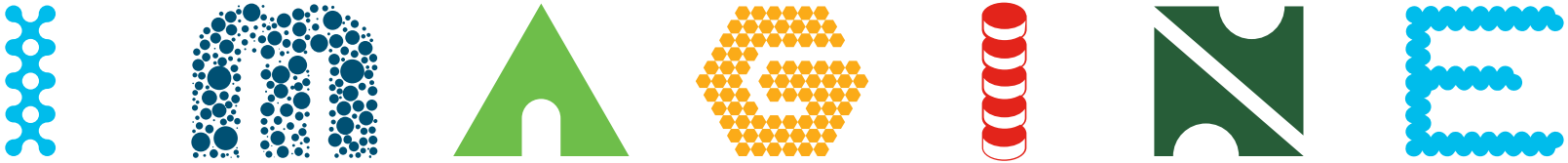
Cisco live!

January 28 - February 1, 2019 • Barcelona

cisco



INTUITIVE

BRKSPG-3965

IOS-XR EVPN Deep Dive

Jiri Chaloupka - Technical Marketing Engineer

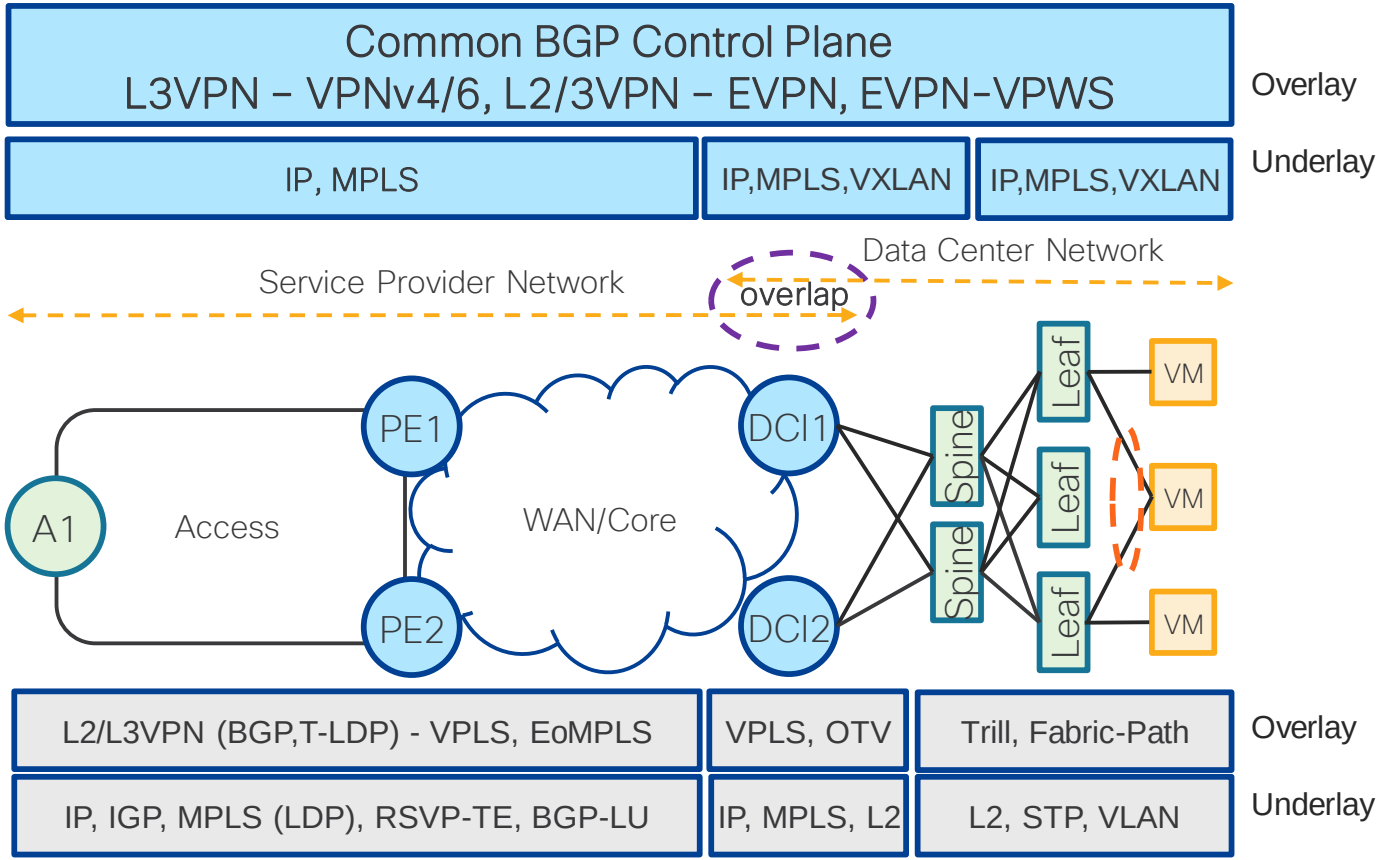
Cisco live!

Agenda

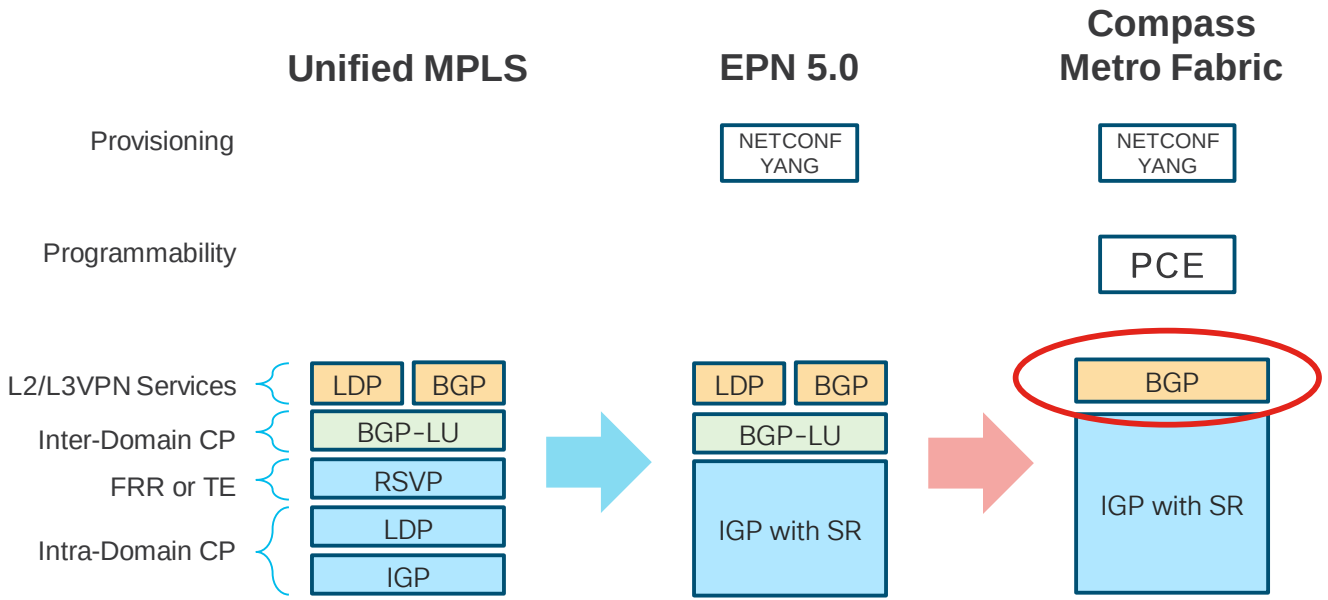
- EVPN Basic Principles
- EVPN L2 All-Active Multihomed Service
- EVPN Distributed L3 Anycast Gateway
- EVPN & VPNv4/6 Interconnect
- EVPN Single-Active
- EVPN Routes - Summary
- EVPN-VPWS All-Active Multihomed Service
- EVPN & VPLS Seamless Integration – Migration
- Conclusion

From Mac Bridging to Mac Routing

Evolution:

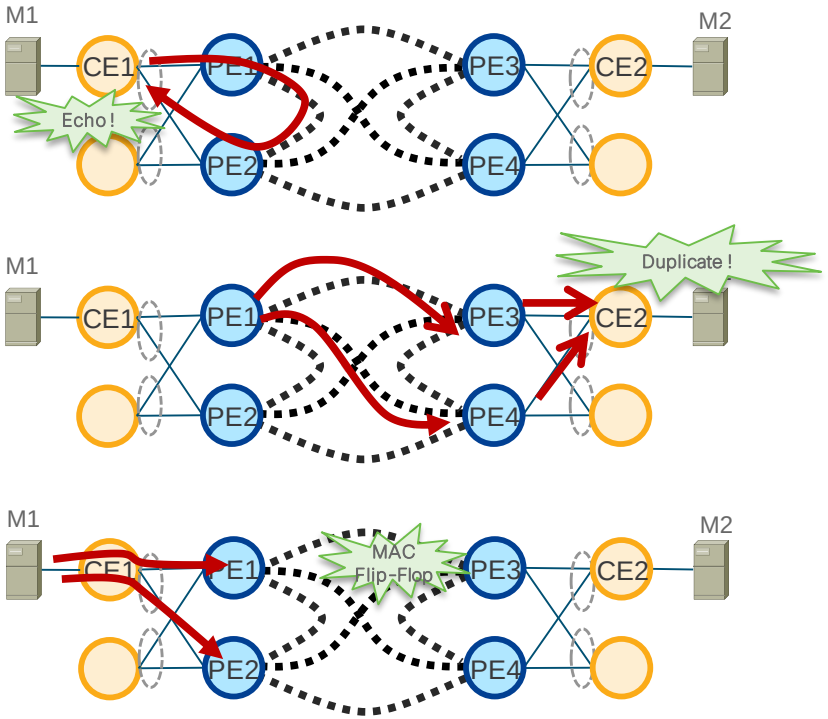


Service Provider Network – Simplification Journey



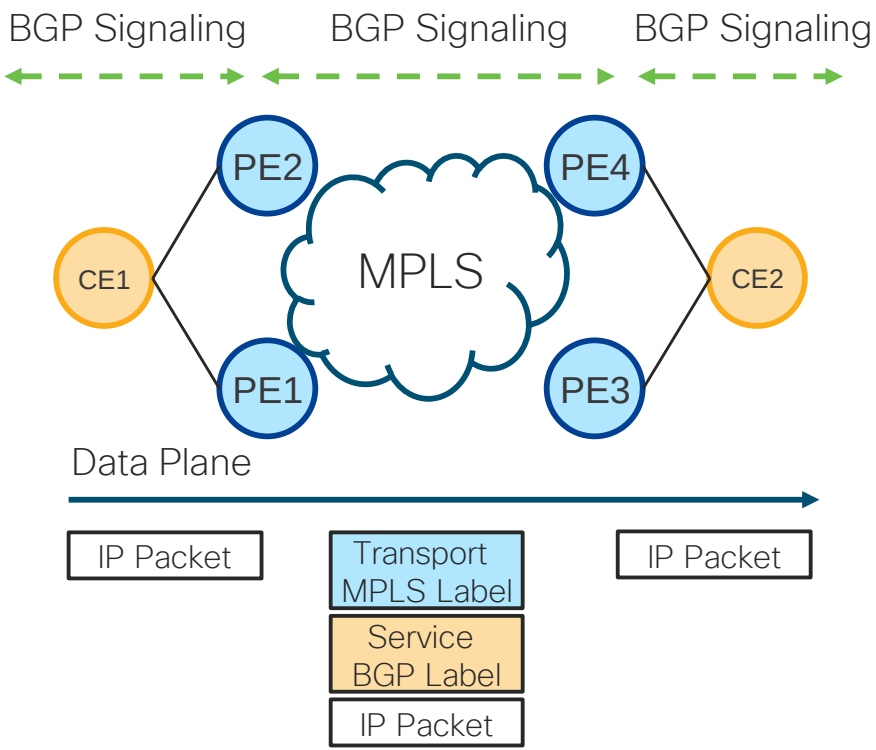
Next-Generation Solutions for L2VPN
Solving VPLS challenges for per-flow Redundancy

- 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)

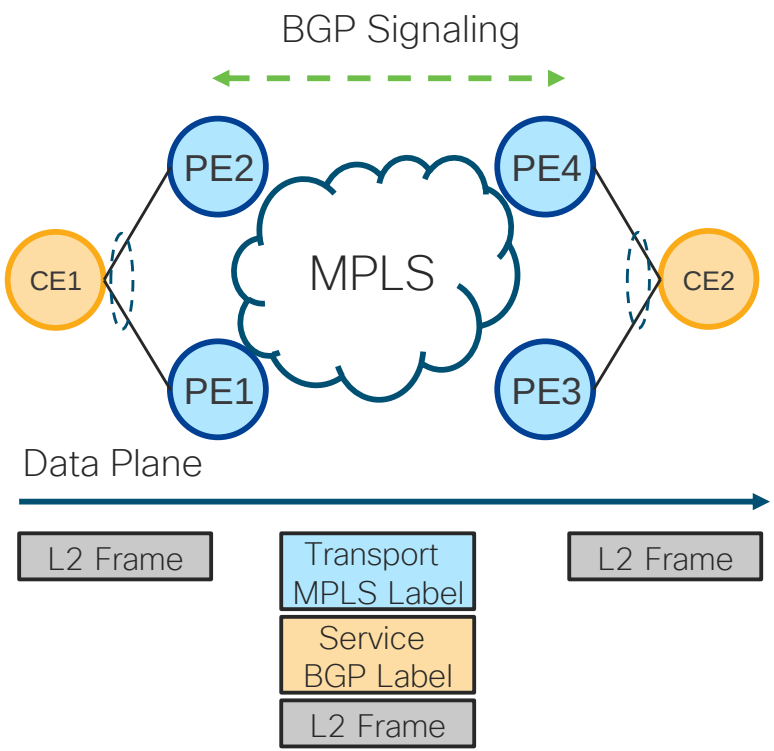


MPLS Transport & BGP Service

BGP L3VPN



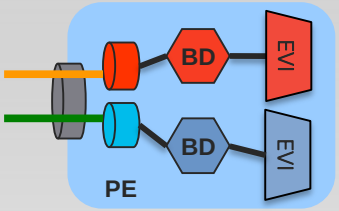
BGP EVPN



EVPN – Basic Principles

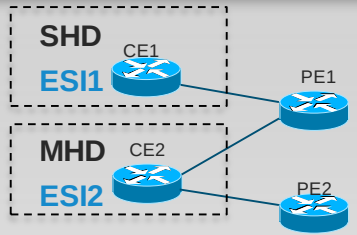
Concepts

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

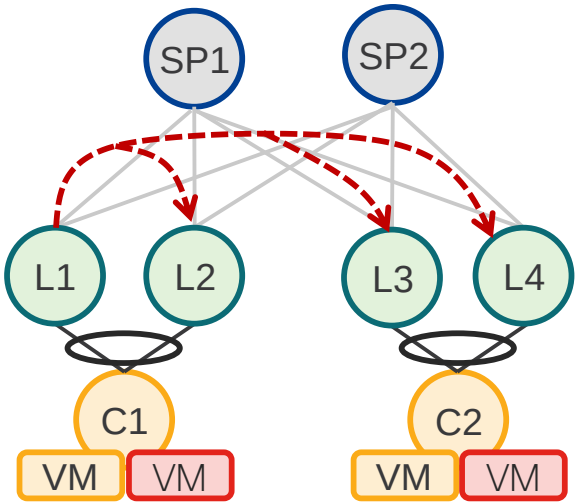
BGP Route Attributes

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

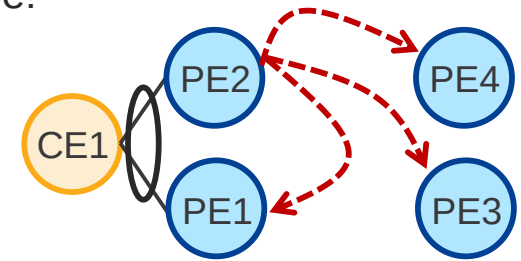
- **New BGP extended communities defined**
- **Expand information carried in BGP routes, including:**
  - MAC address moves
  - Redundancy mode
  - MAC / IP bindings of a GW
  - Split-horizon label encoding
  - Data plane Encapsulation

EVPN – Ethernet VPN

- Concepts are same!!! Pick your side!

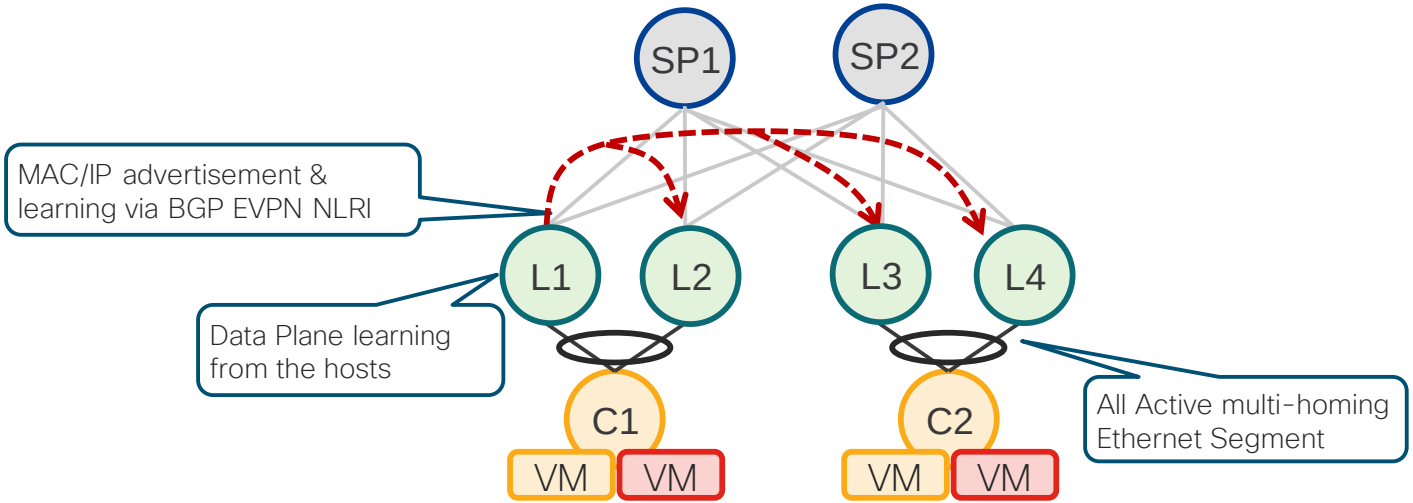


Pick your side!

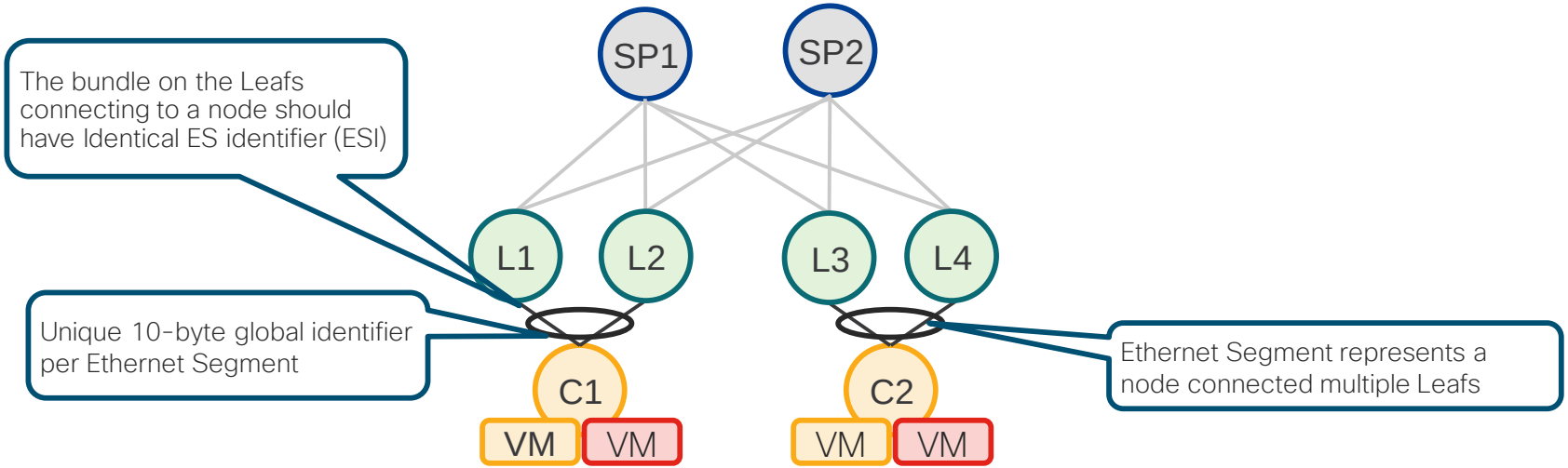


EVPN – Ethernet VPN

- Leafs run Multi-Protocol BGP to advertise & learn MAC/IP addresses over the Network Fabric
- MAC/IP addresses are advertised to rest of Leafs

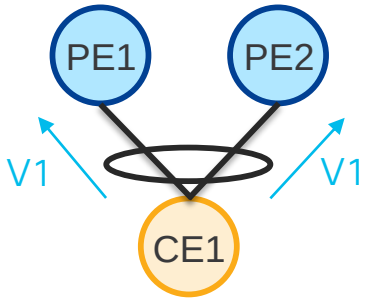


EVPN - Ethernet-Segment for Multi-Homing



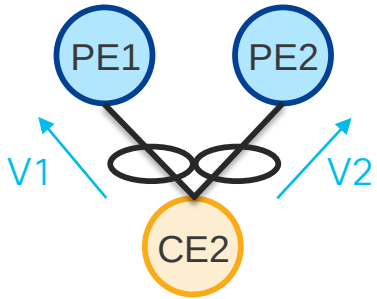
EVPN - load-balancing modes

All-Active (per flow)



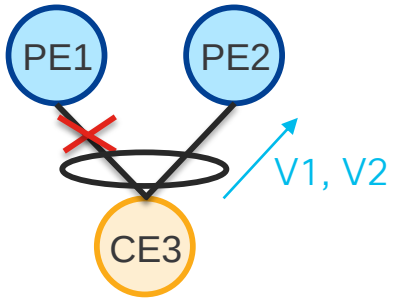
Single LAG at the CE
VLAN goes to both PE
Traffic hashed per flow
Benefits: Bandwidth, Convergence

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)



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

EVPN – Route Distinguisher (RD) and Route Target (RT) Allocation

Per-Node RD – [BGP-routerid]:0,1,2,….

EVPN RT1, RT4

Why more Per-Node RD?

Maximum Route-Targets (RTs) per route is 400

Per-Node/Per-EVI RD – [BGP-RouterID]:[EVI-ID]

EVPN RT1, RT2, RT3

Per-Node/Per-EVI RT – [BGP-AS]:[EVI-ID]

R36 example BGP RouterID 3.3.3.36, BGP-AS: 1, EVI 100:

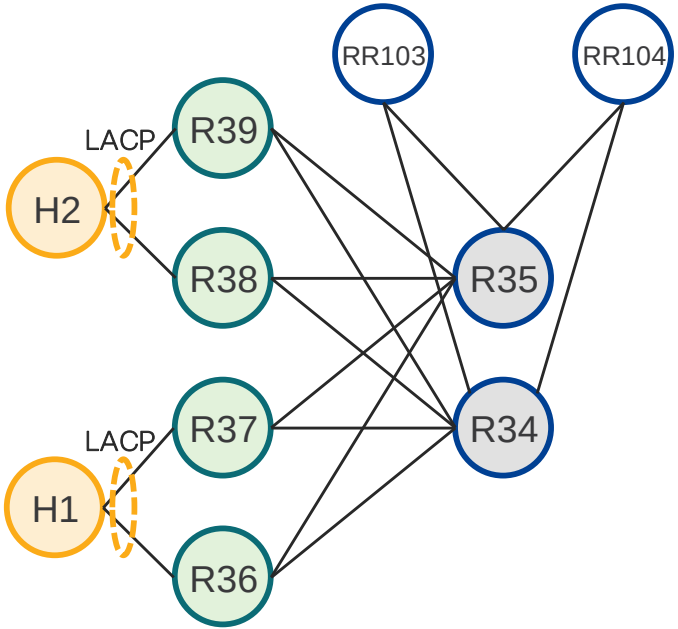
Per-Node RD: 3.3.3.36:0,1,2

Per-Node/Per-EVI RD: 3.3.3.36:100

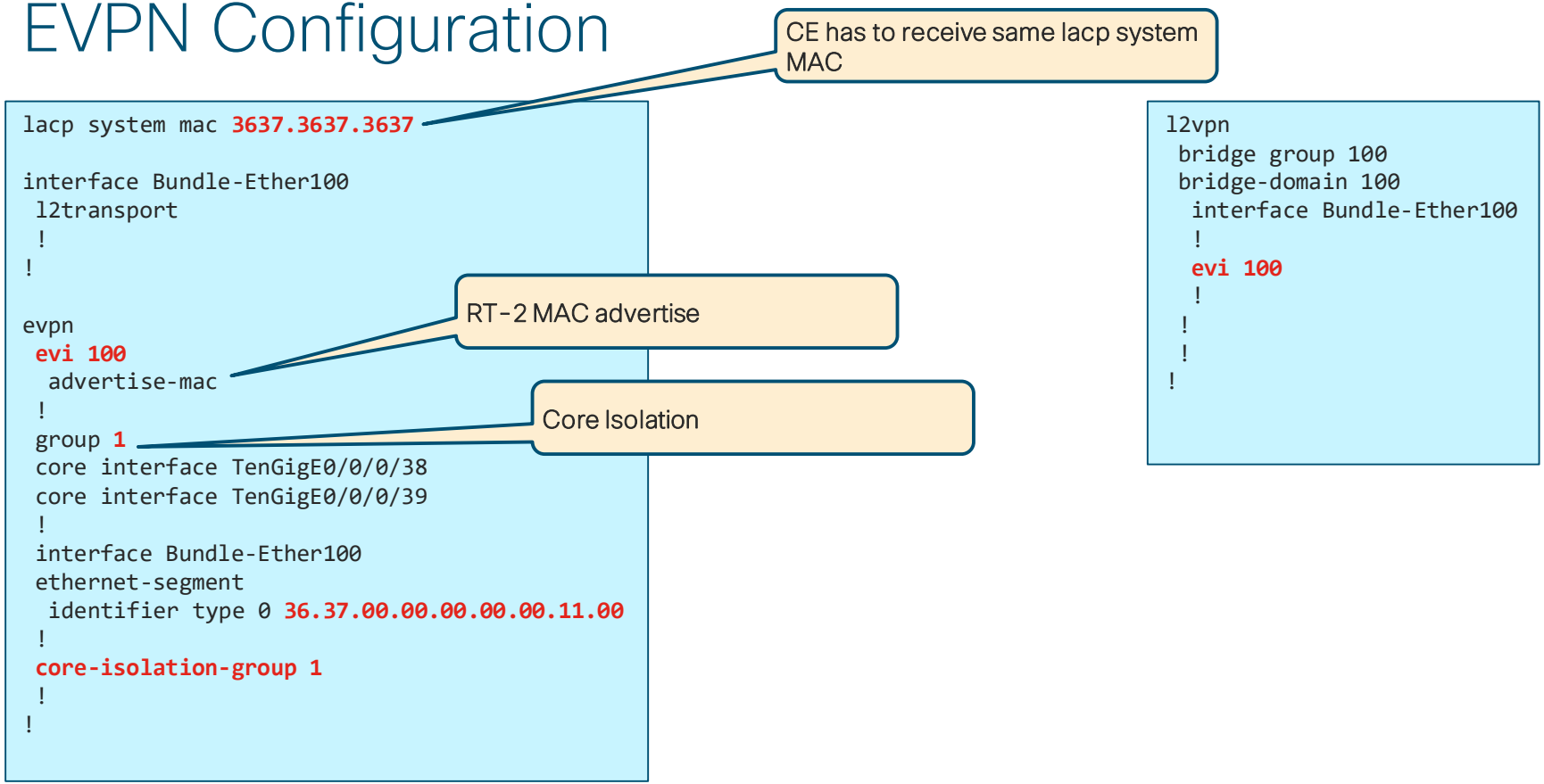
Per-Node/Per-EVI RT: 1:100

EVPN L2 All-Active Multihomed Service

EVPN – Testbed



EVPN Configuration




```
lacp system mac 3637.3637.3637

interface Bundle-Ether100
 l2transport
 !
!

evpn
 evi 100
  advertise-mac
 !
 group 1
 core interface TenGigE0/0/0/38
 core interface TenGigE0/0/0/39
 !
 interface Bundle-Ether100
 ethernet-segment
  identifier type 0 36.37.00.00.00.00.00.11.00
 !
 core-isolation-group 1
 !
!
```

```
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
 !
!
```

Ethernet Segment

```
R36#show evpn ethernet-segment
Mon Oct 15 13:27:44.402 UTC

Ethernet Segment Id      Interface                         Nexthops
------------------------ --------------------------------- --------------------
0036.3700.0000.0000.1100 BE100                              3.3.3.36
                                                            3.3.3.37
```

Ethernet Segment

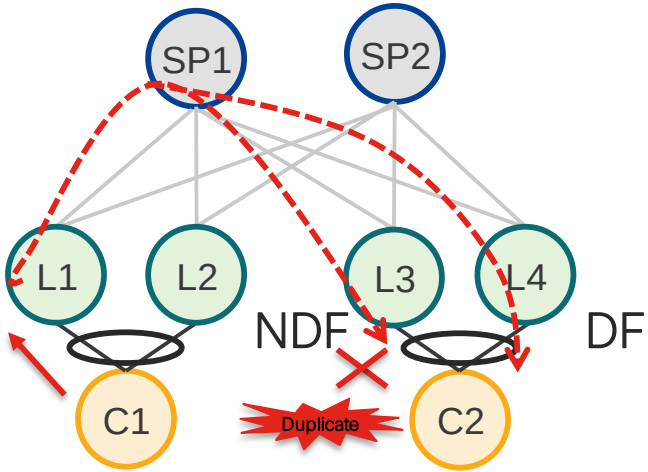
```
R36#show evpn ethernet-segment esi 0036.3700.0000.0000.1100 detail
.....
Ethernet Segment Id     Interface                         Nexthops
----------------------- --------------------------------- --------------------
0036.3700.0000.0000.1100 BE100                             3.3.3.36
         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

EVPN Instance View

```
R36#show evpn evi vpn-id 100 detail

VPN-ID     Encap  Bridge Domain                Type
---------- ------ ---------------------------- -------------------
100        MPLS   100                          EVPN
   Stitching: Regular
   Unicast Label  : 68096
   Multicast Label: 64000
   Flow Label: N
   Control-Word: Enabled
   Forward-class: 0
   Advertise MACs: Yes
   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) 3.3.3.36:100
   RT Auto  : 1:100
   Route Targets in Use           Type
   ------------------------------ --------------------
   1:100                          Import
   1:100                          Export
```

EVPN – Designated Forwarder (DF)

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



DF Election per EVI/ESI - Algorithm Service Carving

Nodes	Position
R36	0
R37	1

+

EVIs
100

EVI-ID modulo Number of Nodes = Position
100 modulo 2 = 0

R36 is DF for EVI-100

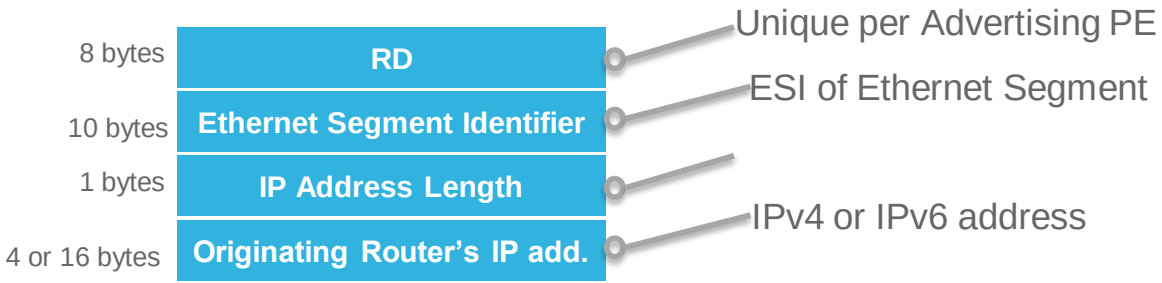
Who will be DF for EVI-101?

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
```

EVPN BGP – Ethernet Segment Router 0x4

- Usage:
  - Auto-discovery of multi-homed Ethernet Segments
  - Designated Forwarder election
- Tagged with ES-Import Extended Community
  - PEs apply route filtering based on ES-Import community. Thus, Ethernet Segment route is imported only by the PEs that are multi-homed to the same Ethernet segment
  - ES-Import extended community is not the same as the Route Target (RT) extended community

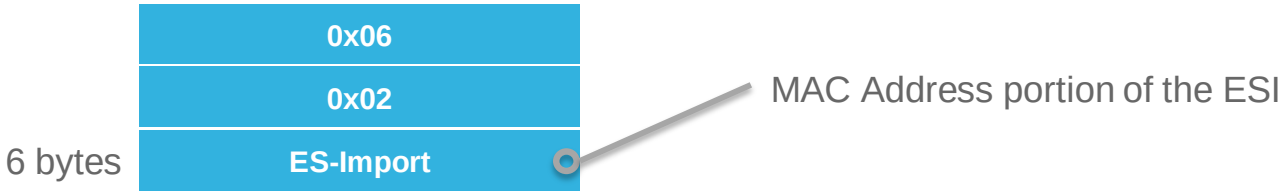


Route Type specific encoding of E-VPN NLRI

ES-Import Extended Community

Usage:

- Used to tag the Ethernet Segment route
- Limits the scope of Ethernet Segment routes distribution to PEs connected to the same multi-homed Segment



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
```

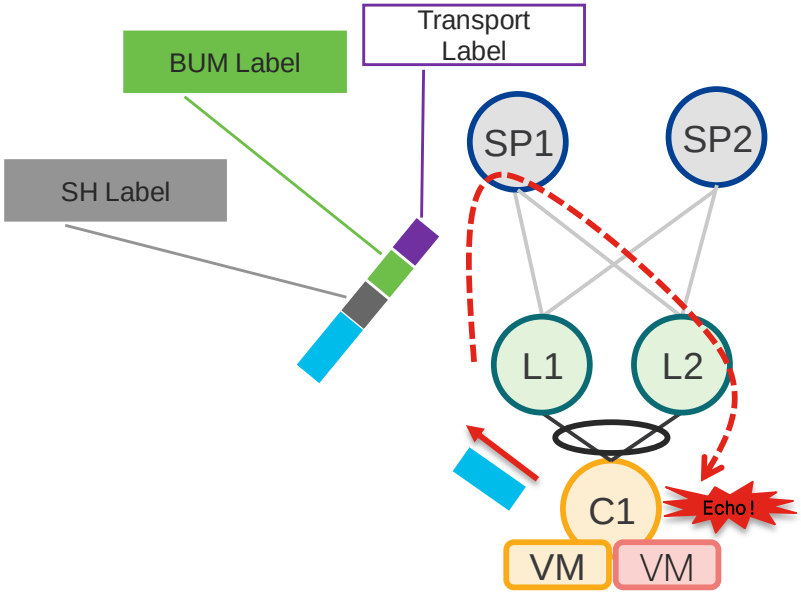
RT-4

Ethernet Segment Identifier (ESI)

Nodes which share same ESI import this route

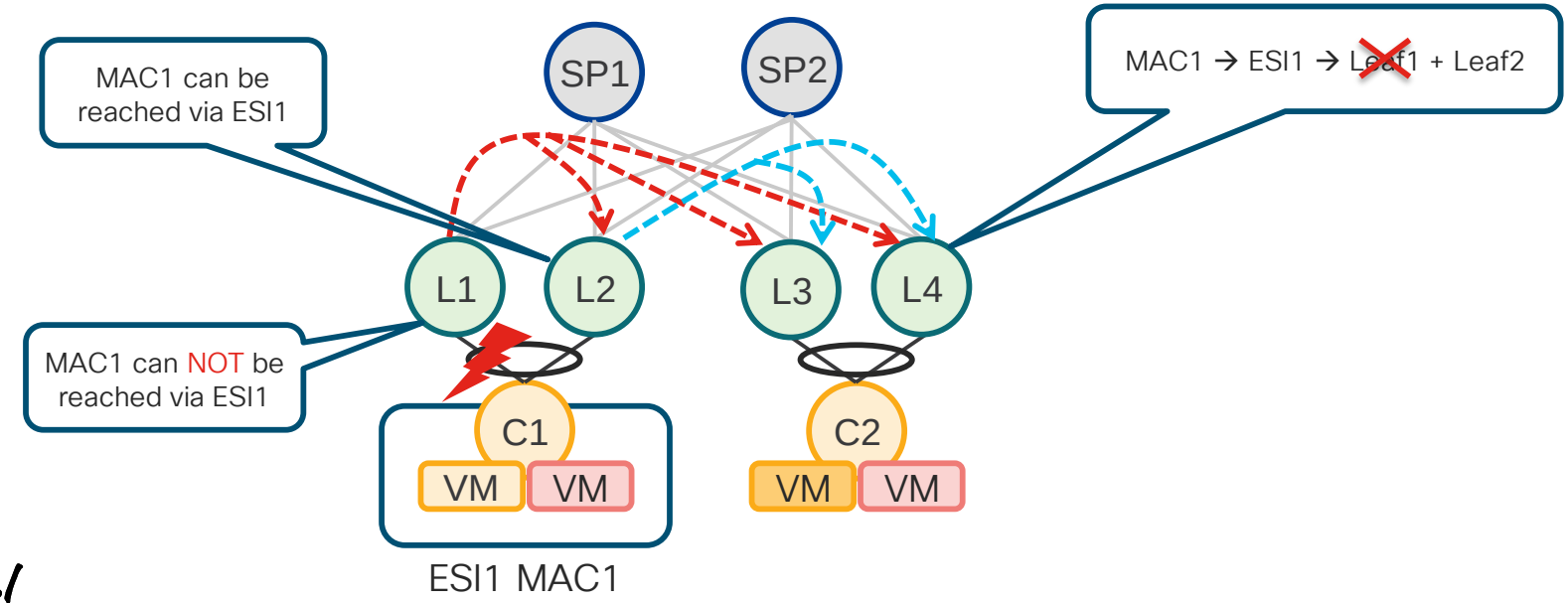
EVPN – Split Horizon

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?



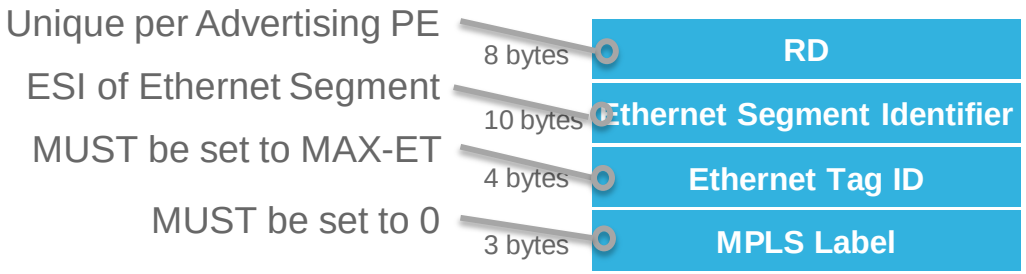
EVPN BGP - Ethernet Auto-discovery Route 0x1

Two flavors:

Per-ESI Ethernet A-D route

- Advertise Split-Horizon Label associated with an Ethernet Segment
- Used for MAC Mass-Withdraw
- Tagged with ESI MPLS Label Extended Community

Per-EVI Ethernet A-D route



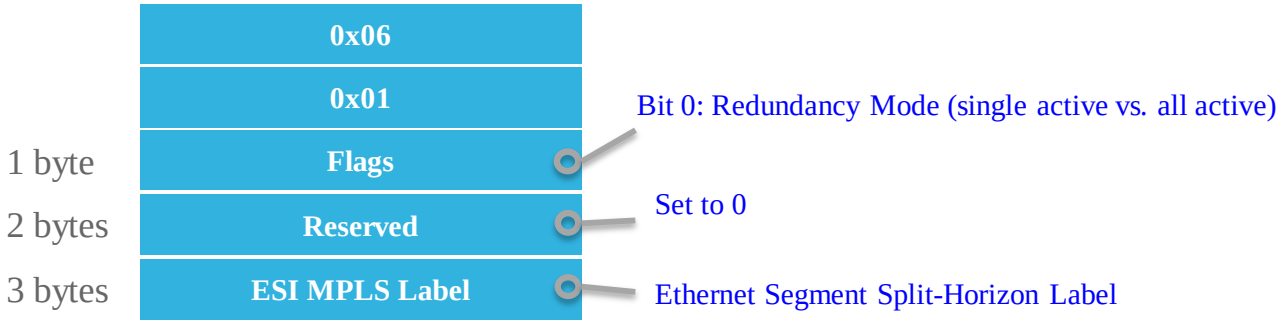
Route Type specific encoding of E-VPN NLRI

MAX-ET=0xFFFFFFFF

ESI Label Extended Community

Usage:

- Used to tag the Ethernet AD Route per ESI
- **Advertises the Split-Horizon Label for the Ethernet Segment**
- **Indicates the Redundancy Mode: Single Active vs. All-Active**



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
   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
```

RT-1

RD – unique per advertising node (R36 unique)

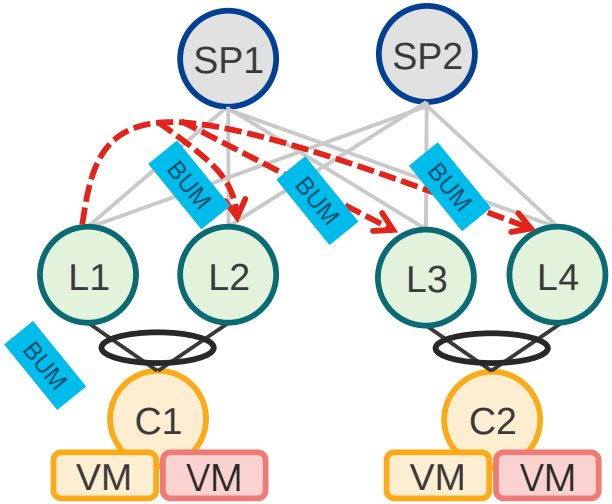
Ethernet Segment Identifier (ESI)

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

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

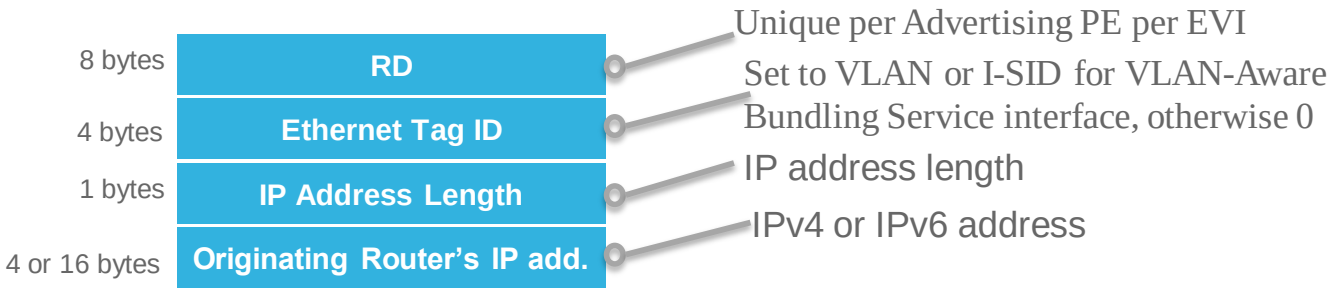
Split-Horizon Label

EVPN – BUM Ingress Replication



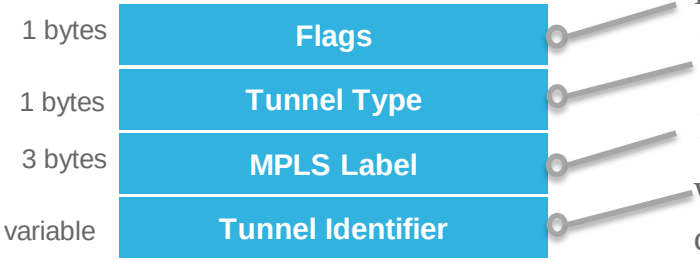
EVPN BGP – Inclusive Multicast Route 0x3

- Usage:
  - Multicast tunnels used to transport Broadcast, Multicast and Unknown Unicast frames (BUM)



Route Type specific encoding of E-VPN NLRI

PMSI Tunnel Attribute – RFC6514



Route Type specific encoding of E-VPN NLRI

Flags based on RFC6514

Ingress Replication/mLDP etc.

Multicast MPLS Label

When the Tunnel Type is set to Ingress Replication, the Tunnel Identifier carries the unicast tunnel endpoint IP address of the local PE that is to be this PE's receiving endpoint address for the tunnel.

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

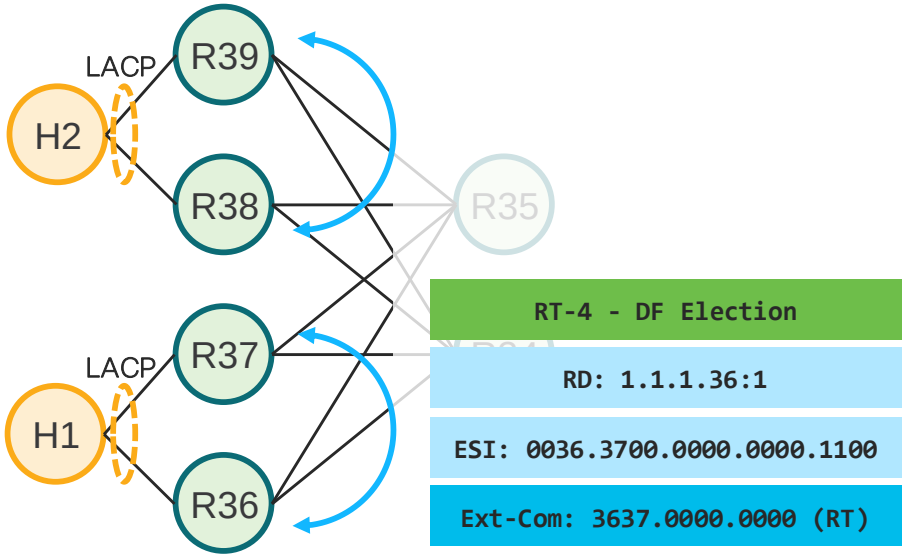
Ingress Replication

Multicast (BUM) Label

R36, R37, R38, R39 – EVPN Startup
R36 – Example

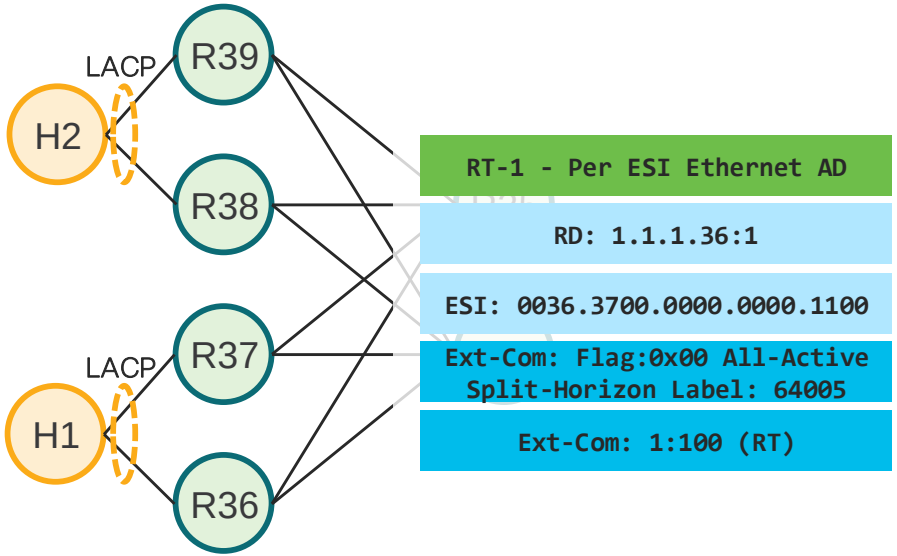
1. RT4: DF Election & Multi-Homed Ethernet Segment Auto-Discovery

Service Carving: 100 modulo 2 = 0
R36 is DF for EVI-100



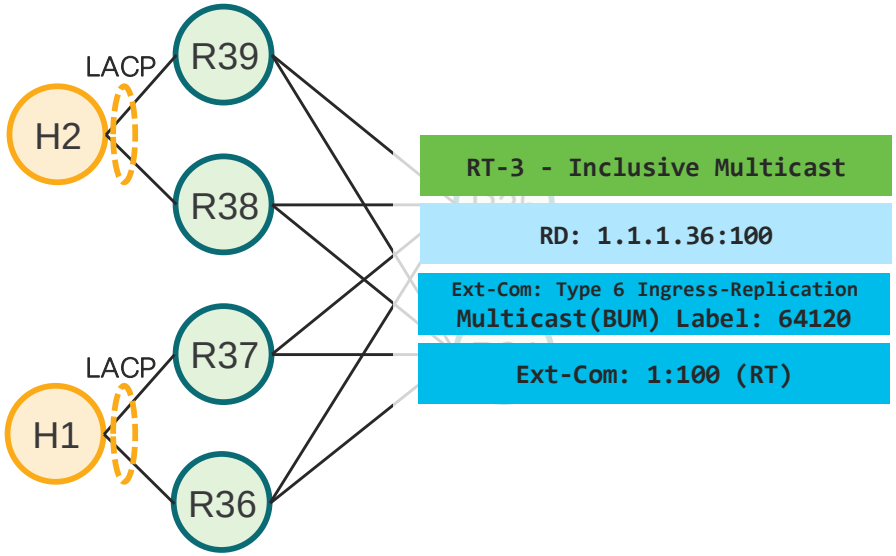
R36, R37, R38, R39 – EVPN Startup
R36 – Example

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



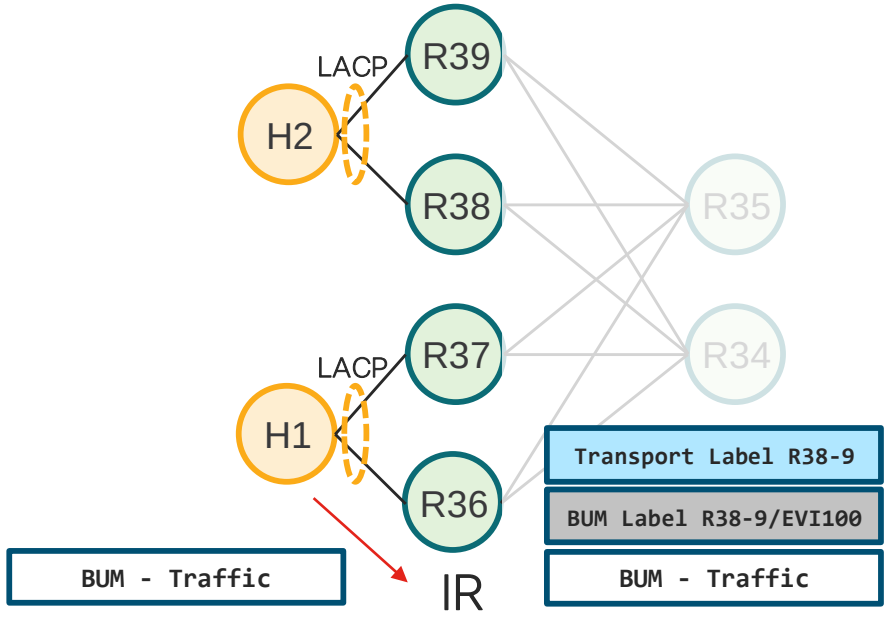
R36, R37, R38, R39 – EVPN Startup
R36 – Example

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



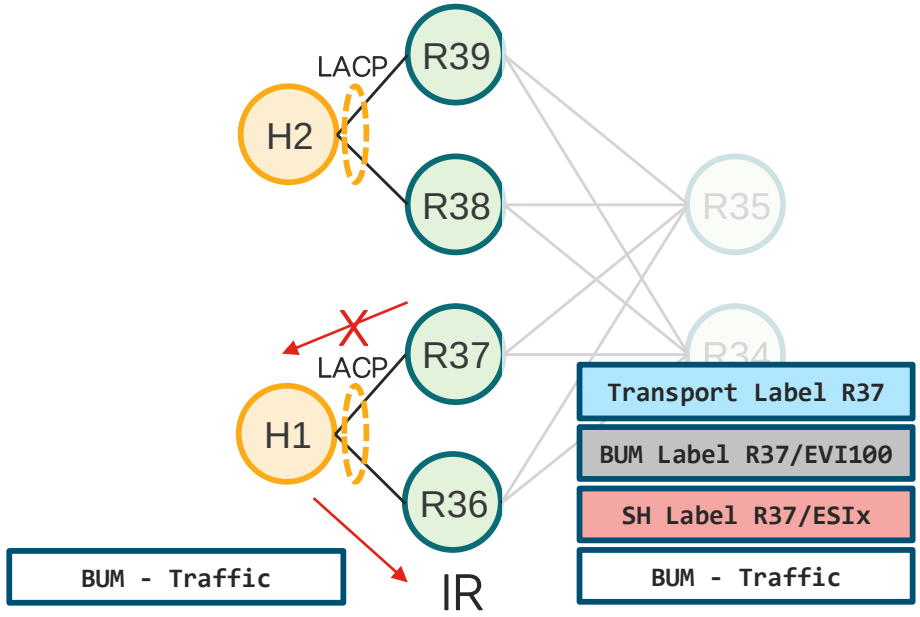
BUM Forwarding

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

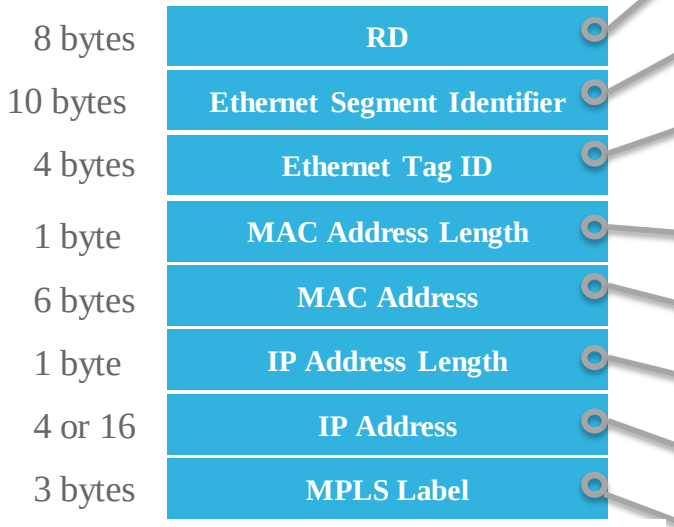


BUM Forwarding

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



EVPN BGP – MAC Advertisement Route 0x2



Route Type specific encoding of EVPN NLRI

Unique per Advertising PE per EVI

ESI of Ethernet Segment on which MAC Address was learnt. All 1s ESI for PBB-EVPN

Set to VLAN or I-SID for VLAN-Aware Bundling Service interface, otherwise 0

Allows for MAC Address 'summarization', i.e. hierarchical MAC Addresses. Typically set to 48

Could be C-MAC Address (EVPN) or B-MAC Address (PBB-EVPN)

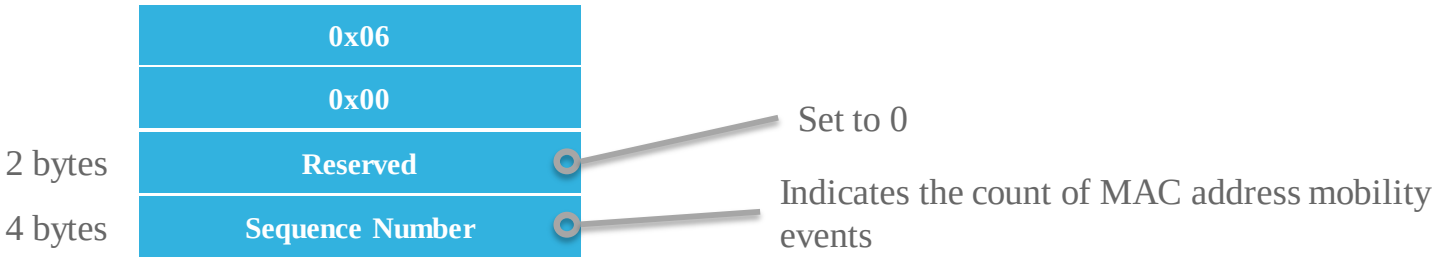
To distinguish IPv4 vs. IPv6 addresses.

Used for ARP flood suppression or for Integrated Routing and Bridging (IRB).

VPN Label downstream assigned

MAC Mobility Extended Community

- Used to tag the MAC Advertisement route
  - **EVPN**: Indicates that a MAC address has moved from one PE to another



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
Last Modified: Oct 15 04:32:31.399 for 00:01:08
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 83317
     Extended community: SoO:3.3.3.37:100 RT:1:100
     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
     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
     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
```

RT-2

Advertised MAC

R36 Re-Advertised

R37 MAC DP Learned and Advertised

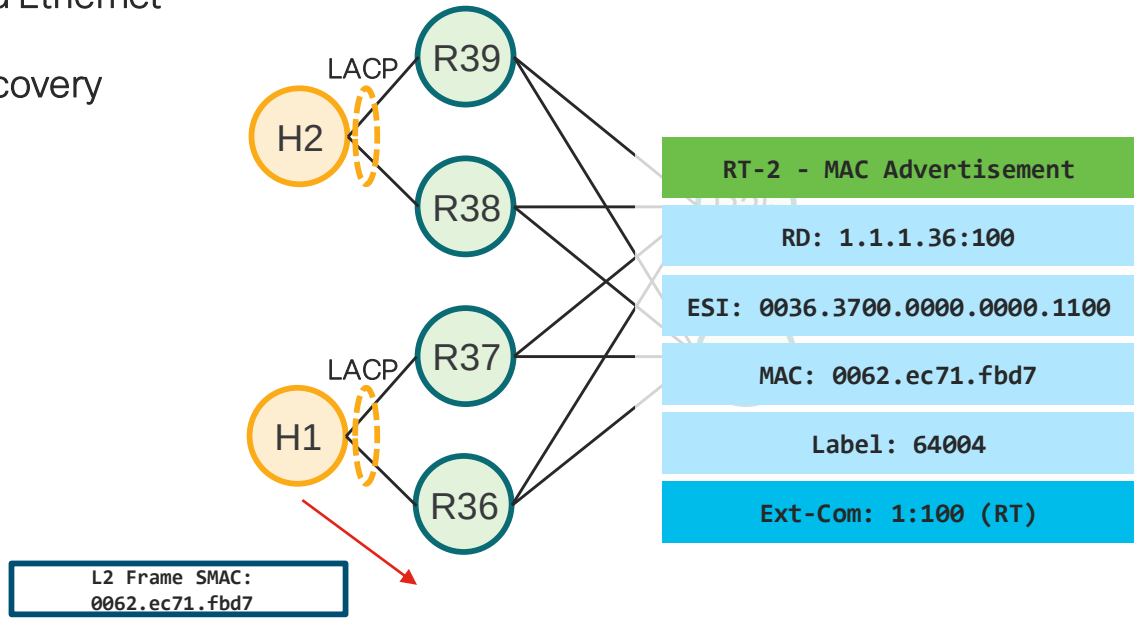
R36: RT-2 MAC Advertisement

```
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

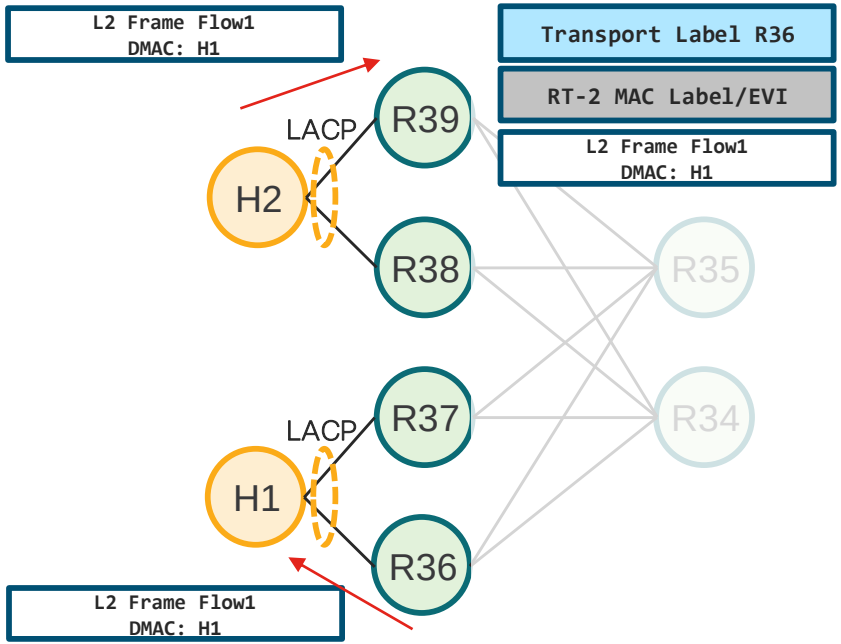
R36, R37, R38, R39 – EVPN Startup
R36 – Example

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



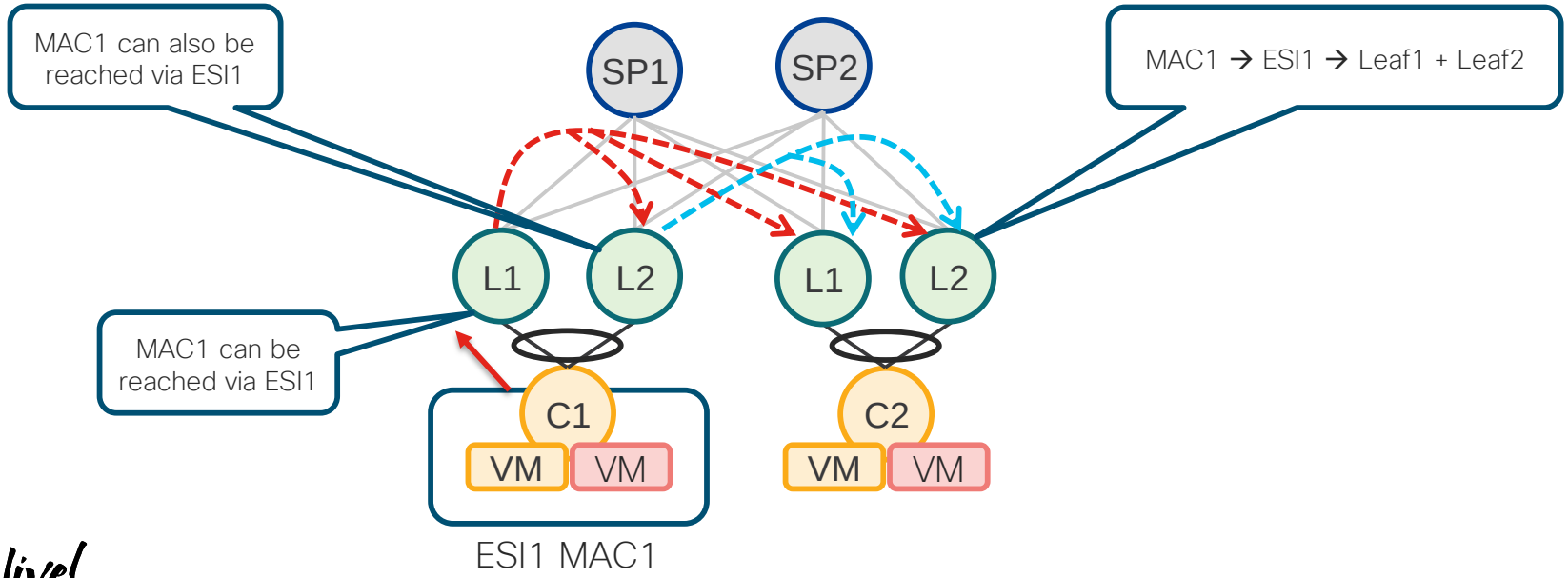
Unicast Forwarding

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?



EVPN BGP - Ethernet Auto-discovery Route 0x1

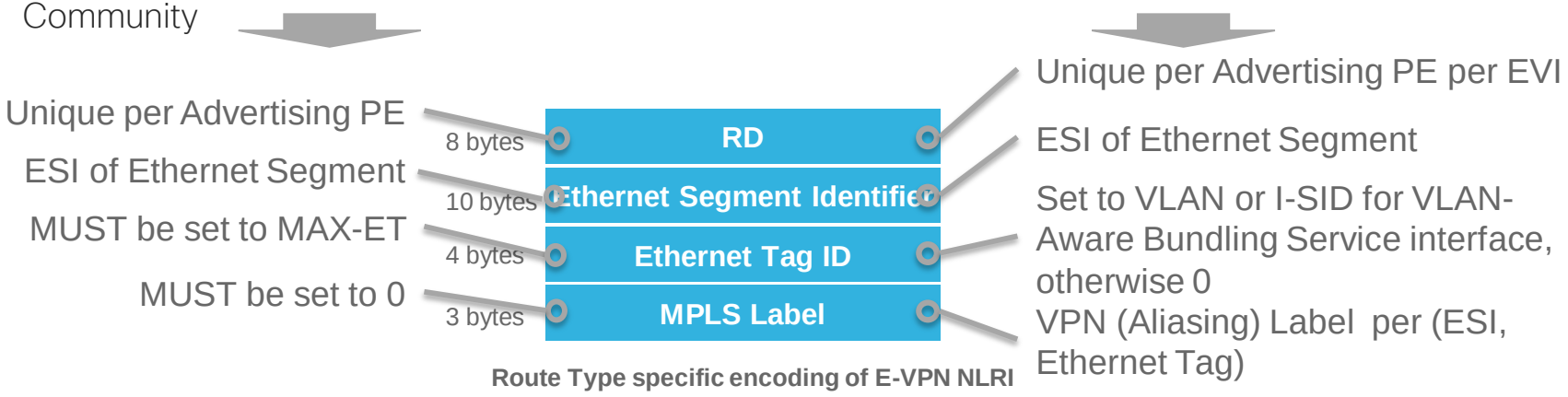
Two flavors:

Per-ESI Ethernet A-D route

- Advertise Split-Horizon Label associated with an Ethernet Segment
- Used for MAC Mass-Withdraw
- Tagged with ESI MPLS Label Extended Community

Per-EVI Ethernet A-D route

- Advertise VPN label used for Aliasing or Backup-Path



Route Type specific encoding of E-VPN NLRI

MAX-ET=0xFFFFFFFF

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
```

RT-1

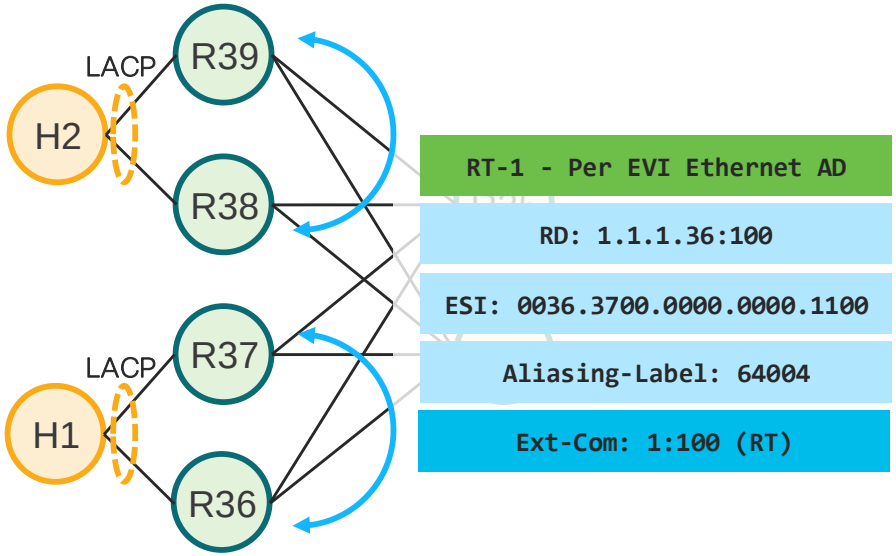
Ethernet Segment Identifier (ESI)

Aliasing Label allocated by R37 for EVI 100

EVI 100 Route-Target

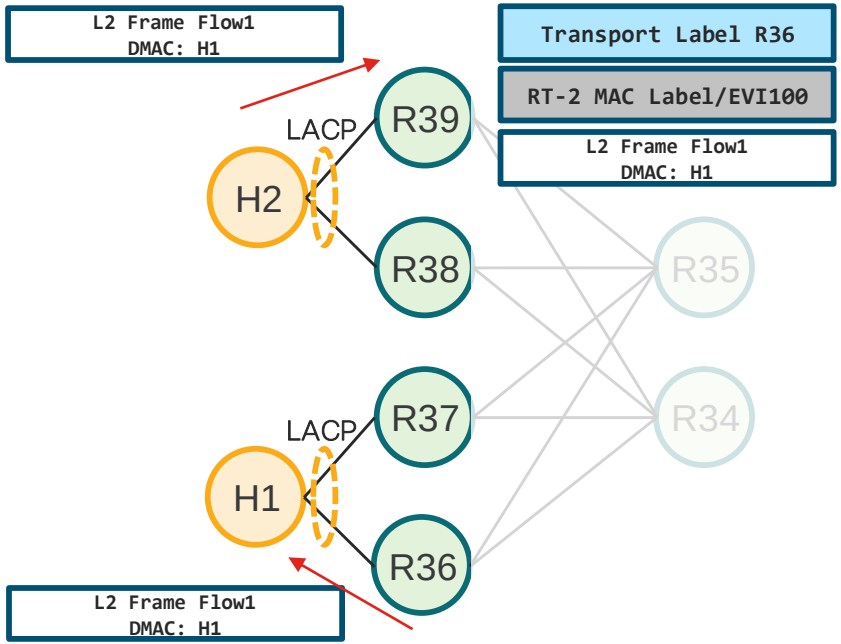
R36, R37, R38, R39 – EVPN Startup
R36 – Example

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



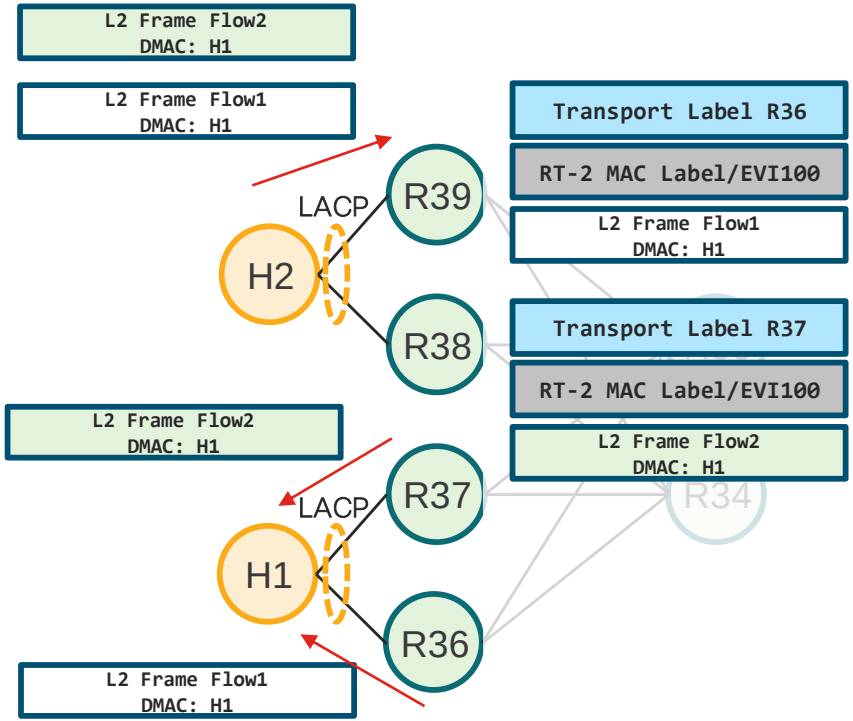
Unicast Forwarding

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

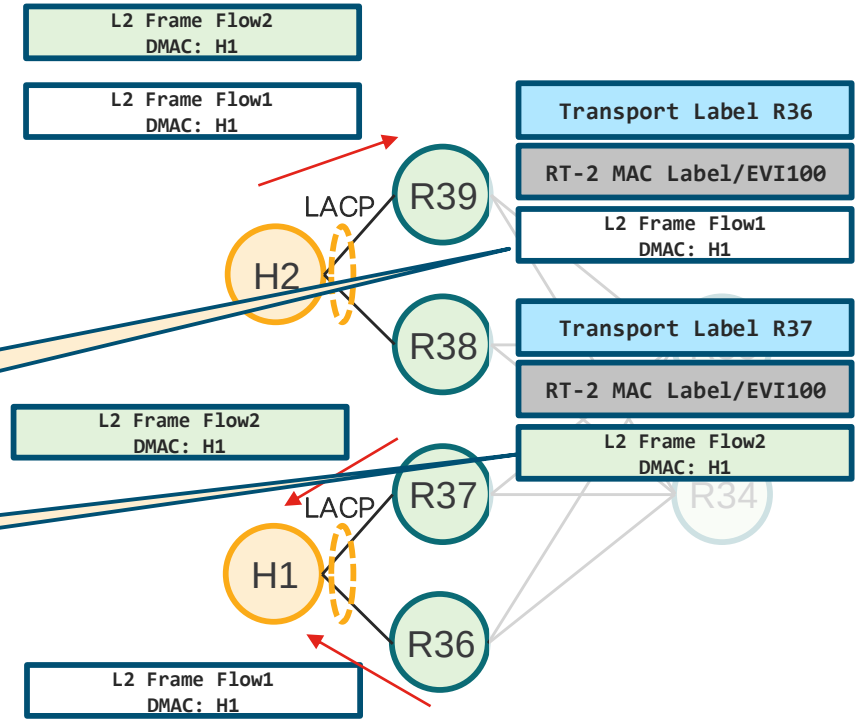


Unicast Forwarding

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

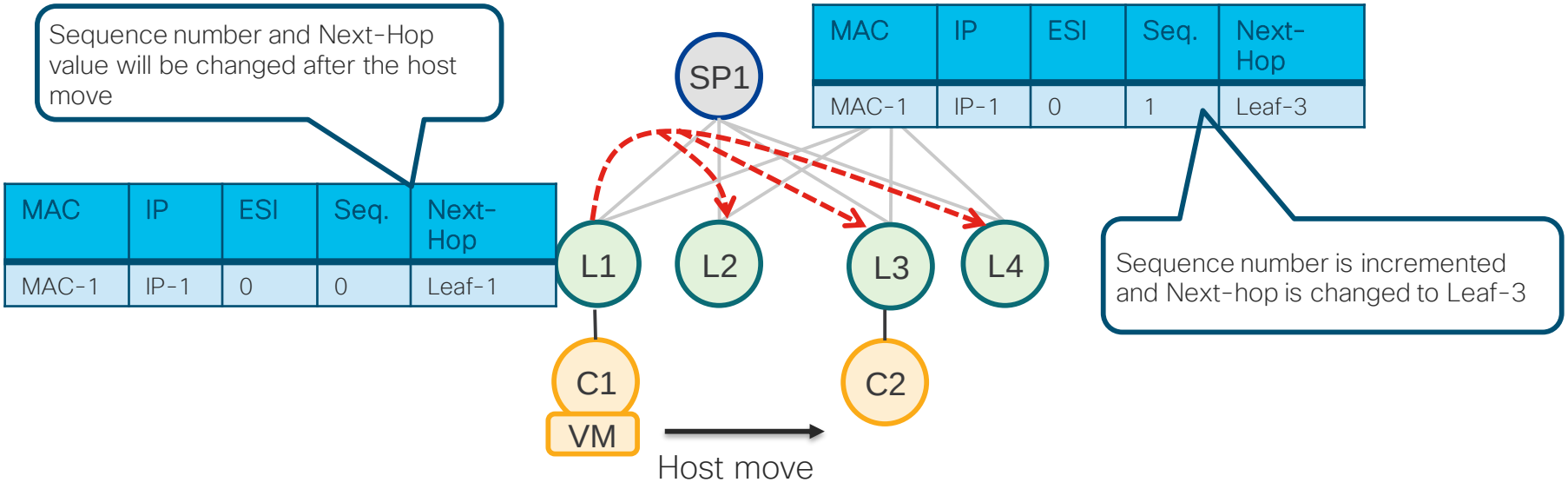
Per Flow Balancing via R36 and R37 - Aliasing

Per Flow Balancing via R36 and R37 - Aliasing

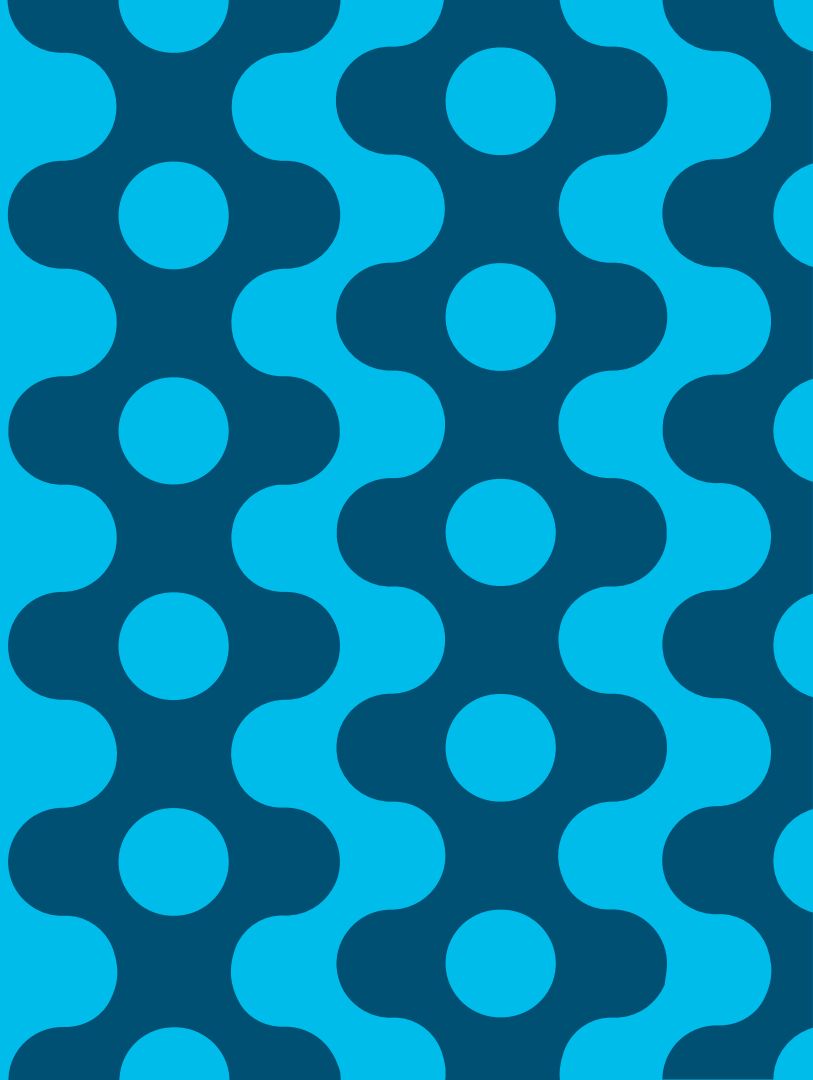


EVPN – MAC Mobility

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



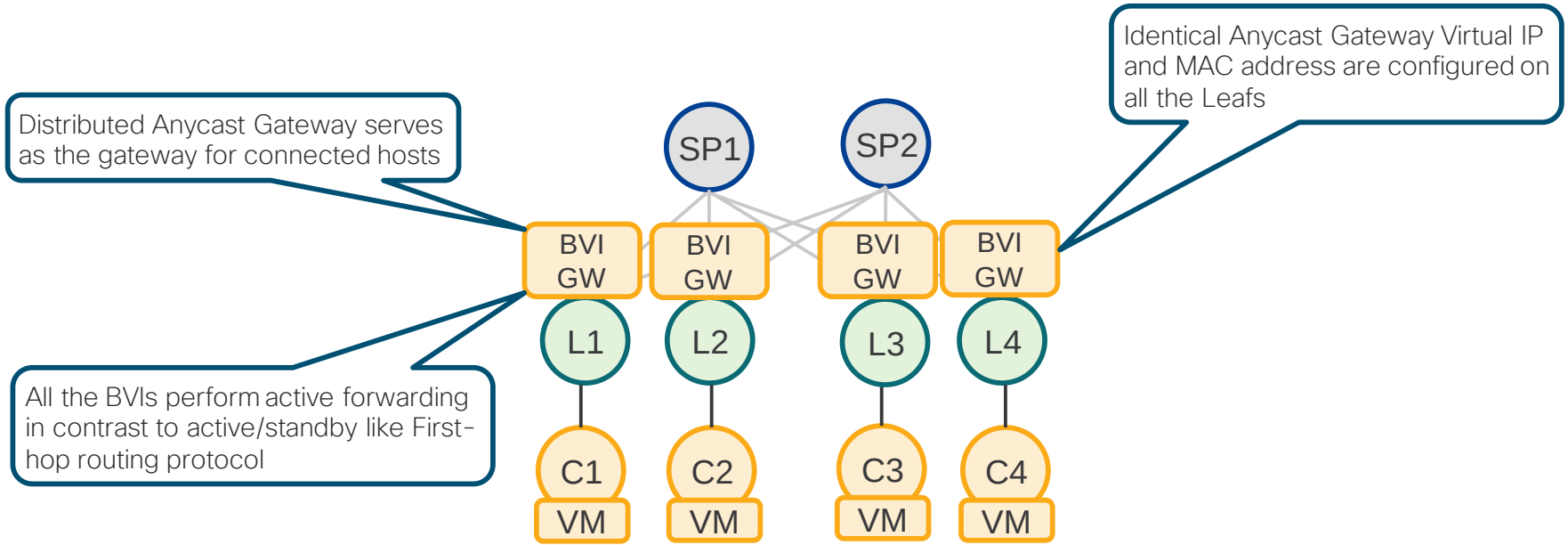
EVPN Distributed L3 Anycast Gateway



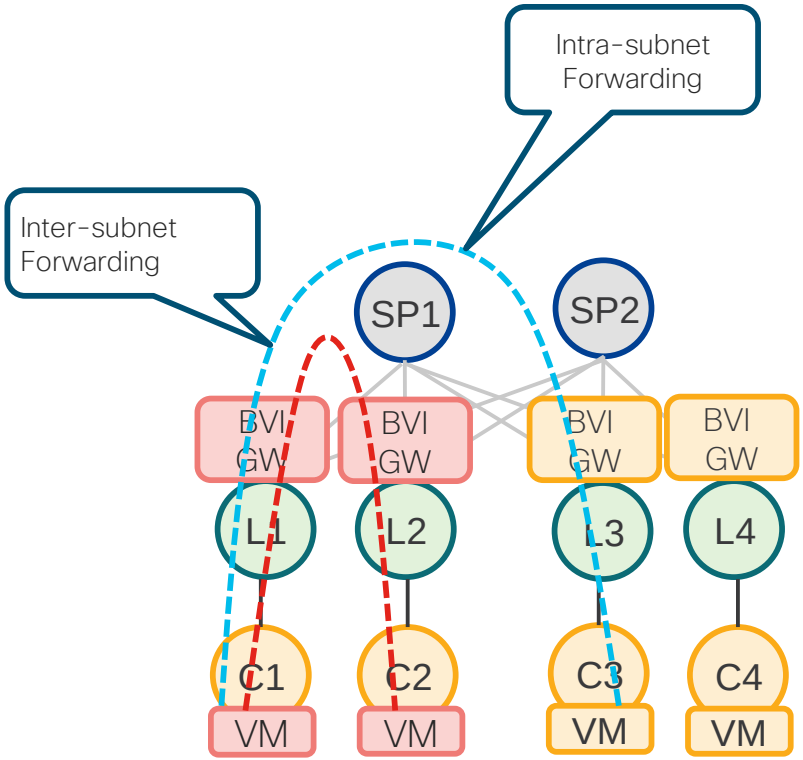
EVPN – Distributed Anycast Gateway

Purpose:

Optimal intra and inter-subnet connectivity with seamless workload mobility

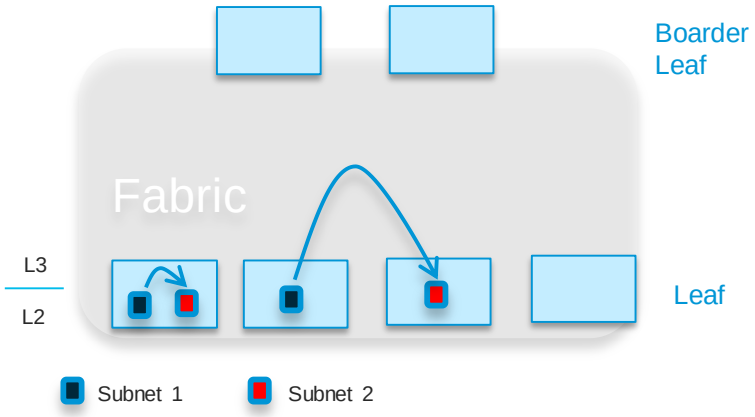


EVPN – IRB in Network Fabric



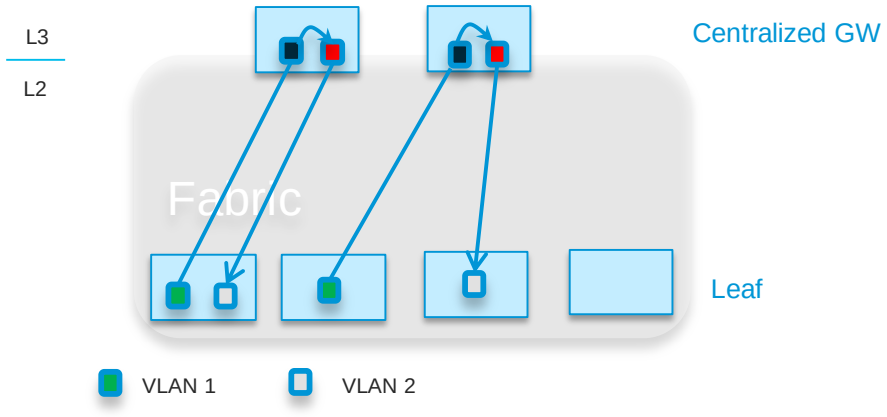
Centralized vs. Distributed Routing

Distributed Routing



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

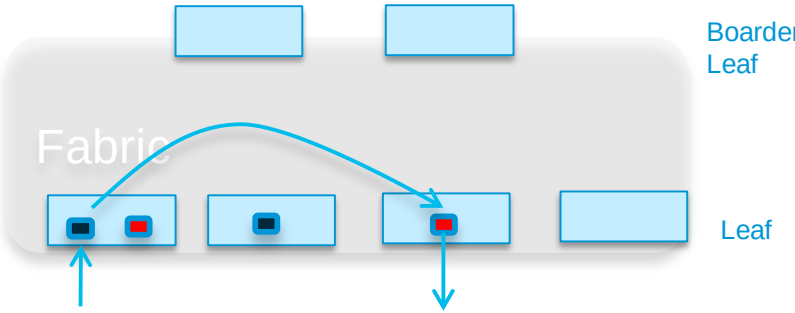
Centralized Routing



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

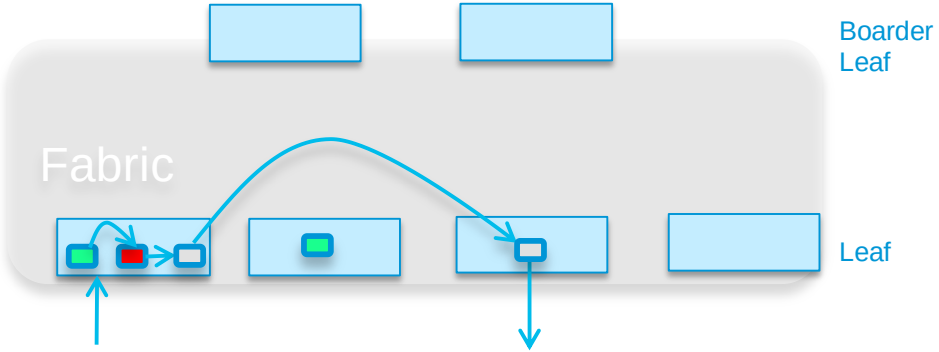
Integrated Routing and Bridging

Symmetric IRB



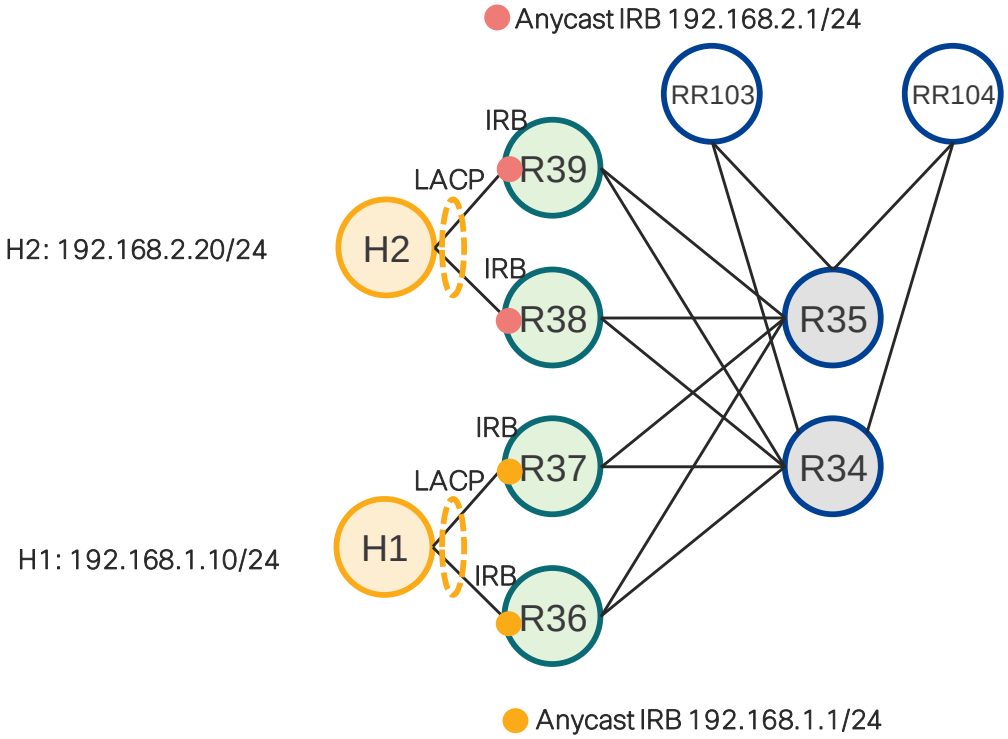
- Flexible workload placement – any subnet anywhere
- ARP/MAC state localized to Leafs
  - Helps with horizontal scaling of DC

Asymmetric IRB



- Egress subnet must be local
- Ingress Leaf needs ARP/MAC state for every egress leaf
  - Limits scale

EVPN Distributed L3 Anycast GW – Symmetric IRB



EVPN Configuration - IRB

```
cef adjacency route override rib


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
!
```

prefer adjacency /32 (ARP) route over RIB

IOS-XR 6.0+
AIB has the lowest priority by default (LSD>RIB>AIB)

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 vpnv4 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:
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#
```

RT-2

Advertised MAC

IP

RT EVI 100 and RT VRF A

RT-2 per-BD label

VRF Agg label

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: AIB preference

```
R36#show cef vrf a 192.168.1.10
Tue Oct 16 02:48:21.376 UTC
192.168.1.10/32, version 9605, internal 0x1020001 0x0 (ptr 0x97c135fc) [1], 0x0 (0x97dda968), 0x0 (0x0)
 Updated Oct 15 23:14:52.111
 local adjacency 192.168.1.10
 Prefix Len 32, traffic index 0, Adjacency-prefix, precedence n/a, priority 3
  via 192.168.1.10/32, BVI100, 3 dependencies, weight 0, class 0 [flags 0x0]
   path-idx 0 NHID 0x0 [0x98750da0 0x0]
   next hop 192.168.1.10/32
   local adjacency
```

Host Available via Local adjancency – AIB

```
cef adjacency route override rib
```

prefer adjacency /32 (ARP) route over RIB

IOS-XR 6.0+
AIB has the lowest priority by default (LSD>RIB>AIB)

R36: VRF A - CEF

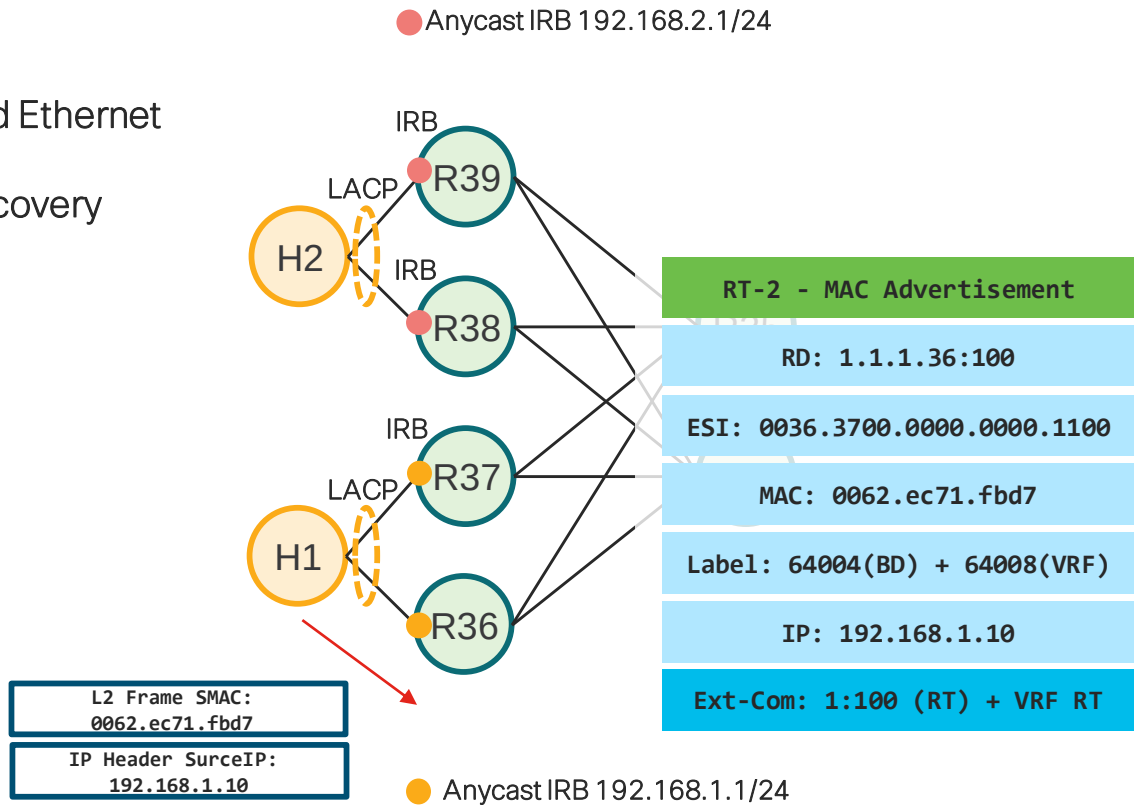
```
R36#show cef vrf a 192.168.2.20/32
Tue Oct 16 03:15:50.092 UTC
192.168.2.20/32, version 9613, internal 0x5000001 0x0 (ptr 0x97c14154) [1], 0x0 (0x0), 0x208 (0x98a06600)
 Updated Oct 15 23:18:06.305
 Prefix Len 32, traffic index 0, precedence n/a, priority 3
  via 3.3.3.38/32, 5 dependencies, recursive, bgp-multipath [flags 0x6080]
   path-idx 0 NHID 0x0 [0x97256420 0x0]
   recursion-via-/32
   next hop VRF - 'default', table - 0xe0000000
   next hop 3.3.3.38/32 via 16038/0/21
    next hop 35.36.1.35/32 Te0/0/0/39 labels imposed {16038 64004}
    next hop 34.36.1.34/32 Te0/0/0/38 labels imposed {16038 64004}
  via 3.3.3.39/32, 5 dependencies, recursive, bgp-multipath [flags 0x6080]
   path-idx 1 NHID 0x0 [0x97257178 0x0]
   recursion-via-/32
   next hop VRF - 'default', table - 0xe0000000
   next hop 3.3.3.39/32 via 16039/0/21
    next hop 35.36.1.35/32 Te0/0/0/39 labels imposed {16039 64004}
    next hop 34.36.1.34/32 Te0/0/0/38 labels imposed {16039 64004}
```

VRF Agg label

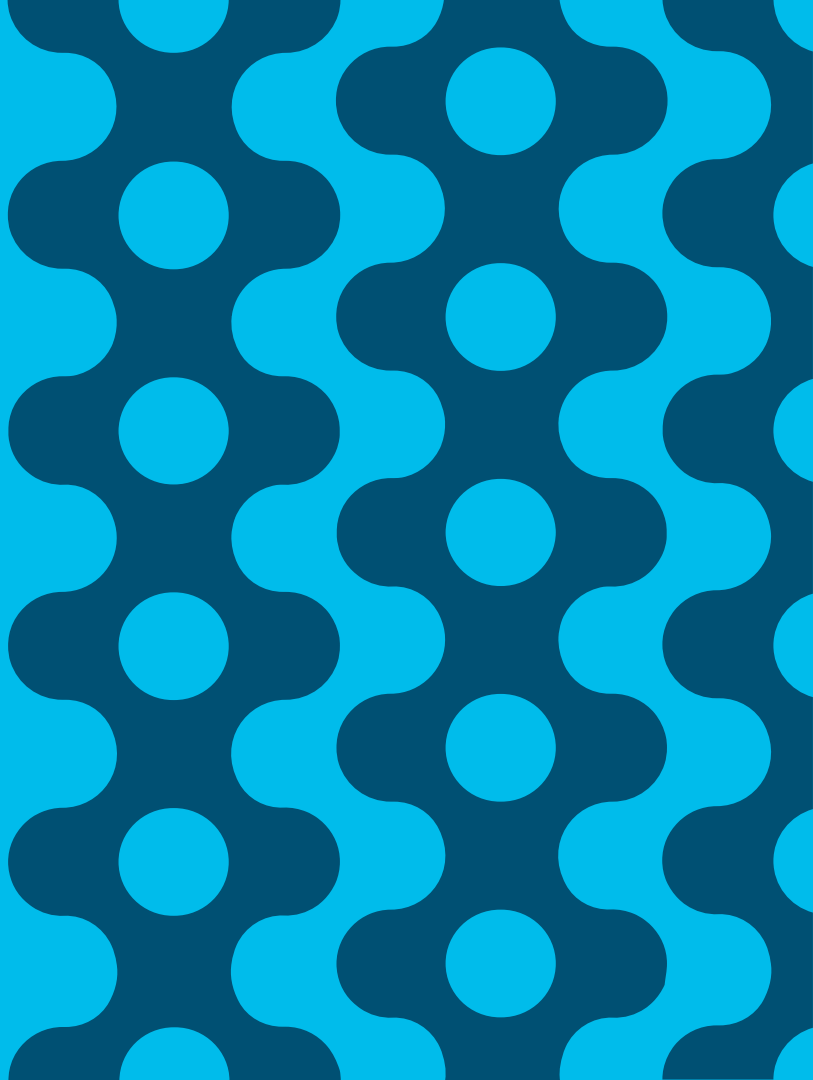
Inter-Subnet Multi-Path and ECMP

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

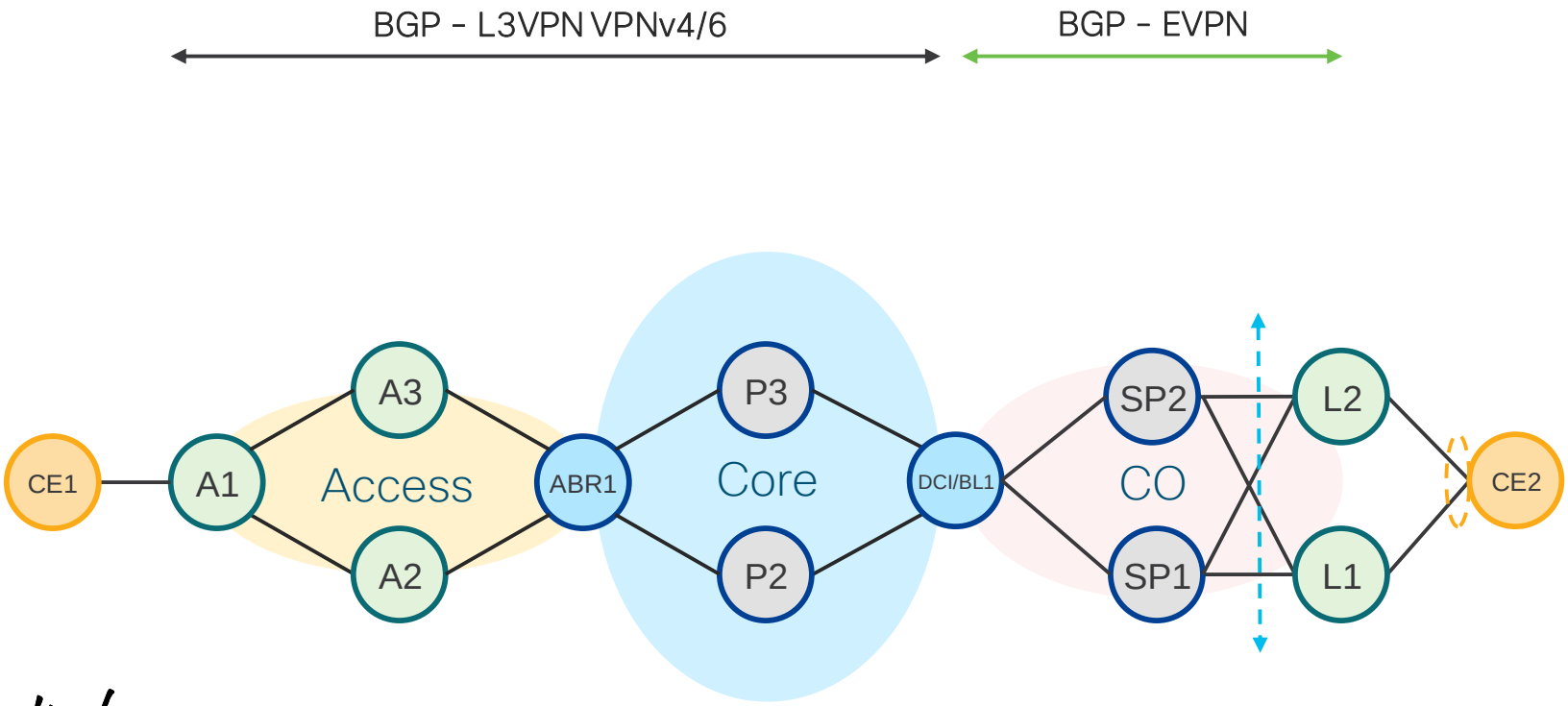


EVPN & VPNv4/6 Interconnect



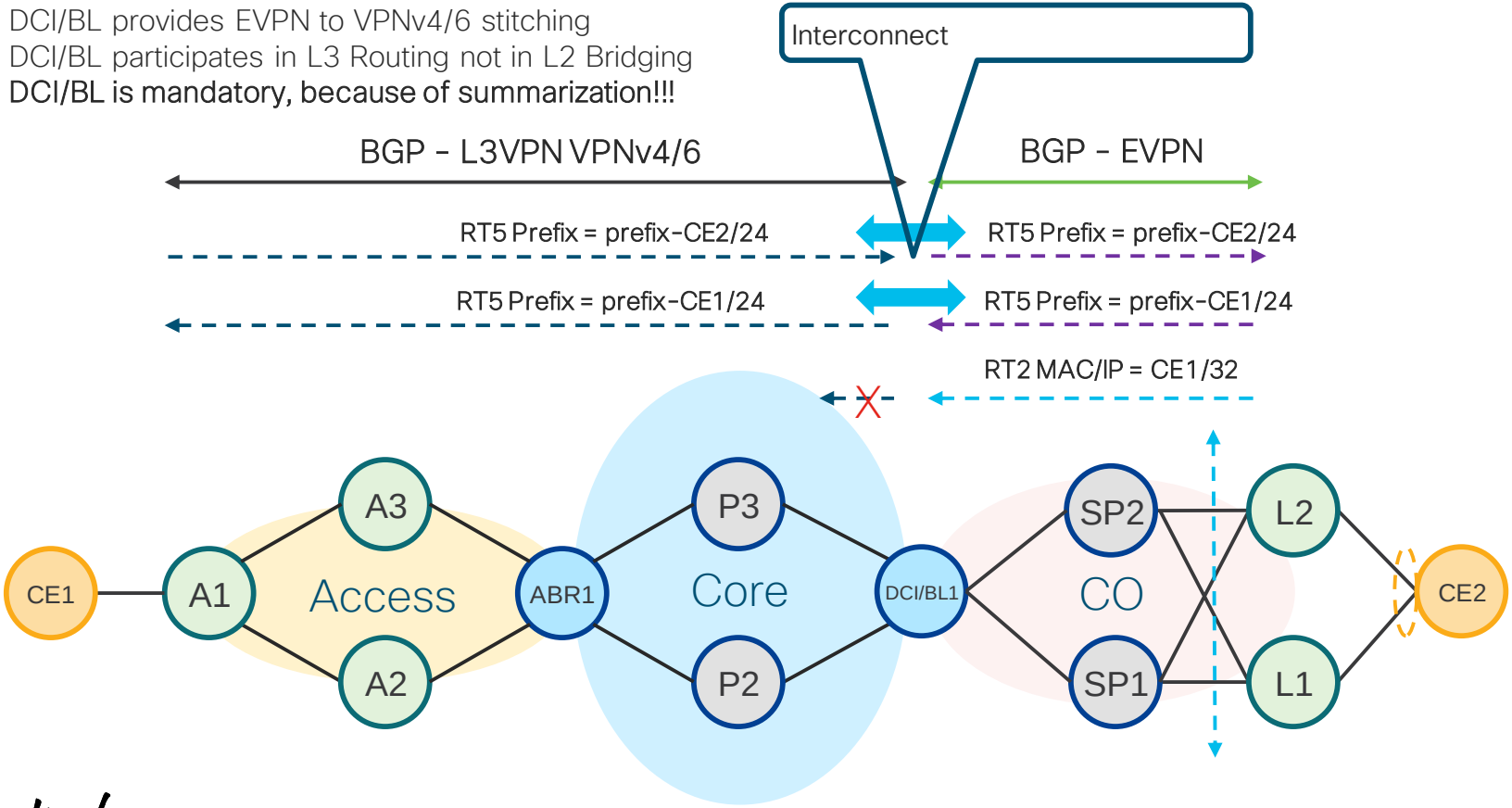
EVPN and VPNv4/6 Interconnect

- DCI/BL provides EVPN to VPNv4/6 stitching
- DCI/BL participates in L3 Routing not in L2 Bridging
- DCI/BL is mandatory, because of summarization!!!

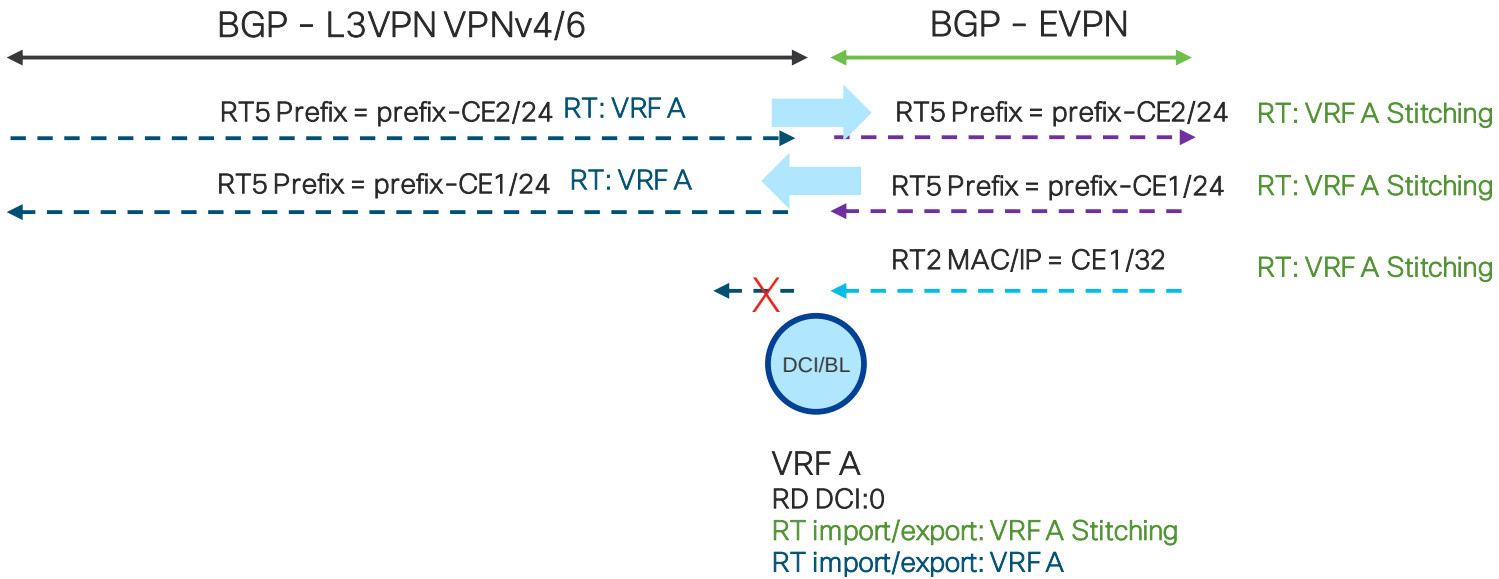


EVPN and VPNv4/6 Interconnect

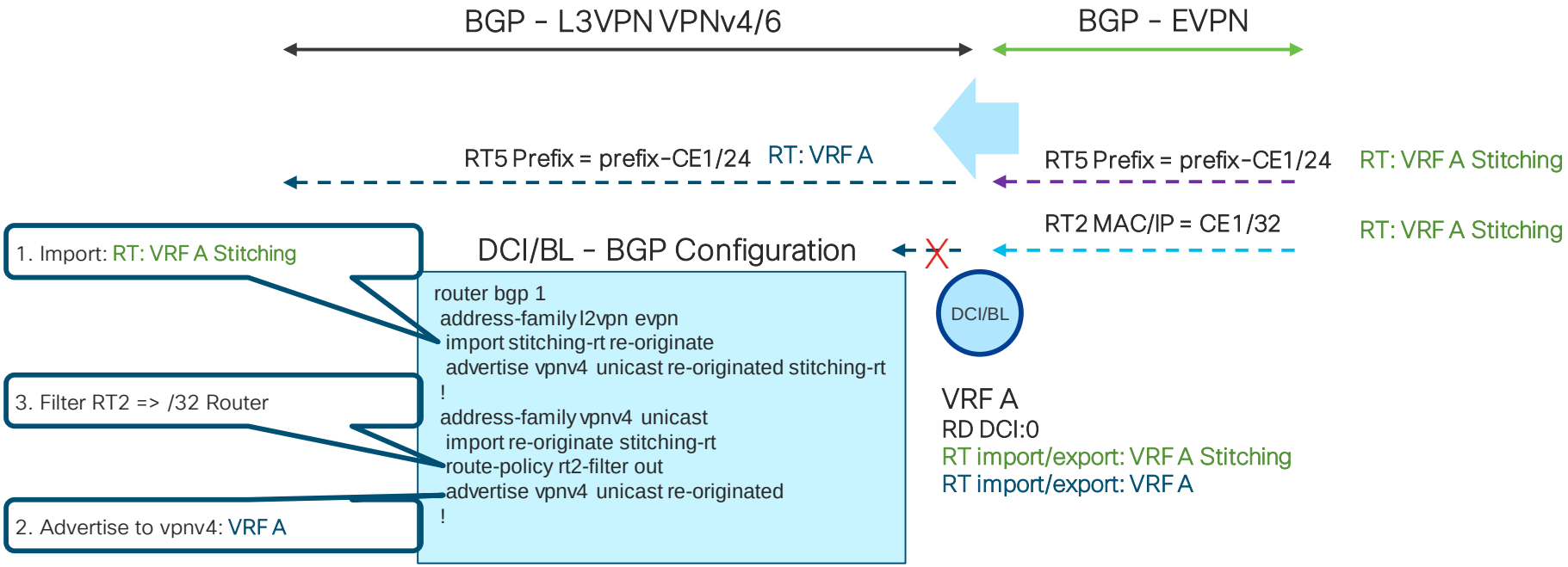
- DCI/BL provides EVPN to VPNv4/6 stitching
- DCI/BL participates in L3 Routing not in L2 Bridging
- DCI/BL is mandatory, because of summarization!!!



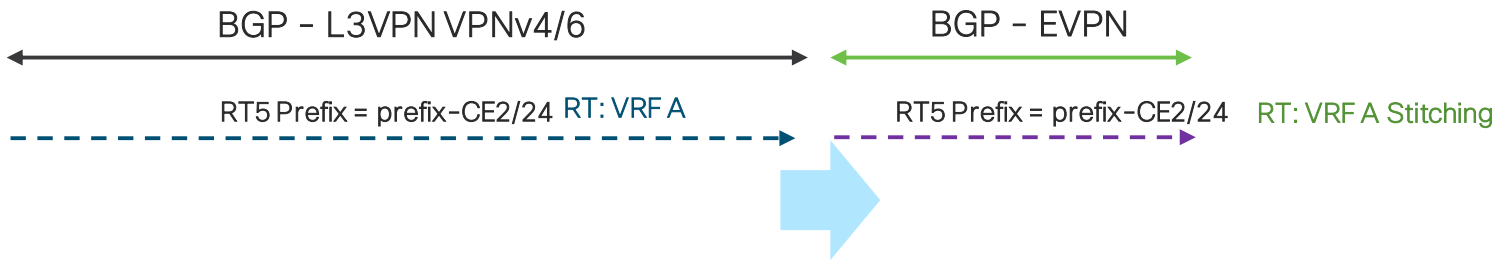
EVPN and VPNv4/6 Interconnect



EVPN and VPNv4/6 Interconnect
EVPN to VPNv4/6 Re-Advertise



EVPN and VPNv4/6 Interconnect
VPNv4/6 to EVPN Re-Advertise



2. Advertise to EVPN: RT: VRF A Stitching

1. Import: VRF A

DCI/BL – BGP Configuration

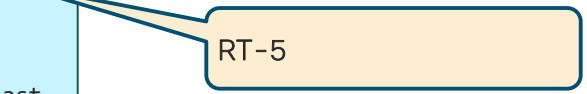
```
router bgp 1
 address-family l2vpn evpn
  import stitching-rt re-originate
  advertise vpnv4 unicast re-originated stitching-rt
 !
 address-family vpnv4 unicast
  import re-originate stitching-rt
  route-policy rt2-filter out
  advertise vpnv4 unicast re-originated
 !
```

DCI/BL

VRF A
RD DCI:0
RT import/export: VRF A Stitching
RT import/export: VRF A

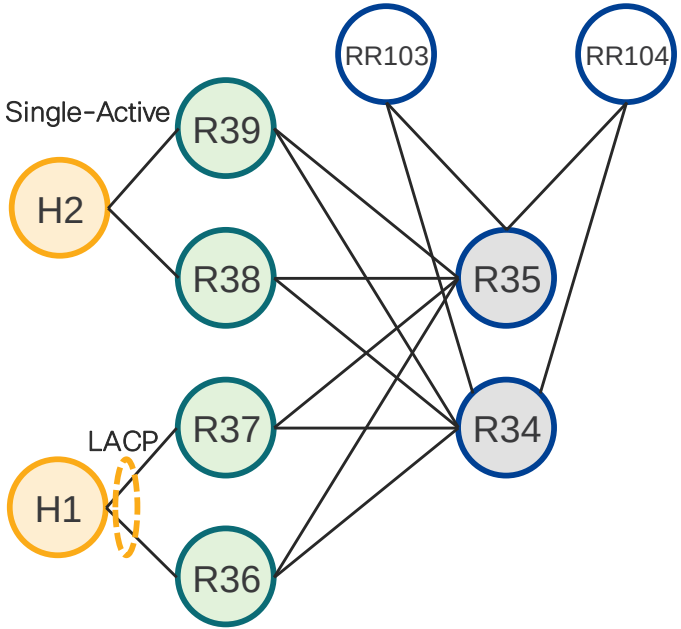
R36: BGP Configuration – RT-5

```
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
 !
```



EVPN Single-Active

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  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    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  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    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

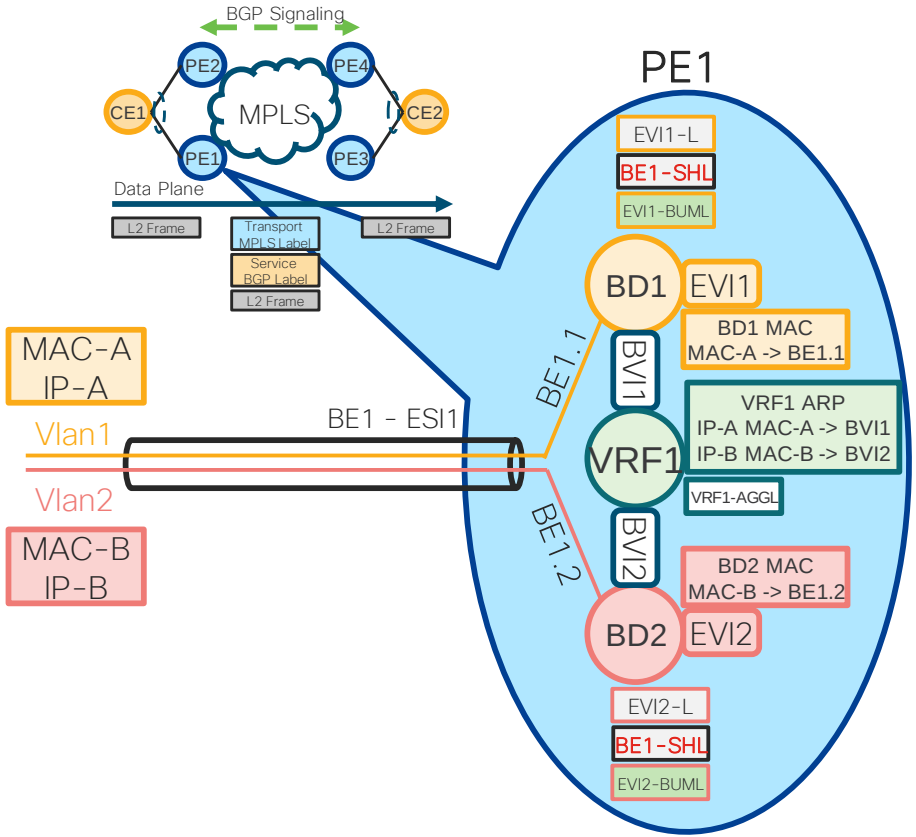
```
evpn
 interface Bundle-Ether100
  ethernet-segment
   load-balancing-mode single-active
  !
  core-isolation-group 1
 !
!
```

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

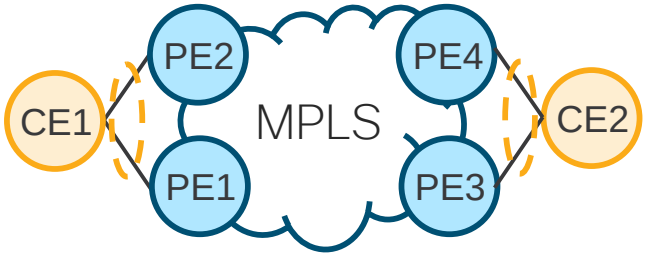
RT-5 Prefix Advertisement Route(s)

- IPv4/6 prefix of **BVI1** in VRF1 via label VRF1-AGGL
- IPv4/6 prefix of **BVI2** in VRF1 via label VRF1-AGGL

EVPN-VPWS All-Active Multihomed Service

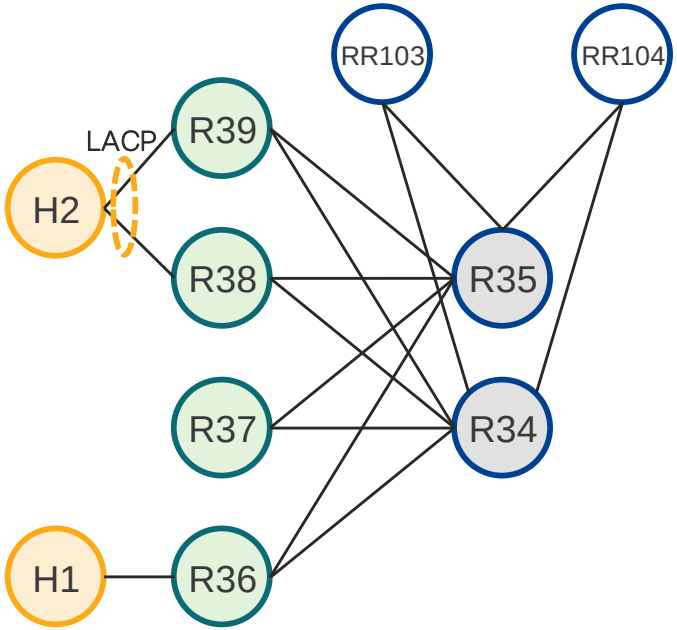
EVPN-VPWS

- Benefits of EVPN applied to point-to-point services
  - No signaling of PWs. Instead signals MP2P LSPs instead (ala L3VPN)
  - All-active CE multi-homing (per-flow LB)
  - Single-active CE multi-homing (per-service LB)
- Relies on a sub-set of EVPN routes to advertise Ethernet Segment and AC reachability
  - PE discovery & signaling via a single protocol – BGP
  - Per-EVI Ethernet Auto-Discovery route



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 36
   !
  !
 !
```

R38/R39

```
l2vpn
 xconnect group 500
  p2p 500
   interface Bundle-Ether100
   neighbor evpn evi 500 target 36 source 3839
   !
  !
 !
```

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

```
R36#show mpls forwarding labels 68106
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    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, group-best, import-candidate, imported, rib-install
      Received Path ID 0, Local Path ID 1, version 314
      Extended community: 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: 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
```

RT-1

ESI R38/R39

AC-ID

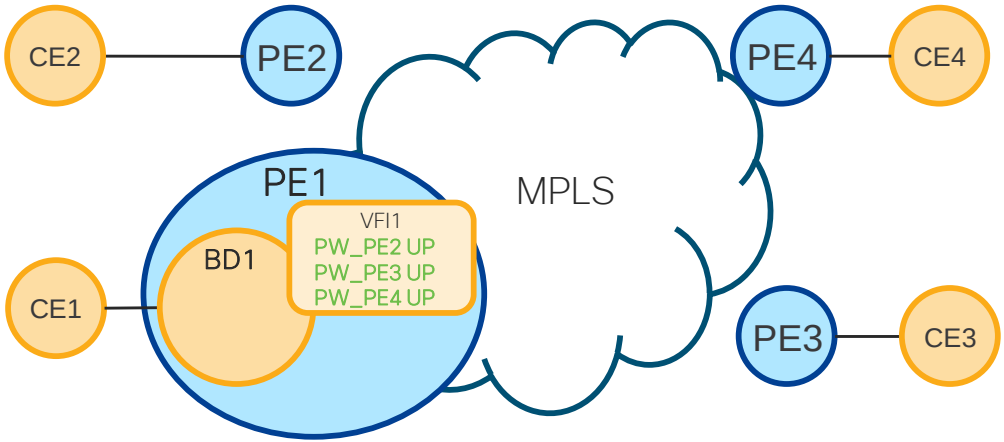
R36: EVPN-VPWS Instance View

```
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 & VPLS Seamless Integration - Migration

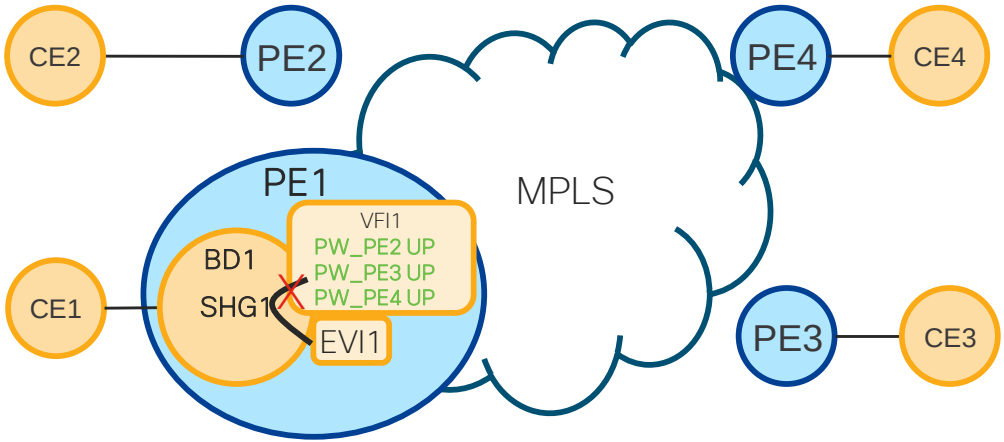
VPLS & EVPN Seamless Integration – Migration



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

VPLS & EVPN Seamless Integration – Migration



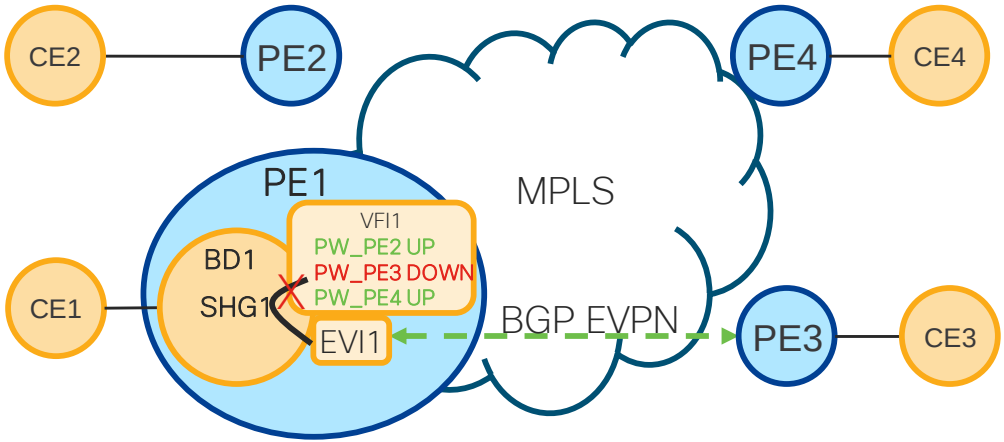
VFI1 is by default in Split Horizon Group 1

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

EVI1 is also by default in Split Horizon Group 1

- PE1 doesn't forward data between VFI1 and EVI1

VPLS & EVPN Seamless Integration – Migration



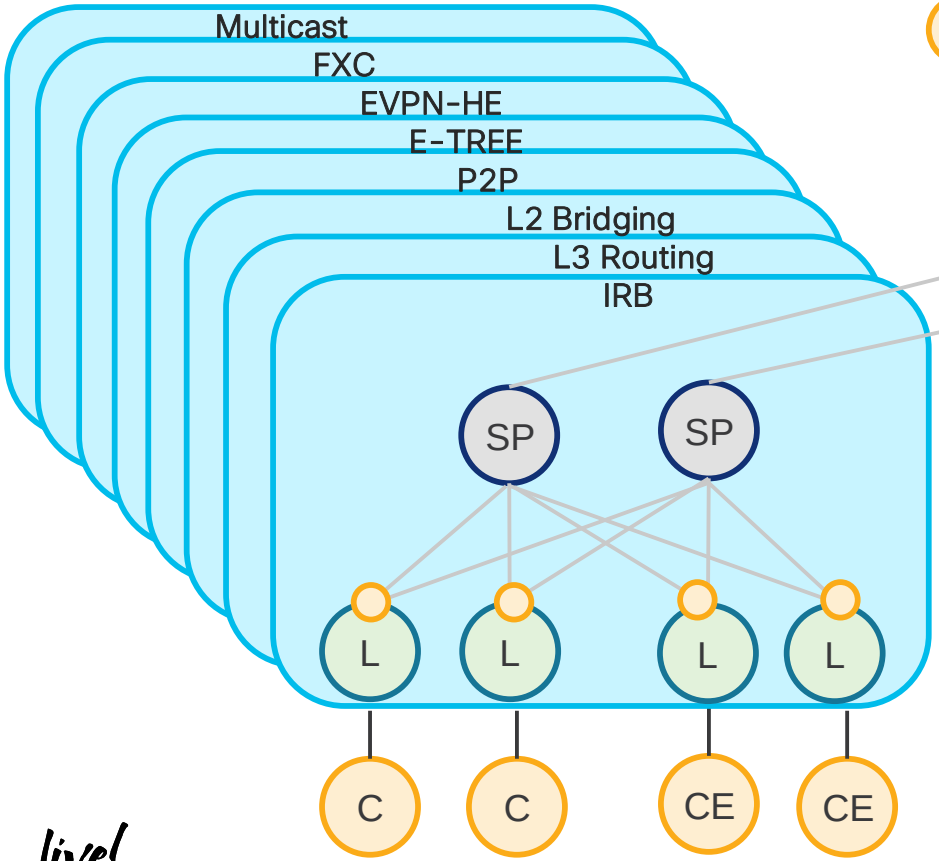
VFI1 is by default in Split Horizon Group 1
- SHG1 protects loops in MPLS Core
- Full Mesh of pseudowires(PW) is required

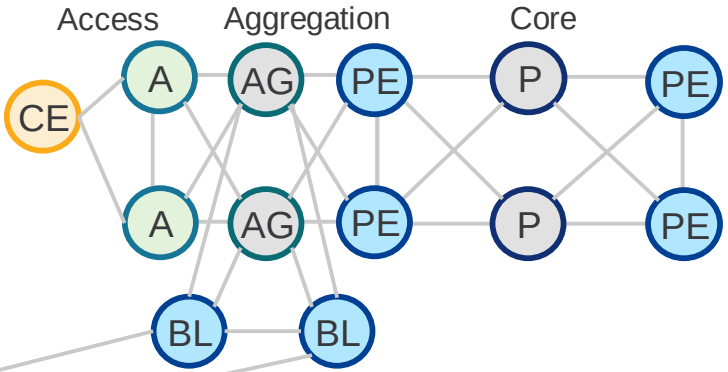
for Any-to-Any forwarding

EVI1 is also by default in Split Horizon Group 1
- PE1 doesn't forward data between VFI1 and EVI1

PE1&PE2 run BGP EVPN
- PW_PE3 goes DOWN
- Data Forwarding between PE1 and PE3 via EVI1

EVPN – Service Layering

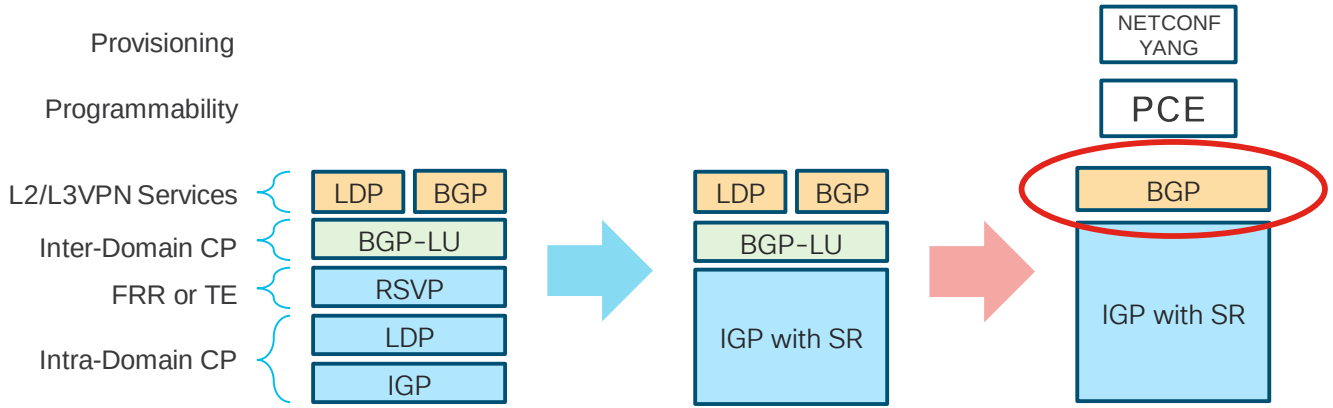




"Shared or single tenant"

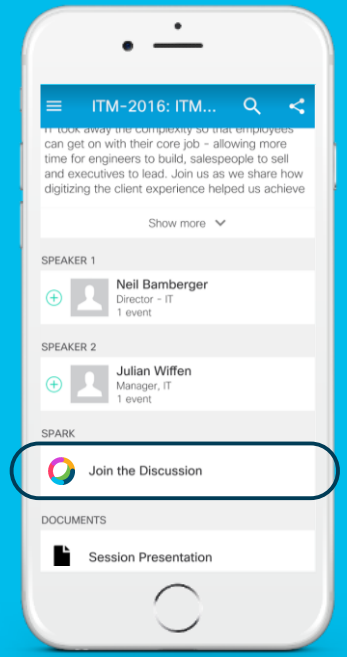
Conclusion

- EVPN is an very important complement to BGP based services
- BGP is Unified Services Control Plane across SP Network
- EVPN All-Active Multihomed Service with Distributed Anycast Gateway & Integration to L3VPN simplifies SPDC/NextGen-CO/WAN Integration



EVPN - Stay Up-To-Date

- https://e-vpn.io/
- Upcoming Conferences: https://e-vpn.io/conferences/
- Cisco Live Barcelona 2019
  - BRKSPG-2322 Service Provider Network Fabric: How to bring Access Services using EVPN
    - Patrice Brissette, Principal Engineer
  - BRKSPG-2518 Service Provider Programmable SDN solution for the Metro Fabric, powered by Segment Routing and EVPN
    - Jiri Chaloupka, Technical Marketing Engineer
  - BRKSPG-3965 EVPN Deep Dive with IOS-XR Configuration examples for Service Provider Metro and Data Center
    - Jiri Chaloupka, Technical Marketing Engineer
  - LTRSPG-2968 IOS-XR EVPN Hands-On Lab
    - Jiri Chaloupka, Technical Marketing Engineer



cs.co/ciscolivebot#BRKSPG-3965

Cisco Webex Teams

Questions?

Use Cisco Webex Teams (formerly Cisco Spark) to chat with the speaker after the session

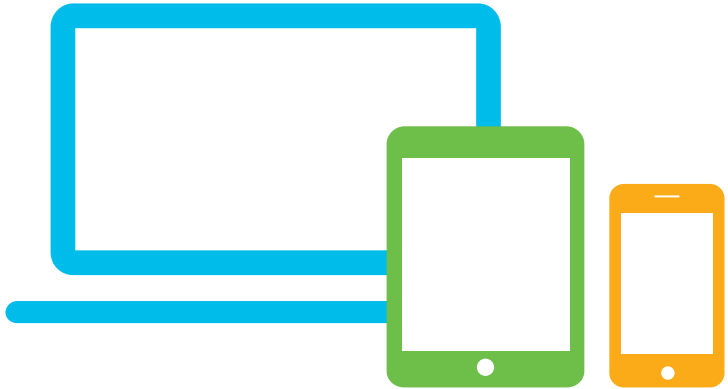
How

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

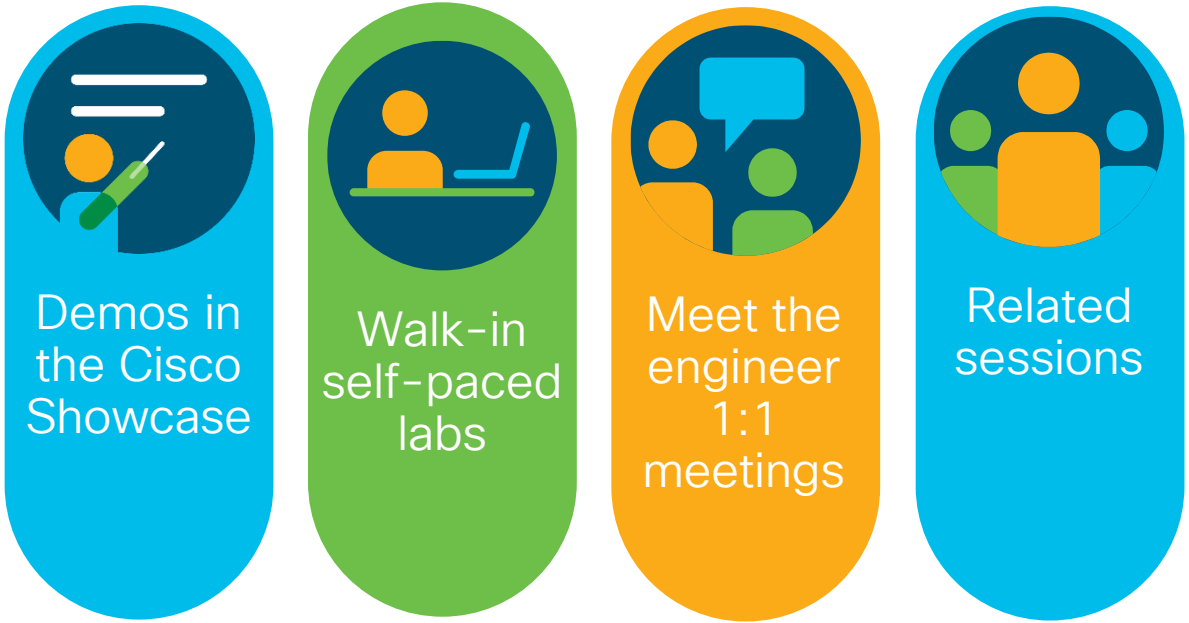
Complete your online session survey

- Please complete your Online Session Survey after each session
- Complete 4 Session Surveys & the Overall Conference Survey (available from Thursday) to receive your Cisco Live T-shirt
- All surveys can be completed via the Cisco Events Mobile App or the Communication Stations

Don't forget: Cisco Live sessions will be available for viewing on demand after the event at ciscolive.cisco.com



Continue Your Education



Thank you

INTUITIVE

CISCO

INTUITIVE