

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC





EVPN IOS XR Deep Dive for Service Providers and Data Center

Jiri Chaloupka - Principal Technical Marketing Engineer
BRKMPL-2253



#CiscoLiveAPJC

Cisco Webex App

Questions?

Use Cisco Webex App to chat with the speaker after the session

How

- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until November 15, 2024.



<https://ciscolive.ciscoevents.com/ciscolivebot/#BRKMPL-2253>

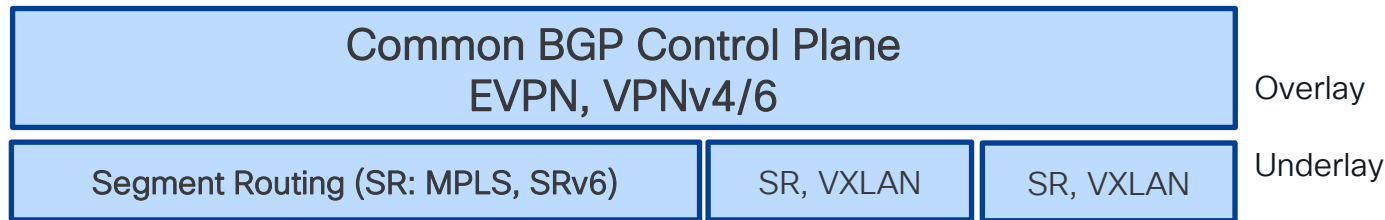
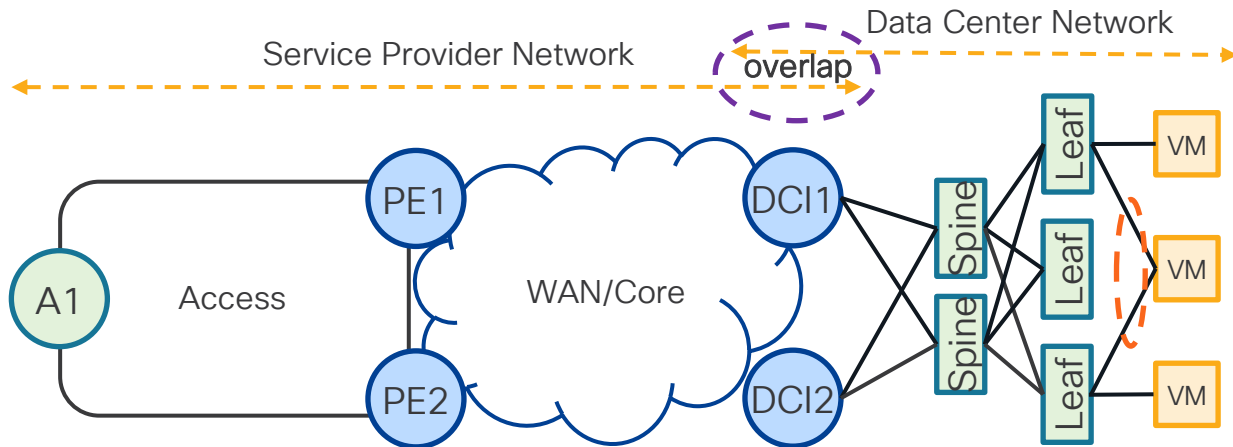


Agenda

- EVPN Basic Principles
- EVPN L2 Multihomed Service
 - All-Active / Single-Active / Port-Active
- EVPN Distributed L3 Anycast Gateway
- EVPN Centralized Gateway
- EVPN L3 Interconnect Options
 - EVPN & VPNv4/6 Interconnect
- EVPN Routes – Summary
- EVPN-VPWS Multihomed Service
 - All-Active / Single-Active
- VPLS to EVPN Seamless Migration
- Pseudowire(PW) to EVPN-VPWS
 - Seamless Migration
- EVPN & VPLS Interconnect
- Summary
- Extra Offline Learning

Unified Control Plane and Data Plane

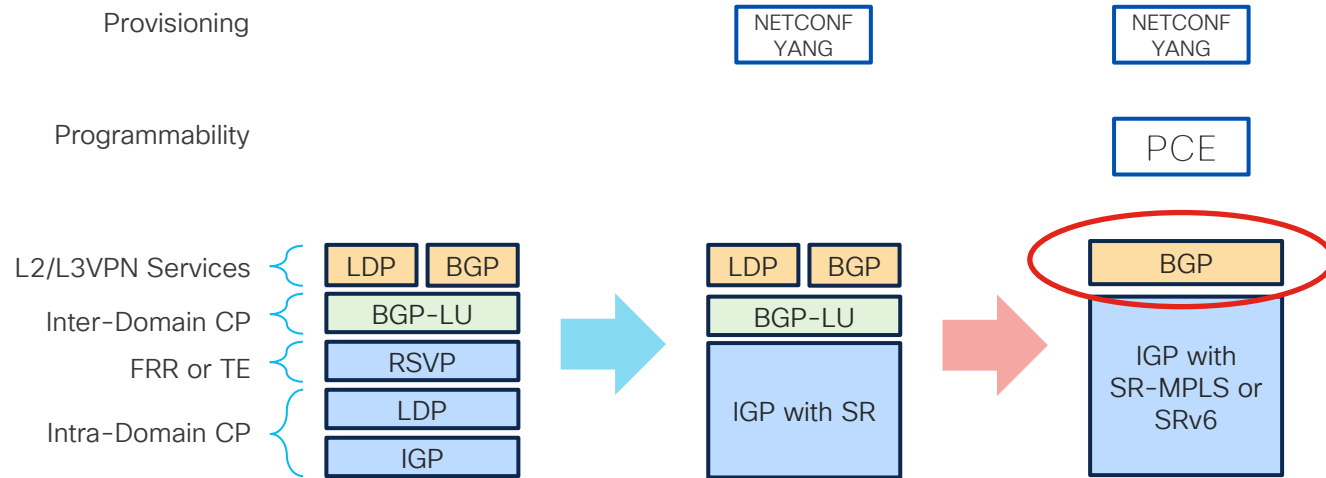
Next Generation
Services Overlay &
Data Plane



Legacy Solution:



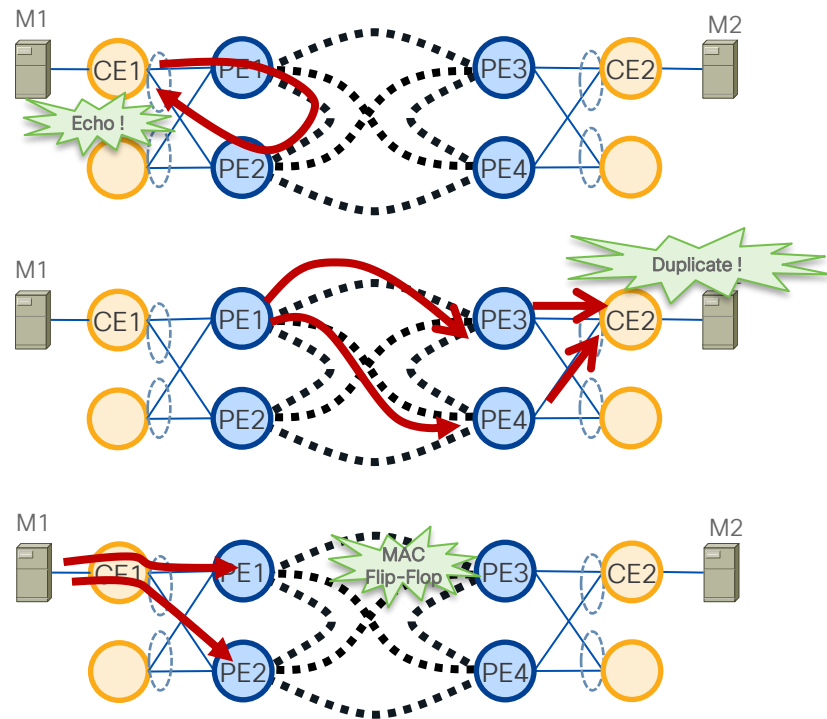
Service Provider Network - Simplification Journey



Next-Generation Solutions for L2VPN

Solving VPLS challenges for per-flow Redundancy

- Existing VPLS solutions do not offer an All-Active per-flow redundancy
- Looping of Traffic Flooded from PE
- Duplicate Frames from Floods from the Core
- MAC Flip-Flopping over Pseudowire
 - E.g. Port-Channel Load-Balancing does not produce a consistent hash-value for a frame with the same source MAC (e.g. non MAC based Hash-Schemes)

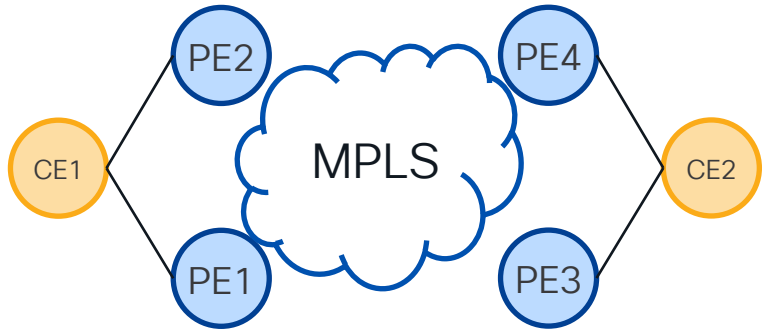


MPLS Transport & BGP Service

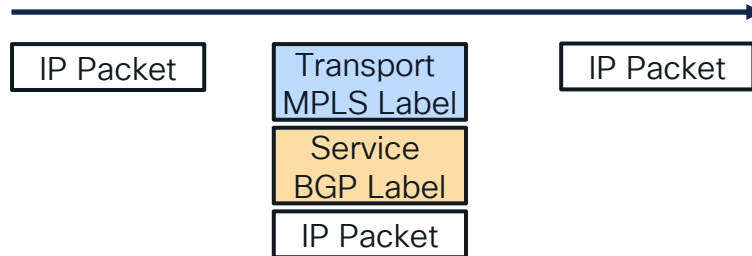
BGP L3VPN/ L3 EVPN

BGP Signaling BGP Signaling BGP Signaling

←-----→ ←-----→ ←-----→



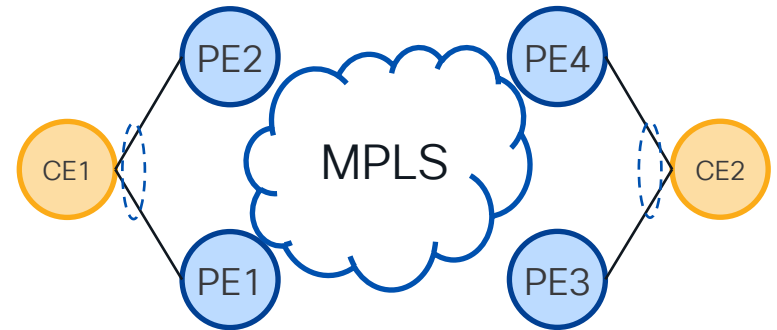
Data Plane



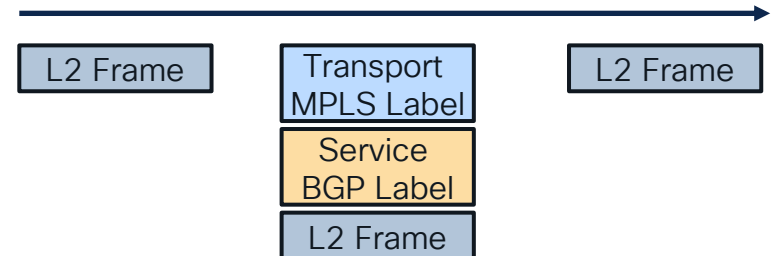
BGP L2VPN EVPN

BGP Signaling

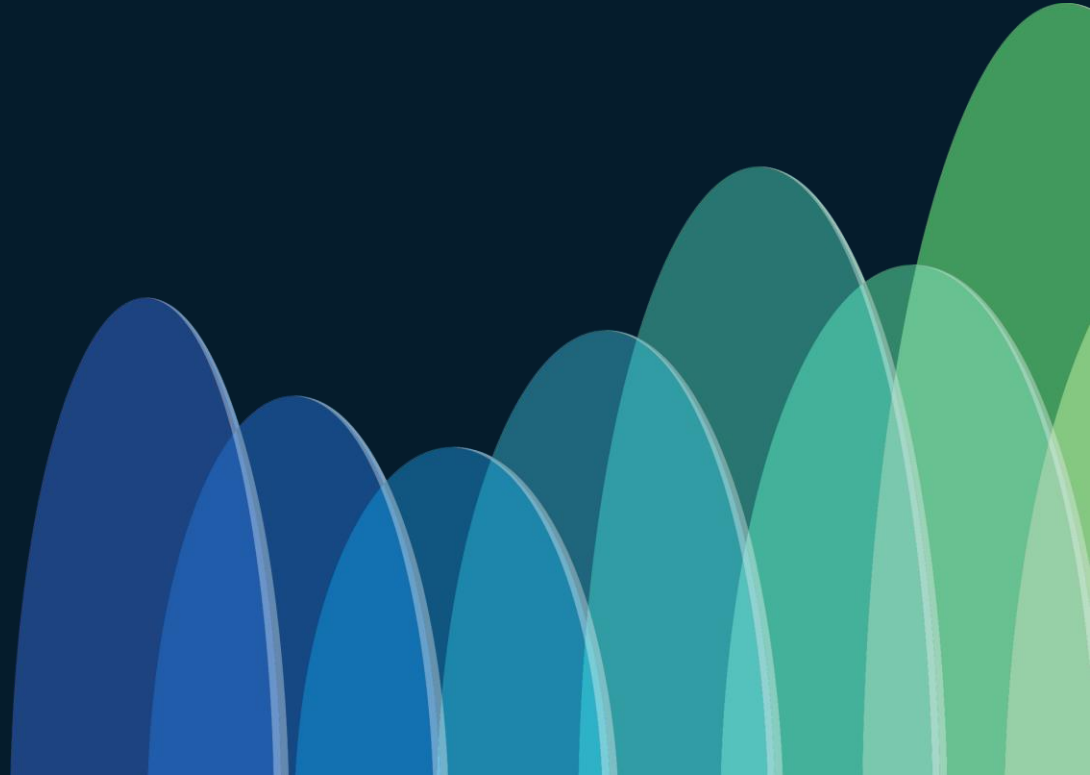
←-----→



Data Plane



EVPN – Basic Principles



EVPN Advantages:

Integrated Services

- Integrated Layer 2 and Layer 3 VPN services
- L3VPN-like principles and operational experience for scalability and control
- All-active Multi-homing & PE load-balancing (ECMP)

Network Efficiency

- Fast convergence (link, node, MAC moves)
- Control-Place (BGP) learning. PWs are no longer used.
- Optimized Broadcast, Unknown-unicast, Multicast traffic delivery

Service Flexibility

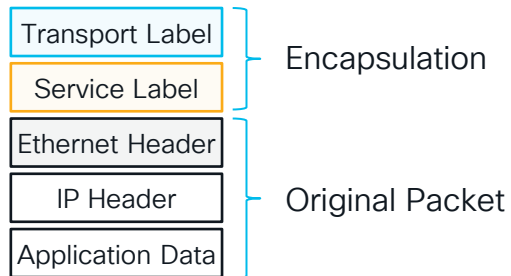
- Choice of MPLS, VxLAN or SRv6 data plane encapsulation
- Support existing and new services types (E-LAN, E-Line, E-TREE)
- Peer PE auto-discovery. Redundancy group auto-sensing

Investment Protection

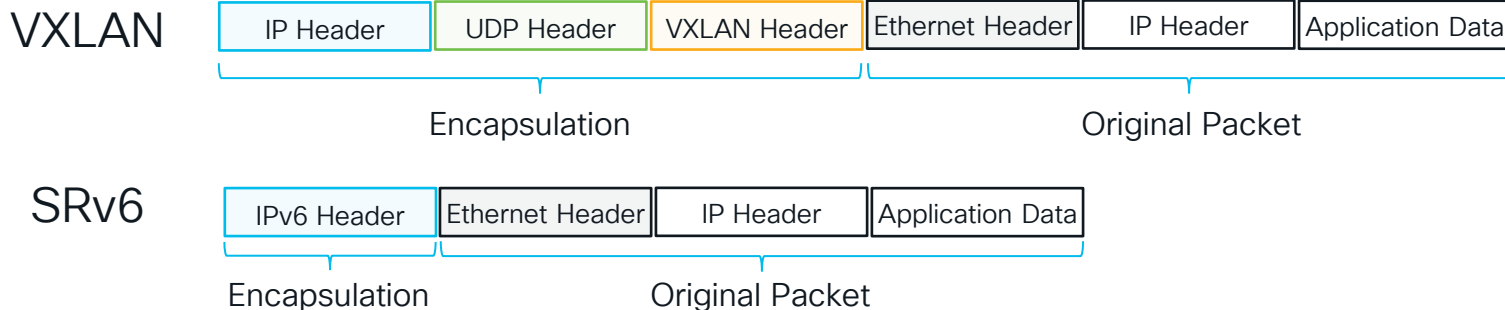
- Fully support IPv4 and IPv6 in the data plane and control plane
- Open-Standard and Multi-vendor support

L2VPN Services Overlay Encapsulation

MPLS Data Plane

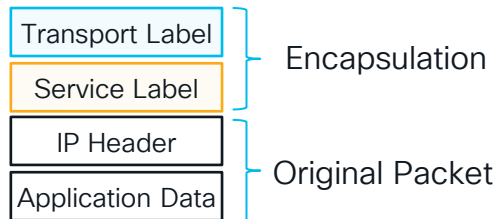


IP Data Plane



L3VPN Services Overlay Encapsulation

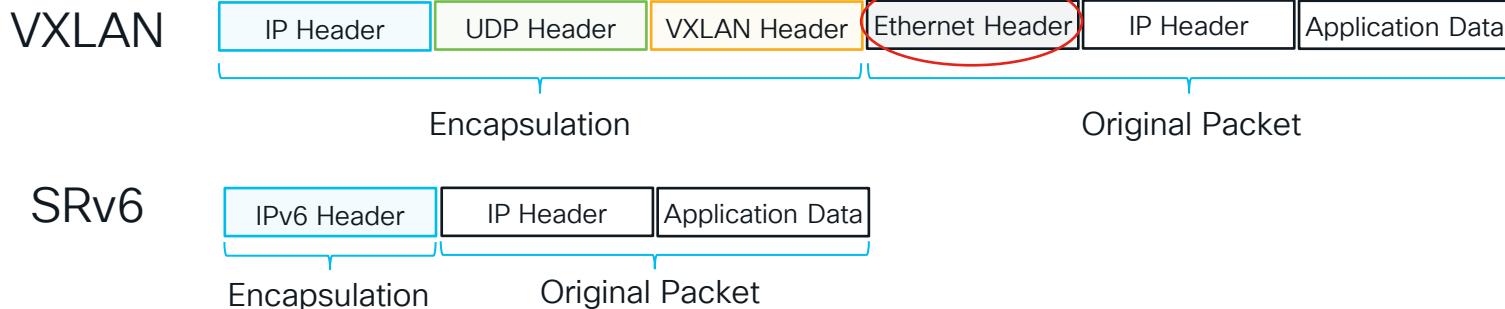
MPLS Data Plane



IP Data Plane

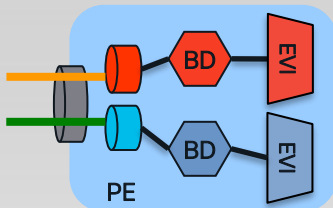
VXLAN RFC7348 - Requires Inner Ethernet

- Additional overhead for L3VPN / IP Forwarding



Concepts

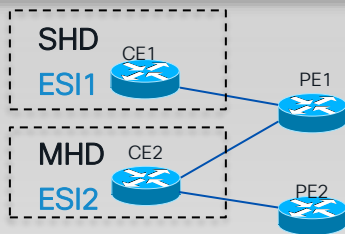
EVPN Instance (EVI)



- EVI identifies a VPN in the network
- Encompass one or more bridge-domains, depending on service interface type

Port-based
VLAN-based (shown above)
VLAN-bundling

Ethernet Segment



- Represents a 'site' connected to one or more PEs
- Uniquely identified by a 10-byte global Ethernet Segment Identifier (ESI)
- Could be a single device or an entire network
 - Single-Homed Device (SHD)
 - Multi-Homed Device (MHD)
 - Single-Homed Network (SHN)
 - Multi-Homed Network (MHN)

BGP Routes

Route Types
[1] Ethernet Auto-Discovery (AD) Route
[2] MAC/IP Advertisement Route
[3] Inclusive Multicast Route
[4] Ethernet Segment Route
[5] IP Prefix Advertisement Route

- New SAFI [70]
- Routes serve control plane purposes, including:
 - MAC address reachability
 - MAC mass withdrawal
 - Split-Horizon label adv.
 - Aliasing
 - Multicast endpoint discovery
 - Redundancy group discovery
 - Designated forwarder election
 - IP address reachability
 - L2/L3 Integration

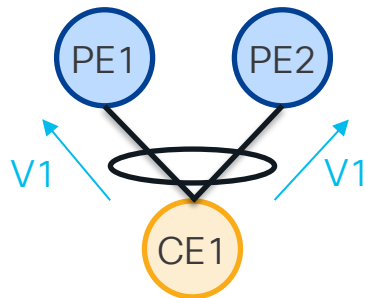
BGP Route Attributes

Extended Communities
ESI MPLS Label
ES-Import
MAC Mobility
Default Gateway
Encapsulation

- New BGP extended communities defined
- Expand information carried in BGP routes, including:
 - MAC address moves
 - Redundancy mode
 - MAC / IP bindings of a GW
 - Split-horizon label encoding
 - Data plane Encapsulation

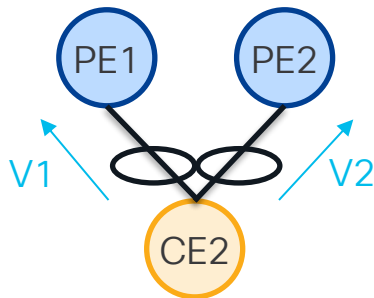
EVPN - Load-Balancing Modes

All-Active
(per flow)



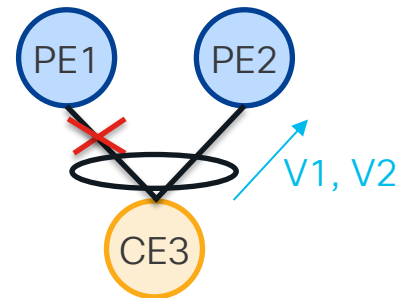
Single LAG at the CE
VLAN goes to both PE
Traffic hashed per flow
Benefits: Bandwidth, Convergence

Single-Active
(per VLAN)



Multiple LAGs at the CE
VLAN active on single PE
Traffic hashed per VLAN
Benefits: Billing, Policing

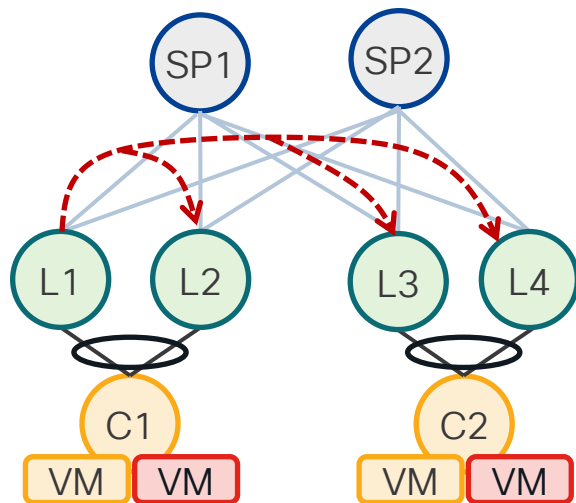
Port-Active
(per port)



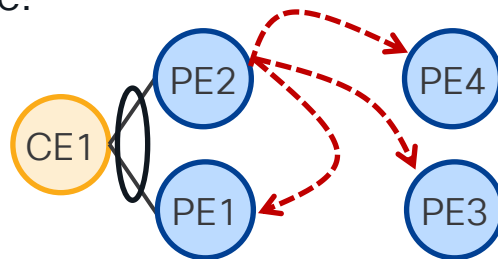
Single LAGs at the CE
Port active on single PE
Traffic hashed per port
Benefits: Protocol Simplification

EVPN – Ethernet VPN

- Concepts are same!!! Pick your side!

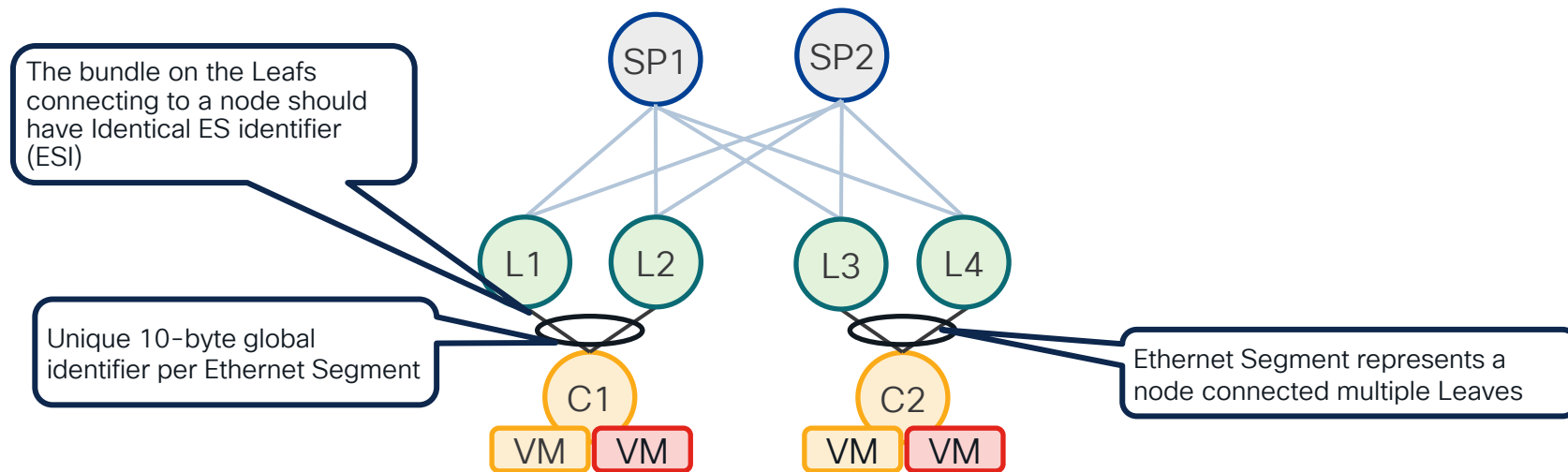


Pick your side!



EVPN – Ethernet-Segment for Multi-Homing

L1 and L2 (L3 and L4) have to know if they multi-home same broadcast domain



EVPN – Ethernet VPN

MAC address advertisement and MAC address table synchronization

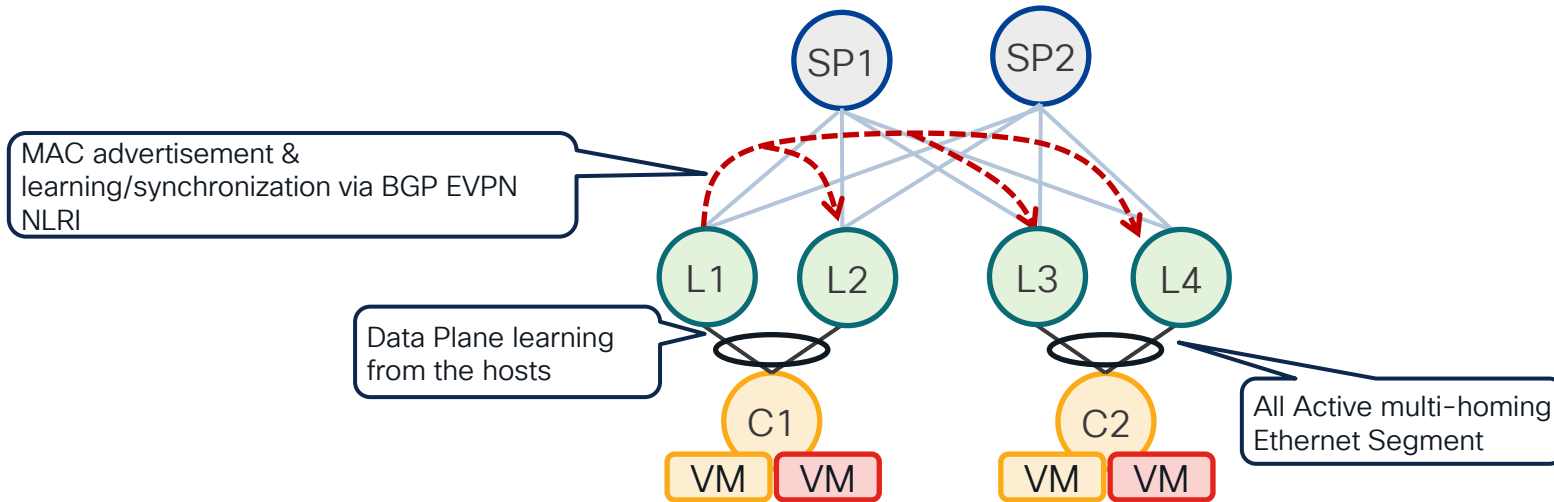
Leaves run Multi-Protocol BGP to advertise & learn MAC addresses over the Network

MAC addresses are advertised to rest of Leaves

L3/4 – Learn MAC address advertised by L1

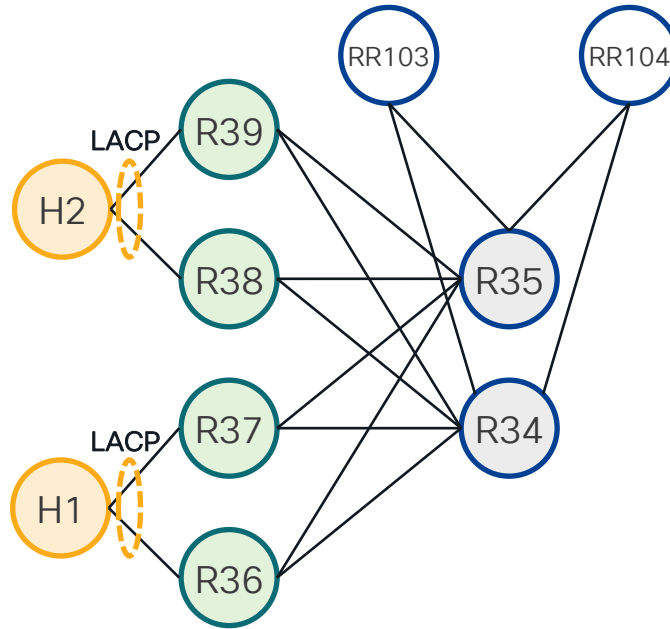
L2 – uses MAC address advertised by L1 to synchronize MAC address table

-> L2 forwards MAC via local ETH interface represented by same Ethernet Segment between L1 and L2



EVPN L2 All-Active Multihomed Service

EVPN - Testbed



EVPN Configuration

CE has to receive same lacp system MAC

```
lacp system mac 3637.3637.3637

interface Bundle-Ether100
  l2transport
  !
!

evpn
  evi 100
    advertise-mac
    !
  interface Bundle-Ether100
    ethernet-segment
      identifier type 0 36.37.00.00.00.00.11.00
    !
  !
```

RT-2 MAC advertise

```
l2vpn
  bridge group 100
  bridge-domain 100
    interface Bundle-Ether100
      !
      evi 100
      !
      !
      !
      !
```

EVPN Configuration - BGP

```
router bgp 1
  bgp router-id 3.3.3.36
  address-family l2vpn evpn
  !
  neighbor-group rr
  remote-as 1
  update-source Loopback0
  address-family l2vpn evpn
  !
  neighbor 3.3.3.103
  use neighbor-group rr
  !
  neighbor 3.3.3.104
  use neighbor-group rr
  !
  !
```

BGP EVPN CP

EVPN – Designated Forwarder (DF)

Challenge:

How to prevent duplicate copies of flooded traffic from being delivered to a multi-homed Ethernet Segment?

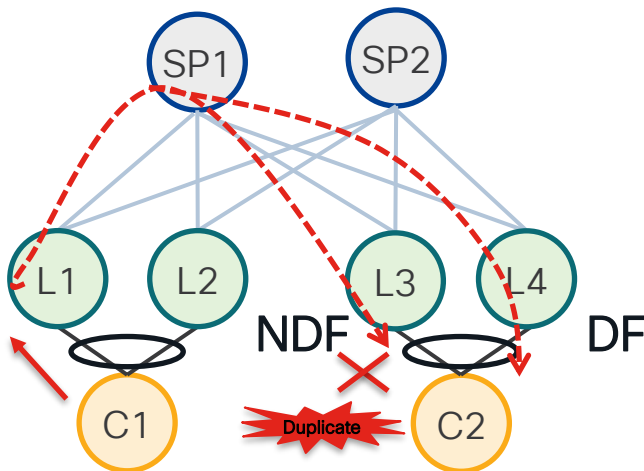
If (L3 and L4) Multi-Homing access via same Ethernet Segment -> only one of them can forward traffic to access
Same for (L1 and L2)

Why extra BUM Label?

What if Unicast Traffic is sent to L3 or L4 (not flooded)? -> DF Election applies only to BUM (from Core to Access)

DF, Redirect, Fast Re-Route (FRR), etc.

Service Label informs egress Leaf if traffic is BUM or Unicast



DF Election per EVI/ESI - Algorithm

Service Carving

Nodes	Position		EVI
R36	0	+	100
R37	1		

EVI-ID modulo Number of Nodes = Position
 $100 \text{ modulo } 2 = 0$

R36 is DF for EVI-100

Who will be DF for EVI-101?

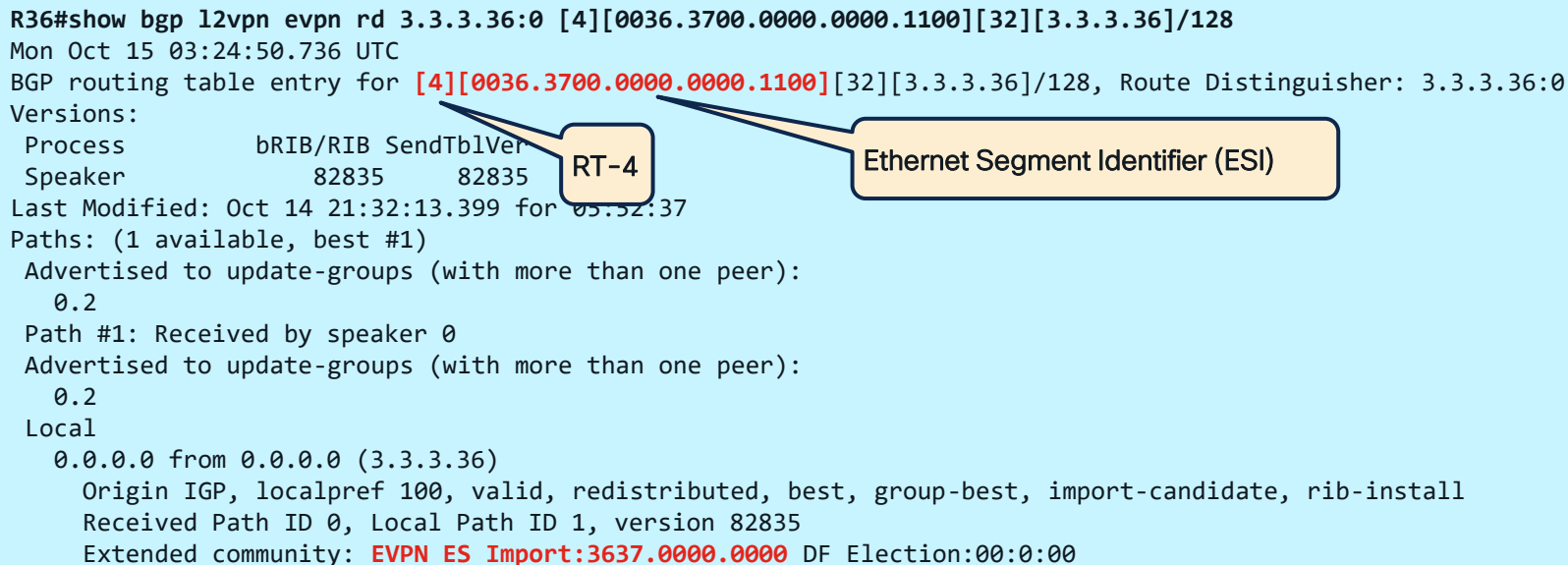
Ethernet Segment - DF Election

```
R36#show evpn ethernet-segment esi 0036.3700.0000.0000.1100 carving detail
.....
Ethernet Segment Id      Interface      Nexthops
-----
-
0036.3700.0000.0000.1100 BE100      3.3.3.36
                                   3.3.3.37

ES to BGP Gates : Ready
ES to L2FIB Gates : Ready
Main port      :
  Interface name : Bundle-Ether100
  Interface MAC  : 008a.9644.d8dd
  IfHandle       : 0x0800001c
  State          : Up
  Redundancy     : Not Defined
ESI type       : 0
  Value          : 36.3700.0000.0000.1100
ES Import RT   : 3637.0000.0000 (from ESI)
Source MAC     : 0000.0000.0000 (N/A)
Topology       :
  Operational    : MH, All-active
  Configured     : All-active (AApF) (default)
Service Carving : Auto-selection
Peering Details : 3.3.3.36[MOD:P:00] 3.3.3.37[MOD:P:00]
Service Carving Results:
  Forwarders    : 1
  Permanent     : 0
  Elected      : 1
    EVI E :      100
  Not Elected  : 0
MAC Flushing mode : STP-TCN
Peering timer    : 3 sec [not running]
Recovery timer   : 30 sec [not running]
Carving timer    : 0 sec [not running]
Local SHG label  : 64005
Remote SHG labels : 1
    64005 : nexthop 3.3.3.37
```

R36: RT-4 Ethernet Segment Router

```
R36#show bgp l2vpn evpn rd 3.3.3.36:0 [4][0036.3700.0000.0000.1100][32][3.3.3.36]/128
Mon Oct 15 03:24:50.736 UTC
BGP routing table entry for [4][0036.3700.0000.0000.1100][32][3.3.3.36]/128, Route Distinguisher: 3.3.3.36:0
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker          82835      82835
Last Modified: Oct 14 21:32:13.399 for 05:52:37
Paths: (1 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.2
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.2
Local
  0.0.0.0 from 0.0.0.0 (3.3.3.36)
  Origin IGP, localpref 100, valid, redistributed, best, group-best, import-candidate, rib-install
  Received Path ID 0, Local Path ID 1, version 82835
  Extended community: EVPN ES Import:3637.0000.0000 DF Election:00:0:00
```



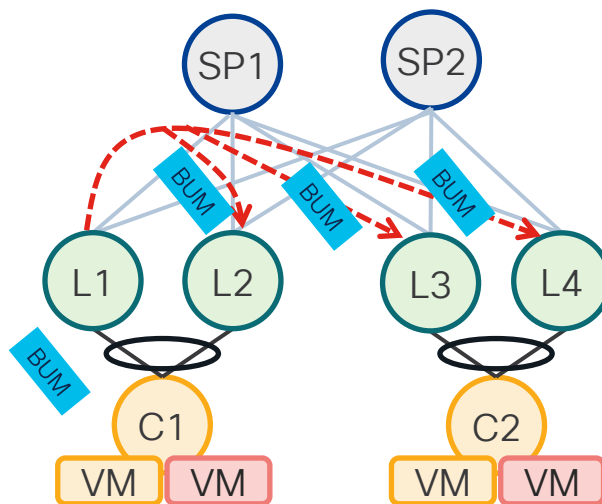
Nodes which share same ESI import this route

EVPN – BUM Ingress Replication

Two service labels per EVPN instance

BUM Label – to forward Broadcast, Unknown Unicast and Multicast

Unicast Label – to forward Unicast



R36: RT-3 Inclusive Multicast

```
R36#show bgp l2vpn evpn rd 3.3.3.36:100 [3][0][32][3.3.3.36]/80
Mon Oct 15 13:10:17.010 UTC
BGP routing table entry for [3][0][32][3.3.3.36]/80, Route Distinguisher: 3.3.3.36:100
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker          39774      39774
Last Modified: Aug 31 01:37:02.399 for 6w3d
Paths: (1 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.2
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.2
Local
  0.0.0.0 from 0.0.0.0 (3.3.3.36)
  Origin IGP, localpref 100, valid, redistributed, best, group-best, import-candidate
  Received Path ID 0, Local Path ID 1, version 39774
  Extended community: RT:1:100
  PMSI: flags 0x00, type 6, label 64120, ID 0x03030324
```

RT-3

EVI 100 Route-Target

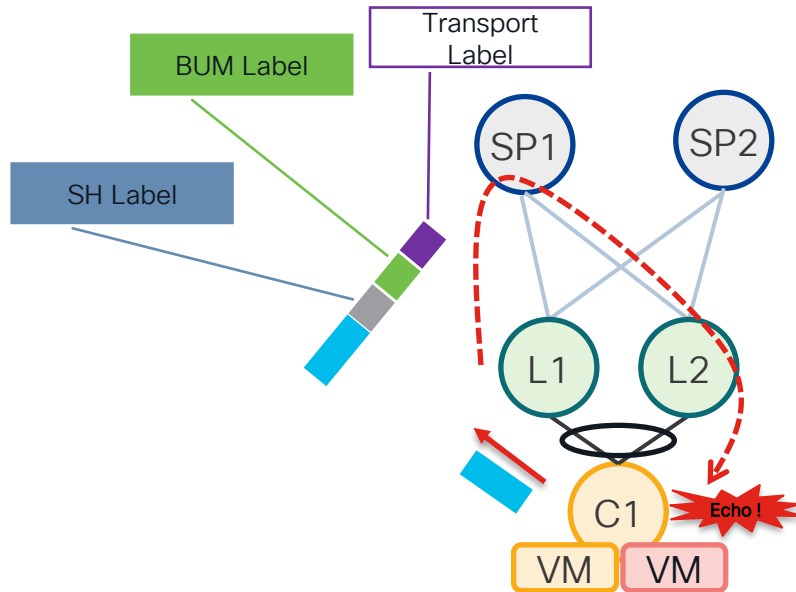
Ingress Replication

Multicast (BUM) Label

EVPN – Split Horizon

Challenge:

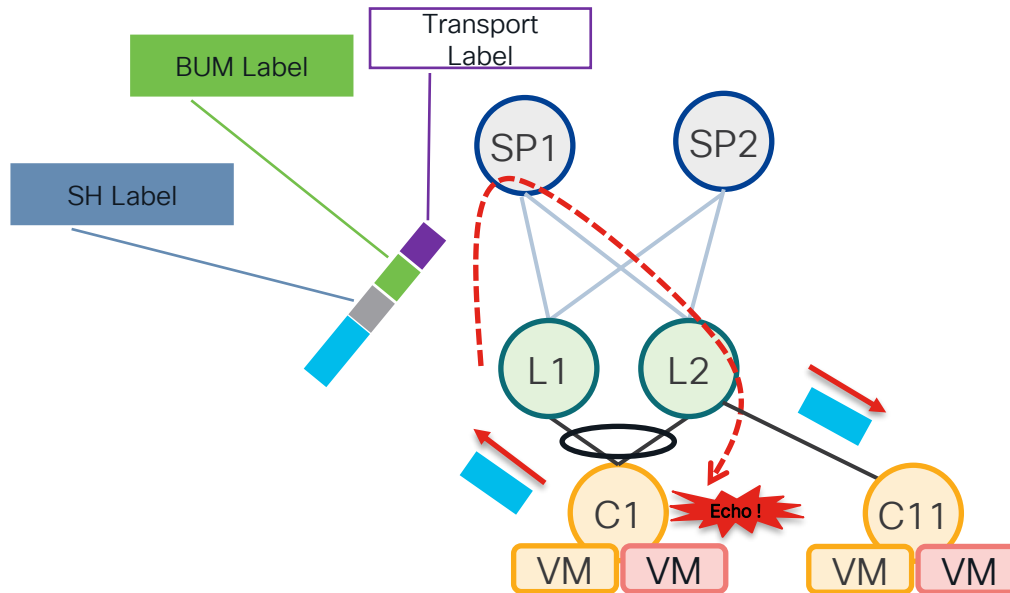
How to prevent flooded traffic from echoing back to a multi-homed Ethernet Segment?



EVPN – Split Horizon

Challenge:

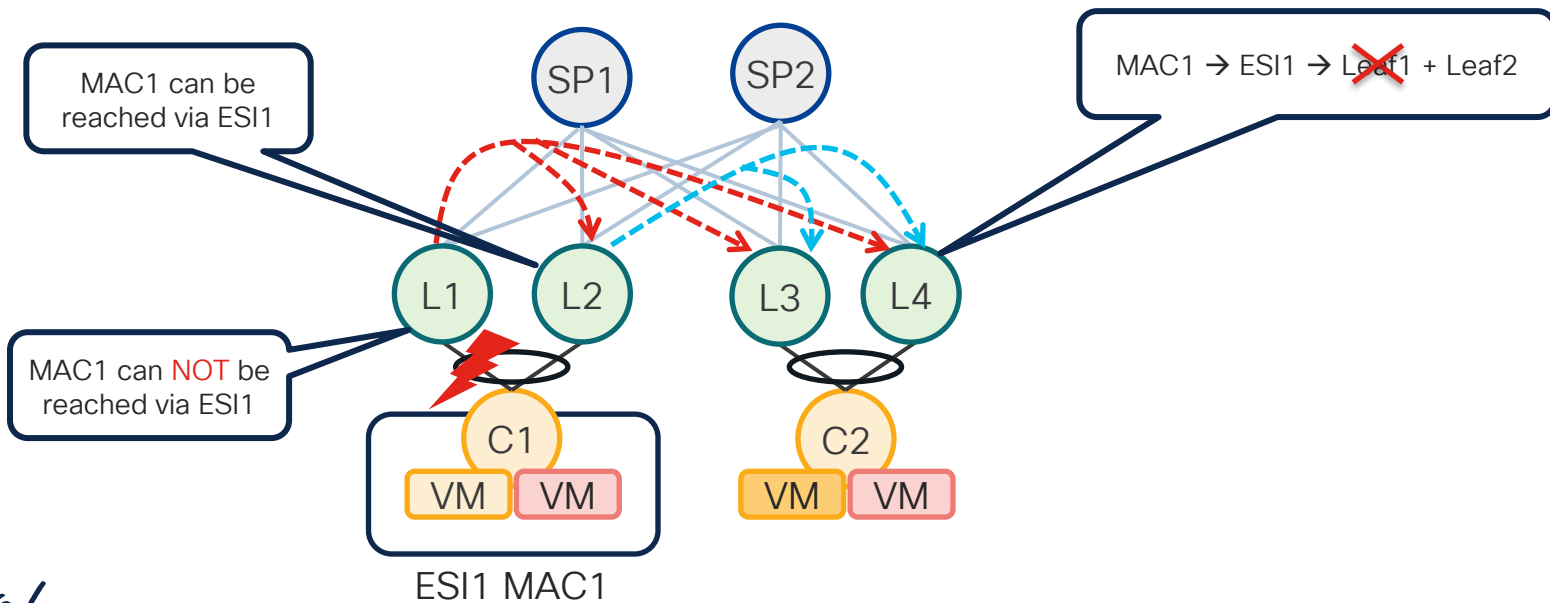
How to prevent flooded traffic from echoing back to a multi-homed Ethernet Segment?



EVPN – MAC Mass-Withdraw

Challenge:

How to inform other Leafs of a failure affecting many MAC addresses quickly while the control-plane re-converges?



R36: RT-1 Per ESI Ethernet Auto-Discovery

```
R36#show bgp l2vpn evpn rd 3.3.3.36:0 [1][3.3.3.36:1][0036.3700.0000.0000.1100][4294967295]/184
```

```
Sun Oct 14 20:56:59.687 UTC
```

```
BGP routing table entry for [1][3.3.3.36:1][0036.3700.0000.0000.1100][4294967295]/184, Route Distinguisher: 3.3.3.36:0
```

```
Versions:
```

```
Process          bRIB/RIB SendTblVer-  
Speaker          76372    76372
```

RT-1

RD - unique per advertising
node (R36 unique)

Ethernet Segment Identifier (ESI)

```
Local Label: 0
```

```
Last Modified: Sep 18 23:02:40.399 for 3w4d
```

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

```
Advertised to update-groups (with more than one peer):
```

```
0.2
```

```
Path #1: Received by speaker 0
```

```
Advertised to update-groups (with more than one peer):
```

```
0.2
```

```
Local
```

```
0.0.0.0 from 0.0.0.0 (3.3.3.36)
```

```
Origin IGP, localpref 100, valid, redistributed, best, group-best, import-candidate, rib-install
```

```
Received Path ID 0, Local Path ID 1, version 76372
```

```
Extended community: EVPN ESI Label:0x00:64005 RT:1:100
```

EVI(s) Route-Target
All EVI(s) which use this ESI

Redundancy mode
All-Active: 0x00
Single-Active: 0x01

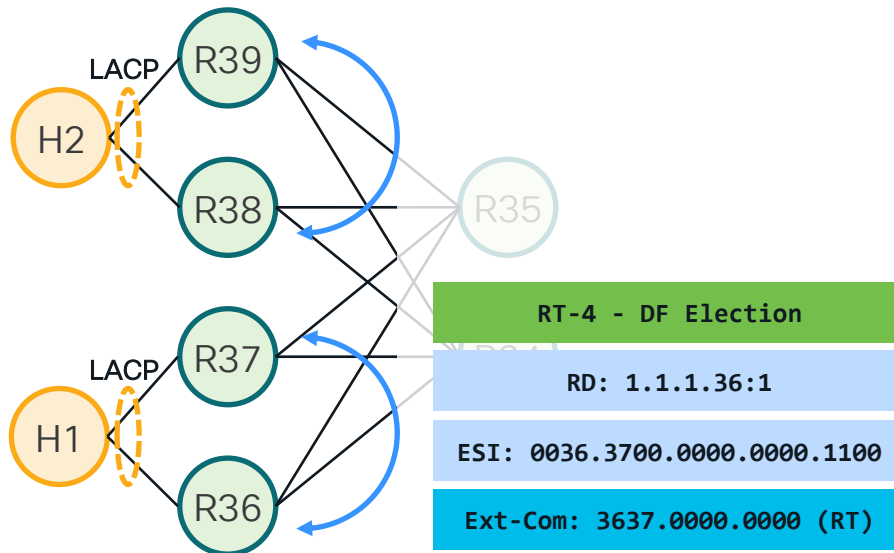
Split-Horizon Label

R36, R37, R38, R39 – EVPN Startup

R36 – Example

1. RT4: DF Election & Multi-Homed Ethernet Segment Auto-Discovery

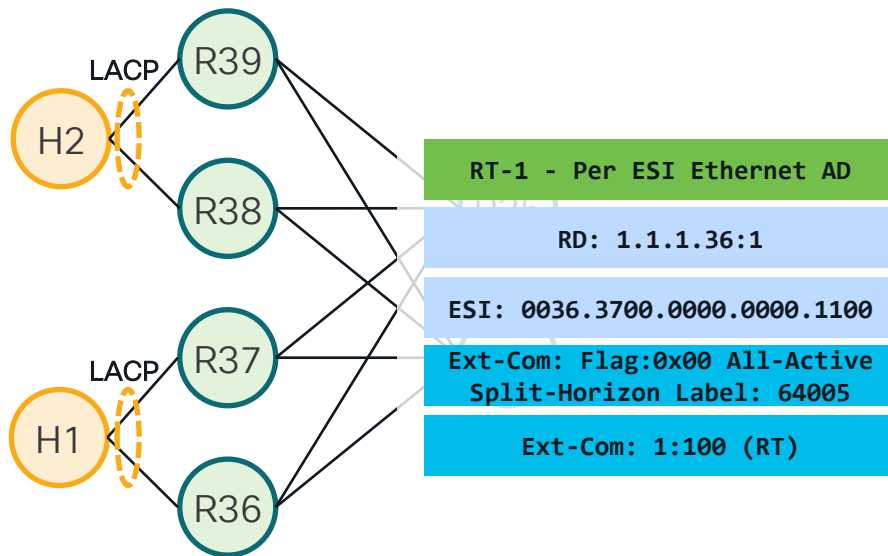
Service Carving: $100 \bmod 2 = 0$
R36 is DF for EVI-100



R36, R37, R38, R39 – EVPN Startup

R36 – Example

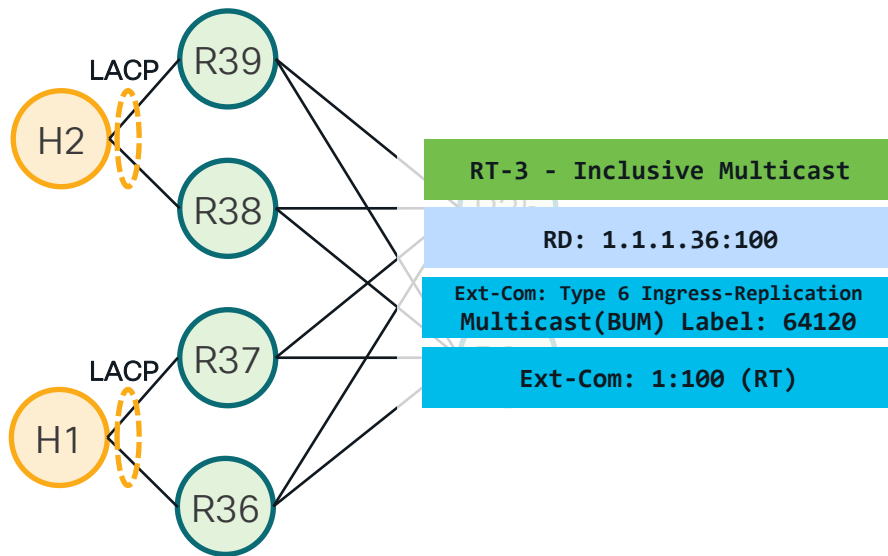
1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)



R36, R37, R38, R39 – EVPN Startup

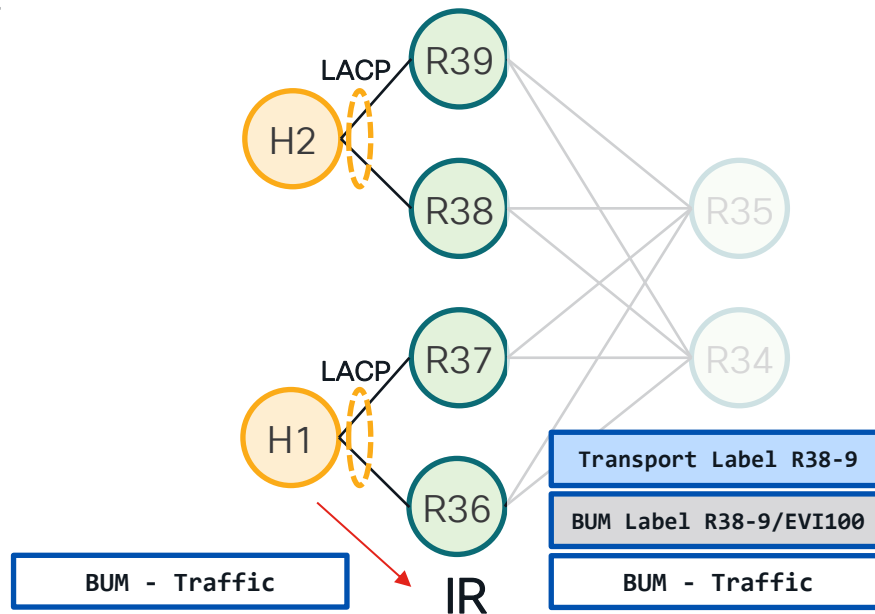
R36 – Example

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast



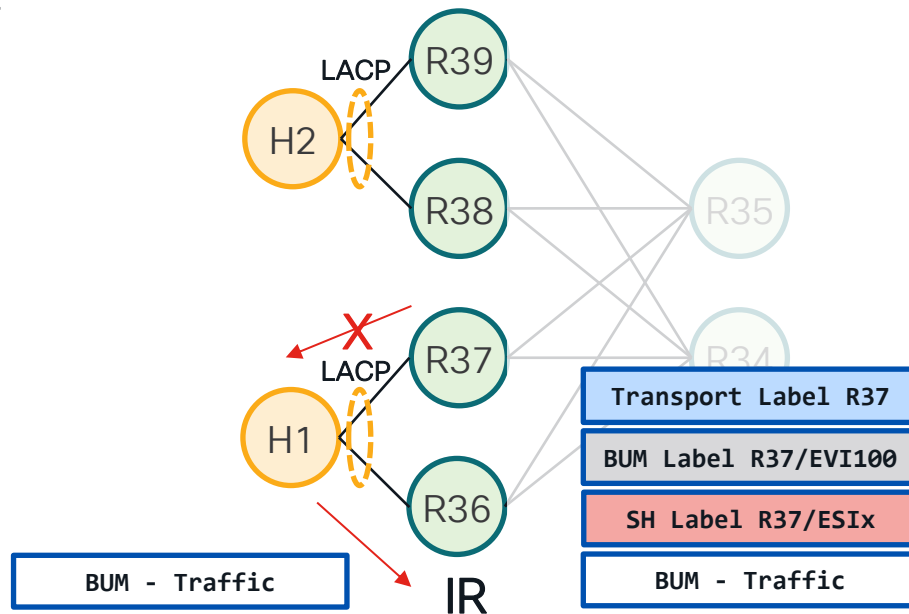
BUM Forwarding

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast



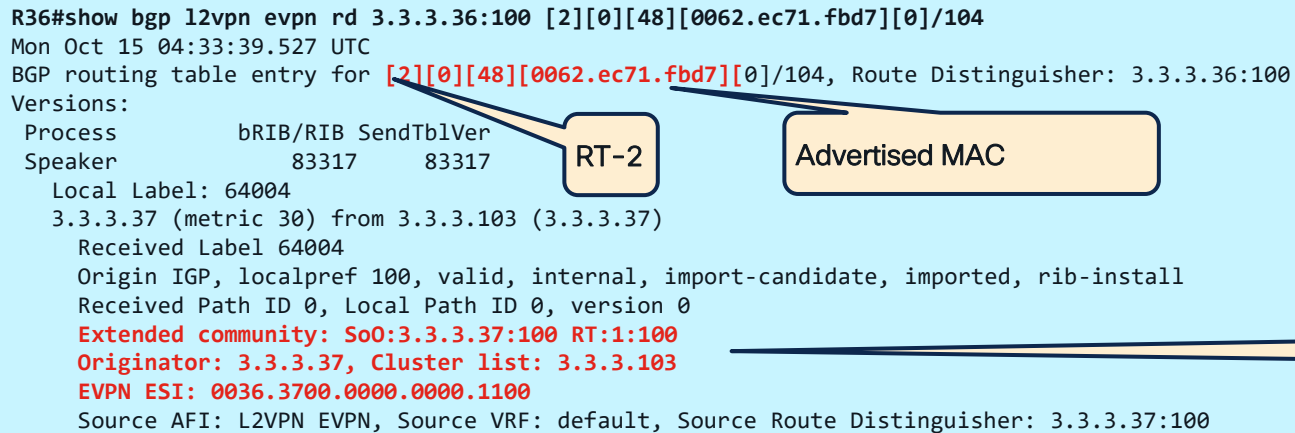
BUM Forwarding

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast



R36: RT-2 MAC Advertisement

```
R36#show bgp l2vpn evpn rd 3.3.3.36:100 [2][0][48][0062.ec71.fbd7][0]/104
Mon Oct 15 04:33:39.527 UTC
BGP routing table entry for [2][0][48][0062.ec71.fbd7][0]/104, Route Distinguisher: 3.3.3.36:100
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker          83317      83317
    Local Label: 64004
    3.3.3.37 (metric 30) from 3.3.3.103 (3.3.3.37)
    Received Label 64004
    Origin IGP, localpref 100, valid, internal, import-candidate, imported, rib-install
    Received Path ID 0, Local Path ID 0, version 0
    Extended community: So0:3.3.3.37:100 RT:1:100
    Originator: 3.3.3.37, Cluster list: 3.3.3.103
    EVPN ESI: 0036.3700.0000.0000.1100
    Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 3.3.3.37:100
```



R36: RT-2 MAC Advertisement

R36#show evpn evi mac

Mon Oct 15 20:57:14.505 UTC

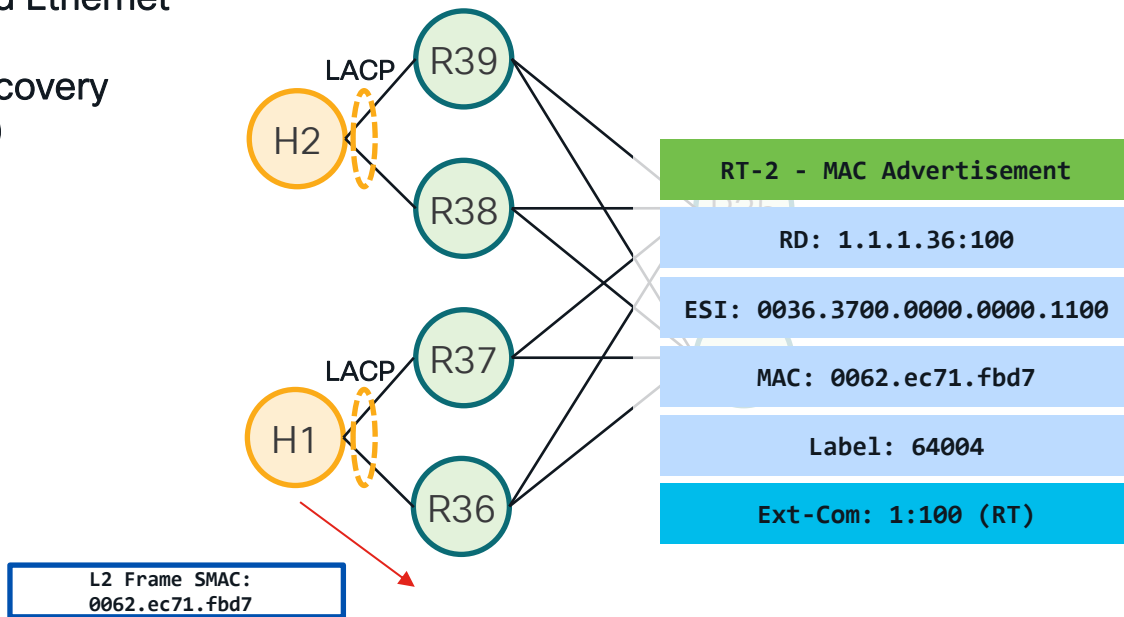
VPN-ID	Encap	MAC address	IP address	Nexthop	Label
100	MPLS	0062.ec71.1000 ::		3.3.3.38	64006
100	MPLS	0062.ec71.1000 ::		3.3.3.39	64006
100	MPLS	0062.ec71.fbd7 ::		3.3.3.37	64004
100	MPLS	0062.ec71.fbd8 ::		Bundle-Ether100	64004
100	MPLS	0062.ec71.fbd9 ::		3.3.3.37	64004
100	MPLS	0062.ec71.fbe0 ::		3.3.3.38	64006
100	MPLS	0062.ec71.fbe0 ::		3.3.3.39	64006
100	MPLS	0062.ec71.fbe1 ::		3.3.3.38	64006
100	MPLS	0062.ec71.fbe1 ::		3.3.3.39	64006

Learned and Advertised
MAC

R36, R37, R38, R39 – EVPN Startup

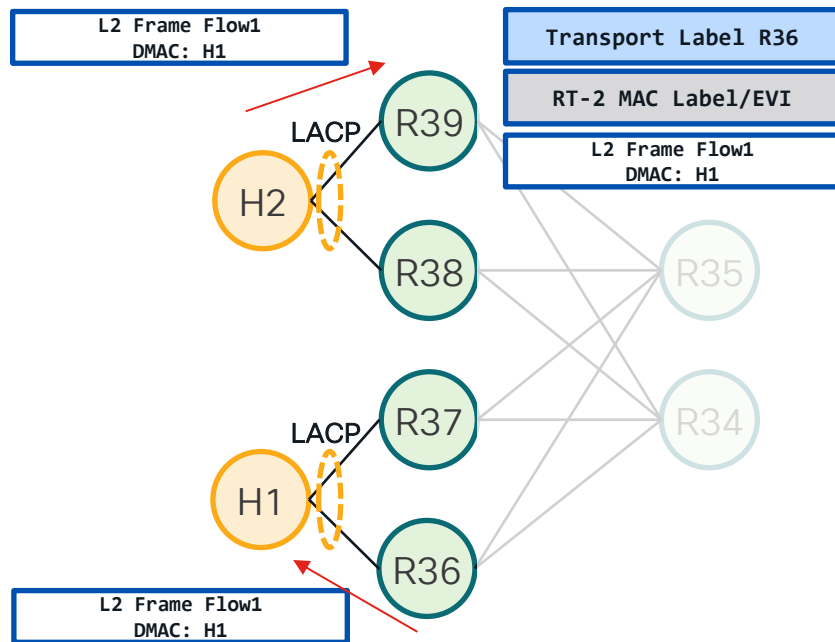
R36 – Example

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast
4. **RT2**: MAC Advertisement



Unicast Forwarding

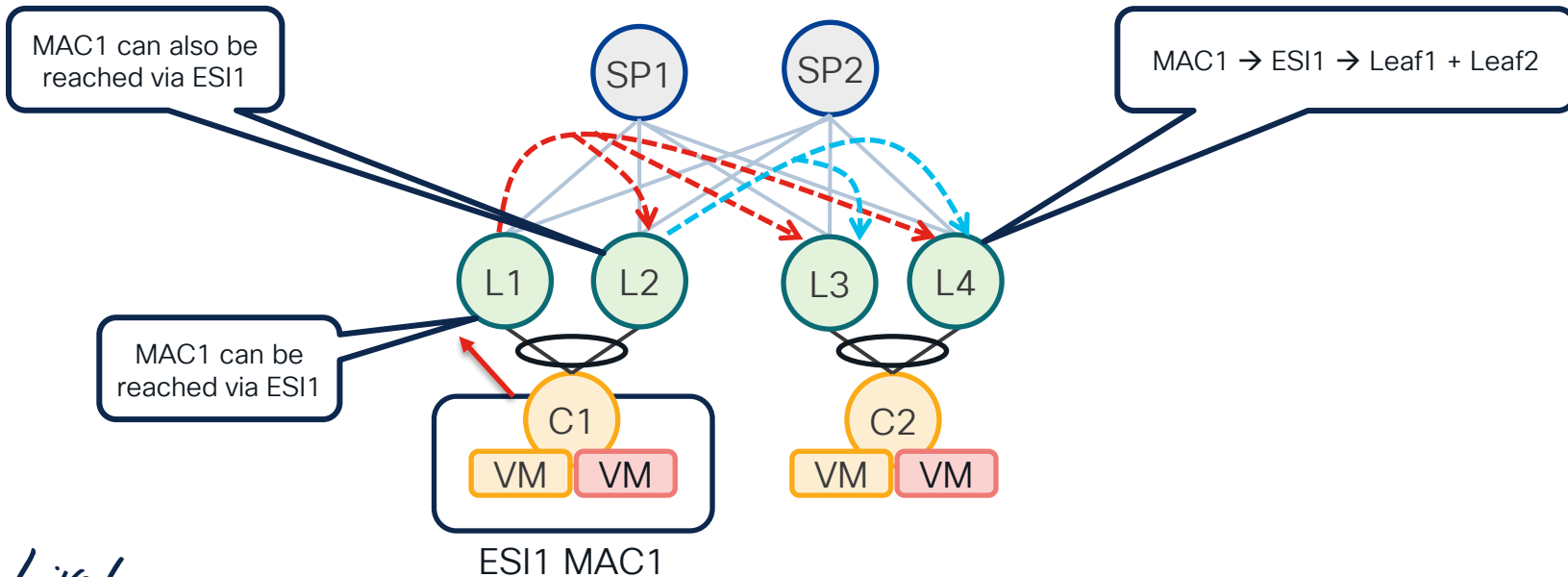
1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast
4. **RT2**: MAC Advertisement



EVPN – Aliasing

Challenge:

How to load-balance traffic towards a multi-homed device across multiple Leafs when MAC addresses are learnt by only a single Leaf?



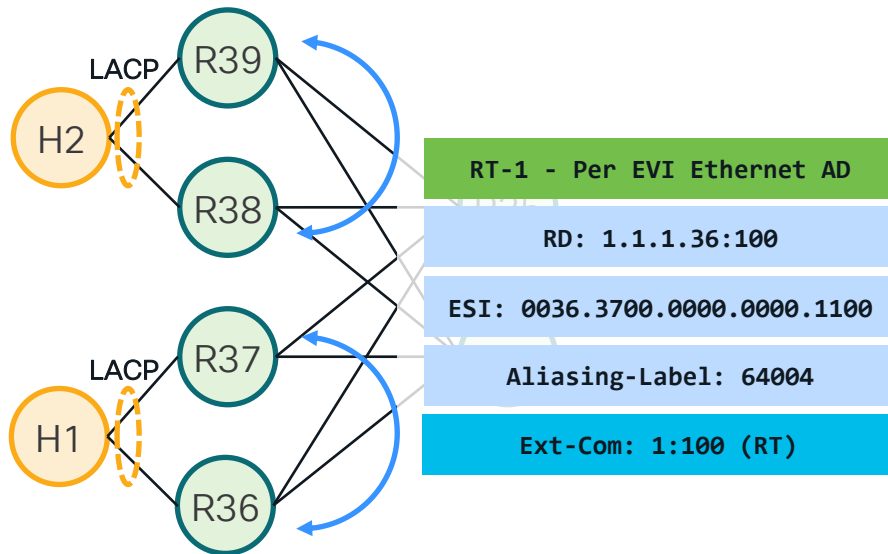
R36: RT-1 Per EVI Ethernet Auto-Discovery

```
RP/0/RP0/CPU0:R36#show bgp l2vpn evpn rd 3.3.3.36:100 [1][0036.3700.0000.0000.1100][0]/120
Mon Oct 15 03:35:13.604 UTC
BGP routing table entry for [1][0036.3700.0000.0000.1100][0]/120, Route Distinguisher: 3.3.3.36:100
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker           79640      7964
Last Modified: Oct 12 17:40:06.399 for 2d09h
Paths: (2 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.2
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.2
  Local
    0.0.0.0 from 0.0.0.0 (3.3.3.36)
      Origin IGP, localpref 100, valid, redistributed, best, group-best, import-candidate, rib-install
      Received Path ID 0, Local Path ID 1, version 39769
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local
    3.3.3.37 (metric 30) from 3.3.3.103 (3.3.3.37)
      Received Label 64004
      Origin IGP, localpref 100, valid, internal, import-candidate, imported, rib-install
      Received Path ID 0, Local Path ID 0, version 0
      Extended community: RT:1:100
      Originator: 3.3.3.37, Cluster list: 3.3.3.103
      Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 3.3.3.37:100
```

R36, R37, R38, R39 – EVPN Startup

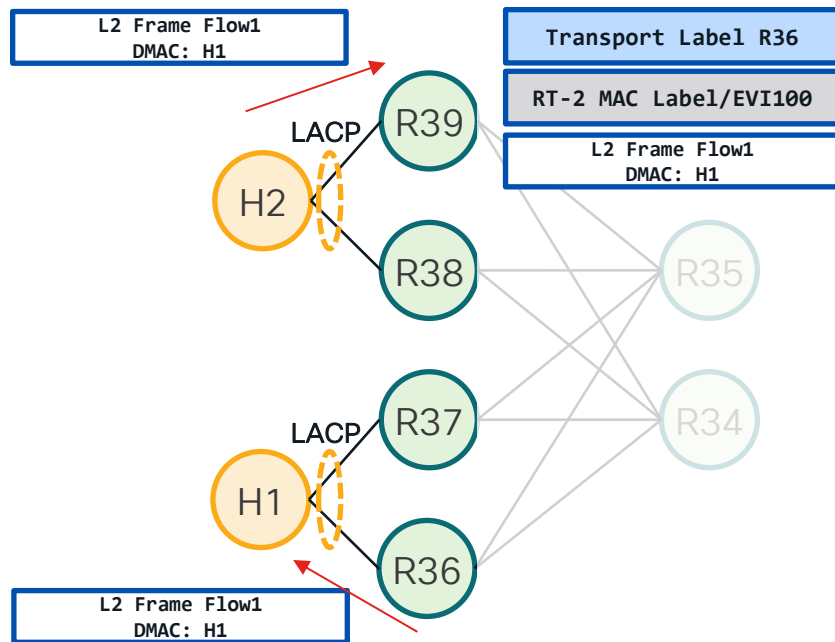
R36 – Example

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast
4. **RT2**: MAC Advertisement
5. **RT1**: Per EVI Ethernet Auto-Discovery



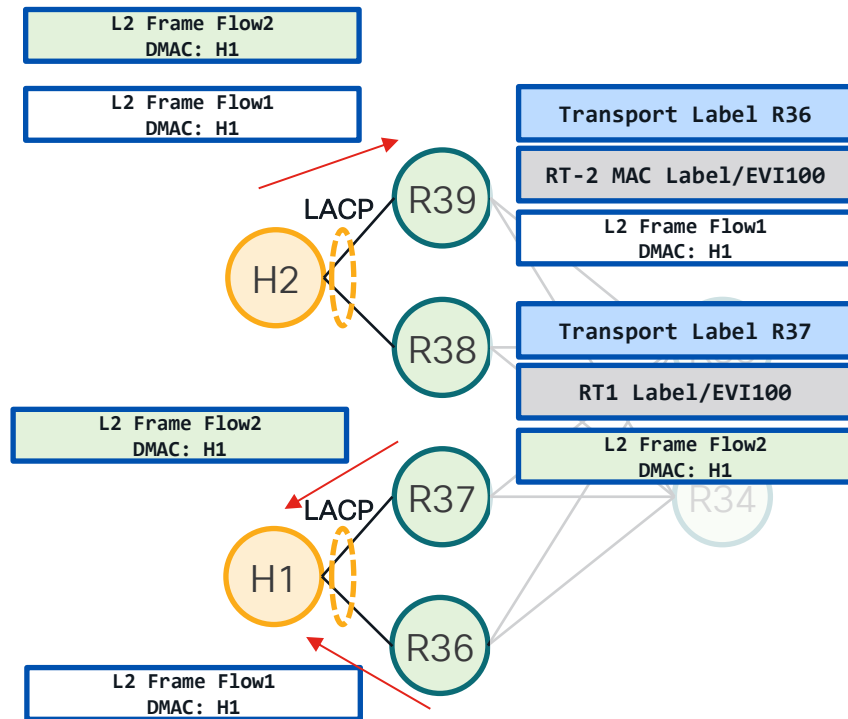
Unicast Forwarding

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast
4. **RT2**: MAC Advertisement
5. **RT1**: Per EVI Ethernet Auto-Discovery



Unicast Forwarding

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast
4. **RT2**: MAC Advertisement
5. **RT1**: Per EVI Ethernet Auto-Discovery

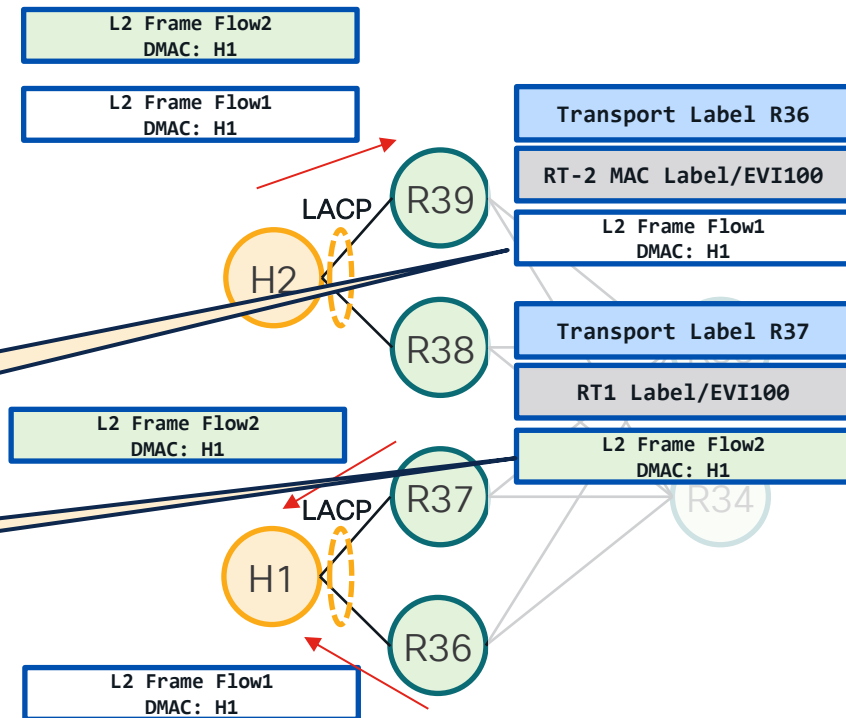


Unicast Forwarding

1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast
4. **RT2**: MAC Advertisement
5. **RT1**: Per EVI Ethernet Auto-Discovery

Per Flow Balancing via R36 and R37 - Aliasing

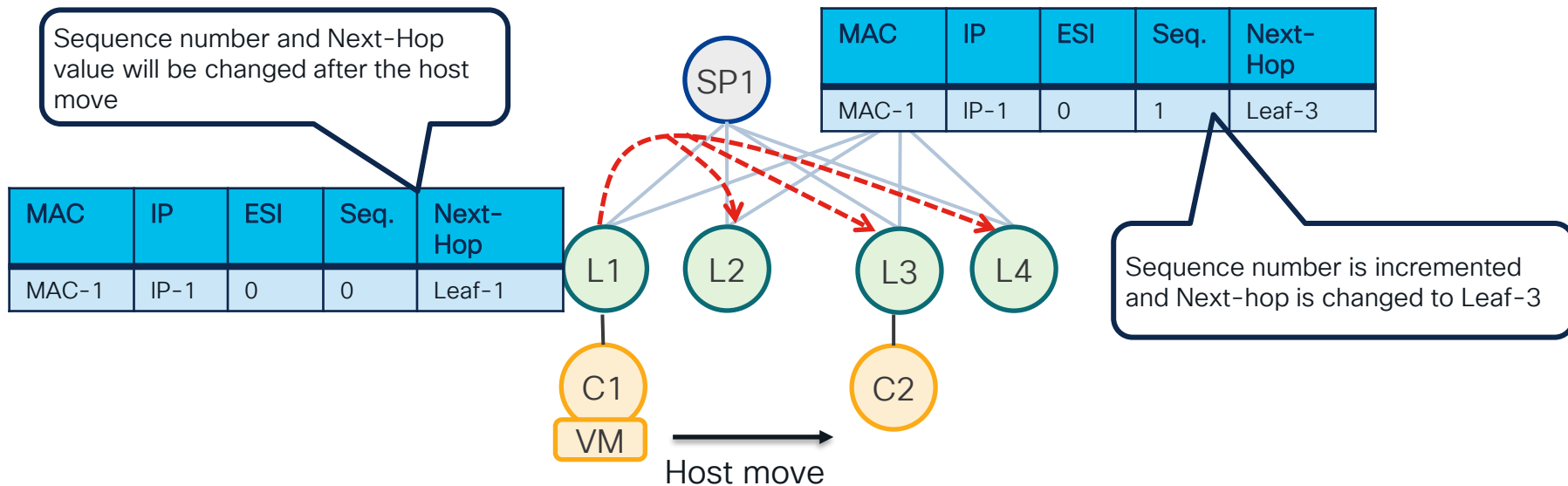
Per Flow Balancing via R36 and R37 - Aliasing



EVPN – MAC Mobility

Challenge:

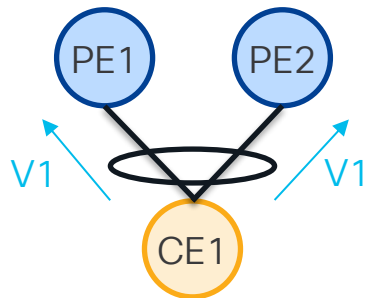
How to detect the correct location of MAC after the movement of host from one Ethernet Segment to another also called “MAC move”?



EVPN Single-Active

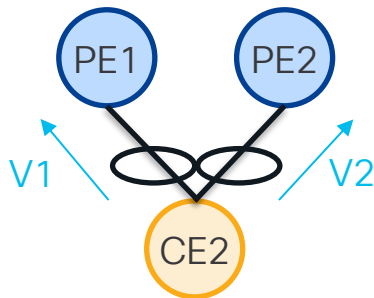
EVPN – Load-Balancing Modes

All-Active
(per flow)



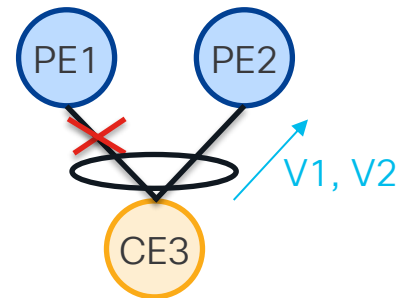
Single LAG at the CE
VLAN goes to both PE
Traffic hashed per flow
Benefits: Bandwidth, Convergence

Single-Active
(per VLAN)



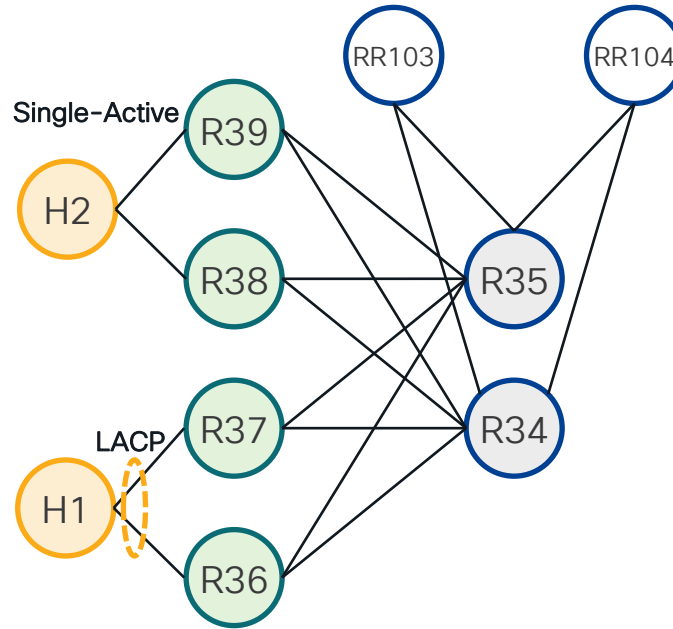
Multiple LAGs at the CE
VLAN active on single PE
Traffic hashed per VLAN
Benefits: Billing, Policing

Port-Active
(per port)



Single LAGs at the CE
Port active on single PE
Traffic hashed per port
Benefits: Protocol Simplification

EVPN - Testbed



All-Active Example

```
R36#show evpn internal-label
```

VPN-ID	Encap	Ethernet Segment Id	EtherTag	Label
100	MPLS	0038.3900.0000.0000.1100	0	68103
Summary pathlist:				
		0x02000001	3.3.3.38	68096
		0x02000002	3.3.3.39	68096

```
R36#show mpls forwarding labels 68103 detail
```

Local Label	Outgoing Label	Prefix or ID	Outgoing Interface	Next Hop	Bytes Switched
68103	68096	EVPN:100		3.3.3.38	0
Updated: Jan 27 07:50:05.582					
Version: 42, Priority: 3					
Label Stack (Top -> Bottom): { 68096 }					
NHID: 0x0, Encap-ID: 0x1386f00000002, Path idx: 0, Backup path idx: 0, Weight: 0					
MAC/Encaps: 0/4, MTU: 0					
Packets Switched: 0					
	68096	EVPN:100		3.3.3.39	0
Updated: Jan 27 07:50:05.582					
Version: 42, Priority: 3					
Label Stack (Top -> Bottom): { 68096 }					
NHID: 0x0, Encap-ID: 0x1387100000002, Path idx: 1, Backup path idx: 0, Weight: 0					
MAC/Encaps: 0/4, MTU: 0					
Packets Switched: 0					

Single-Active – Configuration and Verification

```
R36#show evpn internal-label
```

VPN-ID	Encap	Ethernet Segment Id	EtherTag	Label
100	MPLS	0038.3900.0000.0000.1100	0	68103
Summary pathlist:				
		0x02000001	3.3.3.38	68096
		0x00000000	3.3.3.39 (B)	68096

```
R36#show mpls forwarding labels 68103 detail
```

Sun Jan 27 07:52:03.877 UTC

Local Label	Outgoing Label	Prefix or ID	Outgoing Interface	Next Hop	Bytes Switched
68103	68096	EVPN:100		3.3.3.38	0
Updated: Jan 27 07:51:14.370					
Path Flags: 0x400 [BKUP-IDX:1 (0x0)]					
Version: 47, Priority: 3					
Label Stack (Top -> Bottom): { 68096 }					
NHID: 0x0, Encap-ID: 0x1386f00000002, Path idx: 0, Backup path idx: 1, Weight: 0					
MAC/Encaps: 0/4, MTU: 0					
Packets Switched: 0					
68096		EVPN:100		3.3.3.39	0 (!)
Updated: Jan 27 07:51:14.370					
Path Flags: 0x300 [IDX:1 BKUP, NoFwd]					
Version: 47, Priority: 3					
Label Stack (Top -> Bottom): { 68096 }					
NHID: 0x0, Encap-ID: 0x1387100000002, Path idx: 1, Backup path idx: 0, Weight: 0					
MAC/Encaps: 0/4, MTU: 0					
Packets Switched: 0					
(!): FRR pure backup					

Remote R38/R39

```
evpn
interface Bundle-Ether100
  ethernet-segment
    load-balancing-mode single-active
!
!
```

Single-Active ethernet-segment carving detail

```
R38#show evpn ethernet-segment esi 0038.3900.0000.0000.1100 carving detail
```

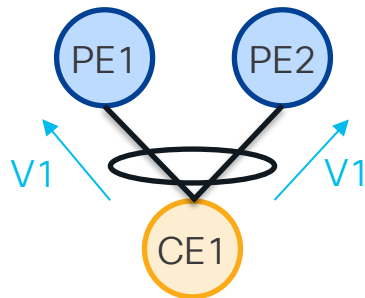
Ethernet Segment Id	Interface	Nexthops
0038.3900.0000.0000.1100	BE100	3.3.3.38 3.3.3.39


```
ES to BGP Gates : Ready
ES to L2FIB Gates : Ready
Main port :
  Interface name : Bundle-Ether100
  Interface MAC : 008a.967f.30dd
  IfHandle : 0x0800002c
  State : Up
  Redundancy : Not Defined
ESI type : 0
  Value : 38.3900.0000.0000.1100
ES Import RT : 3839.0000.0000 (from ESI)
Source MAC : 0000.0000.0000 (N/A)
Topology :
  Operational : MH, Single-active
  Configured : Single-active (AAPS)
Service Carving : Auto-selection
Peering Details : 3.3.3.38[MOD:P:00] 3.3.3.39[MOD:P:00]
Service Carving Results:
  Forwarders : 1
  Permanent : 0
  Elected : 1
  EVI E : 100
  Not Elected : 0
MAC Flushing mode : STP-TCN
Peering timer : 3 sec [not running]
Recovery timer : 30 sec [not running]
Carving timer : 0 sec [not running]
Local SHG label : 68098
Remote SHG labels : 1
  68098 : nexthop 3.3.3.39
```

EVPN Port-Active

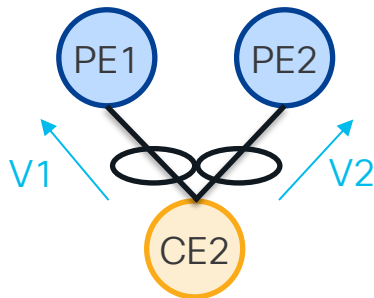
EVPN – Load-Balancing Modes

All-Active
(per flow)



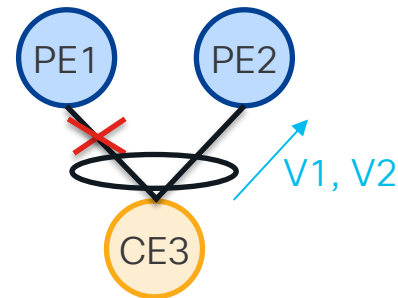
Single LAG at the CE
VLAN goes to both PE
Traffic hashed per flow
Benefits: Bandwidth, Convergence

Single-Active
(per VLAN)



Multiple LAGs at the CE
VLAN active on single PE
Traffic hashed per VLAN
Benefits: Billing, Policing

Port-Active
(per port)

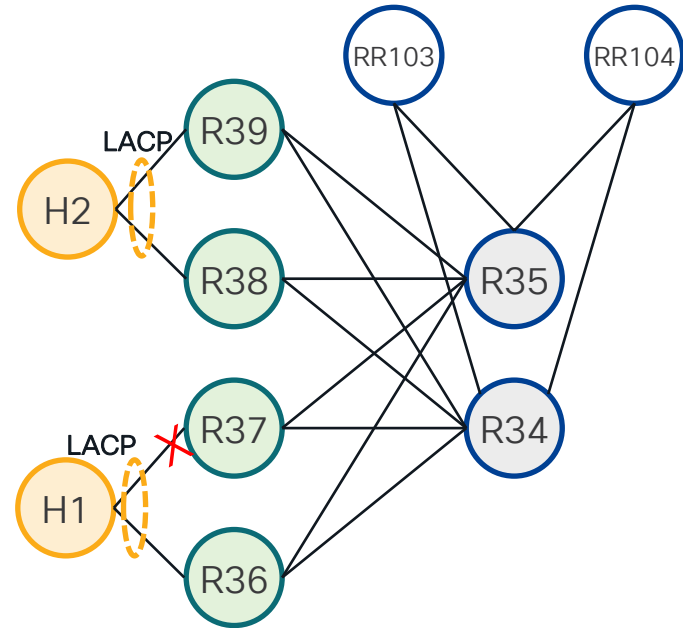


Single LAGs at the CE
Port active on single PE
Traffic hashed per port
Benefits: Protocol Simplification

EVPN - Testbed

R36/R37

```
evpn
interface Bundle-Ether100
  ethernet-segment
    load-balancing-mode port-active
  !
  !
```



Port-Active –Verification

```
R36#show bundle
Bundle-Ether100
```

```
Status: Up
Local links <active/standby/configured>: 1 / 0 / 1
Local bandwidth <effective/available>: 10000000 (10000000) kbps
MAC address (source): 008a.9644.d8de (Chassis pool)
Inter-chassis link: No
Minimum active links / bandwidth: 1 / 1 kbps
Maximum active links: 64
Wait while timer: 2000 ms
Load balancing:
  Link order signaling: Not configured
  Hash type: Default
  Locality threshold: None
LACP: Operational
  Flap suppression timer: Off
  Cisco extensions: Disabled
  Non-revertive: Disabled
mLACP: Not configured
IPv4 BFD: Not configured
IPv6 BFD: Not configured
```

Port	Device	State	Port ID	B/W, kbps
Te0/0/0/0	Local	Active	0x8000, 0x0001	10000000
Link is Active				

```
R37#show bundle
Bundle-Ether100
```

```
Status: LACP 00S (out of service)
Local links <active/standby/configured>: 0 / 1 / 1
Local bandwidth <effective/available>: 0 (0) kbps
MAC address (source): 008a.9644.08de (Chassis pool)
Inter-chassis link: No
Minimum active links / bandwidth: 1 / 1 kbps
Maximum active links: 64
Wait while timer: 2000 ms
Load balancing:
  Link order signaling: Not configured
  Hash type: Default
  Locality threshold: None
LACP: Operational
  Flap suppression timer: Off
  Cisco extensions: Disabled
  Non-revertive: Disabled
mLACP: Not configured
IPv4 BFD: Not configured
IPv6 BFD: Not configured
```

Port	Device	State	Port ID	B/W, kbps
Te0/0/0/0	Local	Standby	0x8000, 0x0001	10000000
Link is in standby due to bundle out of service state				

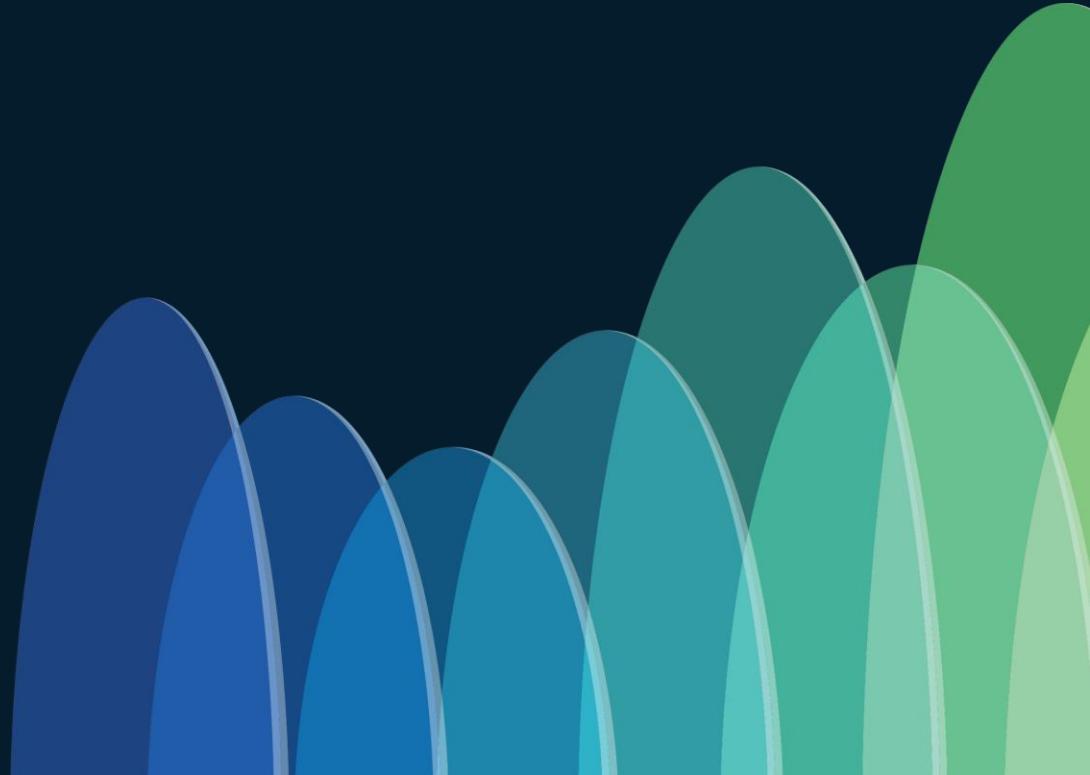
```
R37#show int bundle-ether 100
```

```
Bundle-Ether100 is down, line protocol is down
```

```
R37#show int tenGigE 0/0/0/0
```

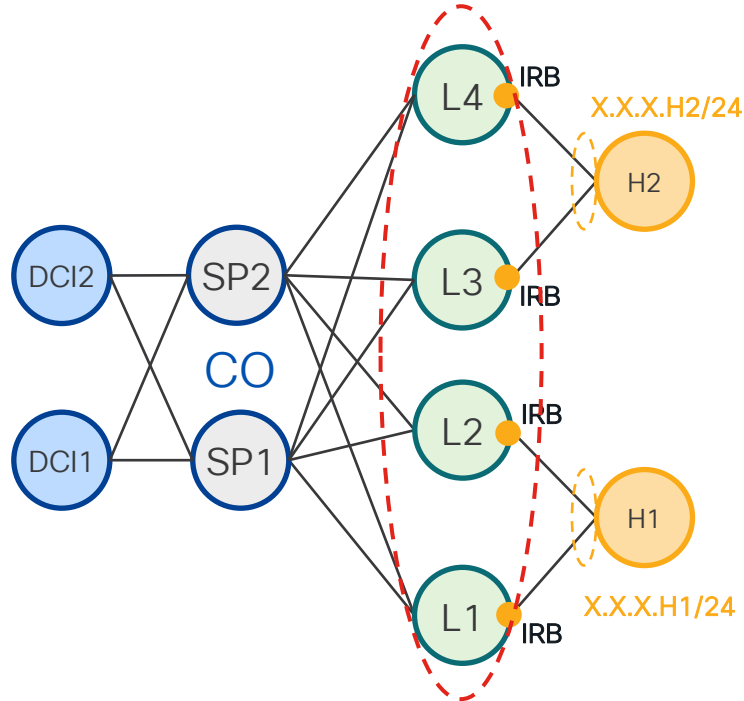
```
TenGigE0/0/0/0 is up, line protocol is up
```

EVPN L2 & L3 Integration



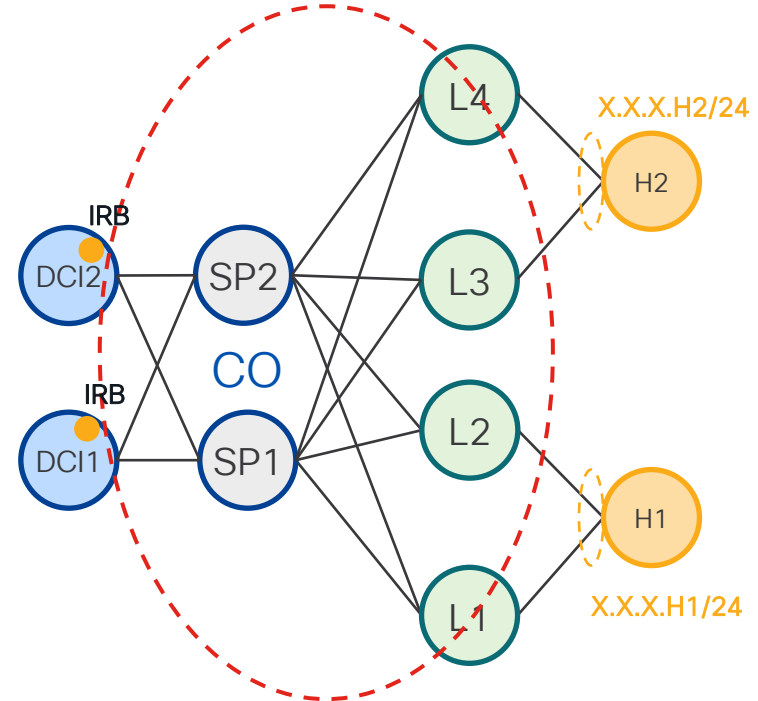
Distributed vs Centralized Routing

Layer2 Bridging mandatory between Leaves only



- Optimized forwarding of east-west traffic
- ARP/MAC state localized to Leafs
- Helps with horizontal scaling of DC

Layer2 Bridging mandatory between Leaves and DCI



- All east<->west routed traffic traverses to centralized gateways
- Centralized gateways have full ARP/MAC state in the DCI
- Scale challenge

EVPN Distributed L3 Anycast Gateway

EVPN – Distributed Symmetric Anycast Gateway

Leaves run Multi-Protocol BGP to advertise & learn MAC + HOST IP addresses over the Network

MAC + IP addresses are advertised to rest of Leaves

L3/4 – Learn MAC + IP HOST address advertised by L1

-> L2/L3 update MAC address table + **IP Forwarding table**

L2 – uses MAC address advertised by L1 to synchronize MAC address table

-> L2 forwards MAC via local ETH interface represented by same Ethernet Segment between L1 and L2

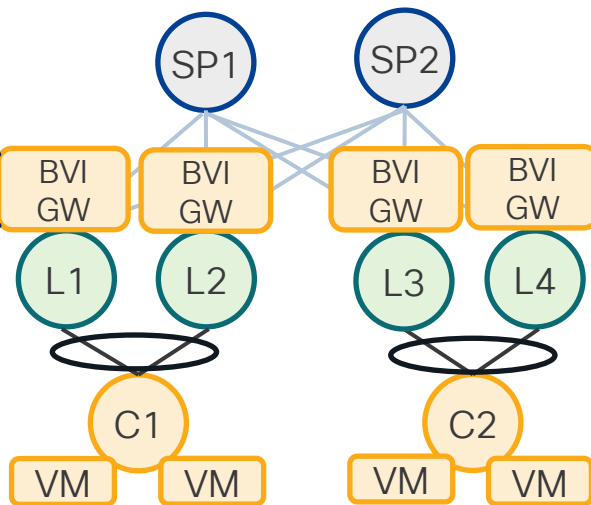
L2 – uses MAC + IP HOST address advertised by L1 to synchronize ARP/ND information

-> L2 forwards IP via local ETH interface

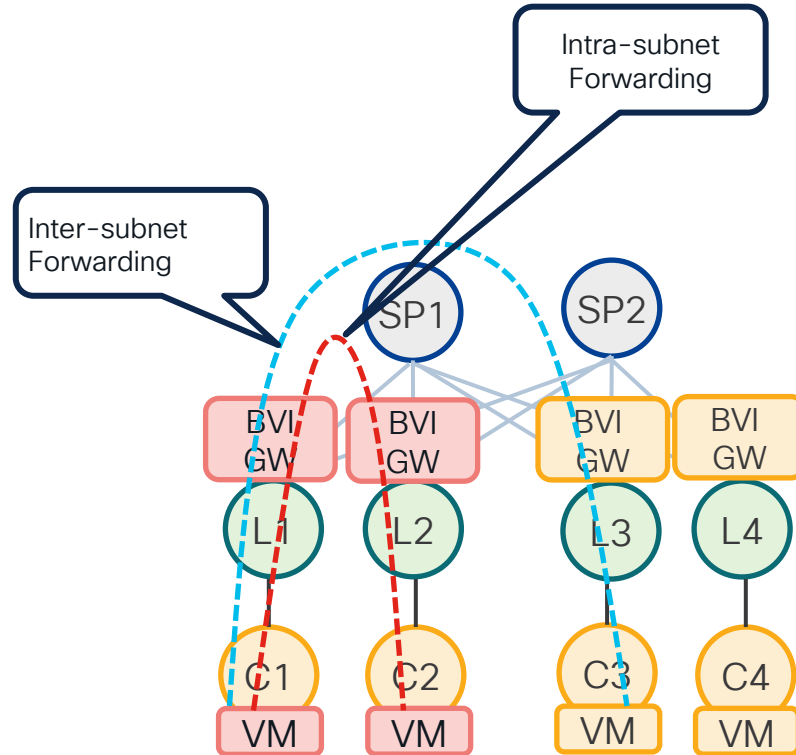
Distributed Anycast Gateway serves as the gateway for connected hosts

All the BVIs perform active forwarding in contrast to active/standby like First-hop routing protocol

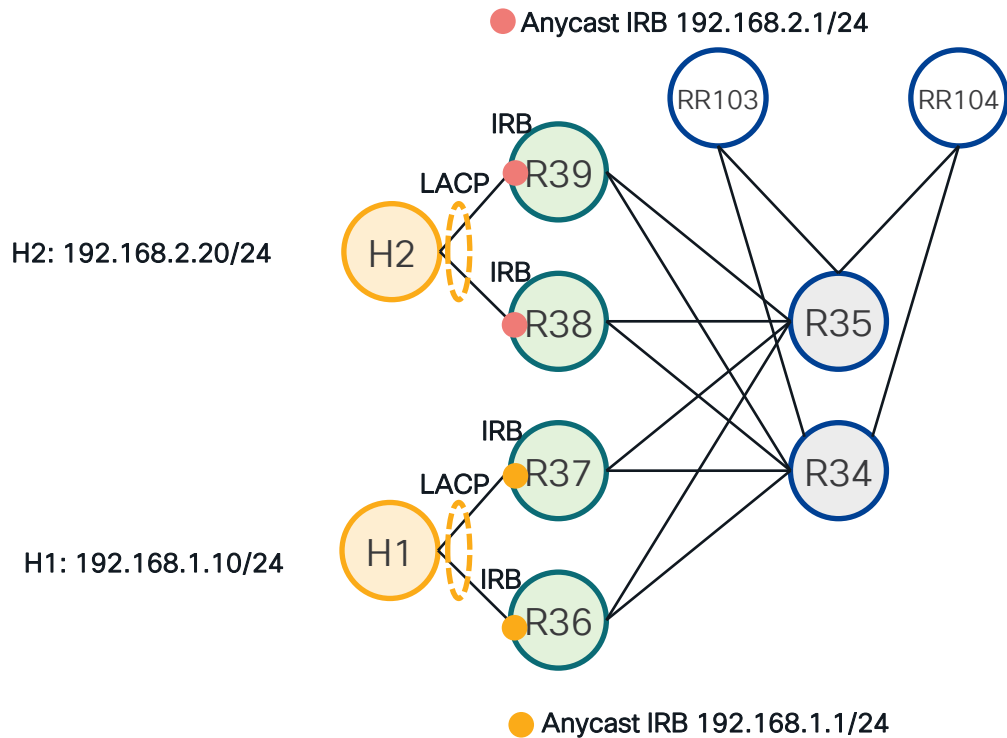
Identical Anycast Gateway Virtual IP and MAC address are configured on all the Leafs



EVPN – IRB in Network Fabric



EVPN Distributed L3 Anycast GW – Symmetric IRB



EVPN Configuration - IRB

```
evpn
no evi 100
  no advertise-mac
!

vrf a
  address-family ipv4 unicast
    import route-target
      100:100
    !
    export route-target
      100:100
    !
    !
    !

interface BVI100
  host-routing
  vrf a
  ipv4 address 192.168.1.1 255.255.255.0
  mac-address 3637.3637.3637
!
```

Not needed! We need MAC/IP RT-2

VRF configuration

MAC/IP RT2

Anycast Distributed IRB: Same IP and MAC
R36,R37

EVPN Configuration - BGP VRF

```
router bgp 1
  bgp router-id 3.3.3.36
  address-family vpv4 unicast
  !
  address-family l2vpn evpn
  !
  neighbor-group rr
  remote-as 1
  update-source Loopback0
  address-family l2vpn evpn
  !
  neighbor 3.3.3.103
  use neighbor-group rr
  !
  neighbor 3.3.3.104
  use neighbor-group rr
  !
  vrf a
  rd auto
  address-family ipv4 unicast
    additional-paths receive
    maximum-paths ibgp 2
    redistribute connected
  !
  !
```

BGP Multi-Path for Inter-subnet forwarding

R36: RT-2 MAC/IP Advertisement

```
R36#show bgp l2vpn evpn rd 3.3.3.36:100 [2][0][48][0062.ec71.fbd7][32][19$
Tue Oct 16 02:47:45.576 UTC
BGP routing table entry for [2][0][48][0062.ec71.fbd7][32][192.168.1.10]/136, Route Distinguisher: 3.3.3.36:100
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker           84847      84847
Last Modified: Oct 15 23:14:52.399 for 03:32:53
Paths: (2 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.2
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.2
  Local
    0.0.0.0 from 0.0.0.0 (3.3.3.36)
      Second Label 64008
      Origin IGP, localpref 100, valid, redistributed, best, group-best, import-candidate, rib-install
      Received Path ID 0, Local Path ID 1, version 84838
      Extended community: SoO:3.3.3.37:100 RT:1:100 RT:100:100
      EVPN ESI: 0036.3700.0000.0000.1100
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local
    3.3.3.37 (metric 30) from 3.3.3.103 (3.3.3.37)
      Received Label 64004, Second Label 64008
      Origin IGP, localpref 100, valid, internal, import-candidate, imported, rib-install
      Received Path ID 0, Local Path ID 0, version 0
      Extended community: SoO:3.3.3.37:100 RT:1:100 RT:100:100
      Originator: 3.3.3.37, Cluster list: 3.3.3.103
      EVPN ESI: 0036.3700.0000.0000.1100
      Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 3.3.3.37:100
RP/0/RP0/CPU0:R36#
```

IP

RT EVI 100 and RT VRF A

RT EVI 100 and RT VRF A

R36: RT-2 MAC/IP

R36#show evpn evi mac

Tue Oct 16 02:52:22.437 UTC

VPN-ID	Encap	MAC address	IP address	Nexthop	Label
100	MPLS	0062.ec71.fbd7	192.168.1.10	3.3.3.37	64004
65535	N/A	008a.9644.d8d8 ::		Local	0

Learned and Advertised
MAC and IP

RT-2 per-BD label

R36: VRF Routes

```
R36#show route vrf a
Tue Oct 16 02:46:34.463 UTC
```

```
Codes: C - connected, S - static, R - RIP, B - BGP, (>) - Diversion path
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - ISIS, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, su - IS-IS summary null, * - candidate default
       U - per-user static route, o - ODR, L - local, G - DAGR, l - LISP
       A - access/subscriber, a - Application route
       M - mobile route, r - RPL, t - Traffic Engineering, (!) - FRR Backup path
```

Gateway of last resort is not set

```
C 192.168.1.0/24 is directly connected, 03:37:59, BVI100
L 192.168.1.1/32 is directly connected, 03:37:59, BVI100
B 192.168.1.10/32 [200/0] via 3.3.3.37 (nexthop in vrf default)
B 192.168.2.20/32 [200/0] via 3.3.3.38 (nexthop in vrf default), 03:28:28
  [200/0] via 3.3.3.39 (nexthop in vrf default), 03:28:28
```

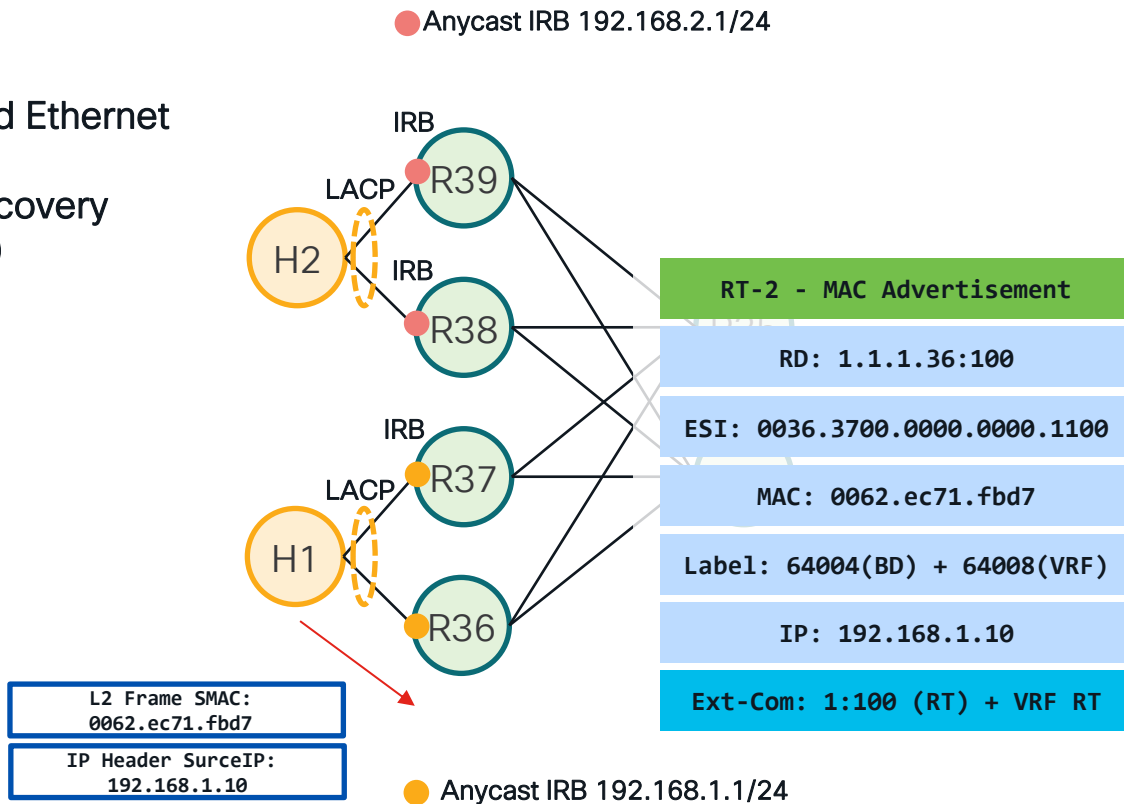
EVPN Learned Route

BGP Multi Path to H2 connected to R38 and R39

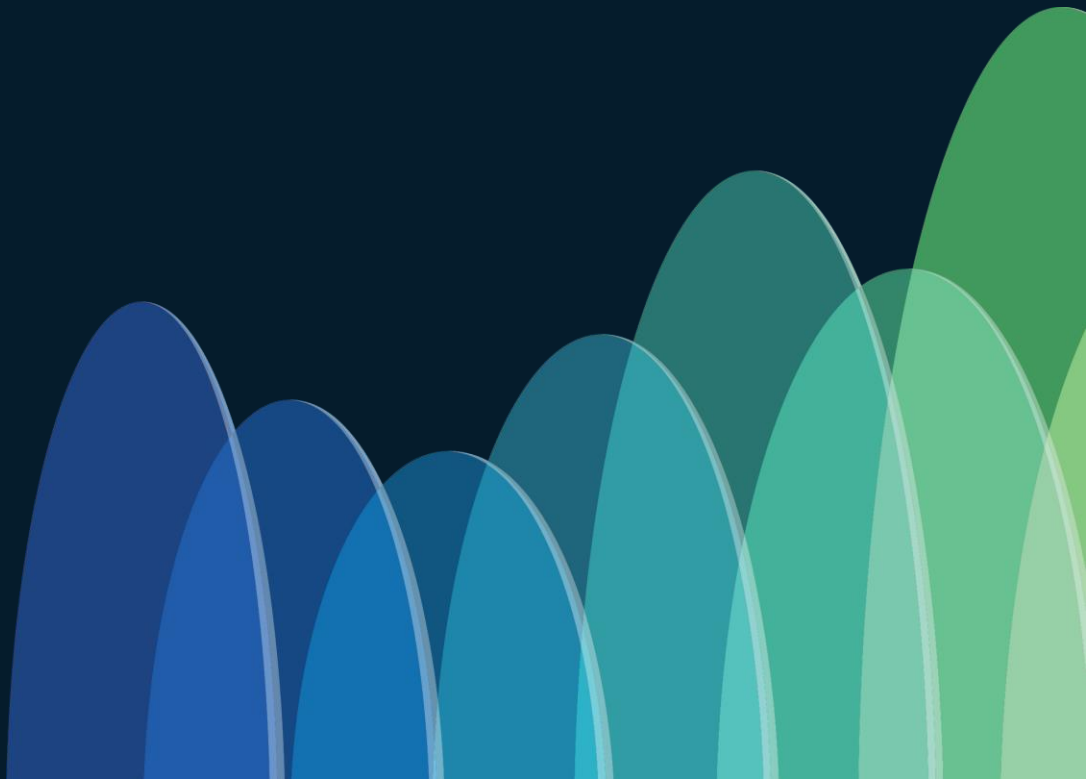
R36, R37, R38, R39 – EVPN Startup

R36 – Example

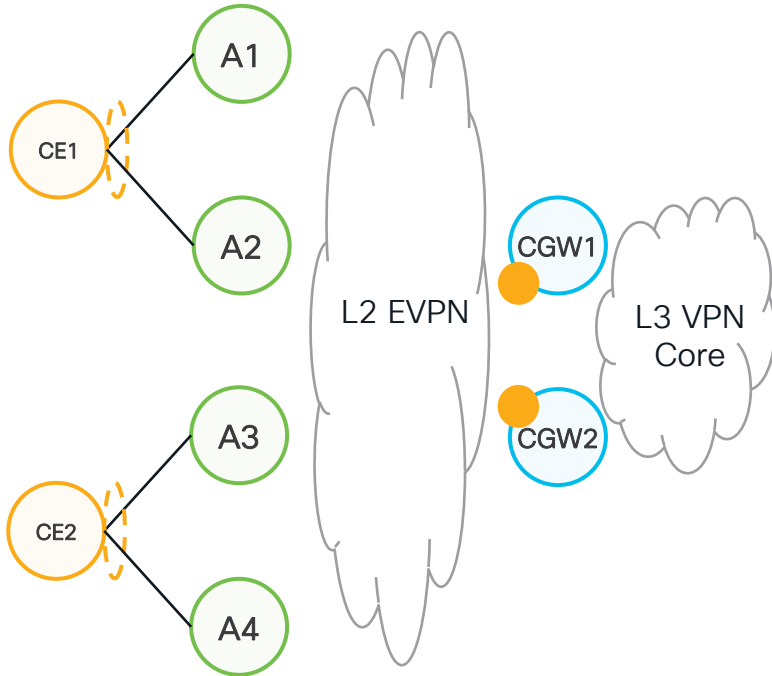
1. **RT4**: DF Election & Multi-Homed Ethernet Segment Auto-Discovery
2. **RT1**: Per ESI Ethernet Auto-Discovery (Split-Horizon, Mass-Withdraw)
3. **RT3**: Inclusive Multicast
4. **RT2**: MAC/IP Advertisement



EVPN Centralized GW CGW



EVPN Centralized Gateway (CGW)



CGW - Configuration

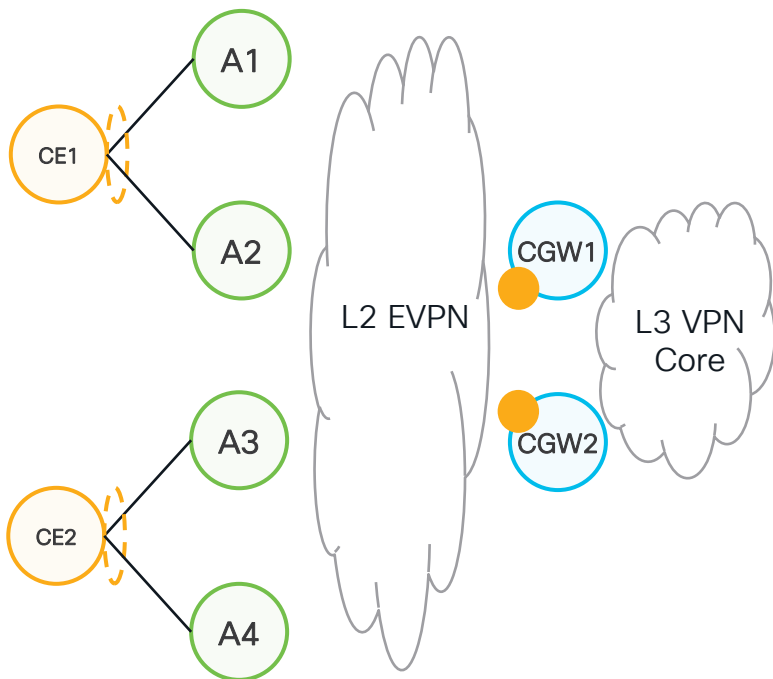
```
evpn
virtual access-evi
  ethernet-segment
    identifier type 0 77.77.77.77.77.77.77.77
l2vpn
bridge group test
bridge-domain test
  access-evi 300
  routed interface BVI300
```

Access - Configuration

```
evpn
evi 300
  advertise-mac

l2vpn
bridge group test
bridge-domain test
  interface Bundle-Ether100
  !
  evi 300
```

EVPN Centralized Gateway (CGW)



```
R28#show evpn ethernet-segment
```

Ethernet Segment Id	Interface	Nexthops
0077.7777.7777.7777	Access-EVI:all	1.1.1.26 1.1.1.28

```
RP/0/RSP0/CPU0:R28#show arp vrf a
```

```
0/0/CPU0
```

Address	Age	Hardware Addr	State	Type	Interface
192.168.250.1	-	a011.1111.1111	Interface	ARPA	BVI300
192.168.250.10	-	28ac.9ea7.d41b	EVPN_SYNC	ARPA	BVI300

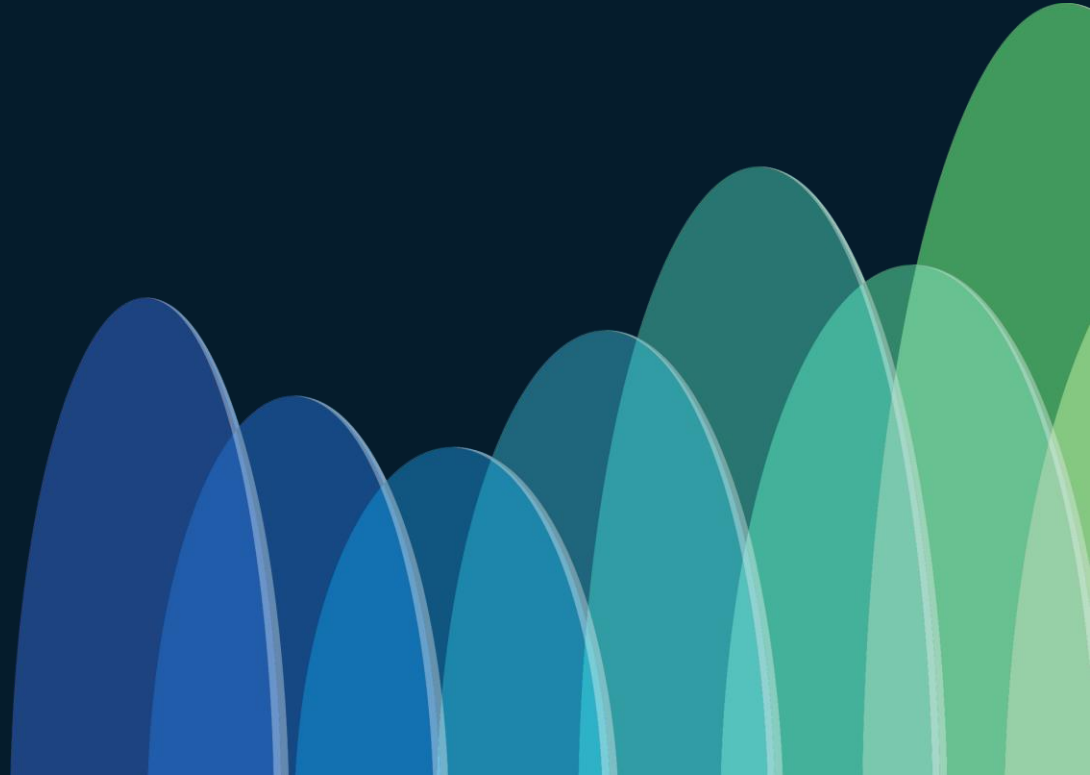
CGW in Single-Active mode from Access-to-CGW (South->North)
Based on Access-EVI DF election NDF CGW BVI is added to Core SHG
prevents traffic from access-EVI go to BVI
allows traffic from BVI to Access-EVI

Single-Active South->North
All-Active North->South

Distributed vs Centralized Gateway

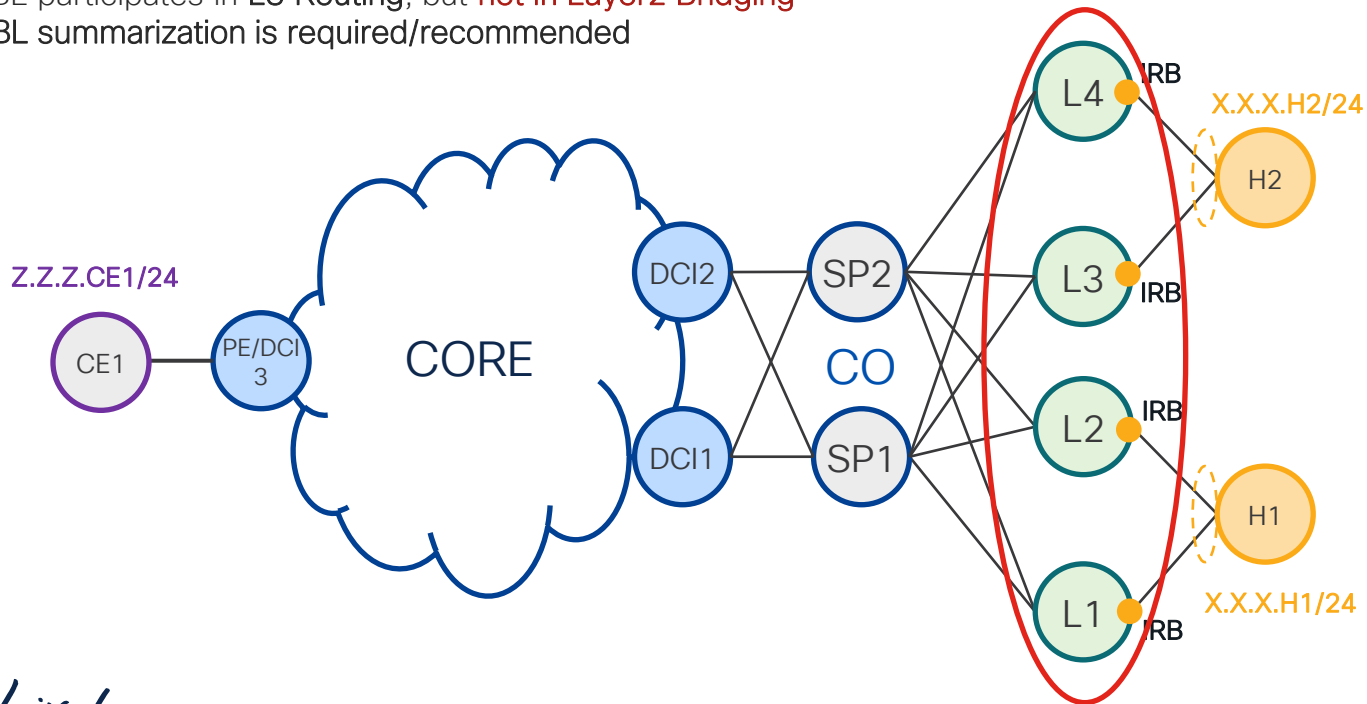
- Distributed Anycast Gateway is our priority!
 - Best Scalable solution
 - Optimal L2/L3 forwarding

BGP Layer3 Interconnect



BGP Layer3 Interconnect Principles

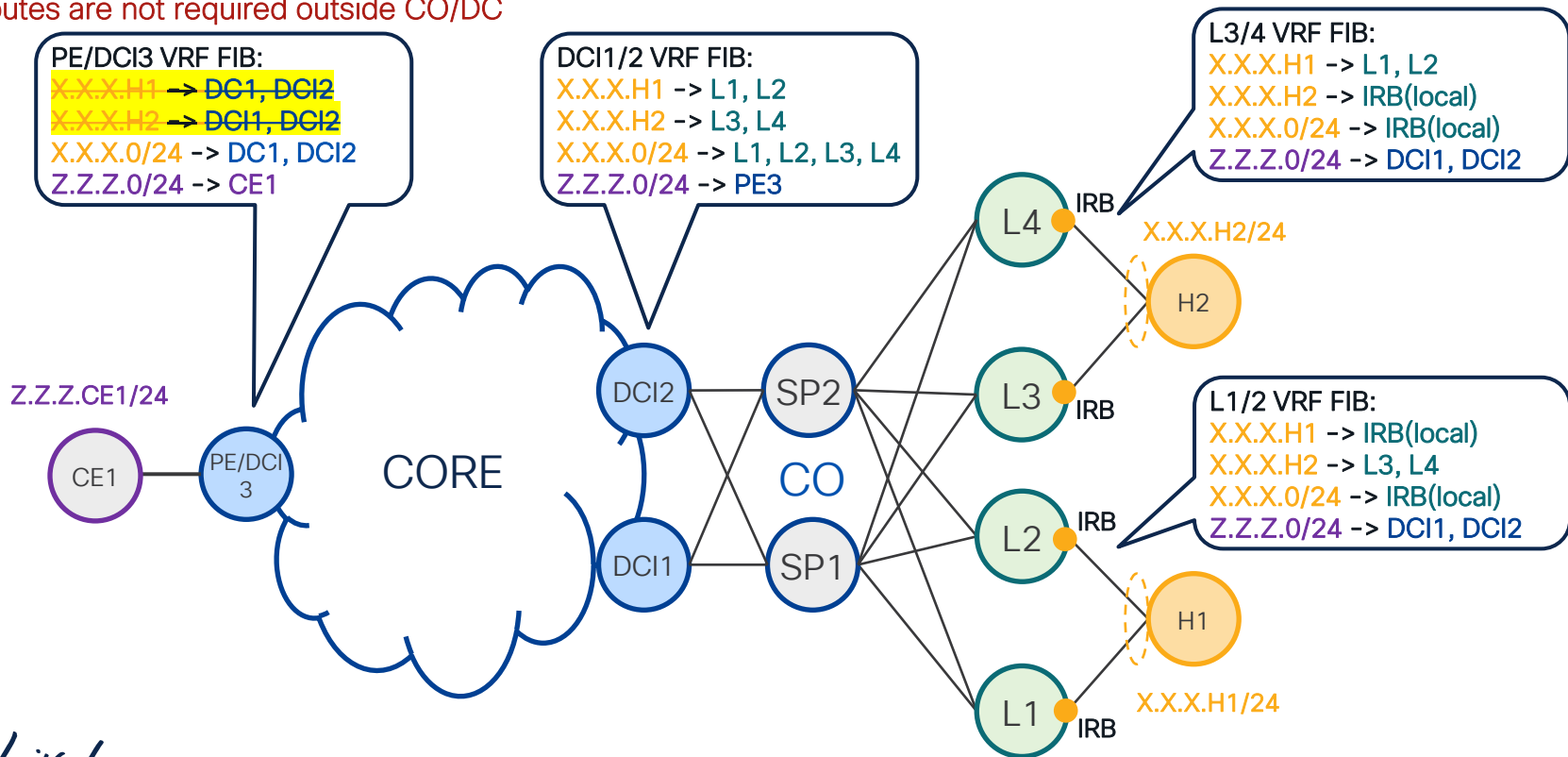
- DCI/BL provides Layer3 Interconnect
- DCI/BL participates in L3 Routing, but **not in Layer2 Bridging**
- DCI/BL summarization is required/recommended



BGP Layer3 Interconnect

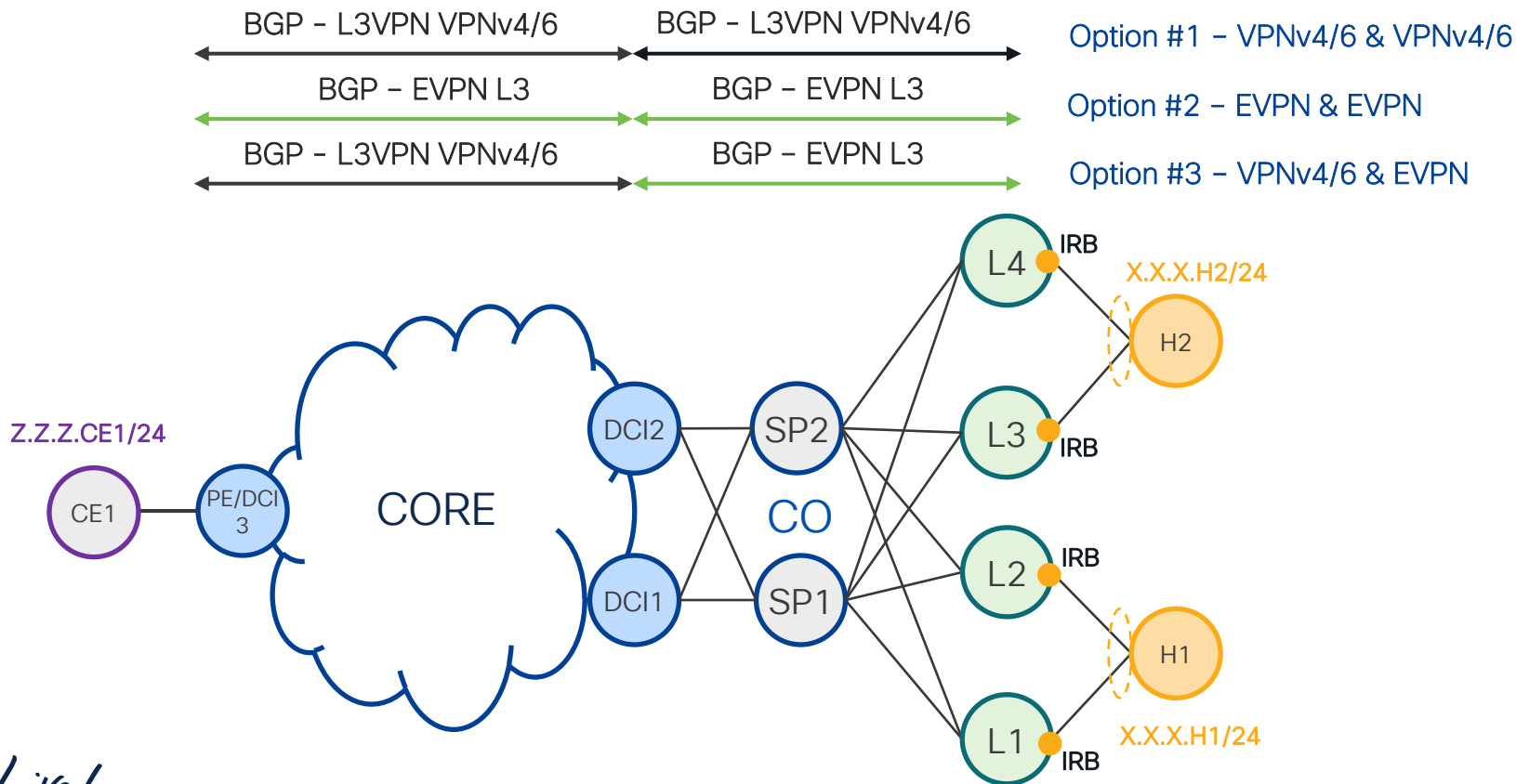
DCI/BL Summarization

Host-Routes are not required outside CO/DC



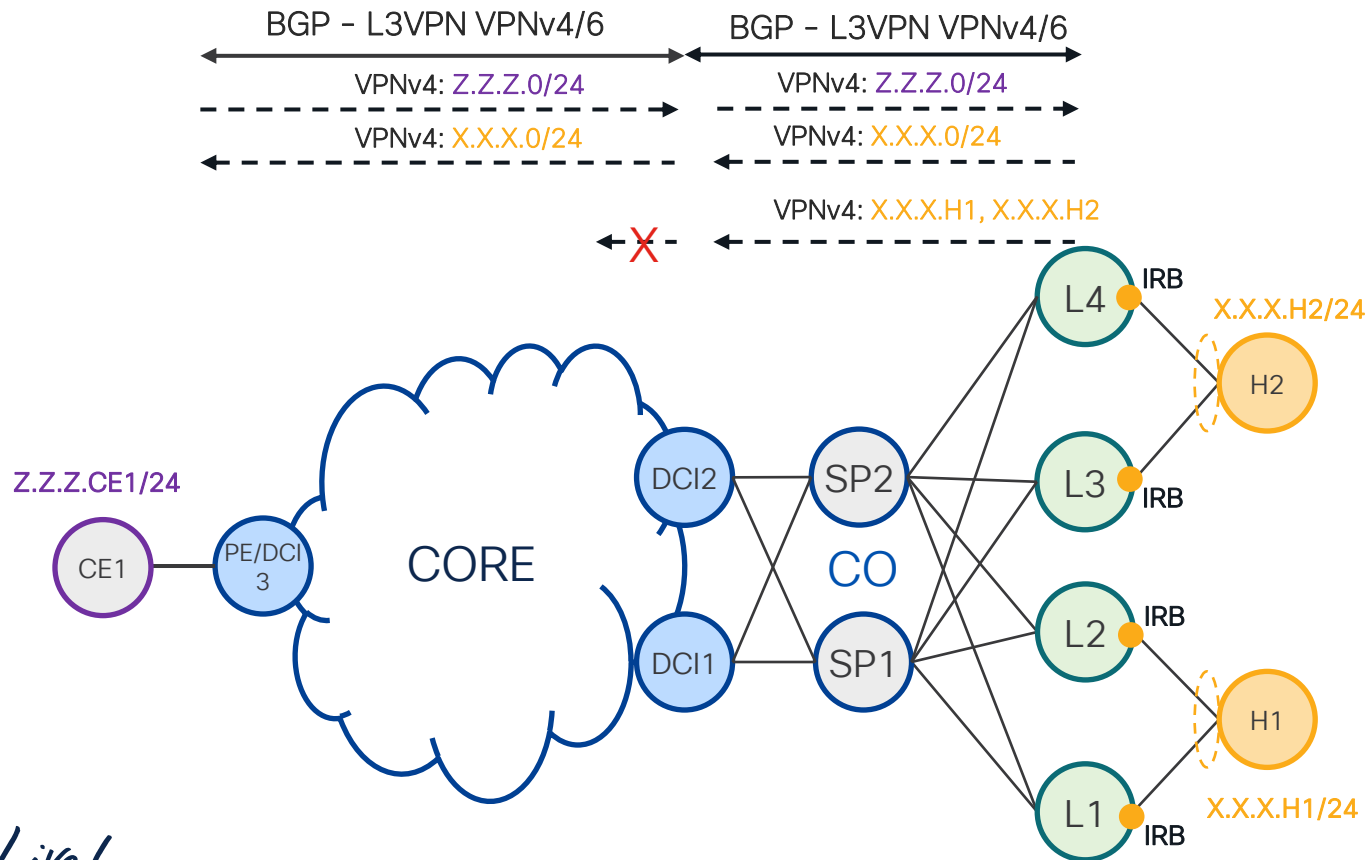
BGP Layer3 Interconnect

Control Plane



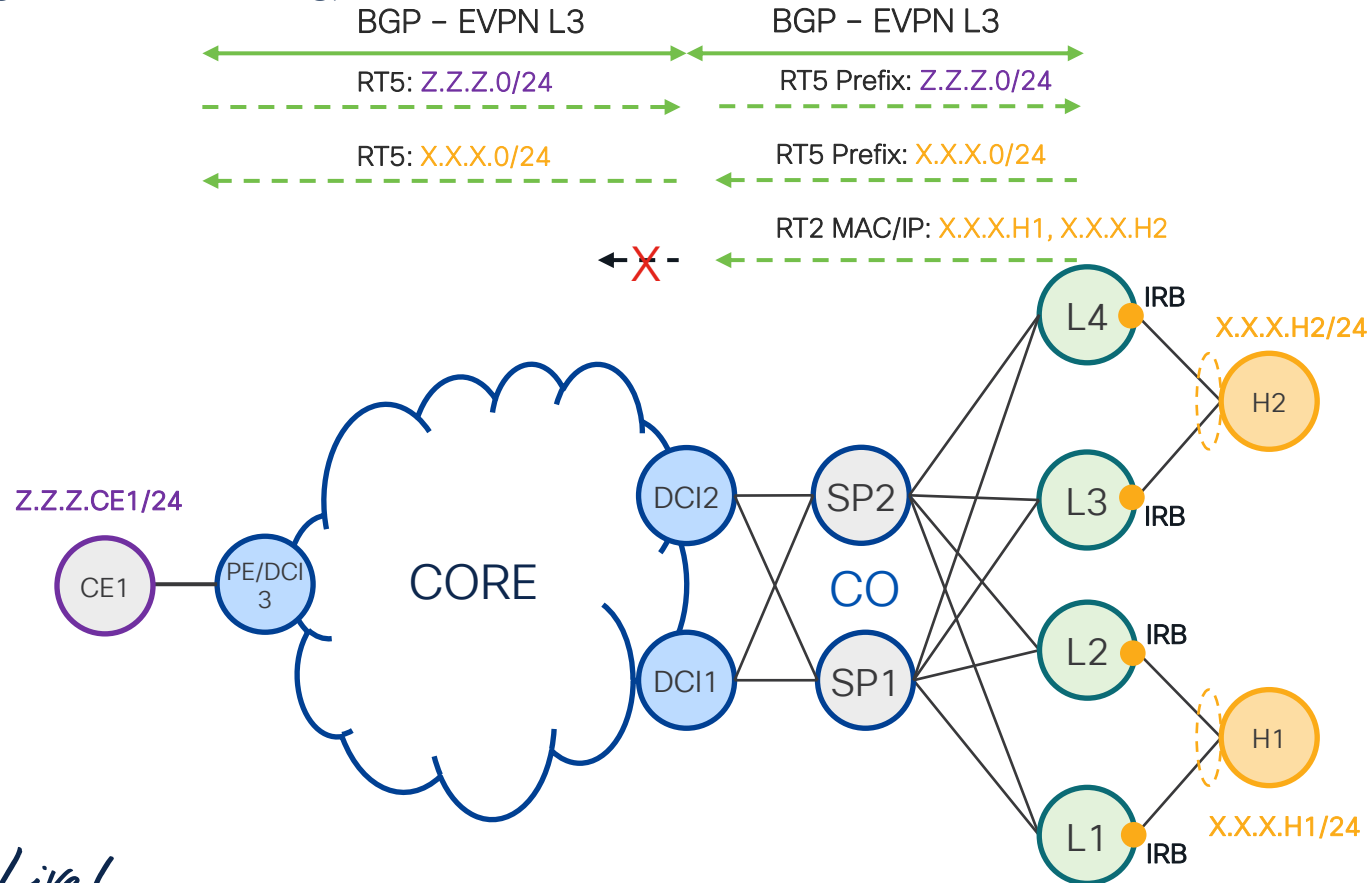
BGP Layer3 Interconnect

Option #1 – VPNv4/6 & VPNv4/6



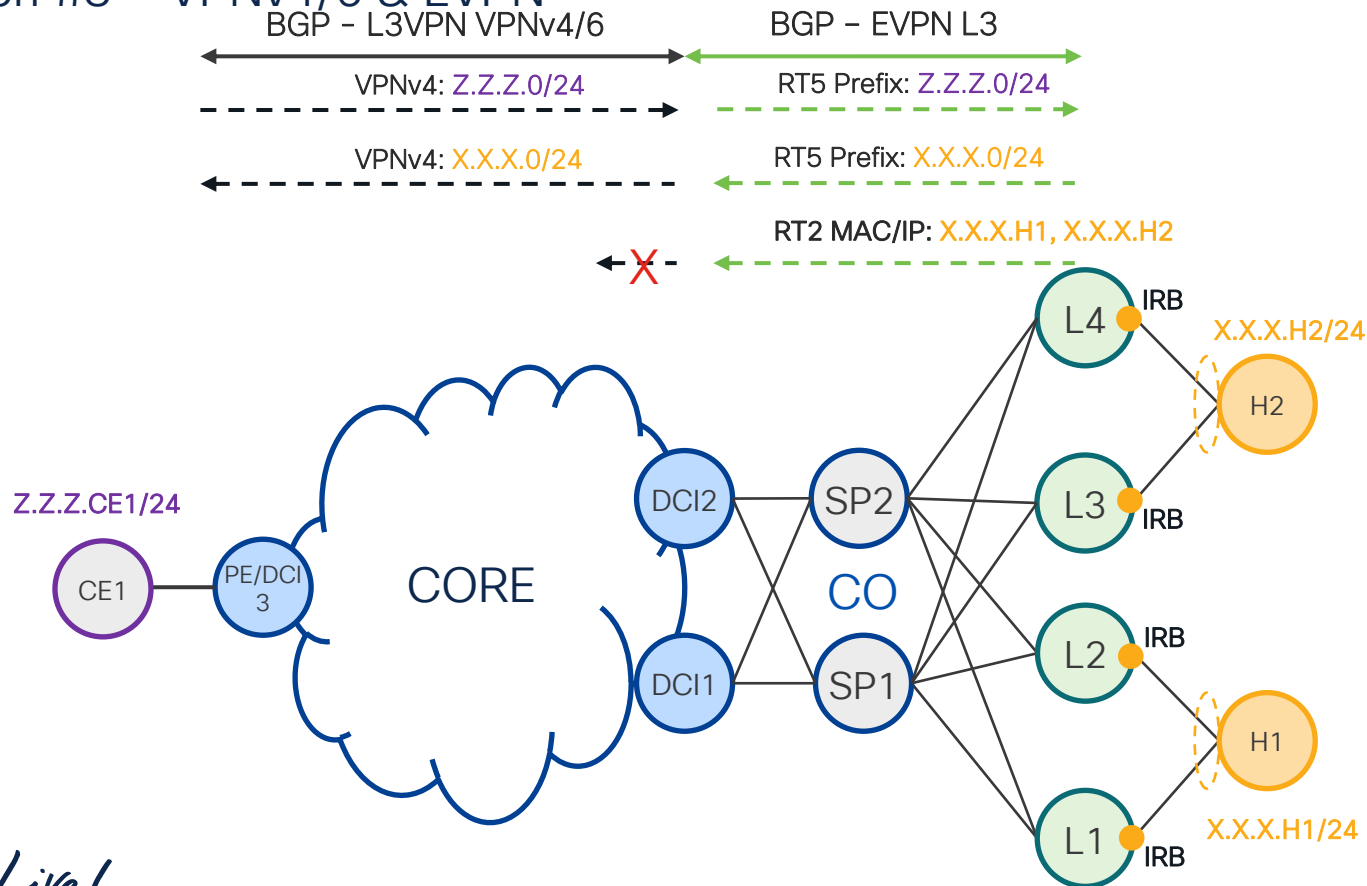
BGP Layer3 Interconnect

Option #2 – EVPN & EVPN



BGP Layer3 Interconnect

Option #3 - VPNv4/6 & EVPN



BGP Layer3 Interconnect

Control Plane Options Highlight

- Option #1 – VPNv4/6 & VPNv4/6
 - + VPNv4/6 Industry proved solution for Layer3 VPN
 - + DCI doesn't need to understand BGP EVPN AF
 - Leaf has to peer with Route-Reflector via both BGP EVPN and VPNv4/6 AF
 - EVPN AF to support L2 stretch (MAC advertisement) across DC/CO between Leaves
 - EVPN AF to sync ARP/ND for Multi-Homed All-Active
 - DC/CO Route-Reflector has to support both BGP EVPN and VPNv4/6 AF
 - Leaf has to advertise VM Host-Routes via VPNv4/6
- Option #2 – EVPN & EVPN
 - + Single BGP Address Family End-To-End in Network
 - Existing L3 VPNv4/6 services has to to migrated to L3 EVPN
 - No technical benefit to migrate existing L3 VPNv4/6 to L3 EVPN
- Option #3 – VPNv4/6 & EVPN
 - + Recommended solution which benefits from both Options #1 and #2
 - + New DC/CO – Leaf, Route-Reflector use single BGP AF EVPN
 - + Existing L3 VPNv4/6 services stay untouched

R36: BGP Configuration - RT-5

```
router bgp 1
  bgp router-id 3.3.3.36
  address-family vpnv4 unicast
  !
  address-family l2vpn evpn
  !
  neighbor-group rr
  remote-as 1
  update-source Loopback0
  address-family l2vpn evpn
    advertise vpnv4 unicast
  !
  vrf a
  rd auto
  address-family ipv4 unicast
    additional-paths receive
    maximum-paths ibgp 2
  !
```

RT-5

R36: RT-5 Route

```
R36#show bgp l2vpn evpn rd 3.3.3.37:0 [5][0][24][192.168.1.0]/80
Tue Oct 16 03:35:06.480 UTC
BGP routing table entry for [5][0][24][192.168.1.0]/80, Route Distinguisher: 3.3.3.37:0
Versions:
  Process          bRIB/RIB SendTblVer
  Speaker          84912      84912
Last Modified: Oct 16 03:23:18.399 for 00.0.0.0
Paths: (2 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  Local
    3.3.3.37 (metric 30) from 3.3.3.103 (3.3.3.37)
      Received Label 64008
      Origin incomplete, metric 0, localpref 100, valid, internal, best, group-best, import-candidate, not-in-vrf
      Received Path ID 0, Local Path ID 1, version 84912
      Extended community: Flags 0x6: RT:100:100
      Originator: 3.3.3.37, Cluster list: 3.3.3.103
      EVPN ESI: 0000.0000.0000.0000.0000, Gateway Address : 0.0.0.0
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local
    3.3.3.37 (metric 30) from 3.3.3.104 (3.3.3.37)
      Received Label 64008
      Origin incomplete, metric 0, localpref 100, valid, internal, not-in-vrf
      Received Path ID 0, Local Path ID 0, version 0
      Extended community: Flags 0x6: RT:100:100
      Originator: 3.3.3.37, Cluster list: 3.3.3.104
      EVPN ESI: 0000.0000.0000.0000.0000, Gateway Address : 0.0.0.0
RP/0/RP0/CPU0:R36#
```

The diagram illustrates the components of the BGP output for the RT-5 route. Callouts point to specific parts of the text:

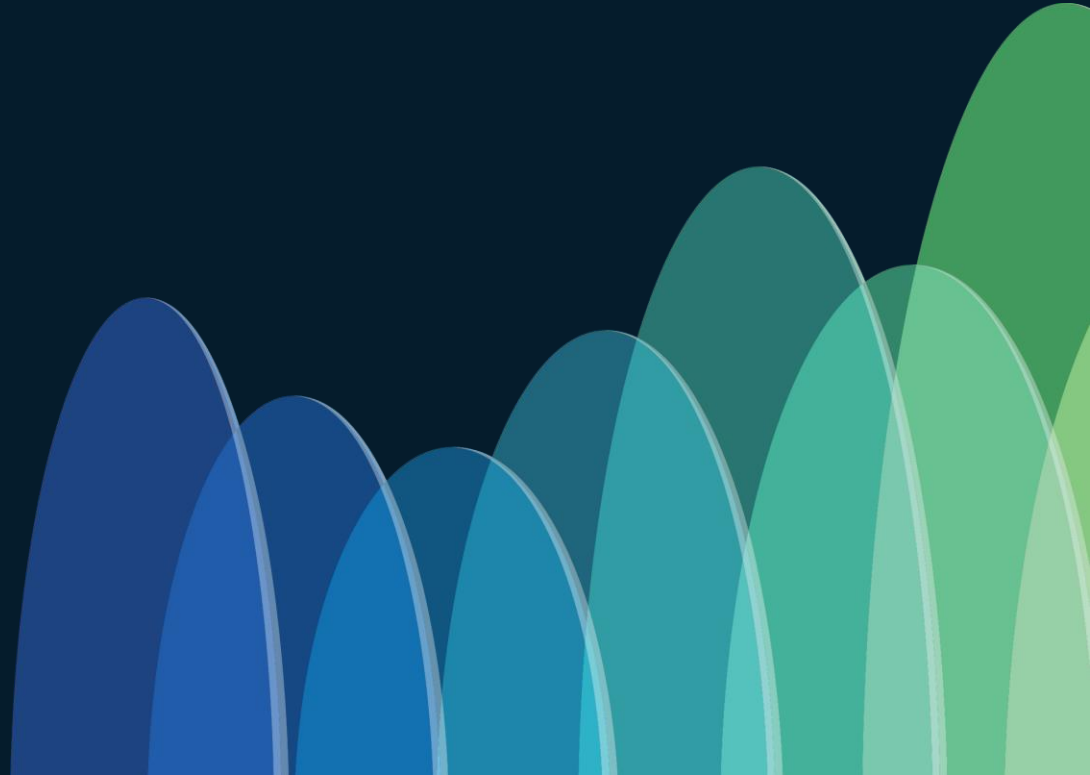
- RT-5**: Points to the route identifier `[5][0][24][192.168.1.0]/80`.
- prefix**: Points to the route identifier `[5][0][24][192.168.1.0]/80`.
- VRF A R37 RD**: Points to the Route Distinguisher `3.3.3.37:0`.
- VRF Agg label**: Points to the `Received Label 64008` for the first path.
- VRF A Route-Target**: Points to the `RT:100:100` extended community for the first path.
- VRF Agg label**: Points to the `Received Label 64008` for the second path.
- VRF A Route-Target**: Points to the `RT:100:100` extended community for the second path.

R36: VRF A - Routing Table

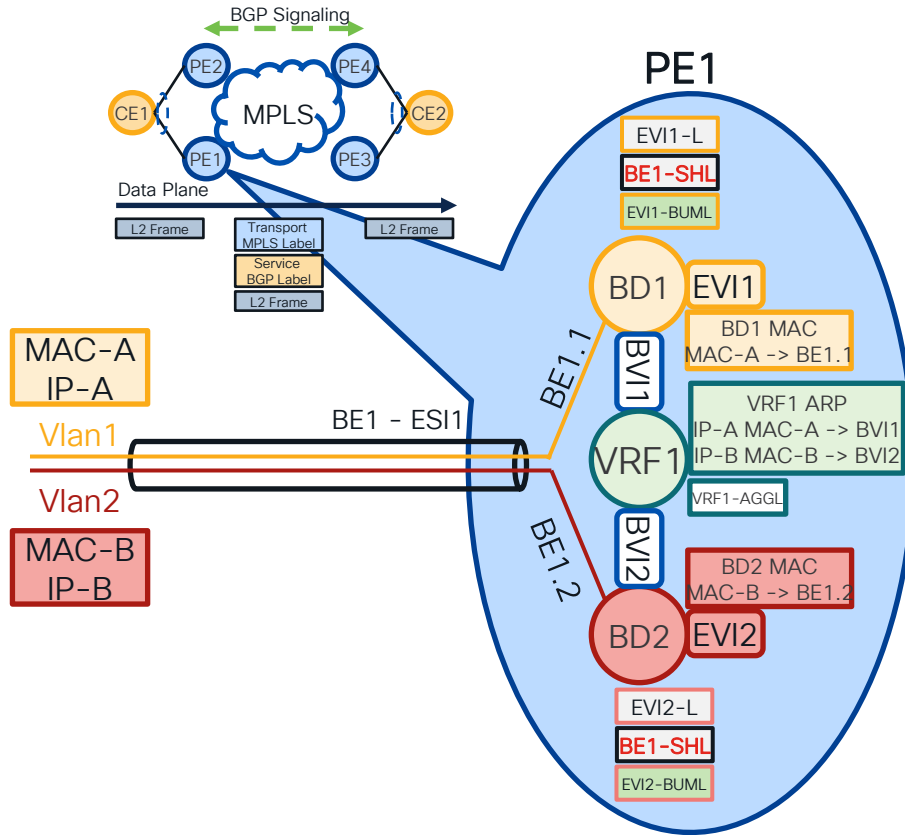
R36#show route vrf a

```
C  192.168.1.0/24 is directly connected, 04:55:09, BVI100
L  192.168.1.1/32 is directly connected, 04:55:09, BVI100
B  192.168.1.10/32 [200/0] via 3.3.3.37 (nexthop in vrf default)
B  192.168.2.0/24 [200/0] via 3.3.3.38 (nexthop in vrf default), 00:40:26
    [200/0] via 3.3.3.39 (nexthop in vrf default), 00:40:26
B  192.168.2.20/32 [200/0] via 3.3.3.38 (nexthop in vrf default), 00:40:26
    [200/0] via 3.3.3.39 (nexthop in vrf default), 00:40:26
RP/0/RP0/CPU0:R36
```

EVPN Routes Summary



EVPN Routes – Cheat Sheet



PE1 – Advertises:

RT-4 Ethernet Segment Route

- I have **ESI1** in case when someone needs this information for Designated Forwarder(DF) Election

RT-1 Per **ESI** Ethernet Auto-Discovery (AD) Route

- I have **ESI1**
- ESI1 is All-Active
- AC with ESI1 is connected to **EVI1** and **EVI2**
- My Split Horizon Label for ESI1 is **BE1-SHL**

RT-1 Per **EVI** Ethernet Auto-Discovery (AD) Route(s)

- EVI1** per-EVI (Aliasing) Label is **EVI1-L**
- EVI2** per-EVI (Aliasing) Label is **EVI2-L**

RT-3 Inclusive Multicast Route(s)

- EVI1** Label for BUM traffic is **EVI1-BUML**
- EVI2** Label for BUM traffic is **EVI2-BUML**

RT-2 MAC/IP Advertisement Route(s)

- MAC-A** in **EVI1** via label **EVI1-L** and **IP-A** in **VRF1** via label **VRF1-AGGL**
- MAC-B** in **EVI2** via label **EVI2-L** and **IP-B** in **VRF1** via label **VRF1-AGGL**

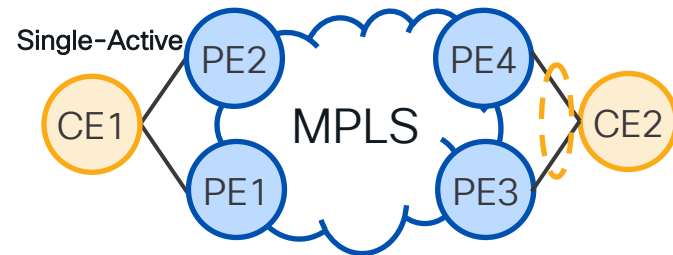
RT-5 Prefix Advertisement Route(s)

- IPv4/6 prefix of **BV11** in **VRF1** via label **VRF1-AGGL**
- IPv4/6 prefix of **BV12** in **VRF1** via label **VRF1-AGGL**

EVPN-VPWS Multihomed Service

EVPN vs EVPN-VPWS – Balancing Mode

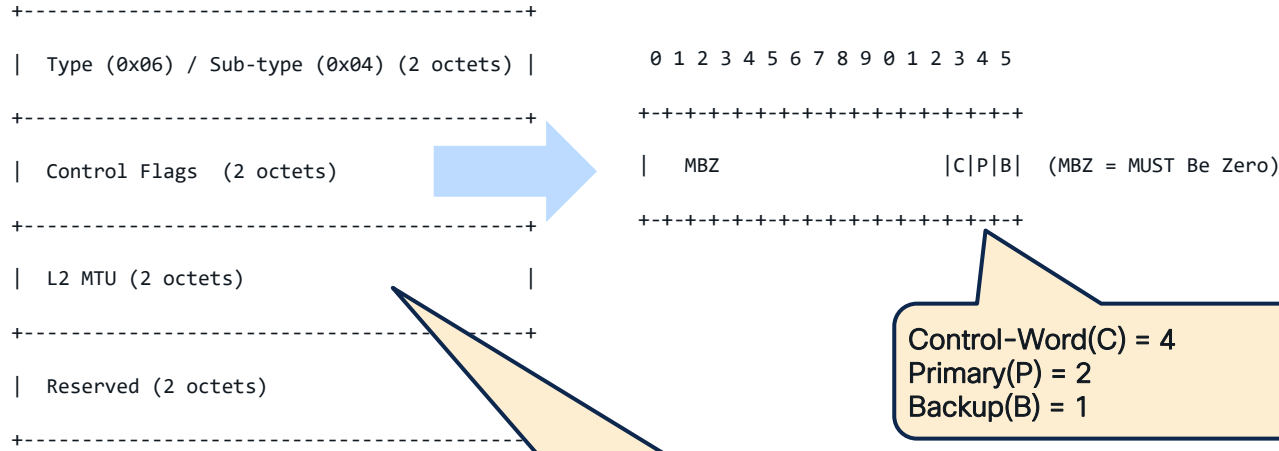
- Both EVPN and EVPN-VPWS advertise RT1(per-ESI)
 - Signal All-Active or Single-Active



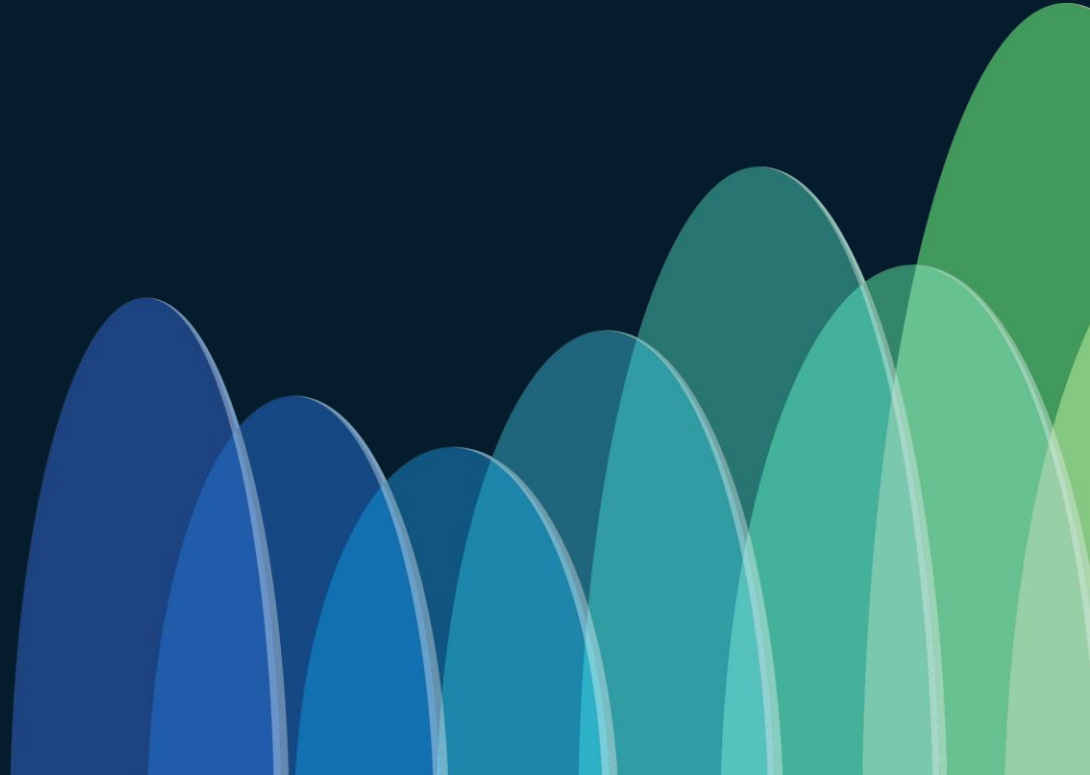
- Remote node performs per-flow load-balancing -> All-Active mode
- How remote node knows who is Active in Single-Active mode?
 - EVPN
 - Remote node follows MAC (RT2) advertisement -> node advertising MAC is active
 - EVPN-VPWS
 - Additional signaling per-service is required to inform remote node who is Active

EVPN-VPWS Layer 2 Attributes

Extended Community

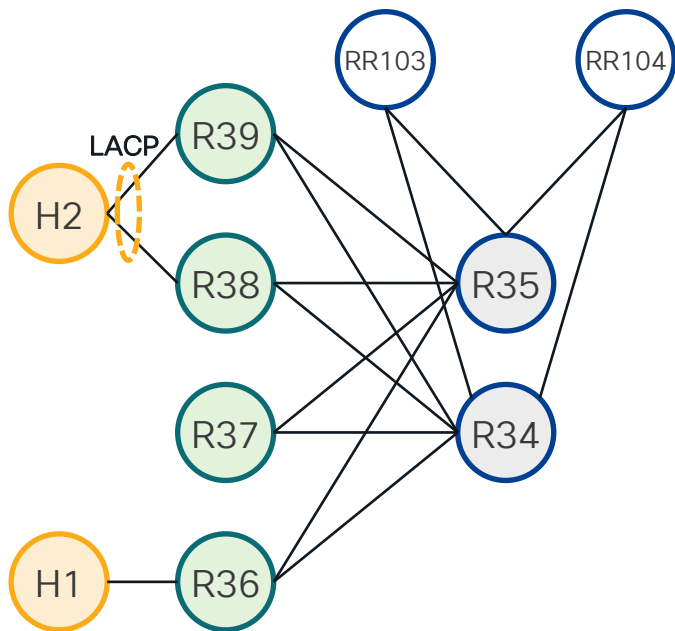


EVPN-VPWS All-Active



EVPN-VPWS - Testbed

Startup Sequence is almost identical with EVPN except:
RT3 and RT2 are not required



R38/R39

```
l2vpn
xconnect group 500
  p2p 500
    interface Bundle-Ether100
      neighbor evpn evi 500 service 333
    !
  !
!
```

R36

```
l2vpn
xconnect group 500
  p2p 500
    interface Bundle-Ether100
      neighbor evpn evi 500 service 333
    !
  !
!
```

R36: L2vpn xconnect status & Data Plane verification

R36#show l2vpn xconnect

Legend: ST = State, UP = Up, DN = Down, AD = Admin Down, UR = Unresolved,
SB = Standby, SR = Standby Ready, (PP) = Partially Programmed

XConnect Group	Name	ST	Segment 1 Description	ST	Segment 2 Description	ST
500	500	UP	BE100	UP	EVPN 500,3839,68106	UP

R36#show mpls forwarding labels 68106

Local Label	Outgoing Label	Prefix or ID	Outgoing Interface	Next Hop	Bytes Switched
68106	68107	EVPN:500		3.3.3.38	0
	68107	EVPN:500		3.3.3.39	0

R36: RT-1 Per EVI Ethernet Auto-Discovery

```
R36#show bgp l2vpn evpn rd 3.3.3.36:500 [1][0038.3900.0000.0000.1100][3839]/120
```

BGP routing table entry for [1][0038.3900.0000.0000.1100][3839]/120, Route Distinguisher: 3.3.3.36:500

Versions:

Process	bRIB/RIB	SendTblVer
Speaker	316	316

Last Modified: Jan 27 08:24:37.527 for 00:01:42

Paths: (2 available, best #1)

Not advertised to any peer

Path #1: Received by speaker 0

Not advertised to any peer

Local

3.3.3.38 (metric 30) from 3.3.3.103 (3.3.3.38)

Received Label 68107

Origin IGP, localpref 100, valid, internal, best, sup-best, import-candidate, imported, rib-install

Received Path ID 0, Local Path ID 1, version 314

Extended community: EVPN L2 ATTRS:0x06:1500 RT:1:500

Originator: 3.3.3.38, Cluster list: 3.3.3.103

Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 3.3.3.38:500

Path #2: Received by speaker 0

Not advertised to any peer

Local

3.3.3.39 (metric 30) from 3.3.3.103 (3.3.3.39)

Received Label 68107

Origin IGP, localpref 100, valid, internal, import-candidate, imported, rib-install

Received Path ID 0, Local Path ID 0, version 0

Extended community: EVPN L2 ATTRS:0x06:1500 RT:1:500

Originator: 3.3.3.39, Cluster list: 3.3.3.103

Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 3.3.3.39:500

Control-Word(C) = 4

Primary(P) = 2

Backup(B) = 1

Control-Word + Primary
MTU 1500B

Control-Word + Primary
MTU 1500B

R36: EVPN-VPWS Instance View

```
R36#show evpn evi vpn-id 500 detail
```

VPN-ID	Encap	Bridge Domain	Type
500	MPLS	VPWS:500	VPWS (vlan-unaware)

Stitching: Regular

Unicast Label : 0

Multicast Label: 0

Flow Label: N

Control-Word: Enabled

Forward-class: 0

Advertise MACs: No

Advertise BVI MACs: No

Aliasing: Enabled

UUF: Enabled

Re-origination: Enabled

Multicast source connected: No

Statistics:

Packets	Sent	Received
Total	: 0	0
Unicast	: 0	0
BUM	: 0	0

Bytes	Sent	Received
Total	: 0	0
Unicast	: 0	0
BUM	: 0	0

RD Config: none

RD Auto : (auto) 3.3.3.36:500

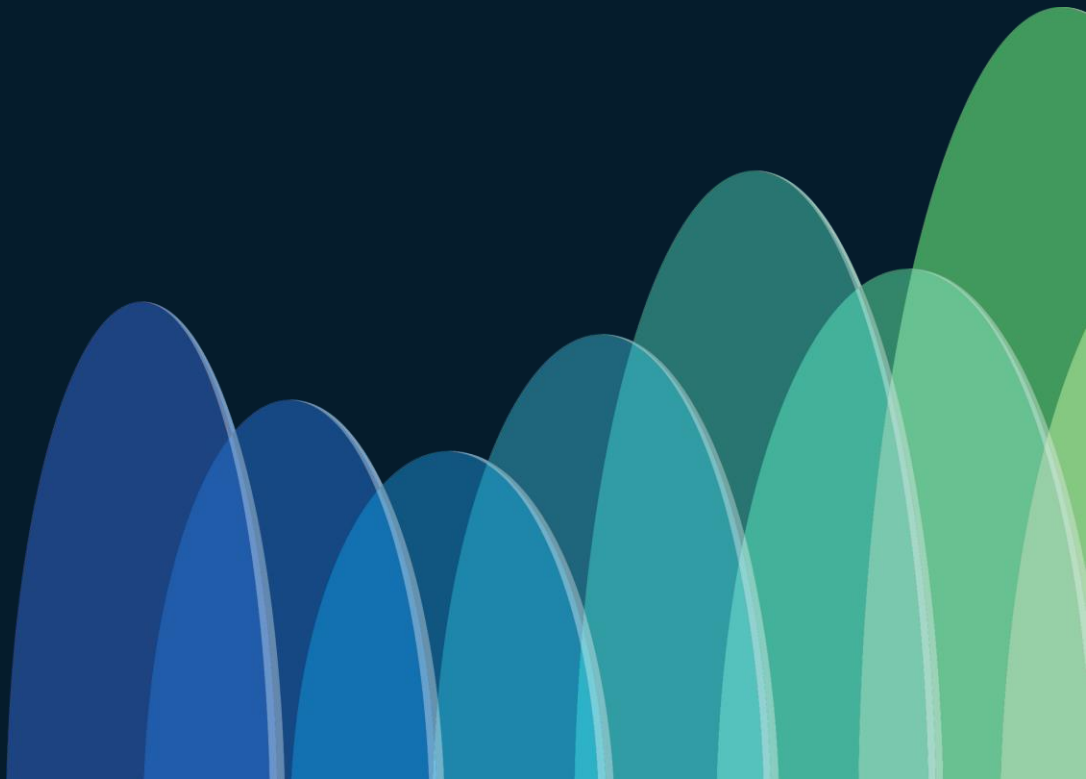
RT Auto : 1:500

Route Targets in Use	Type
1:500	Import
1:500	Export

EVPN-VPWS

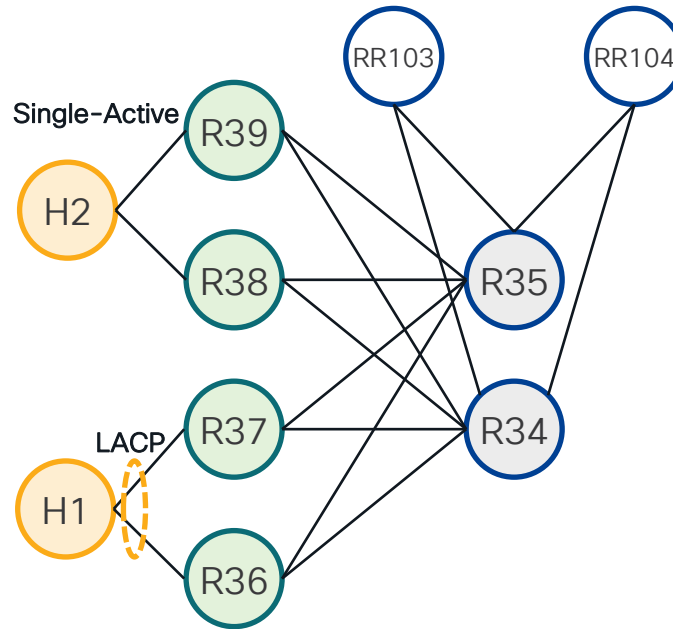
- No RT2 - MAC
- No RT3 - BUM

EVPN-VPWS Single-Active



EVPN-VPWS - Testbed

Startup Sequence is almost identical with EVPN except:
RT3 and RT2 are not required



Config: EVPN-VPWS

R36

```
l2vpn
xconnect group 500
p2p 500
 interface Bundle-Ether100
  neighbor evpn evi 500 target 3839 source 3637
!
!
!
```

R38/R39

```
l2vpn
xconnect group 500
p2p 500
 interface Bundle-Ether100
  neighbor evpn evi 500 target 3637 source 3839
!
!
!
```

R36: L2vpn xconnect status & Data Plane verification

```
R36#show l2vpn xconnect
```

XConnect Group	Name	ST	Segment 1 Description	ST	Segment 2 Description	ST
500	500	UP	BE100	UP	EVPN 500,3839,24004	UP

```
R36#show mpls forwarding labels 24004
```

Local Label	Outgoing Label	Prefix or ID	Outgoing Interface	Next Hop	Bytes Switched
24004	28127	EVPN:500		3.3.3.39	0
	28127	EVPN:500		3.3.3.38	0

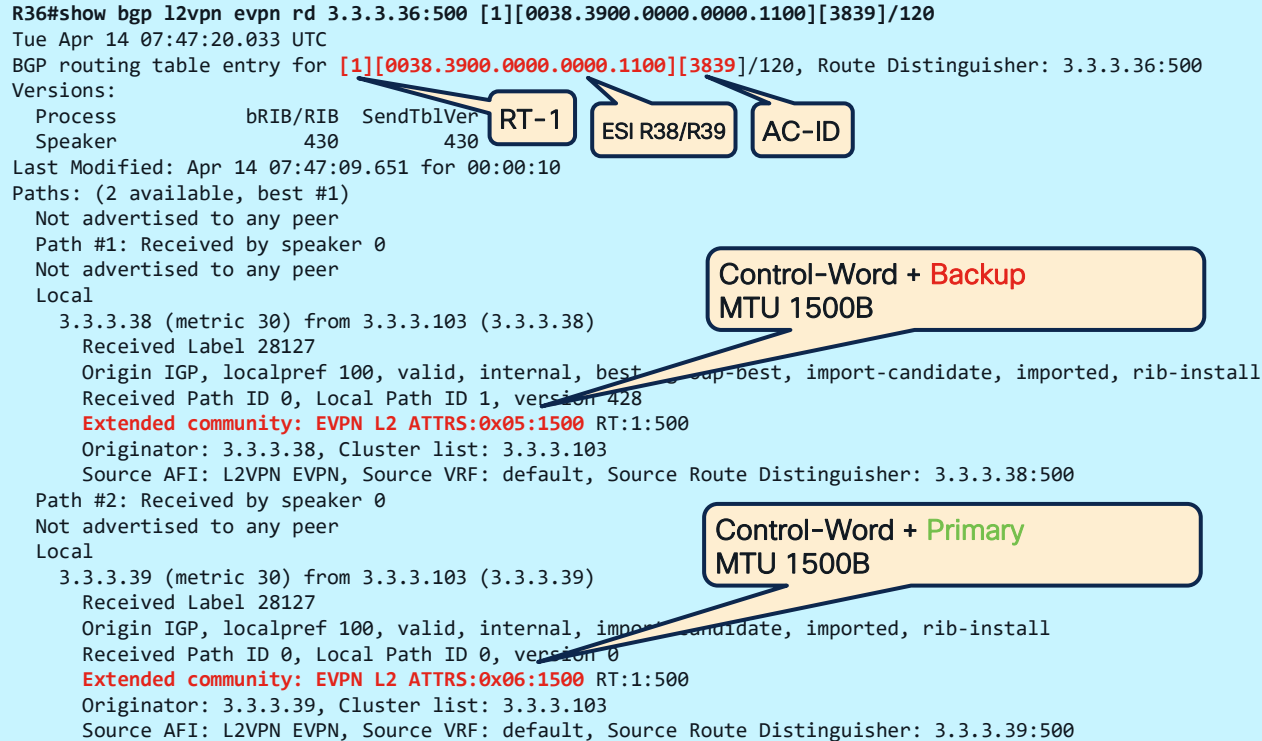
Active

Standby

(!)

R36: RT-1 Per EVI Ethernet Auto-Discovery

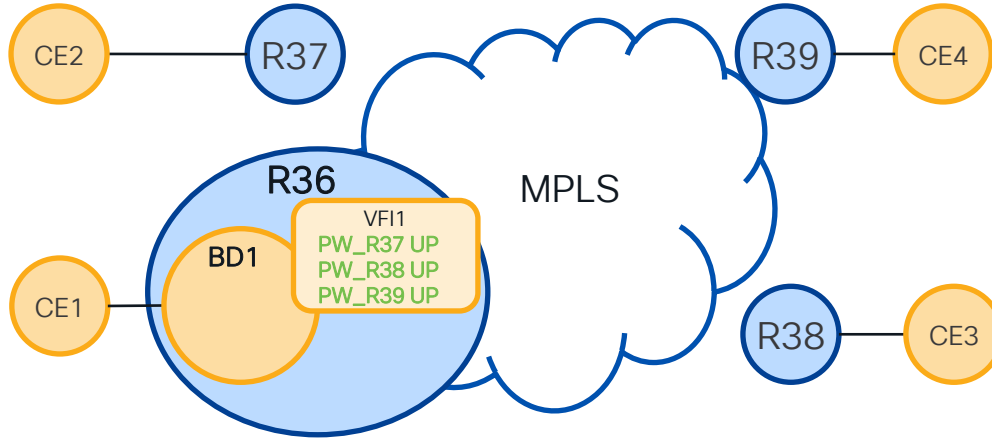
```
R36#show bgp l2vpn evpn rd 3.3.3.36:500 [1][0038.3900.0000.0000.1100][3839]/120
Tue Apr 14 07:47:20.033 UTC
BGP routing table entry for [1][0038.3900.0000.0000.1100][3839]/120, Route Distinguisher: 3.3.3.36:500
Versions:
  Process          bRIB/RIB  SendTblVer
  Speaker          430      430
Last Modified: Apr 14 07:47:09.651 for 00:00:10
Paths: (2 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  Local
    3.3.3.38 (metric 30) from 3.3.3.103 (3.3.3.38)
    Received Label 28127
    Origin IGP, localpref 100, valid, internal, best-external, import-candidate, imported, rib-install
    Received Path ID 0, Local Path ID 1, version 428
    Extended community: EVPN L2 ATTRS:0x05:1500 RT:1:500
    Originator: 3.3.3.38, Cluster list: 3.3.3.103
    Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 3.3.3.38:500
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local
    3.3.3.39 (metric 30) from 3.3.3.103 (3.3.3.39)
    Received Label 28127
    Origin IGP, localpref 100, valid, internal, import-candidate, imported, rib-install
    Received Path ID 0, Local Path ID 0, version 0
    Extended community: EVPN L2 ATTRS:0x06:1500 RT:1:500
    Originator: 3.3.3.39, Cluster list: 3.3.3.103
    Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 3.3.3.39:500
```



Control-Word(C) = 4
Primary(P) = 2
Backup(B) = 1

VPLS to EVPN Seamless Migration

VPLS & EVPN Seamless Integration – Migration

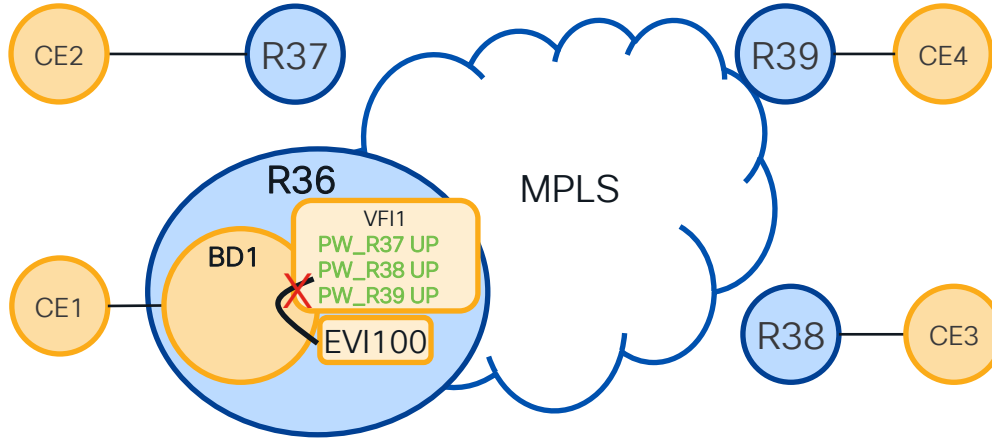


VFI1 is by default in Split Horizon Group 1

- SHG1 protects loops in MPLS Core
- Full Mesh of pseudowires(PW) is required for Any-to-Any forwarding

```
12vpn
bridge group 100
bridge-domain 100
vfi 1
neighbor x.x.x.37 pw-id 37
!
neighbor x.x.x.38 pw-id 38
!
neighbor x.x.x.39 pw-id 39
!
!
```

VPLS & EVPN Seamless Integration – Migration



VFI1 is by default in Split Horizon Group 1

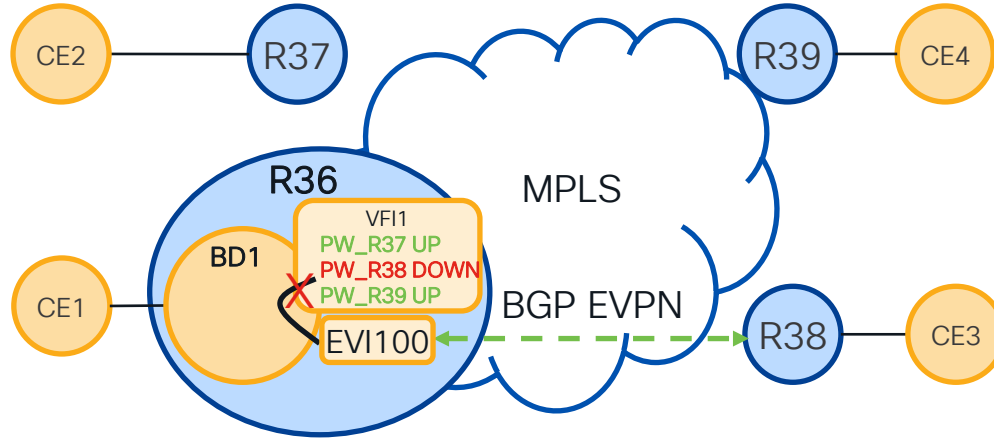
- SHG1 protects loops in MPLS Core
- Full Mesh of pseudowires(PW) is required for Any-to-Any forwarding

EVI100 is also by default in Split Horizon Group 1

- R36 doesn't forward data between VFI1 and EVI100

```
12vpn
bridge group 100
bridge-domain 100
vfi 1
neighbor x.x.x.37 pw-id 37
!
neighbor x.x.x.38 pw-id 38
!
neighbor x.x.x.39 pw-id 39
!
evi 100
!
```

VPLS & EVPN Seamless Integration – Migration



```
12vpn
bridge group 100
bridge-domain 100
vfi 1
neighbor x.x.x.37 pw-id 37
!
neighbor x.x.x.38 pw-id 38
!
neighbor x.x.x.39 pw-id 39
!
evi 100
!
```

VFI1 is by default in Split Horizon Group 1

- SHG1 protects loops in MPLS Core
- Full Mesh of pseudowires(PW) is required for Any-to-Any forwarding

EVI1 is also by default in Split Horizon Group 1

- R36 doesn't forward data between VFI1 and EVI100

R36&R38 run BGP EVPN

- **PW_R38 goes DOWN**
- **Data Forwarding between R36 and R38 via EVI100**

PW to EVPN-VPWS Seamless Migration

EVPN-VPWS/Legacy-PW Seamless Migration

Supported Modes

Discovery: Static/BGP-AD

Signaling: LDP, BGP



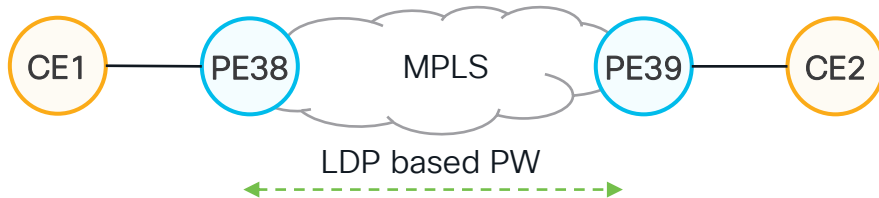
R38 Configuration

```
12vpn
xconnect group test
p2p test
interface TenGigE0/0/0/0
neighbor ipv4 3.3.3.39 pw-id 10
```

R38#show 12vpn xconnect

XConnect			Segment 1		Segment 2		
Group	Name	ST	Description	ST	Description		ST
test	test	UP	Te0/0/0/0	UP	3.3.3.39	10	UP

EVPN-VPWS/Legacy-PW Seamless Migration



R38 Configuration

```
l2vpn
 xconnect group test
  p2p test
   vpws-seamless-integration
   interface TenGigE0/0/0/0
    neighbor ipv4 3.3.3.39 pw-id 10

 p2p test-new
  interface TenGigE0/0/0/0
   neighbor evpn evi 1000 service 10
```

Allows Tengig0/0/0/0 to be migrated

Existing LDP based PW is UP and forwarding data
New EVPN-VPWS service is ready and is signaled via BGP EVPN AF

R38#show l2vpn xconnect

XConnect Group		Name	ST	Segment 1 Description	ST	Segment 2 Description	ST
test		test	UP	Te0/0/0/0	UP	3.3.3.39 10	UP
test		test-new	DN	Te0/0/0/0	UP	EVPN 1000,10,None	DN

```
R38#show bgp l2vpn evpn rd 3.3.3.38:1000
Route Distinguisher: 3.3.3.38:1000 (default for vrf VPWS:1000)
*> [1][0000.0000.0000.0000.0000][10]/120
0.0.0.0
```

EVPN-VPWS/Legacy-PW Seamless Migration



LDP based PW - DOWN
 EVPN-VPWS - UP

R38 Configuration

```

12vpn
xconnect group test
p2p test
  vpws-seamless-integration
  interface TenGigE0/0/0/0
  neighbor ipv4 3.3.3.39 pw-id 10
p2p test-new
  interface TenGigE0/0/0/0
  neighbor evpn evi 1000 service 10
  
```

R39 Configuration

```

12vpn
xconnect group test
p2p test
  vpws-seamless-integration
  interface TenGigE0/0/0/0
  neighbor ipv4 3.3.3.38 pw-id 10
p2p test-new
  interface TenGigE0/0/0/0
  neighbor evpn evi 1000 service 10
  
```

EVPN-VPWS is UP
 LDP PW is Down and service is in "Seamless Inactive" mode
 p2p test can be removed

R38#show 12vpn xconnect

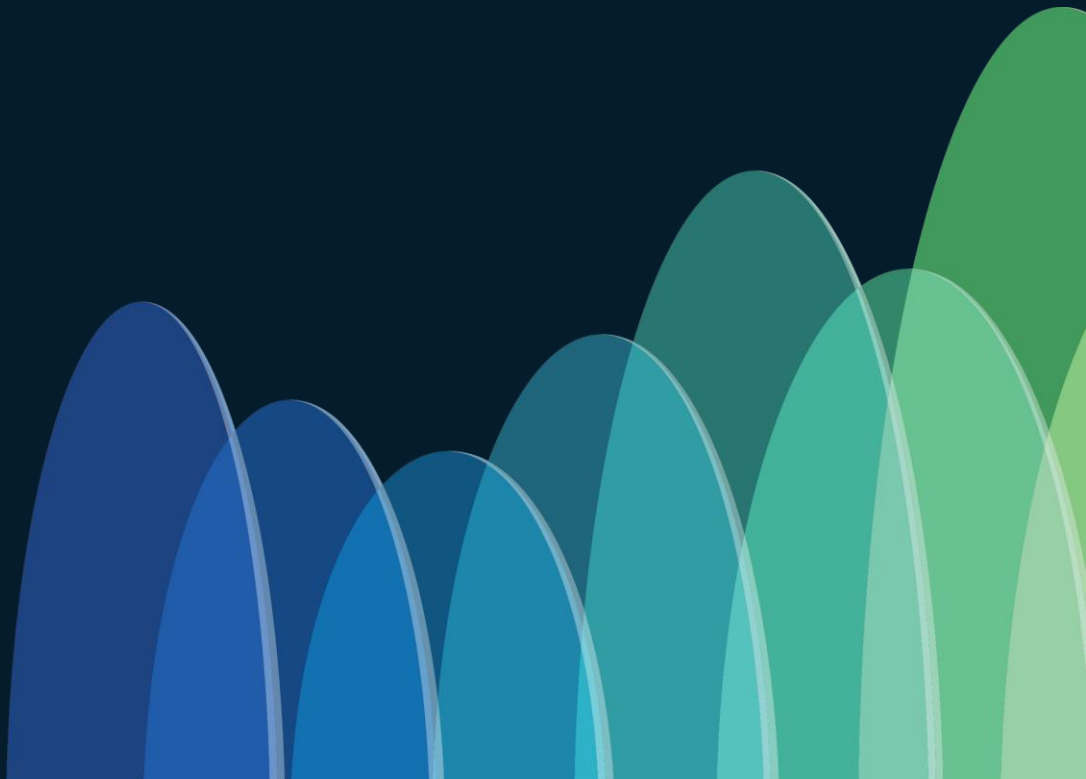
XConnect Group	Name	ST	Segment 1 Description	ST	Segment 2 Description	ST
test	test	DN	Te0/0/0/0	SB(SI)	3.3.3.39 10	UP
test	test-new	UP	Te0/0/0/0	UP	EVPN 1000,10,3.3.3.39	UP

R38#show bgp 12vpn evpn rd 3.3.3.38:1000

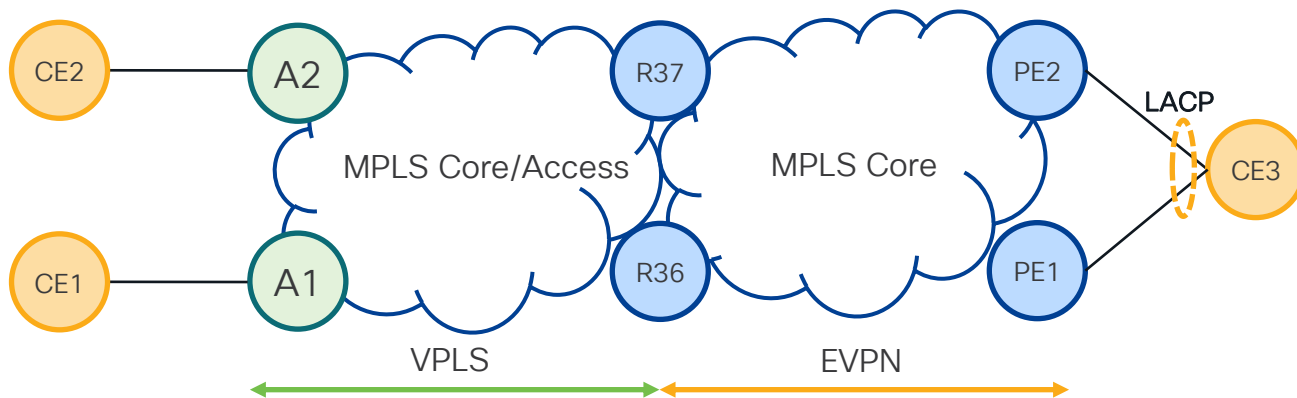
```

Network          Next Hop          Metric LocPrf Weight Path
Route Distinguisher: 3.3.3.38:1000 (default for vrf VPWS:1000)
*> [1][0000.0000.0000.0000.0000][10]/120
          0.0.0.0          0 i
* i          3.3.3.39          100 0 i
  
```

EVPN & VPLS Interconnect



EVPN & VPLS Interconnect



R36/R37 Configuration

```
evpn
 evi 100
  advertise-mac
  !
  virtual vfi 1
   ethernet-segment
    identifier type 0 11.11.11.11.11.11.11.11
```

Virtual Ethernet Segment (vES)
• VPLS is Single-Active Access to EVPN

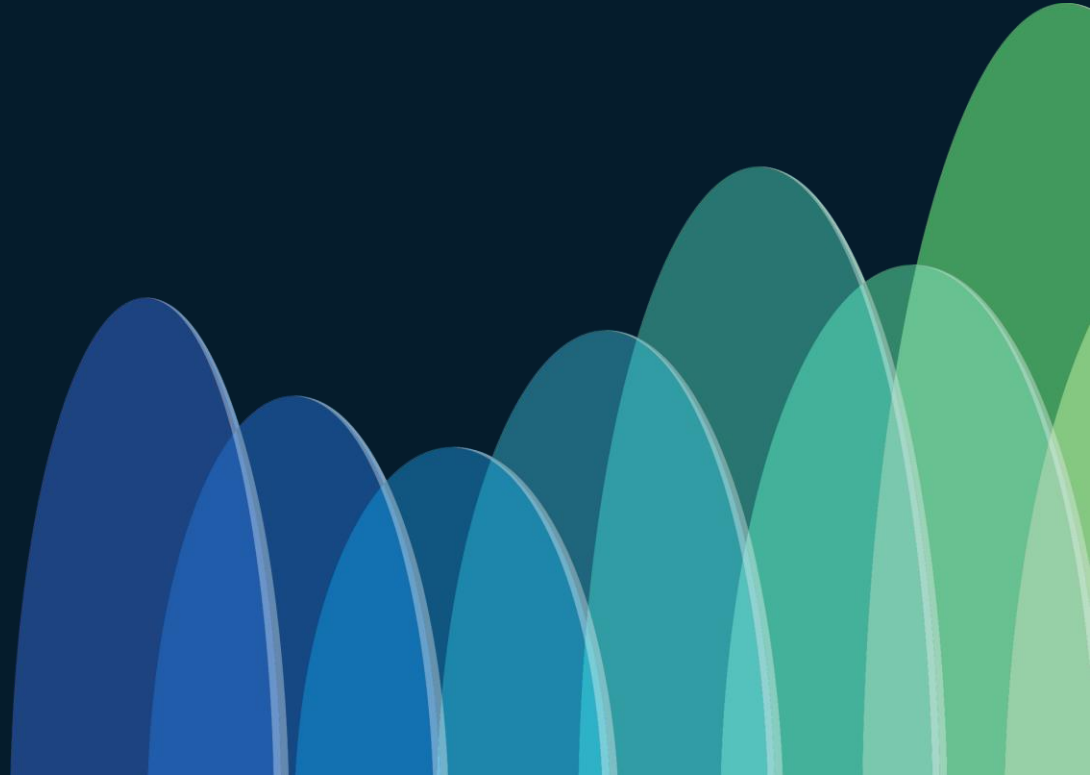
R36 Configuration

```
12vpn
 bridge group 100
 bridge-domain 100
 access-vfi 1
  neighbor x.x.x.A1 pw-id 1
  !
  neighbor x.x.x.A2 pw-id 2
  !
  !
 evi 100
```

R37 Configuration

```
12vpn
 bridge group 100
 bridge-domain 100
 access-vfi 1
  neighbor x.x.x.A1 pw-id 10
  !
  neighbor x.x.x.A2 pw-id 20
  !
  !
 evi 100
```

Summary



Summary

- Extra Offline Learning:
 - ETREE
 - EVPN Fast Re-Route (FRR)
- EVPN is an very important complement to BGP based services
- BGP is Unified Services Control Plane across Network
- EVPN All-Active Multihomed Service with Distributed Anycast Gateway & Integration to L3VPN simplifies SPDC/NextGen-CO/WAN Integration
- EVPN is not strictly a replacement of “traditional” VPNv4/6
 - EVPN and VPNv4/6 can coexist
- Service Layer is Data Plane independent, but the right Data Plane (encapsulation) selection decreases complexity and provides additional capabilities
- Stay up to date <https://e-vpn.io/>

Keynote Deep Dives

Wednesday
10:30am – 11:30am



Experiences Amplified:
How AI Can Fuel
Better Employee
and Customer
Experiences

Level 1
Room 106



Smart, Secure,
Seamless:
Transforming
Experiences with
Next-Generation
Networking

Level 2
Room 204



Harness a Bold
New Era:
Transform Data
Centre and
Service Provider
Connectivity

Level 2
Room 203



Securing User to
Application and
Everything in
Between

Level 2
Melbourne Room 2



Unlocking
Digital Resilience
through Unified
Observability

The HUB
Centre Stage

Complete Your Session Evaluations



Complete a minimum of 4 session surveys and the Overall Event Survey to claim a **Cisco Live T-Shirt**.



Complete your surveys in the **Cisco Live mobile app**.



Continue your education



- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

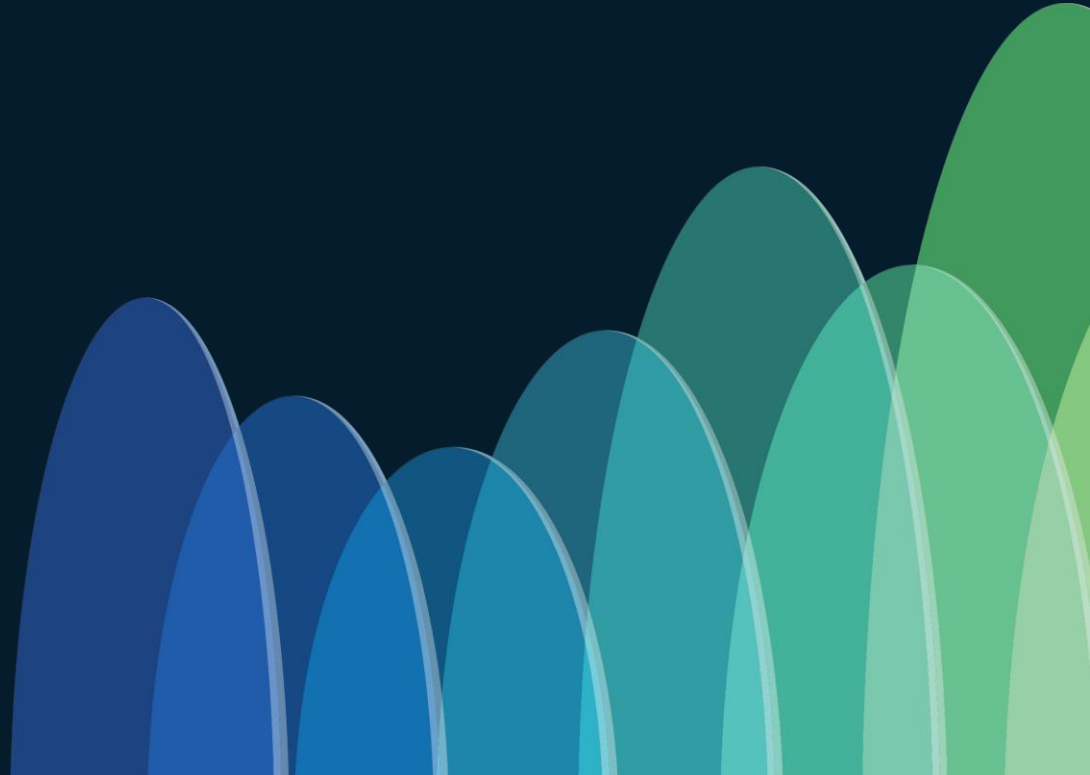


Thank you

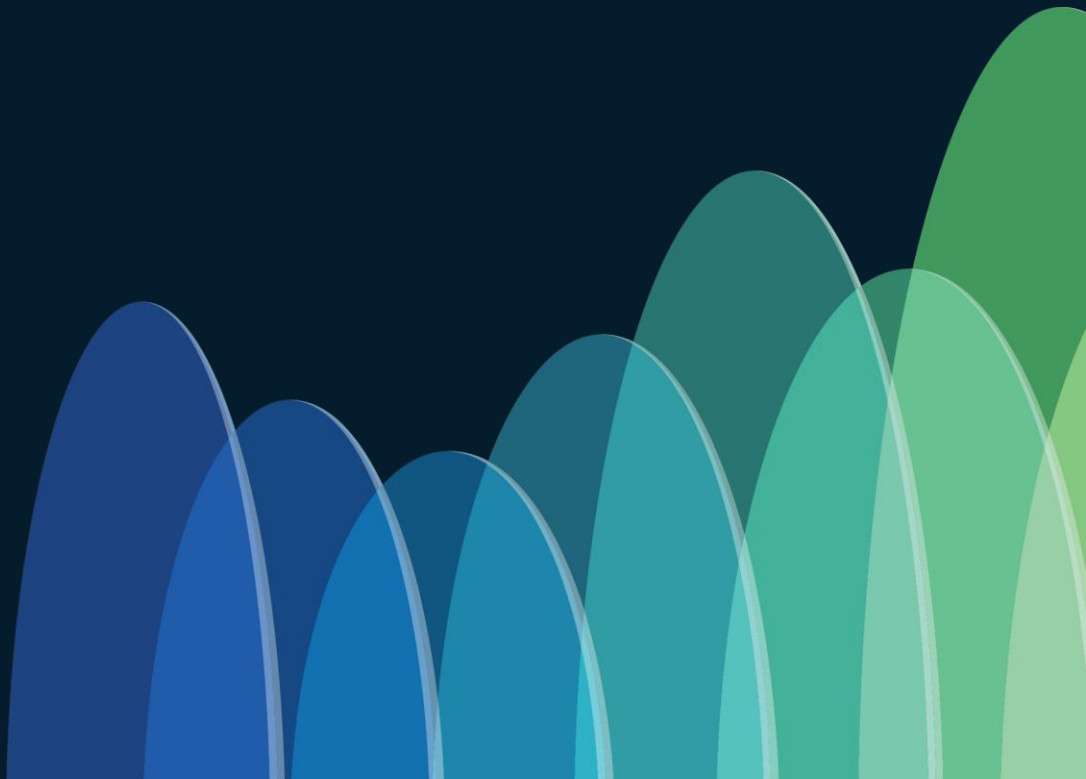
CISCO *Live!*

#CiscoLiveAPJC

Extra Offline Learning

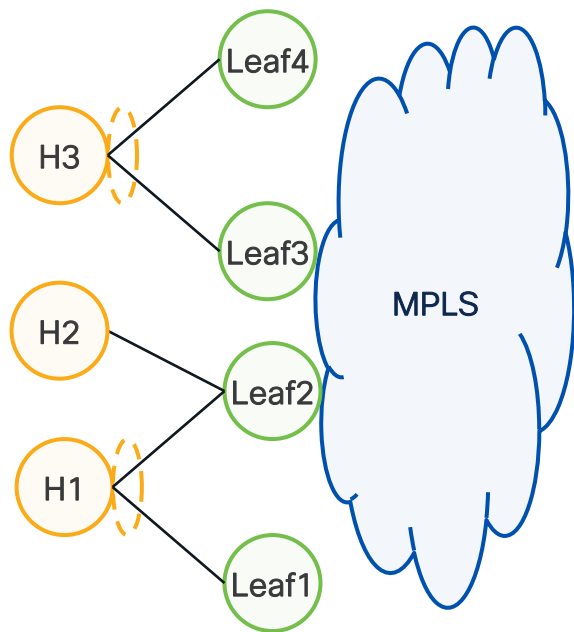


EVPN ETREE



EVPN ETREE – RT Constrains (Scenario 1a)

- Host connected to Leaf can talk **ONLY** to device connected to Root
- H1, H2, H3 can talk to H4
- **H1, H2, H3 CANNOT talk to each other**



Root Configuration

```
evpn
 evi 100
  bgp
    route-target export 1:1000
    route-target import 1:1000
    route-target import 1:100
  !
!
```

Leaf Configuration

```
evpn
 evi 100
  bgp
    route-target export 1:100
    route-target import 1:1000
  !
 etree
  rt-leaf <- MAC Synchronization
  !
!
```

Leaf Additional Configuration Prevents H1 and H2 to talk locally

```
l2vpn
 bridge-group evpn
 bridge-domain evpn100
  interface TenGigE0/0/0/0
    split-horizon group
  !
  interface Bundle-Ether100
    split-horizon group
  !
!
```

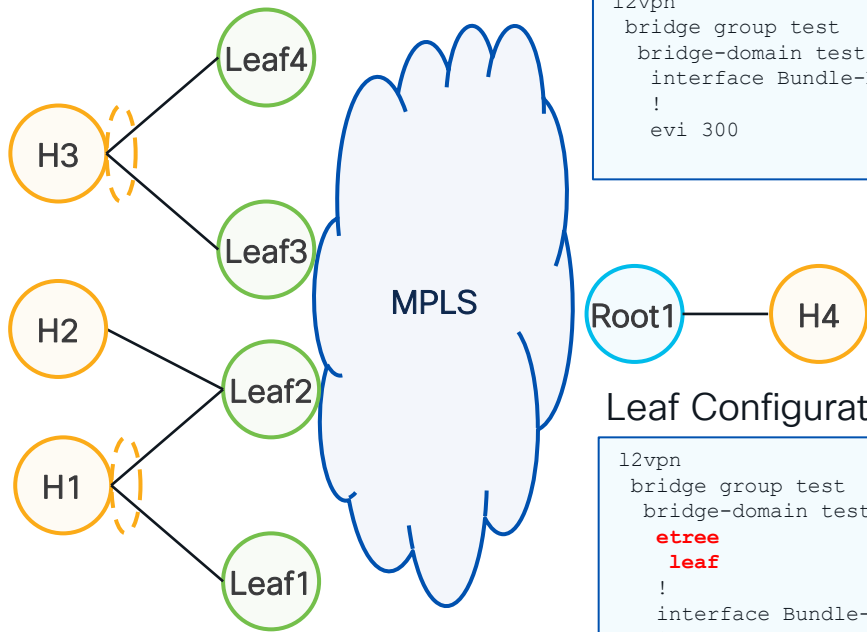
EVPN ETREE Leaf Label (Scenario 1b)

Root Configuration

No specific Root Configuration

```
l2vpn
bridge group test
bridge-domain test
interface Bundle-Ether100
!
evi 300
```

- ASR9k/NCS add Leaf ACs to SHG2 automatically
=> Prevents local Leaf to Leaf AC forwarding



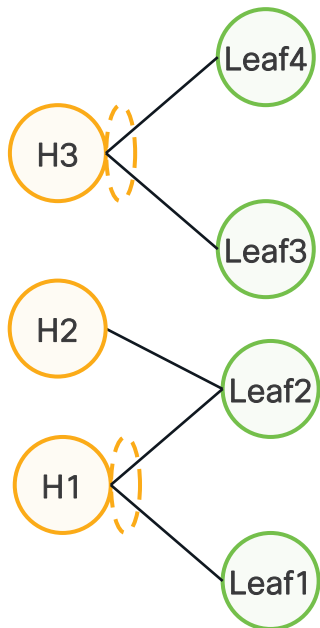
Leaf Configuration

```
l2vpn
bridge group test
bridge-domain test
etree
leaf
!
interface Bundle-Ether100
!
evi 300
```

EVPN ETREE Leaf Label (Scenario 1b) - BUM

Leaf Configuration

```
l2vpn
bridge group test
bridge-domain test
  etree
  leaf
!
interface Bundle-Ether100
!
evi 300
```



Each Leaf (device with at least one Leaf AC) advertises RT1 per-ESI with ESI 0 with ETREE extended community to distribute ETREE Label

```
R28#show bgp l2vpn evpn rd 1.1.1.28:0 [1][1.1.1.28:1][0000.0000.0000.0000][4294967295]/184
Wed Mar 23 03:41:36.734 UTC
BGP routing table entry for [1][1.1.1.28:1][0000.0000.0000.0000][4294967295]/184, Route Distinguisher: 1.1.1.28:0
Versions:
  Process          bRIB/RIB  SendTblVer
  Speaker          1481327   1481327
    Local Label: 0
Last Modified: Mar 23 03:21:20.580 for 00:20:17
Paths: (1 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.2
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.2
  Local
    0.0.0.0 from 0.0.0.0 (1.1.1.28)
    Origin IGP, localpref 100, valid, redistributed, best, group-best, import-candidate, rib-install
    Received Path ID 0, Local Path ID 1, version 1481327
    Extended community: EVPN E-TREE:0x00:24010 RT:1:3000
```

ETREE Label works same as Split-Horizon Label (SHL)

SHL prevents BUM forwarding between two ACs with the same ESI

ETREE Label prevents forwarding between Leaves ACs

Leaf to Leaf BUM traffic has ETREE Label

If Traffic with ETREE label is received cannot be forwarded to Leaf AC

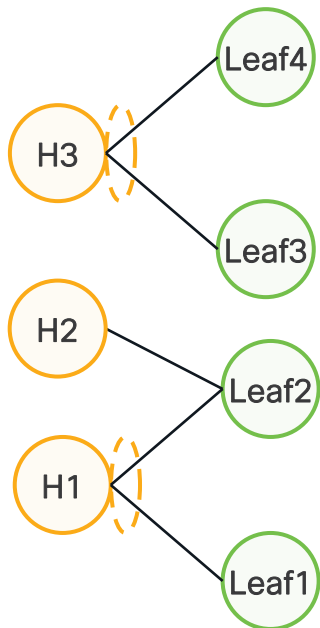
Root to Leaf or Leaf to Root BUM traffic doesn't have ETREE label

BUM between Root <-> Leaf is allowed

EVPN ETREE Leaf Label (Scenario 1b) - Unicast

Leaf Configuration

```
l2vpn
bridge group test
bridge-domain test
  etree
  leaf
!
interface Bundle-Ether100
!
evi 300
```



Leaf Advertises local MAC with ETREE extended community
Same extended community was used to distribute ETREE Label

```
RP/0/RSP0/CPU0:R28#show bgp l2vpn evpn bridge-domain test [2][0][48][682c.7b24.c63d][0]/104
Wed Mar 23 04:13:10.244 UTC
BGP routing table entry for [2][0][48][682c.7b24.c63d][0]/104, Route Distinguisher: 1.1.1.28:300
Versions:
  Process          bRIB/RIB  SendTblVer
  Speaker          1481349   1481349
    Local Label: 24012
Last Modified: Mar 23 03:21:48.580 for 00:51:22
Paths: (1 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.2
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.2
  Local
    0.0.0.0 from 0.0.0.0 (1.1.1.28)
    Origin IGP, localpref 100, valid, redistributed, best, group-best, import-candidate, rib-install
    Received Path ID 0, Local Path ID 1, version 1481349
    Extended community: So0:1.1.1.28:300 EVPN E-TREE:0x01:0 RT:1:300
    EVPN ESI: 0026.2826.2826.2802
```

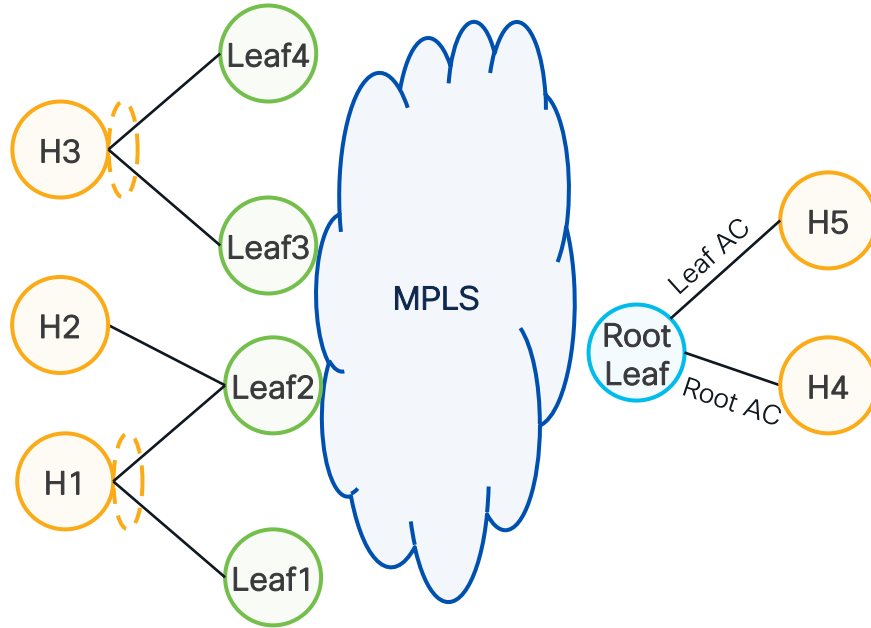
ETREE Label is set to 0, but Leaf Flag is set to 1

Unicast traffic is filtered by ingress node

If traffic is originated from Leaf AC and destination is local/remote Leaf AC frame is dropped

EVPN ETREE Leaf Label (Scenario 2) per-AC

Root/Leaf Configuration



```
l2vpn
bridge group test
bridge-domain test
  interface Bundle-Ether100 <- interface to H4
  interface Bundle-Ether200 <- interface to H5
  etree
  leaf
  !
  !
  evi 300
```

Leaf Configuration Same as Scenario 1b

```
l2vpn
bridge group test
bridge-domain test
  etree
  leaf
  !
  interface Bundle-Ether100
  !
  evi 300
```

EVPN ETREE Summary

Scenario 1a: RT Constrains is simple and HW “friendly”

Unicast/BUM filtering by ingress node => scale benefit

Scenario 1b: Simple configuration, but additional ETREE label must be imposed for BUM

BUM filtered by egress node

Support IRB

Scenario 2: Same principle as Scenario 1b also compatible with Scenario 1b

ASR9k allows to combine Root/Leaf ACs in the same Bridge-Domain

Support IRB

EVPN Fast Re Route (FRR)

Fast Convergence (FRR Data Plane) – Core

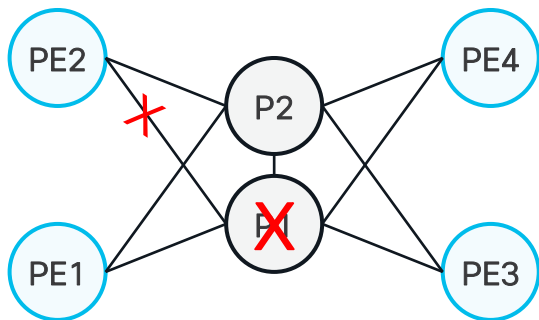
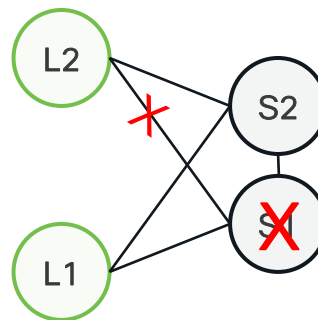
Core Failure (Link/Node) – PIC Core

Technology: RSVP-TE/LFA/rLFA/TI-LFA

Transport: IGP -> MPLS, SRv6

Overlay Service: Service Independent

Device: P-Router, Spine



Fast Convergence (FRR Control Plane) – DC Leaf/TOR MAC Mobility

VM/MAC Move

Technology: EVPN Mac Mobility (EVPN RT-2)

Transport: Transport Independent

Overlay Service: EVPN

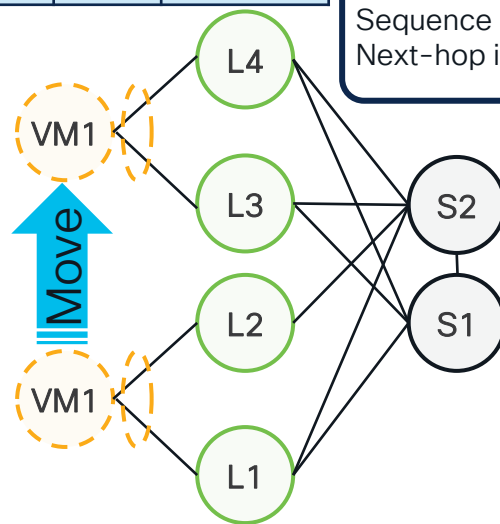
Device: Leaf/TOR

MAC	IP	ESI	Seq.	Next-Hop
MAC-1	IP-1	0	1	Leaf-3/4

Sequence number is incremented and Next-hop is changed to Leaf-3/4

Sequence number and Next-Hop value will be changed after the host move

MAC	IP	ESI	Seq.	Next-Hop
MAC-1	IP-1	0	0	Leaf-1/2



Fast Convergence (FRR CP/DP) – Edge/Leaf/TOR

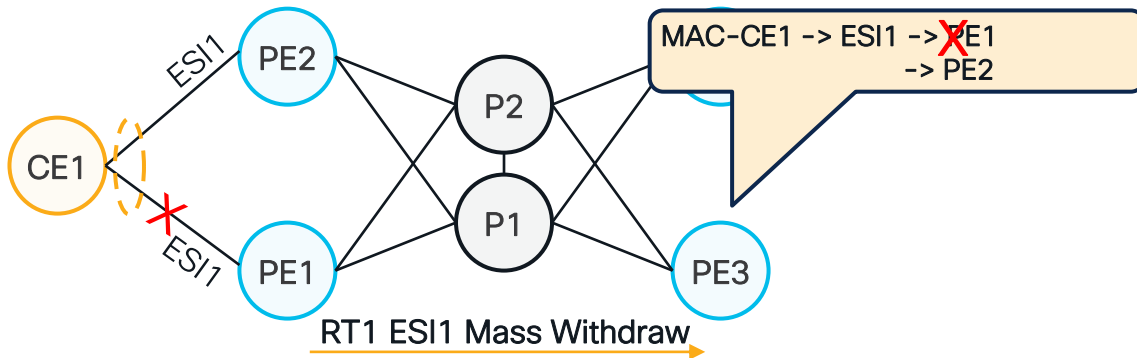
Leaf/TOR Failure (Link) – EVPN Mass Withdraw

Technology: EVPN RT1 Mass Withdraw

Transport: Transport Independent

Overlay Service: EVPN

Device: Leaf/TOR/Access/Edge



Fast Convergence (FRR Data Plane) – Edge L3VPN

Edge Failure (Link) – BGP PIC Edge

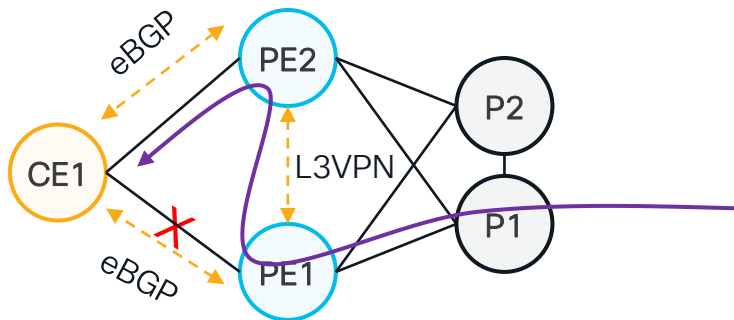
Technology: BGP PIC Edge

Transport: MPLS, SRv6 (Transport Independent)

Overlay Service: L3VPN

Device: Access/PE

BGP CE-PE is mandatory!!!



Fast Convergence (FRR Data Plane) – Edge L2VPN

Edge Failure (Link) – EVPN FRR

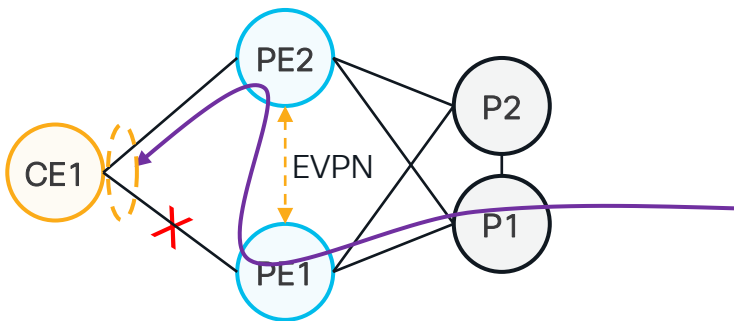
Technology: EVPN FRR

Transport: Transport Independent

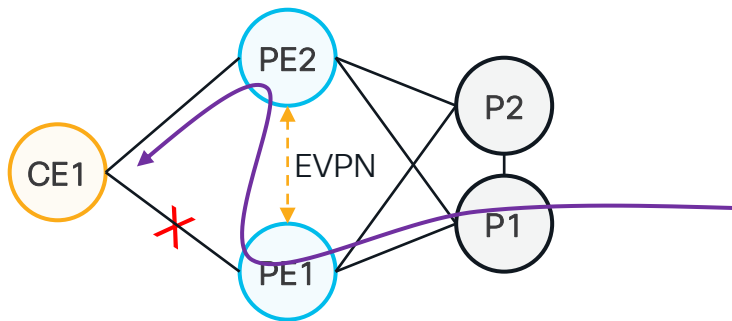
Overlay Service: EVPN

Device: Access/PE/Leaf/TOR

All-Active



Single-Active

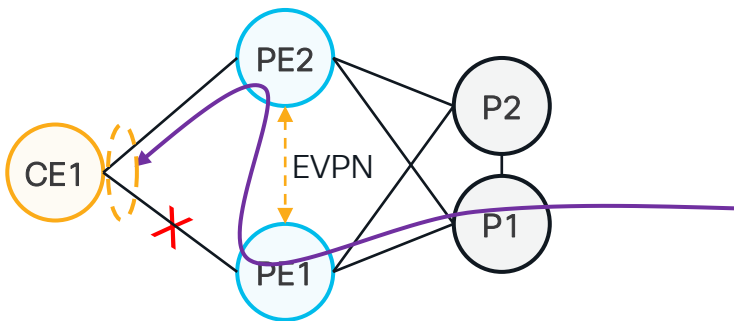


Fast Convergence (EVPN FRR Data Plane) – Edge

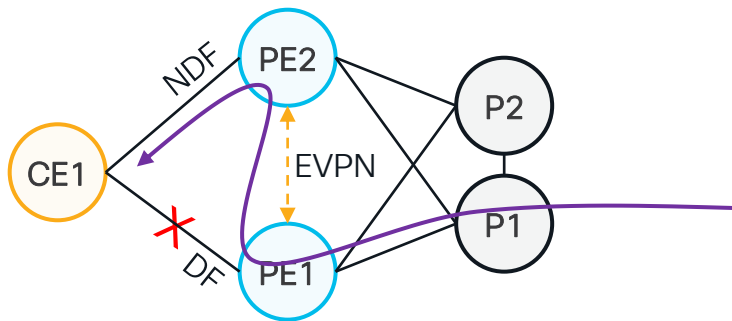
- Single-Active NDF filter traffic in both directions
- Re-Directed traffic will be re-directed back to PE1 (L3 Loop) or dropped

- Solution is to bypass NDF => Only redirected packet can bypass NDF!
 - Extra FRR label is used to bypass NDF
 - FRR Label is used for both All-Active and Single-Active access

All-Active



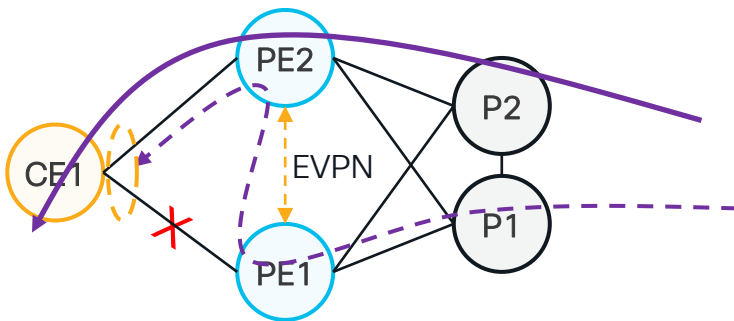
Single-Active



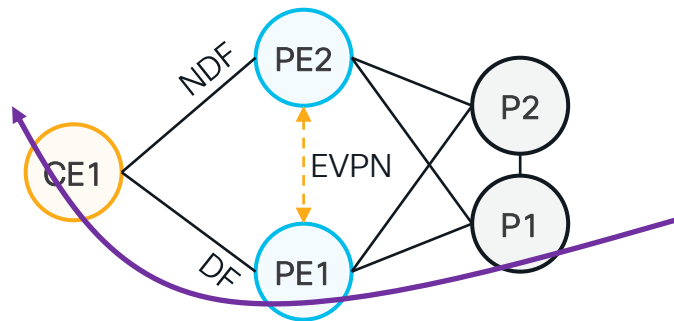
Solving the EVPN DF Election Problem on Recovery

- NTP based solution
- Clocks are synchronized, stratum 3
- Clock timestamping exchange between peering PE
- Service Carving Synchronization

Failure



Recovery



EVPN FRR - Configuration

All-Active

```
evpn
interface Bundle-Ether100
  ethernet-segment
  identifier type 0 36.37.36.37.36.37.36.37.01
  convergence
  reroute
```

Single-Active

```
evpn
interface Bundle-Ether100
  ethernet-segment
  identifier type 0 36.37.36.37.36.37.36.37.01
  load-balancing-mode single-active
  convergence
  reroute
```

DF Election Convergence Improvements

```
evpn
interface Bundle-Ether100
  ethernet-segment
    identifier type 0 36.37.36.37.36.37.01
    load-balancing-mode single-active
    convergence
      nexthop-tracking
    reroute
```

BGP Next-Hop Tracking for RT4
Node Failure Convergence
Improvement

NTP Timestamping for RT4

```
R37#show evpn ethernet-segment carving detail
```

Service Carving Synchronization:

```
Mode           : NTP_SCT
Peer Updates   :
3.3.3.36 [SCT: 2020-10-28 12:57:47:456146]
3.3.3.37 [SCT: 2020-10-28 12:57:47:451599]
```

NTP Timestamping for RT4

```
R37#show ntp status
```

Clock is synchronized, stratum 3, reference is 10.255.11.1

```
R37#show bgp 12vpn evpn rd 3.3.3.36:0 [4][0036.3736.3736.3701][32]
```

```
3.3.3.36 (metric 30) from 3.3.3.103 (3.3.3.36)
  Origin IGP, localpref 100, valid, internal, best, group-best, import-candidate, not-in-vrf
  Received Path ID 0, Local Path ID 1, version 1359
  Extended community: EVPN ES Import:3637.3637.3637 DF Election:0:0x0008:0 EVPN NTP: 3812880149.4488
  Originator: 3.3.3.36, Cluster list: 3.3.3.103
```



Thank you

CISCO *Live!*

#CiscoLiveAPJC

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC