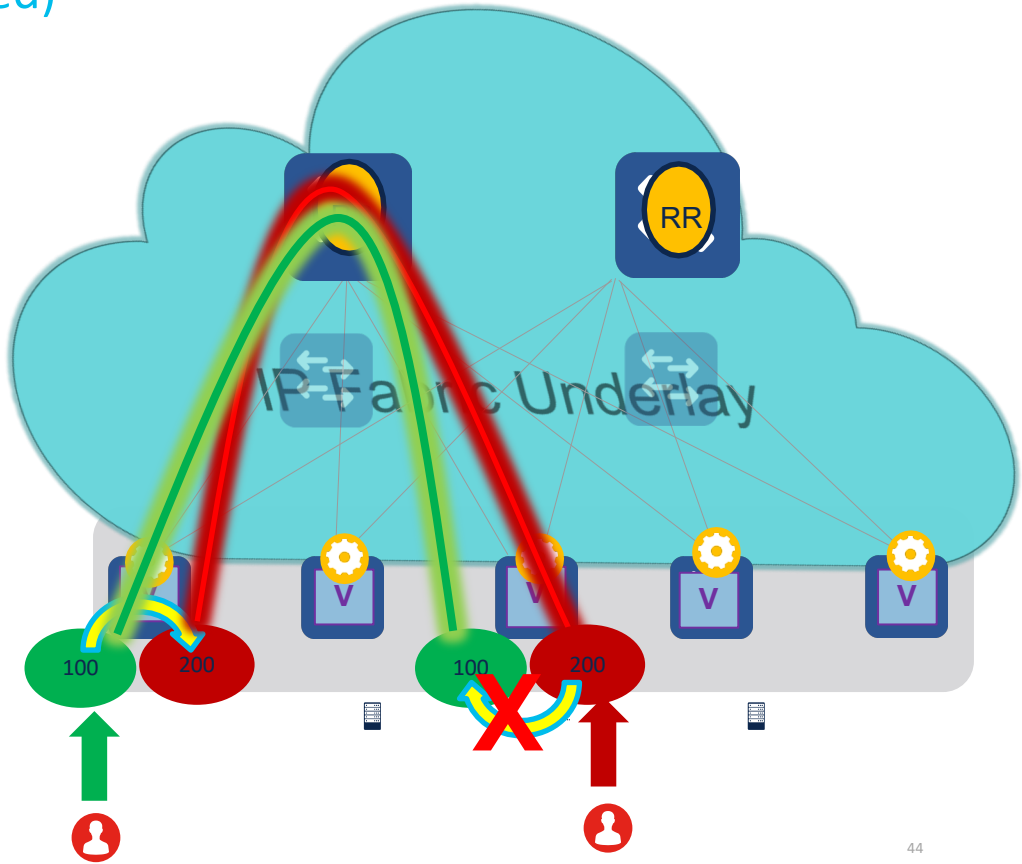




Asymmetric IRB (Not Supported)

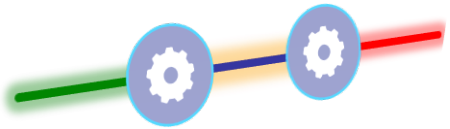
- What if you don't have consistent configuration across all the VTEPs?

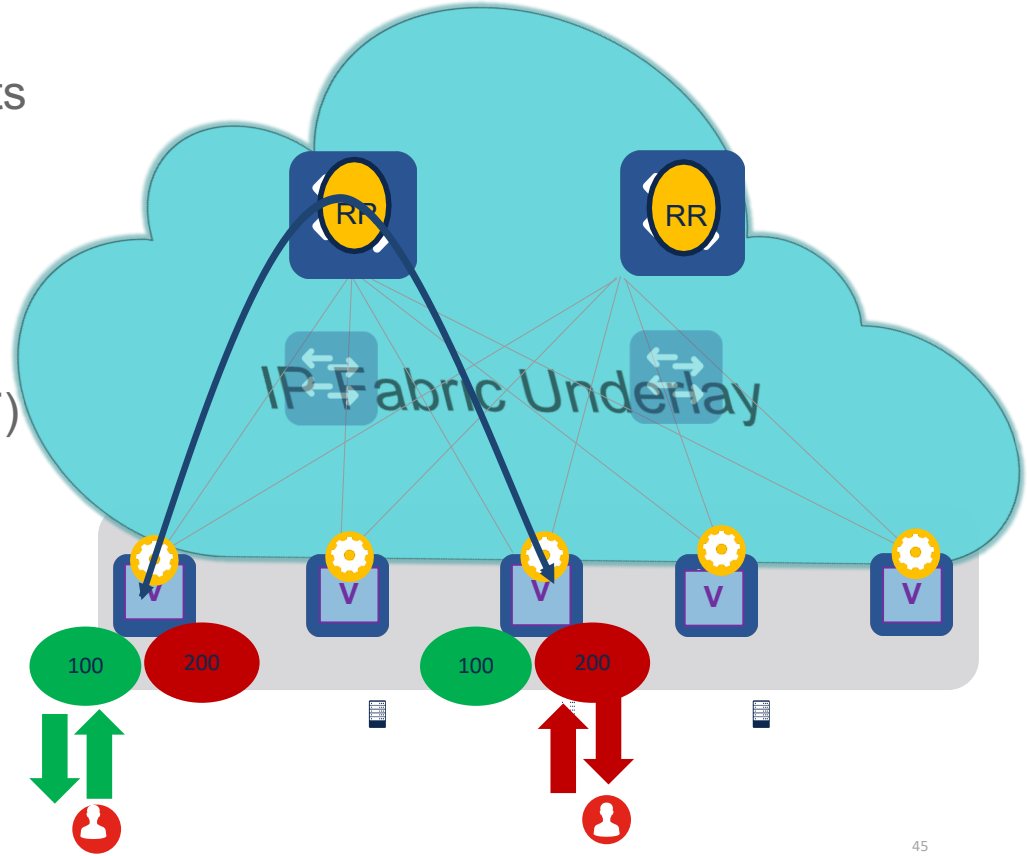




Symmetric IRB

- Similar to Transit Routing Segments
- Scoped Configuration of VLAN/L2VNI; only required where End-Points (Server) reside
- New VNI (L3VNI) introduced per virtual routing and forwarding (VRF) context
- Routed traffic uses transit VNI (L3VNI), while bridged traffic uses L2VNI

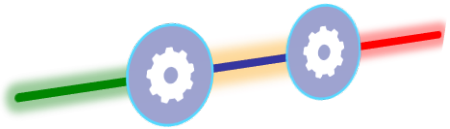


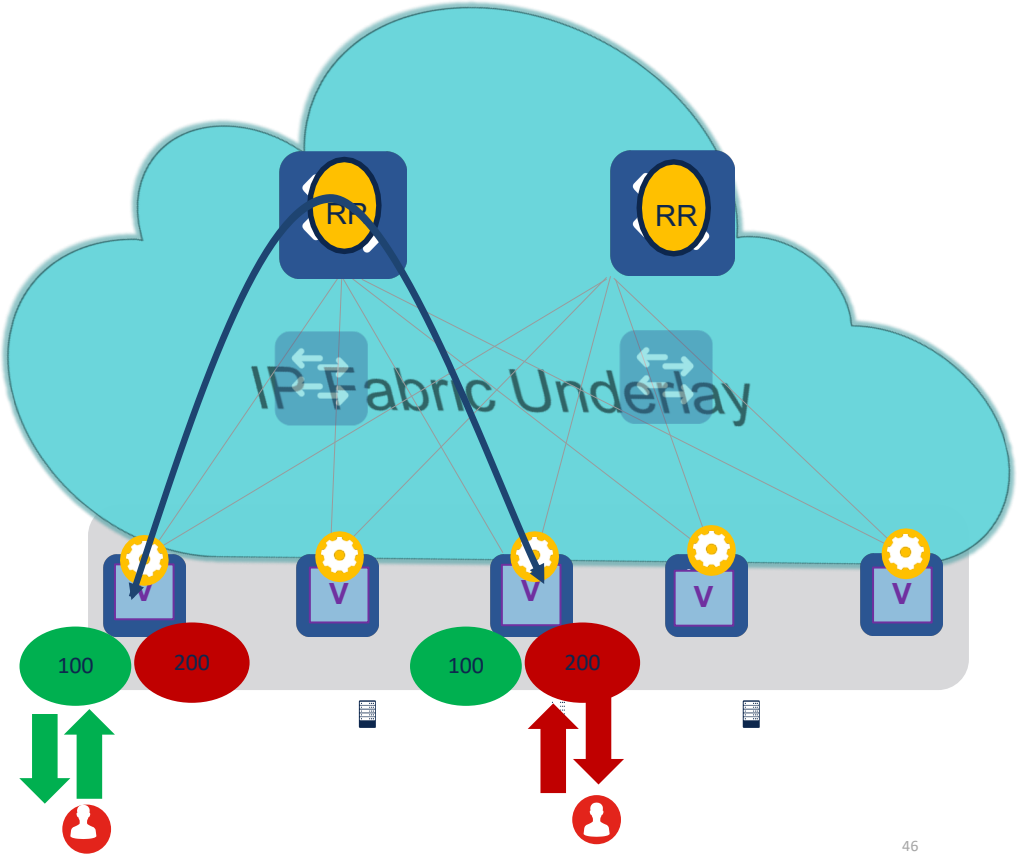


Symmetric IRB

- What if you don't have consistent configuration across all the VTEPs?

No Changes, since we use dedicated L3 VNI for Routed Traffic





VXLAN with BGP EVPN:
Multi-Tenancy

What is Multi-Tenancy

- A mode of operation, where multiple independent instances (tenant) operate in a shared environment.
- Each instance (i.e. VRF/VLAN) is logically isolated, but physically integrated.

Where can we apply Multi-Tenancy

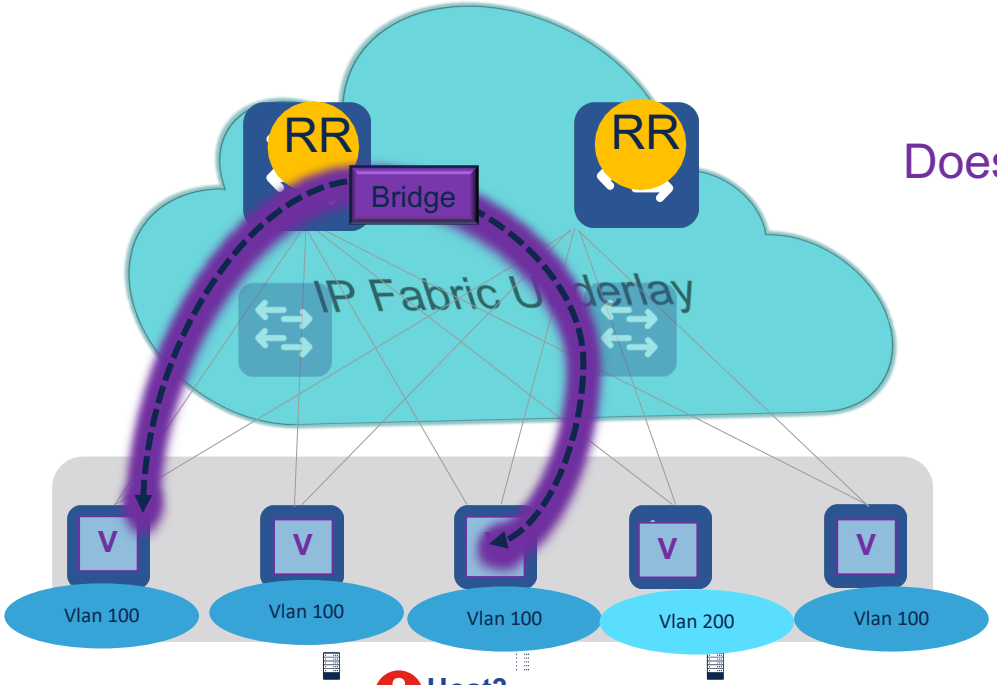
Multi-Tenancy at Layer-2

- Per-Switch VLAN-to-VNI mapping
- Per-Port VLAN Significance

Multi-Tenancy at Layer-3

- VRF-to-VNI mapping
- MP-BGP for scaling with VPNs

Layer-2 Multi-Tenancy

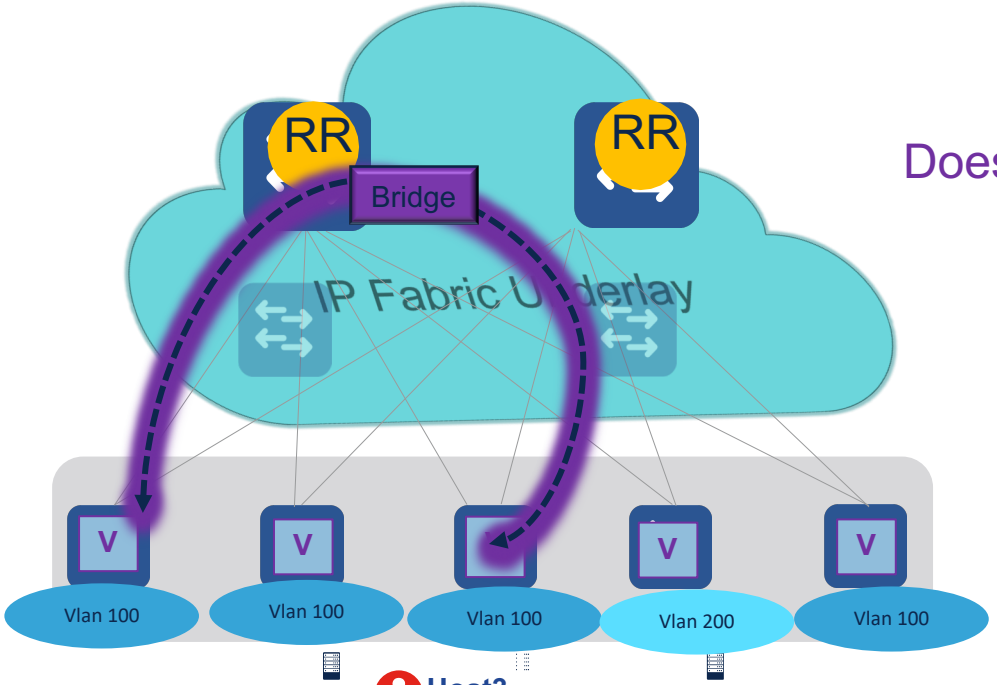


Does not require a VRF

Host1
MAC: AA:AA:AA:AA:AA:AA
IP: 192.168.1.11
VLAN 100
VXLAN VNI 30001

Host3
MAC: CC:CC:CC:CC:CC:CC
IP: 192.168.1.33
VLAN 100
VXLAN VNI 30001

Layer-2 Multi-Tenancy

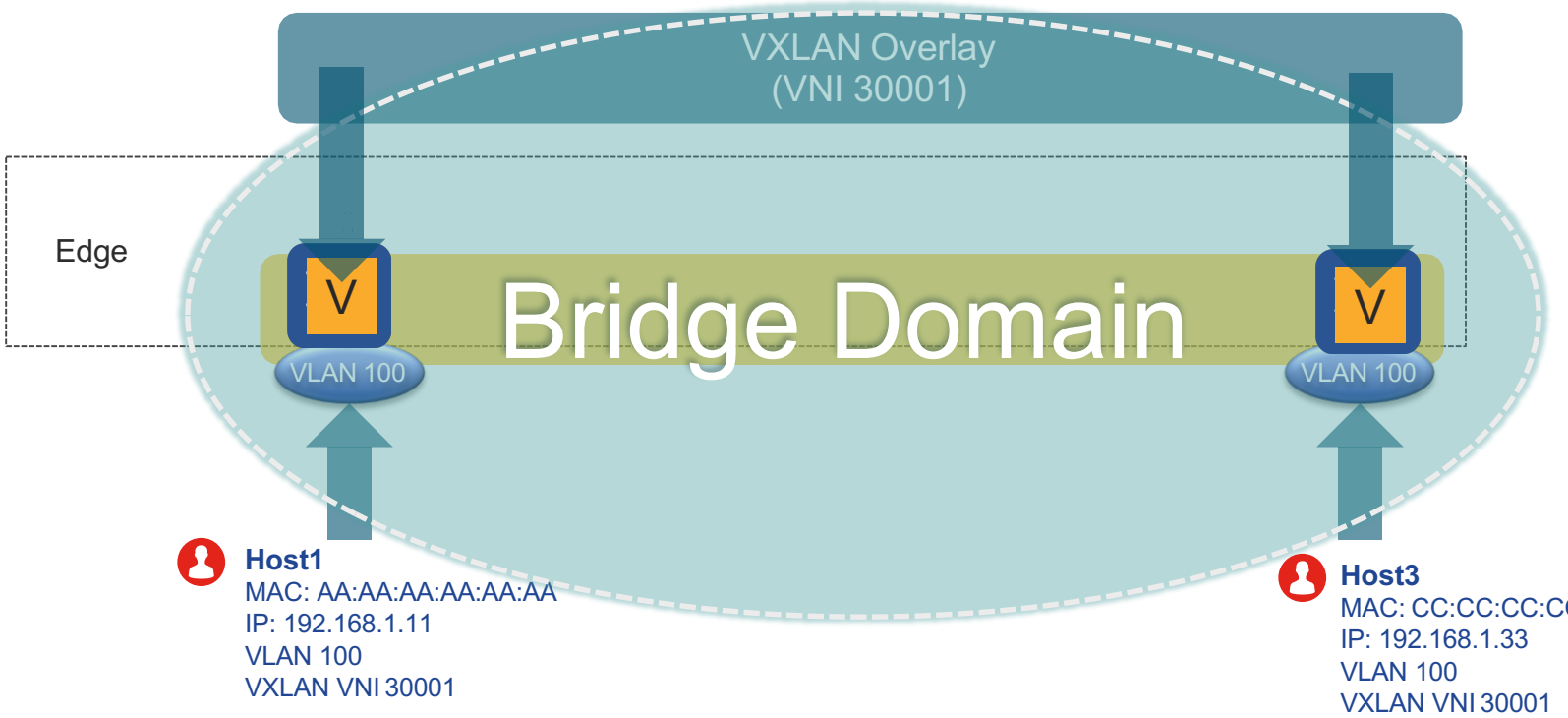


Does not require a VRF

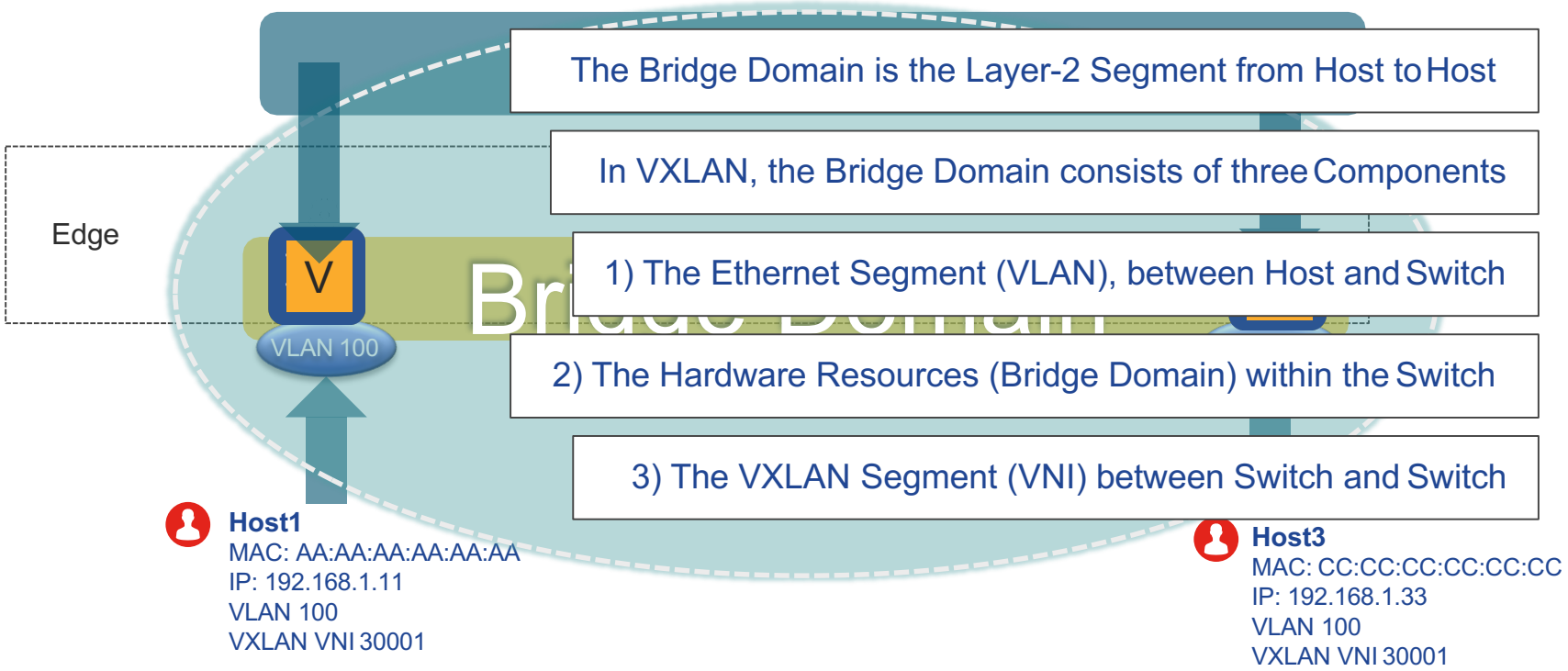
Host1
MAC: AA:AA:AA:AA:AA:AA
IP: 192.168.1.11
VLAN 100
VXLAN VNI 30001

Host3
MAC: CC:CC:CC:CC:CC:CC
IP: 192.168.1.33
VLAN 100
VXLAN VNI 30001

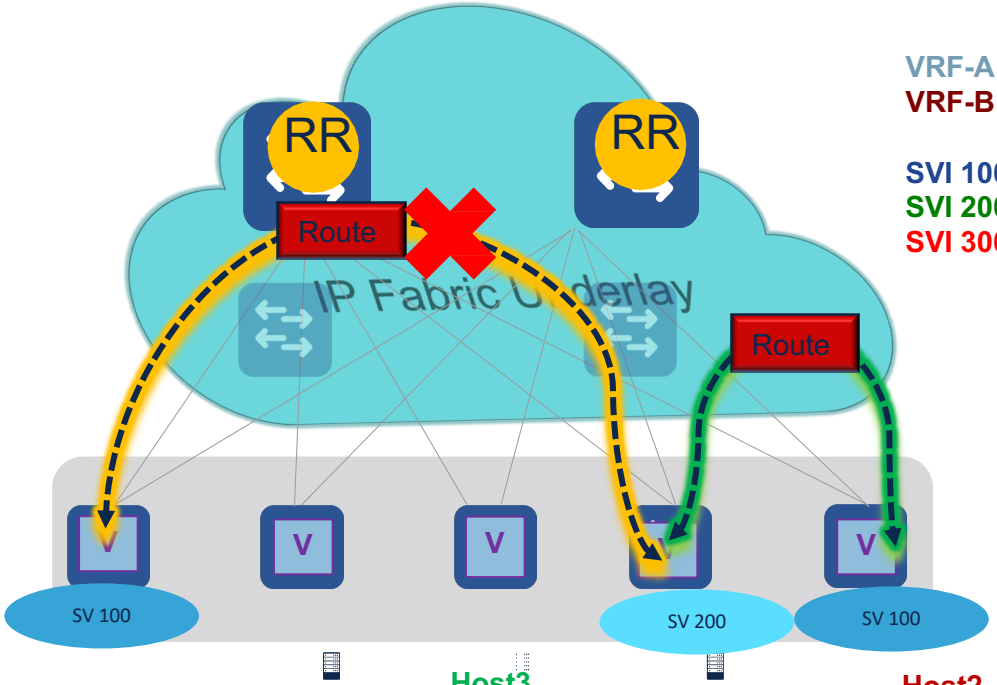
Layer-2 Multi-Tenancy – Bridge Domains



Layer-2 Multi-Tenancy – Bridge Domains



Layer-3 Multi-Tenancy



VRF-A (VNI 50001)
VRF-B (VNI 50002)

SVI 100, Gateway IP: 192.168.1.1 (VRF-A)
SVI 200, Gateway IP: 172.26.200.11 (VRF-B)
SVI 300, Gateway IP: 10.10.10.22 (VRF-B)

Host1
IP: 192.168.1.11 **(VRF A)**
VLAN 100
VXLAN VNI 50001

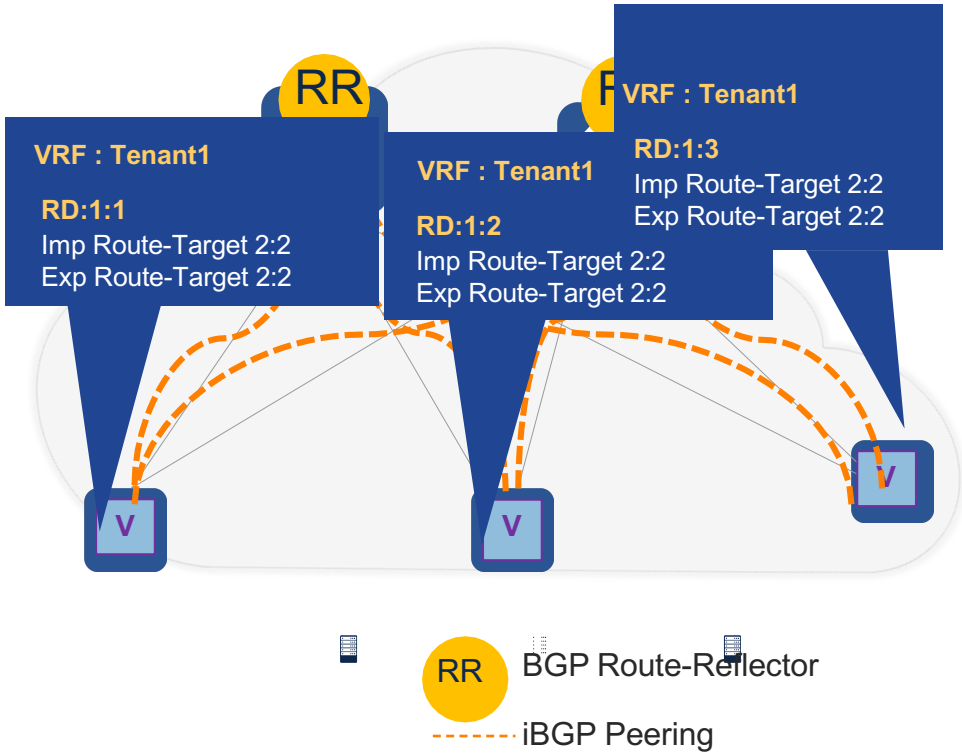
Host3
IP: 172.26.200.11 **(VRF B)**
VLAN 200
VXLAN VNI 50002

Host2
IP: 10.10.10.22 (VRF B)
VLAN 300
VXLAN VNI 50002

Multiprotocol BGP (MP-BGP) Primer

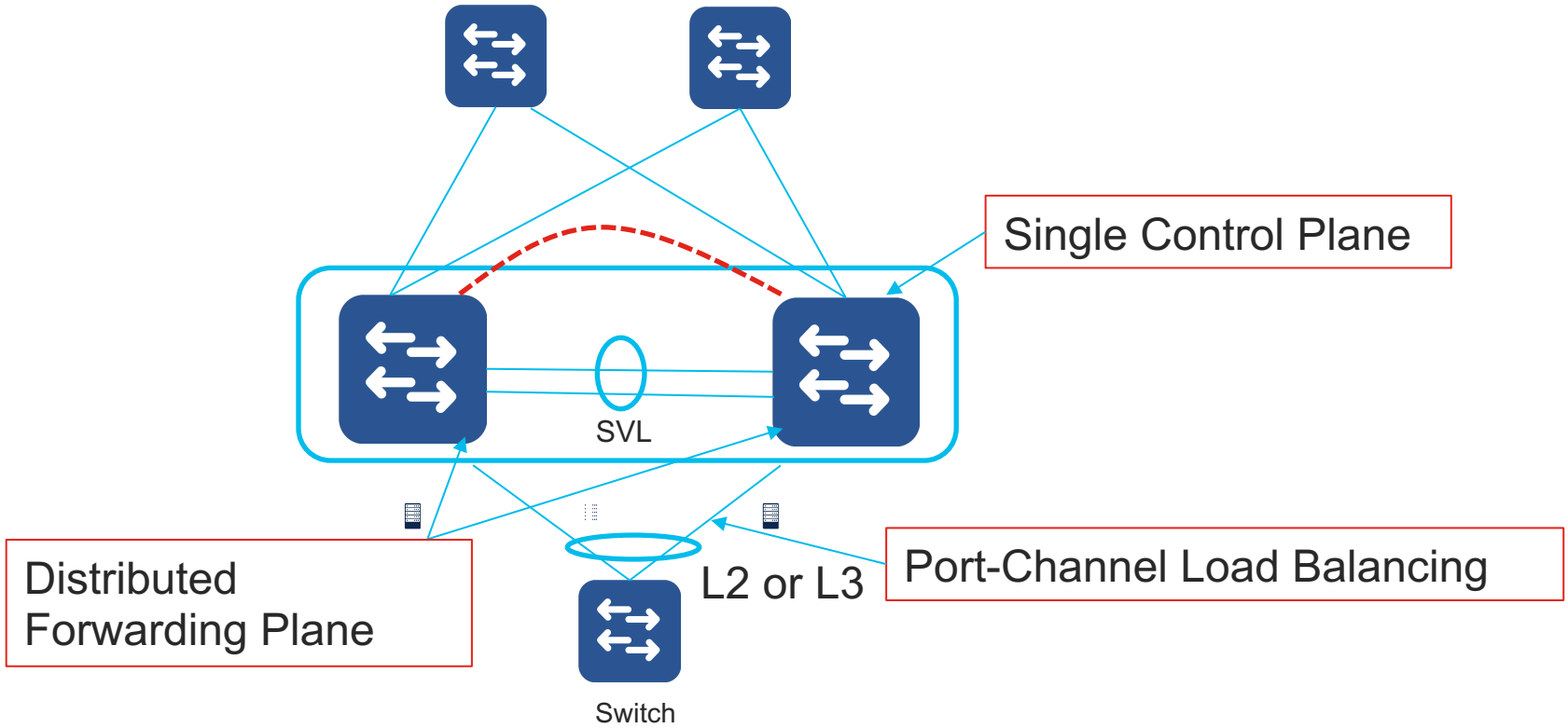
- VPN segmentation for tenant routing (Multi-Tenancy)
  - Route Distinguisher (RD)
  - 8-byte field of VRF parameters
  - value to make VPN prefix unique:
    - RD + VPN prefix

- Selective distribute VPN routes - Route Target (RT)
  - 8-byte field of VRF parameter
  - unique value to define the import/export rules for VPN prefix



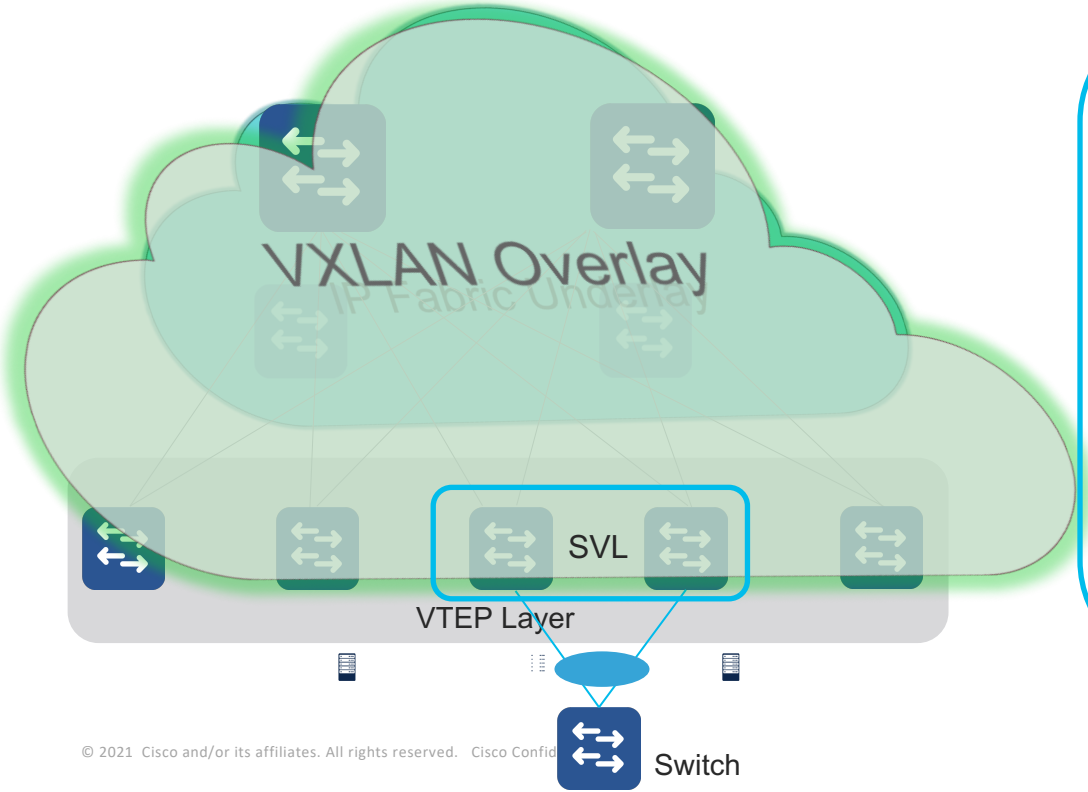
VXLAN with BGP EVPN:
StackWise Virtual Redundancy

Multi-Homing Support via StackWise Virtual



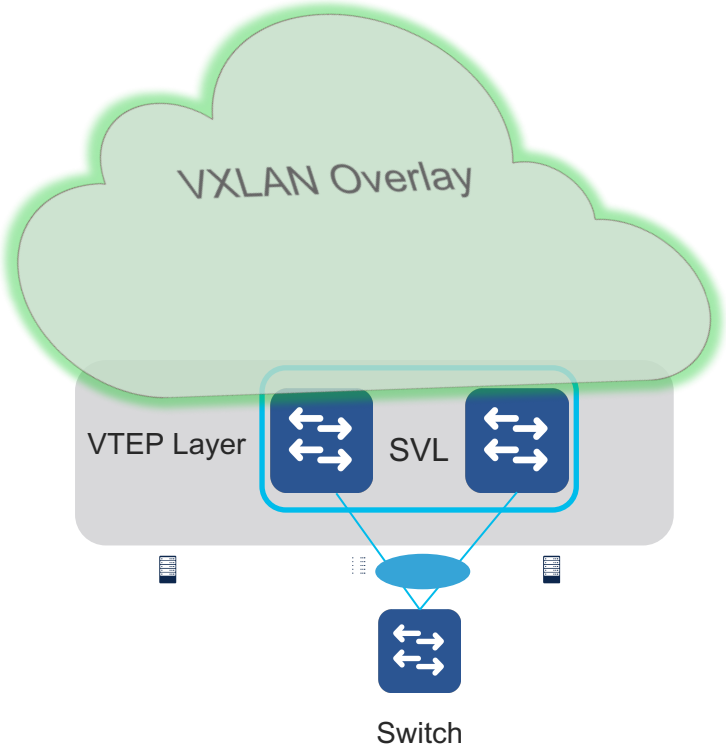
No additional configuration required for VTEP Role

StackWise Virtual Gateway Redundancy



- StackWise Virtual
  - Multi-Chassis Link Aggregation
  - Extended for VXLAN
- Access/Host Side
  - Dual-Connect Access or Hosts
  - Using Port-channels
- Fabric Side
  - Seen as one VTEP from remote nodes

StackWise Virtual Gateway Redundancy



- Independent Device in the EVPN Control-Plane
  - Single Protocol Peering
  - Single Route Distinguisher(RD)
- Single VXLAN Device
  - No special configuration needed
  - Underlay Port-channel Load Sharing to SVL VTEP

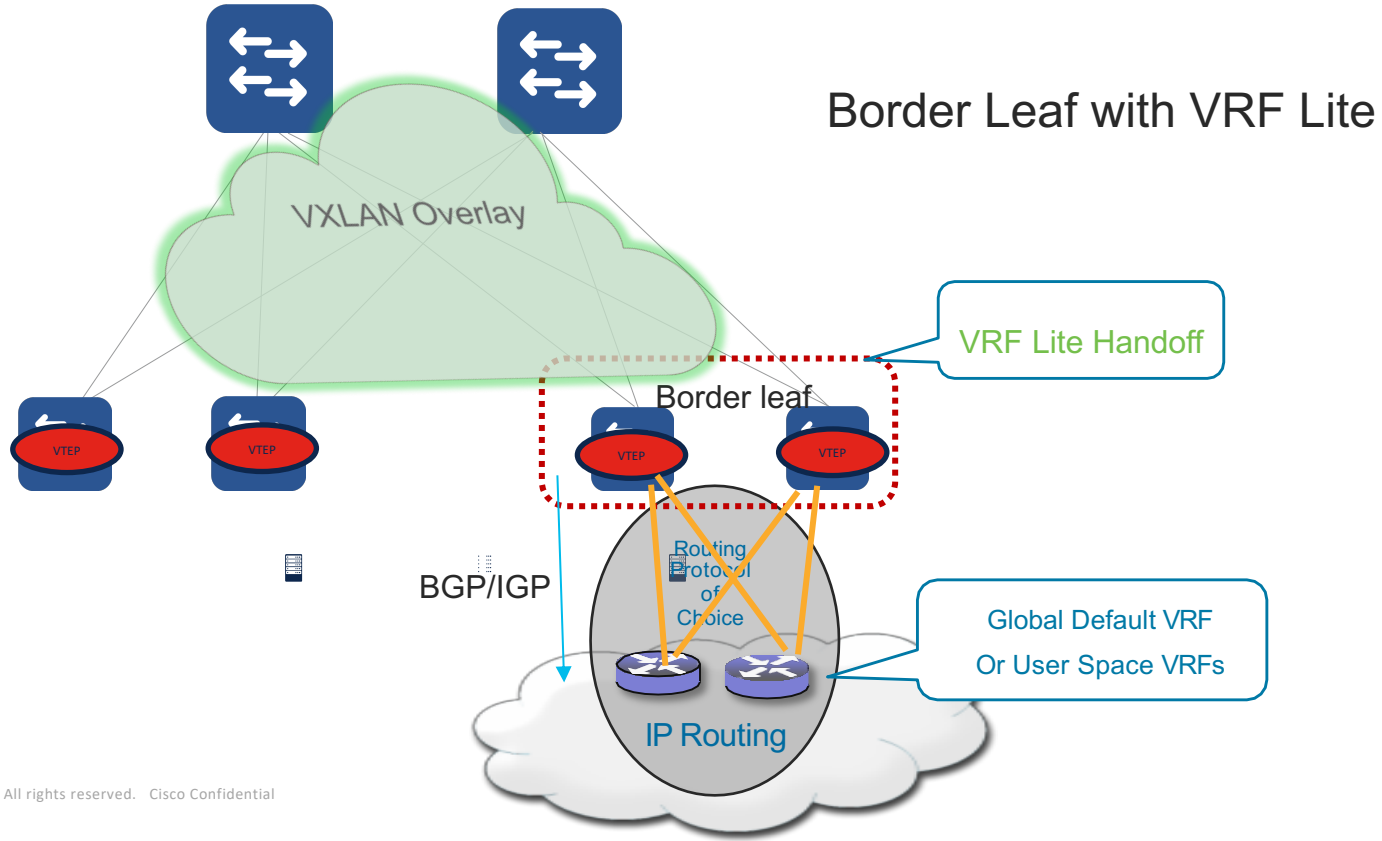
External Connectivity Options

Border Leaf Deployment Options

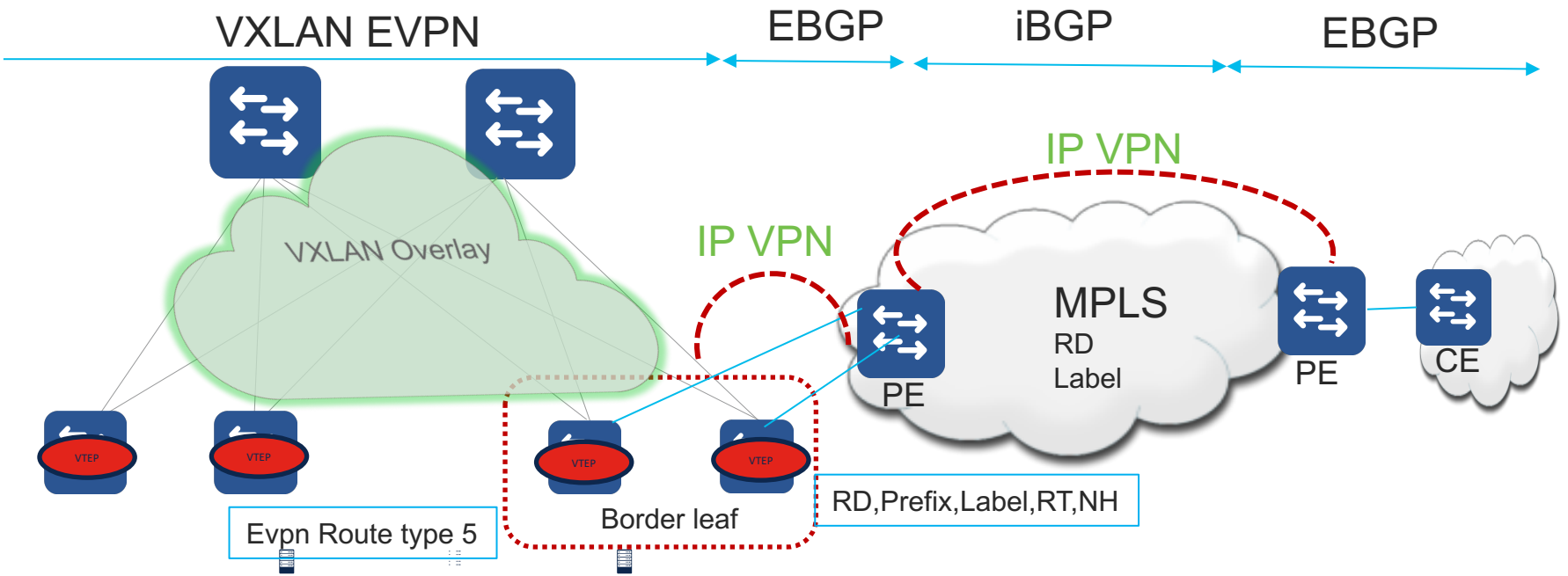
- VRF Lite Handoff
- L3VPN Hanoff
- VPLS Handoff

Border Functionality is supported with Edge/Leaf Switches and also with Spine

EVPN VXLAN Fabric External Routing

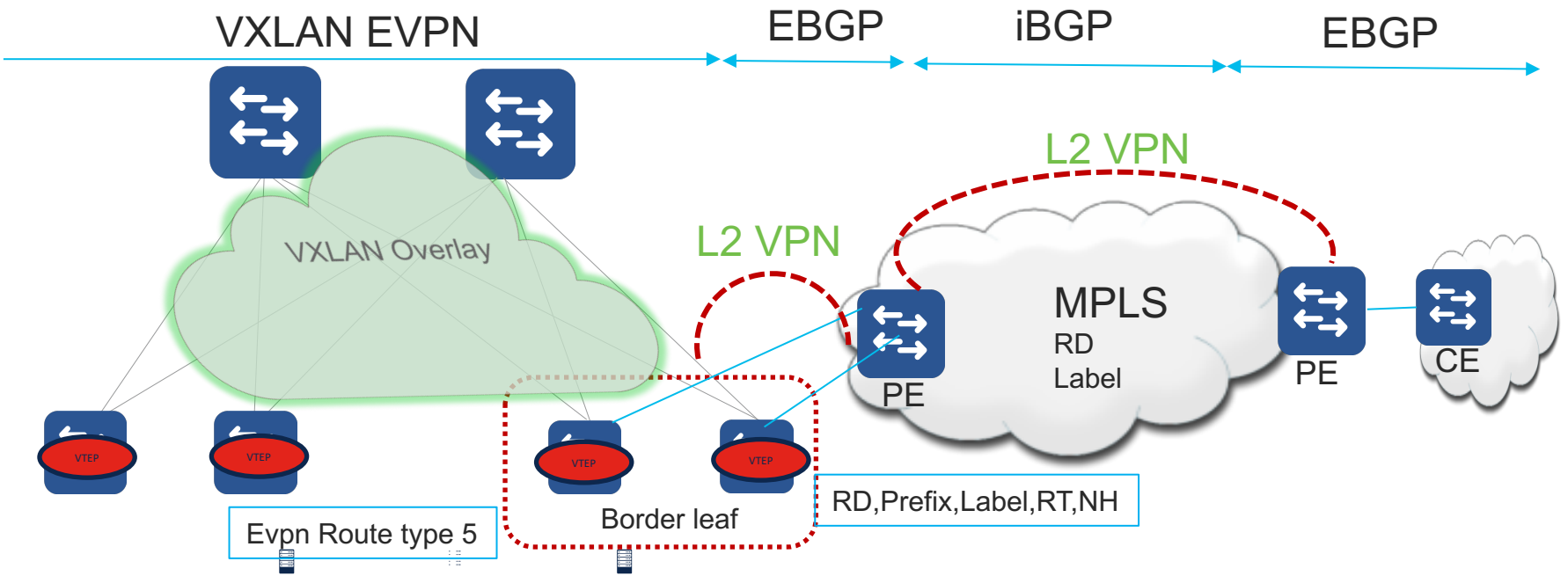


EVPN VXLAN Fabric <> L3VPN Handoff



Single Box Solution – Border Leaf interconnecting EVPN with MPLS L3VPN

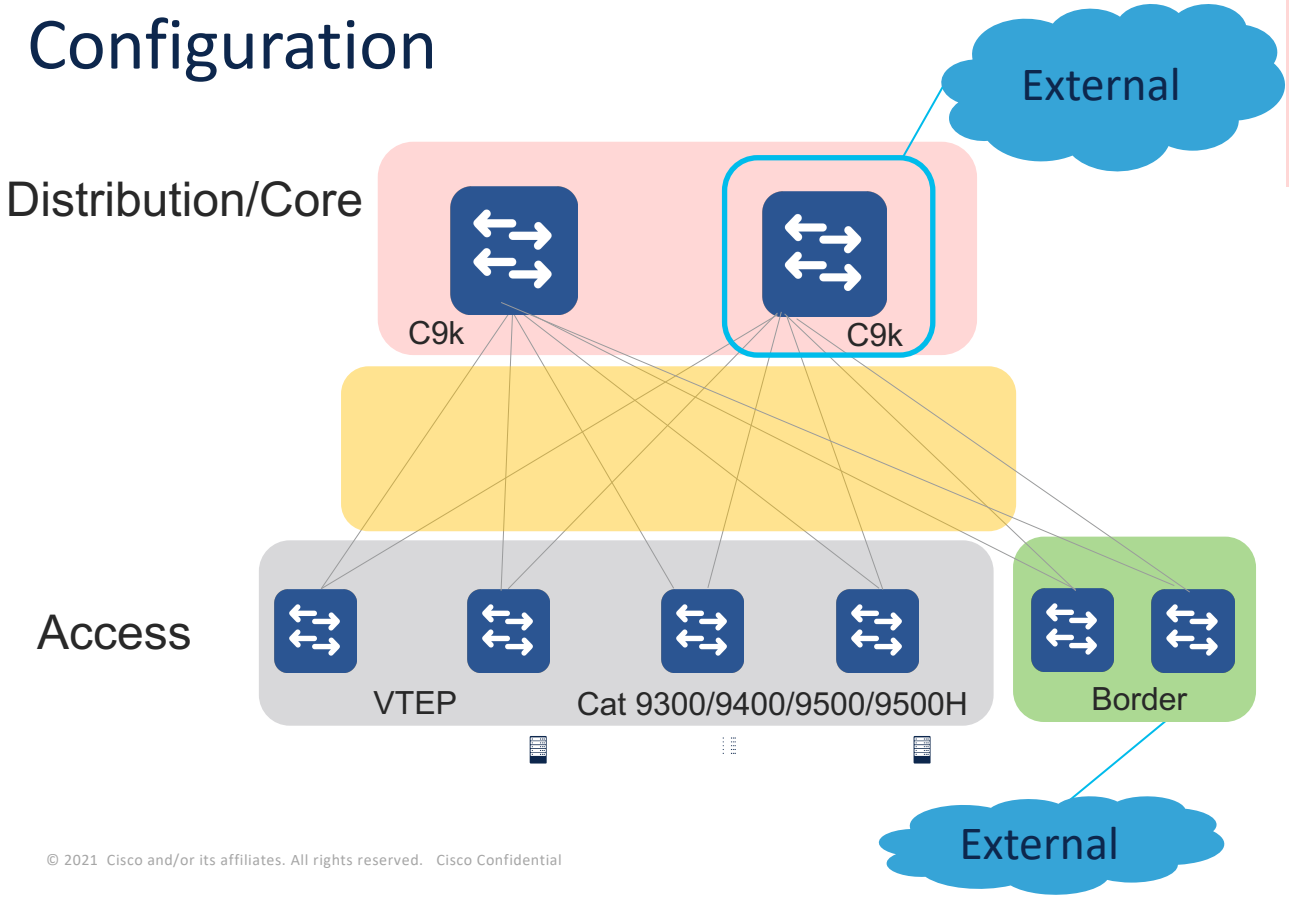
EVPN VXLAN Fabric <> L2VPN Handoff



Single Box Solution – Border Leaf interconnecting EVPN with VPLS

Underlay Designs and Configurations

Design Consideration and Configuration



Supported Core/Spine as Route Reflector

Nexus 9k
Nexus 7k
C9k(16.12.1)

VTEP Only

16.9.1/1 6.10.1

Catalyst 9300
Catalyst 9400
Catalyst 9500/9500H/9600

16.12.1

Border Leaf Support
9500H/9600

17.2.1

Border Spine Support
9500H/9600

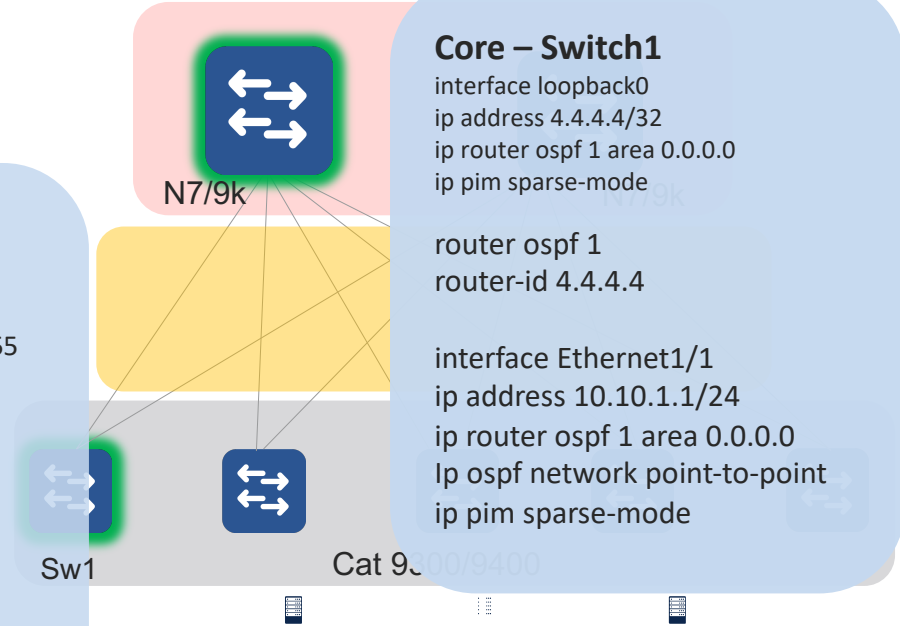
Platforms supported and Scale

- Platforms supported include Catalyst 9300/9400/9500 in 16.9.1
- Catalyst 9500-H is supported from 16.10.1
- Targeted Scale is listed in Table below:

Feature	Scale
Number of VTEPs	**200 (Tested #)**
Number of L2 + L3 VNI	**225**
Number of Access SVIs/Vlans	**225**
Number of TOTAL MAC and IPv4 host routes per VTEP (local + remote)	**32k/64k** **(subject to platform limitation and sdm template, use "show sdm prefer" command to identify Cat9300/9400/9500/9500H/9600 limit)**

Underlay Configuration with OSPF

Distribution/Core



Access Layer – Switch1

```
interface Loopback0
ip address 3.3.3.3 255.255.255.255
ip pim sparse-mode
ip ospf 1 area 0

router ospf 1
router-id 3.3.3.3

interface TenGigabitEthernet1/1/1
description To Core1
no switchport
ip address 10.10.1.2 255.255.255.0
ip pim sparse-mode
ip ospf 1 area 0
Ip ospf network point-to-point
```

Core – Switch1

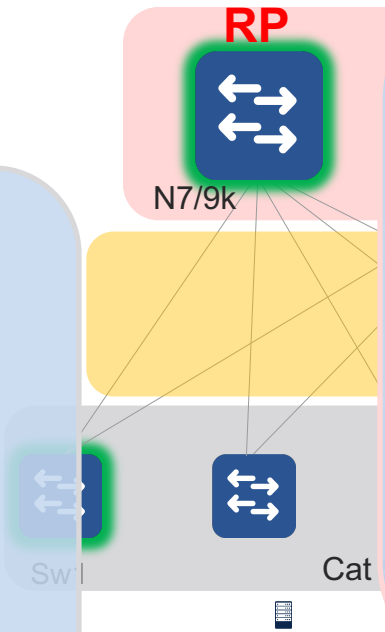
```
interface loopback0
ip address 4.4.4.4/32
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode

router ospf 1
router-id 4.4.4.4

interface Ethernet1/1
ip address 10.10.1.1/24
ip router ospf 1 area 0.0.0.0
Ip ospf network point-to-point
ip pim sparse-mode
```

Underlay Configuration for BUM – PIM ASM

Distribution/Core



Access Layer – Switch1

interface Loopback0
ip pim sparse-mode

interface TenGigabitEthernet1/1/1
description To Core1
ip pim sparse-mode

ip pim rp-address 4.5.4.5

Core – Switch1

interface loopback0
ip pim sparse-mode

interface Ethernet1/1
ip pim sparse-mode

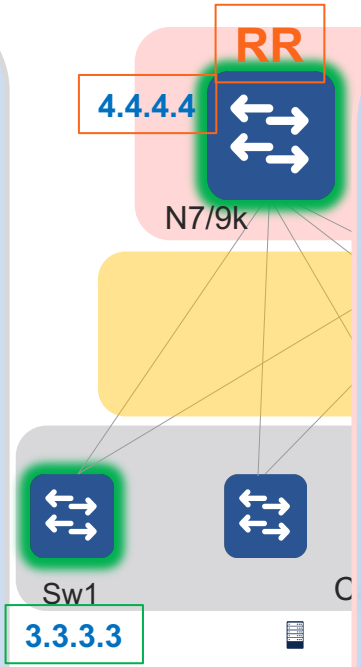
interface loopback1---- **Anycast RP (Seperate Loopback)**
ip address 4.5.4.5/32
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode

ip pim rp-address 4.5.4.5 group-list 224.0.0.0/4 ip pim ssm range 232.0.0.0/8

Overlay Configuration- iBGP Control Plane

Access Layer – Switch1

```
router bgp 1
 bgp router-id 3.3.3.3
 bgp log-neighbor-changes
 neighbor 4.4.4.4 remote-as 1
 neighbor 4.4.4.4 update-source Loopback0
 !
 address-family ipv4
  redistribute connected
  neighbor 4.4.4.4 activate
 exit-address-family
 !
 address-family l2vpn evpn
  neighbor 4.4.4.4 activate
  neighbor 4.4.4.4 send-community both
  neighbor 4.4.4.4 soft-reconfiguration
inbound
 exit-address-family
 !
 address-family ipv4 vrf tenant1
  advertise l2vpn evpn
  redistribute connected
 exit-address-family
```



Core – Switch1

```
router bgp 1
 bgp log-neighbor-changes
 neighbor 3.3.3.3 remote-as 1
 neighbor 3.3.3.3 update-source Loopback0
 neighbor 4.4.4.4 remote-as 1
 neighbor 4.4.4.4 update-source Loopback0
 !
 address-family ipv4
  redistribute connected
  neighbor 3.3.3.3 activate
  neighbor 4.4.4.4 activate
 exit-address-family
 !
 address-family l2vpn evpn
  neighbor 3.3.3.3 activate
  neighbor 3.3.3.3 send-community both
  neighbor 3.3.3.3 route-reflector-client
  neighbor 4.4.4.4 activate
  neighbor 4.4.4.4 send-community both
  neighbor 4.4.4.4 route-reflector-client
  maximum-paths 2
 exit-address-family
```

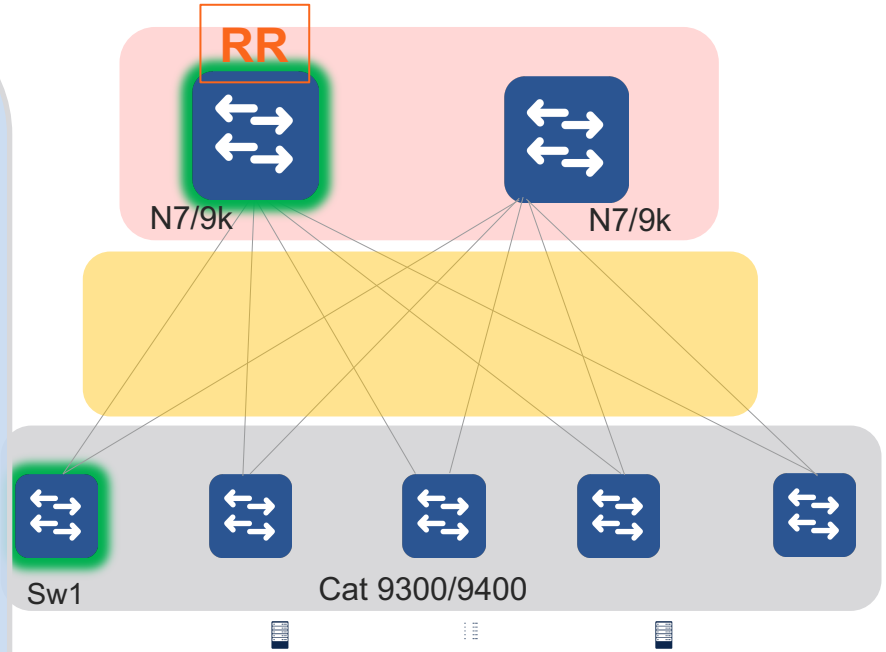
EVPN – Multitenancy Configuration

Access Layer – Switch1

```
vrf definition tenant1
 rd 1:1 – This should be unique per VTEP/vrf
 !
 address-family ipv4
  route-target export 2:2 stitching
  route-target import 2:2 stitching
 exit-address-family
```

Access Layer – Switch2

```
vrf definition tenant1
 rd 2:2
 !
 address-family ipv4
  route-target export 2:2 stitching
  route-target import 2:2 stitching
 exit-address-family
```



Overlay Configuration- Vxlan Data Plane

Switch1/2 - VTEP Configuration

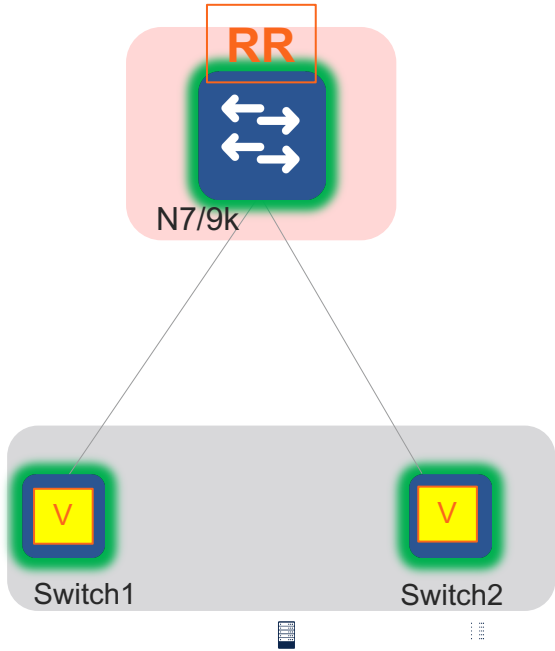
```
l2vpn evpn  -- # EVPN Control Plane
 replication-type static
 router-id Loopback1 – Loopback 1 used for EVPN

! Single Bridge Table/Broadcast Domain ( Vlan to Vni mapping)
l2vpn evpn instance 1 vlan-based
 encapsulation vxlan
!
l2vpn evpn instance 2 vlan-based
 encapsulation vxlan

vlan configuration 11
 member evpn-instance 1 vni 11001 - Vlan <> VNI Mapping for Bridging
vlan configuration 12
 member evpn-instance 2 vni 11002 - Vlan <> VNI Mapping for Bridging

vlan configuration 901 - Vlan <> VNI Mapping for Routing
 member vni 900001

interface nve1 - Nve is logical interface where VXLAN packets are encapsulated and decapsulated.
 no ip address
 source-interface Loopback1
 host-reachability protocol bgp - This means BGP control plane is used to exchange updates
 member vni 11001 mcast-group 239.0.0.1 – vni is mapped to mcast groups for Arp Handling
member vni 11002 mcast-group 239.0.0.1 -          vni is mapped to mcast groups for Arp Handling
 member vni 900001 vrf tenant1 - associate-vrf is used for for layer3 vni
```



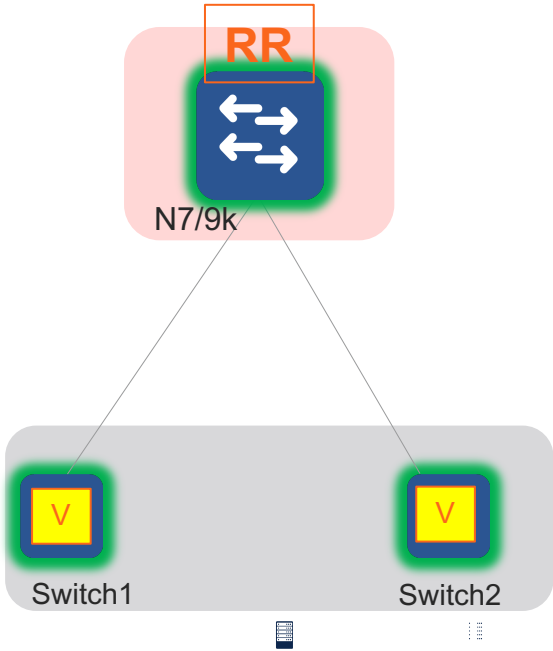
Distributed Anycast Gateway Configuration

Switch1/2 – Anycast Gateway and Host Vlans (To be configured on all VTEPS)

```
interface Vlan10----------------------------access SVI
description connected to 100010 mac-address 0001.0001.0001
vrf forwarding tenant1
ip address 192.168.10.1 255.255.255.0
!
interface Vlan11-----------------------access SVI
description connected to 100011 mac-address 0001.0001.0001
vrf forwarding tenant1
ip address 192.168.11.1 255.255.255.0

interface Vlan501---------------------core SVI
description connected to 50000 vrf forwarding l3vni50000
vrf forwarding tenant1
ip unnumbered Loopback0

interface GigabitEthernet1/0/1------access trunk port
switchport mode trunk !
interface GigabitEthernet1/0/2----access vlan port
 switchport access vlan 10
switchport mode access
```



EVPN VXLAN- “show commands”

EVPN Mgr Show commands

```
# show l2vpn evpn evi [detail]
# show l2vpn evpn mac [detail]
# show l2vpn evpn mac ip [detail]
# show l2vpn evpn summary
# show l2vpn capabilities
```

L2RIB Show commands for EVPN VXLAN-IRB

```
# show l2route evpn summary
# show l2rib topologies [detail]
# show l2route evpn mac [detail]
# show l2route evpn mac ip [detail]
# show  l2rib clients [detail]
# show  l2rib producers [detail]
# show  l2rib registrations [detail]
```

SISF show commands

```
# show device-tracking database
# show device-tracking database mac
```

Multicast RIB show commands

```
# show ip mroute
# show ip mfib
```

IP RIB show commands

```
# show ip route vrf xx (xx is VRF name)
```

IP CEF show commands

```
# show ip cef vrf xx (xx is VRF name)
```

ARP show commands

```
# show arp vrf xx (xx is VRF name)
```

EVPN VXLAN- “show commands”

BGP Show commands for EVPN VXLAN-IRB

```
# show bgp l2vpn evpn
# show bgp l2vpn evpn route-type 2 <filters>
# show bgp l2vpn evpn route-type 5 <filters>
# show bgp l2vpn evpn evi context
# show bgp l2vpn evpn local-vtep vrf <vrf-name>
```

L2FIB Show commands for EVPN VXLAN-IRB

```
# show l2fib bridge-domain <ID> detail
# show l2fib bridge-domain <ID> address unicast <H.H.H>
# show l2fib bridge-domain <ID> vxlan encap
# show l2fib bridge-domain <ID> vxlan decap
# show l2fib bridge-domain <ID> vxlan mac_oce
```

NVE show commands

```
# show nve vni
# show nve vni <id> detail
# show nve peers
```

VLAN MAC Table command in IOS-MATM, FED MATM, SISF

IOS-MATM: # show mac address-table vlan <id>

FED-MATM: # show platform software fed switch active matm macTable vlan xx

(xx is vlan #)

SISF: # show device-tracking database mac

Cont.

EVPN VXLAN- “debug/trace commands”

EVPN Mgr debug/trace commands

```
# debug l2vpn evpn event
# debug l2vpn evpn event detail
# show monitor event-trace evpn event all
# show monitor event-trace evpn error all
```

L2RIB debug/trace commands

```
# debug l2rib event # debug l2rib event detail
# show monitor event-trace l2rib event all
# show monitor event-trace l2rib error all
```

Note: MAC and MAC/IP event traces are not enabled by default in L2RIB. To enable them, use the following configuration:

```
(config)# monitor event-trace l2rib event include detail
(config)# monitor event-trace l2rib event size 1000000
```

BGP debug/trace commands

```
# debug bgp l2vpn evpn evi context detail
# debug bgp l2vpn evpn evi event detail
# debug bgp l2vpn evpn nve detail
# show monitor event-trace bgp l2vpn evpn
```

SISF debug/trace commands

```
# debug device-tracking
# show device-tracking events
```

NVE debug/trace commands

```
# show evpn log event
# show nve log event
# debug nve all
```

EVPN VXLAN – Troubleshooting steps

Step	Verification show command(s)
0- Verify Underlay IGP/BGP or eBGP is configured properly	IGP and BGP show commands to show BGP peers, IGP adjacencies established
1- Verify Underlay Multicast is configured properly	- sh ip mroute - sh ip mfib - sh ip pim rp
2- Verify L2 VNI is provisioned properly in NVE	sh nve vni
3- Verify EVPN Instance is provisioned properly in EVPN Mgr	sh l2vpn evpn evi xx detail
4- Verify L2 Topology for the Access VLAN is properly provisioned in L2RIB	show l2rib topologies detail
5- Verify EVI context is properly added to BGP	show bgp l2vpn evpn evi context
6a- Verify MAC Table in IOS-MATM (local MACs only) **6b- Verify MAC Table in FED-MATM** **6c- Verify MAC entries in SISF**	- show mac address-table vlan xx - show platform software fed switch active matm macTable vlan xx - show device-tracking database mac
7- Verify MAC entries in EVPN Mgr	show l2vpn evpn mac/mac ip (MAC only or MAC/IP route)
8- Verify MAC routes in L2RIB	show l2route evpn mac/mac ip (MAC only or MAC/IP route)

EVPN VXLAN – Troubleshooting steps

Step	Verification show command(s)
9- Verify MAC/IP, Prefix routes in BGP	show bgp l2vpn evpn evi xx (xx is evi #) show bgp l2vpn evpn evi 1 route-type 2 show bgp l2vpn evpn route-type 5
10- Verify MAC routes check in L2FIB	show l2fib bridge-domain xx detail (xx is bridge domain #)
11- Verify Access SVIs, Core SVIs and NVE Interface are all UP	show ip interface brief
12- Verify EVPN Mgr got all L2 and IRB attributes from NVE	show l2vpn evpn evi detail
13- Verify Remote L3 VNIs are received by NVE from BGP	show nve peers
14- Verify Remote MAC/IP route in IP VRF xx	show bgp vpnv4 unicast vrf xx (xx is IP-VRF name)
15- Verify RNH in BGP	show bgp l2vpn evpn rnh vrf xx (xx is IP-VRF name)
16- Verify Remote IP route in RIB	show ip route vrf xx (xx is IP-VRF name)
Execute show tech command	show tech-support evpn show tech-support evpn | redirect xx (xx is location to collect info)

Cont.

The bridge to possible