



# Under the hood of IOS-XE EVPN/VXLAN on Catalyst 9000

Dmytro Vishchuk - TAC Technical Leader  
BRKENS-3840

CISCO *Live!*



# Goal of the Session



- Understand the dependencies
- Build a solid picture
- Specific for cat9k details
- Knowledge in EVPN are expected

# Who I am?



- Start to work in industry in 2003
- Join Cisco in 2014
- Main domains – Routing, SP
- CCIE RS/SP
- Last 5 years – EVPN on Cat9k

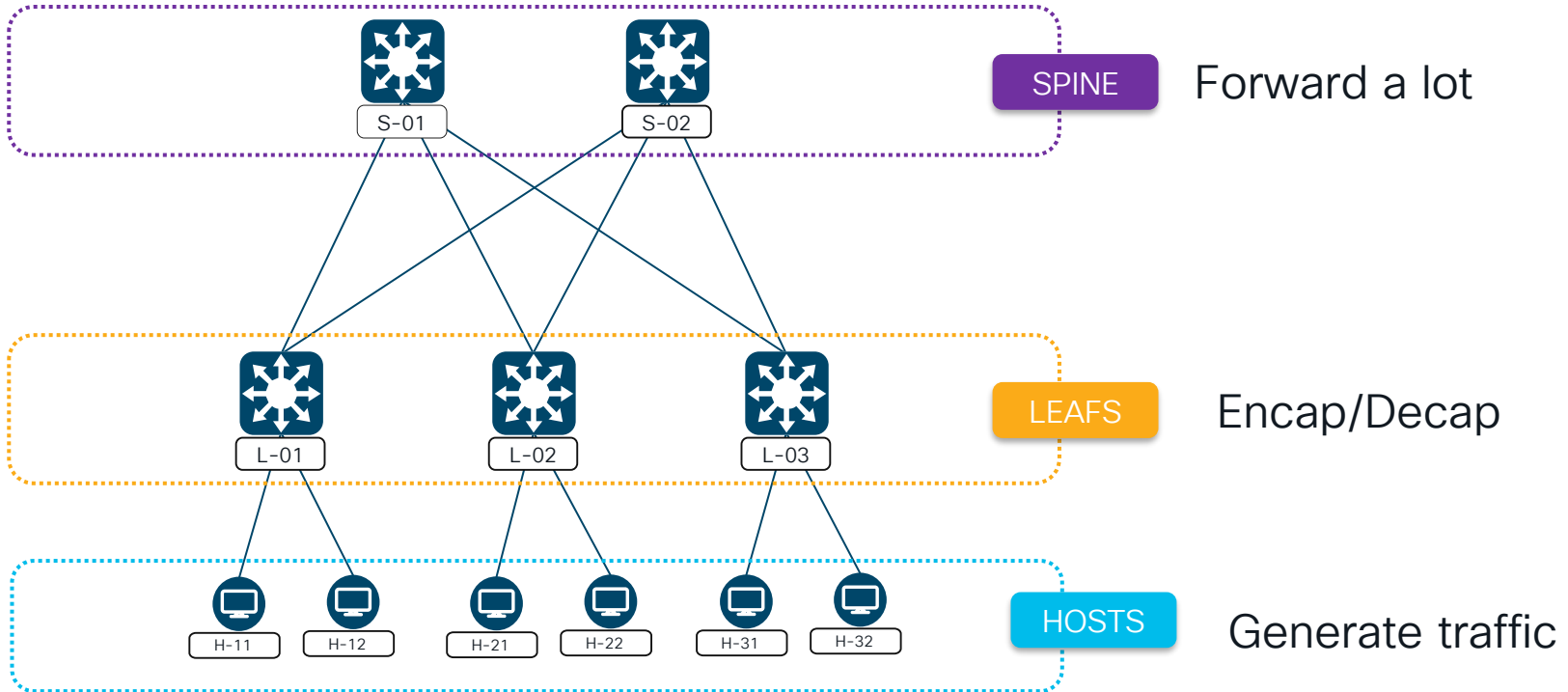
# Agenda

- Recap
- Underlay
- Bridging and Routing
- Distributed Anycast Gateway
- Mobility and loop prevention
- Conclusion

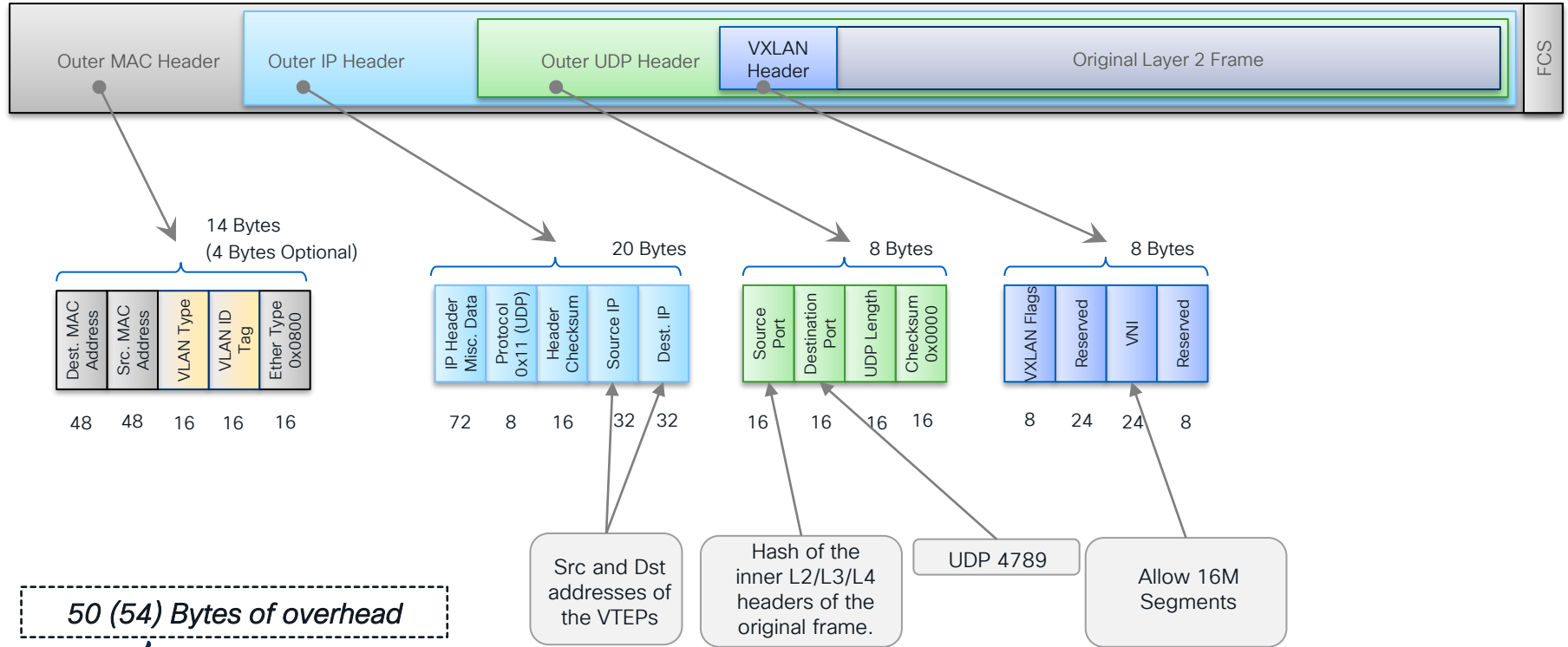
# Recap



# EVPN Fabric and Device Roles

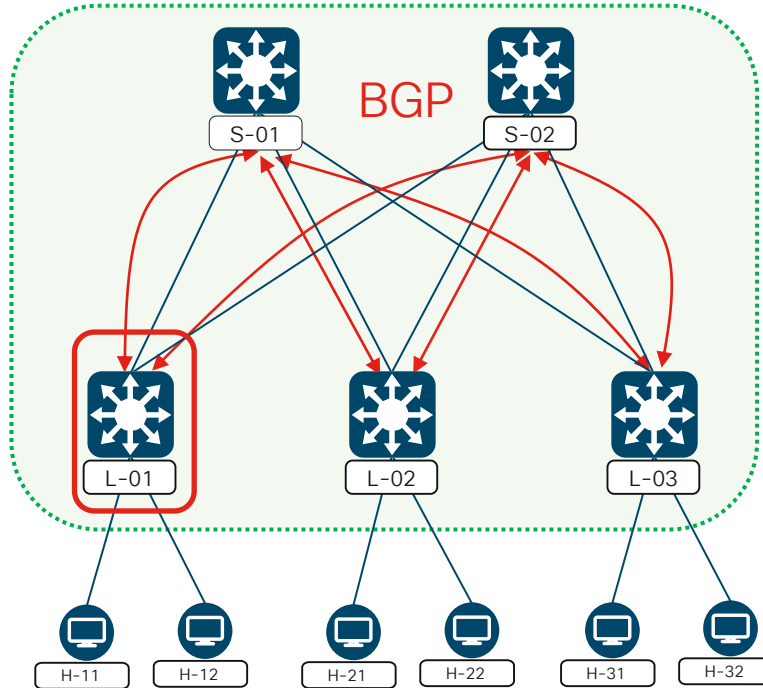


# VXLAN packet format



*CISCO Live!*

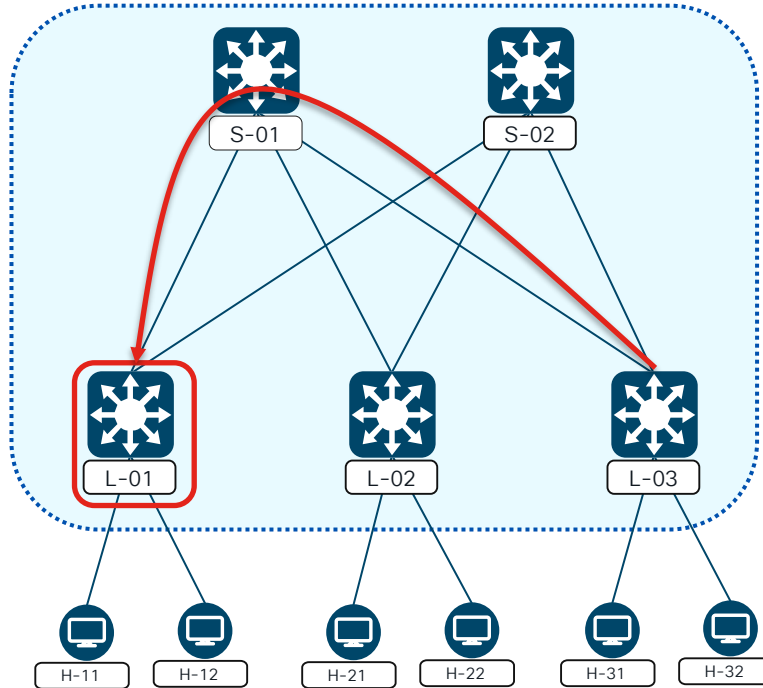
# EVPN Signaling Overlay



Overlay signaling:

- Hosts/network reachability (find the NH)
- Encap attributes
- BGP

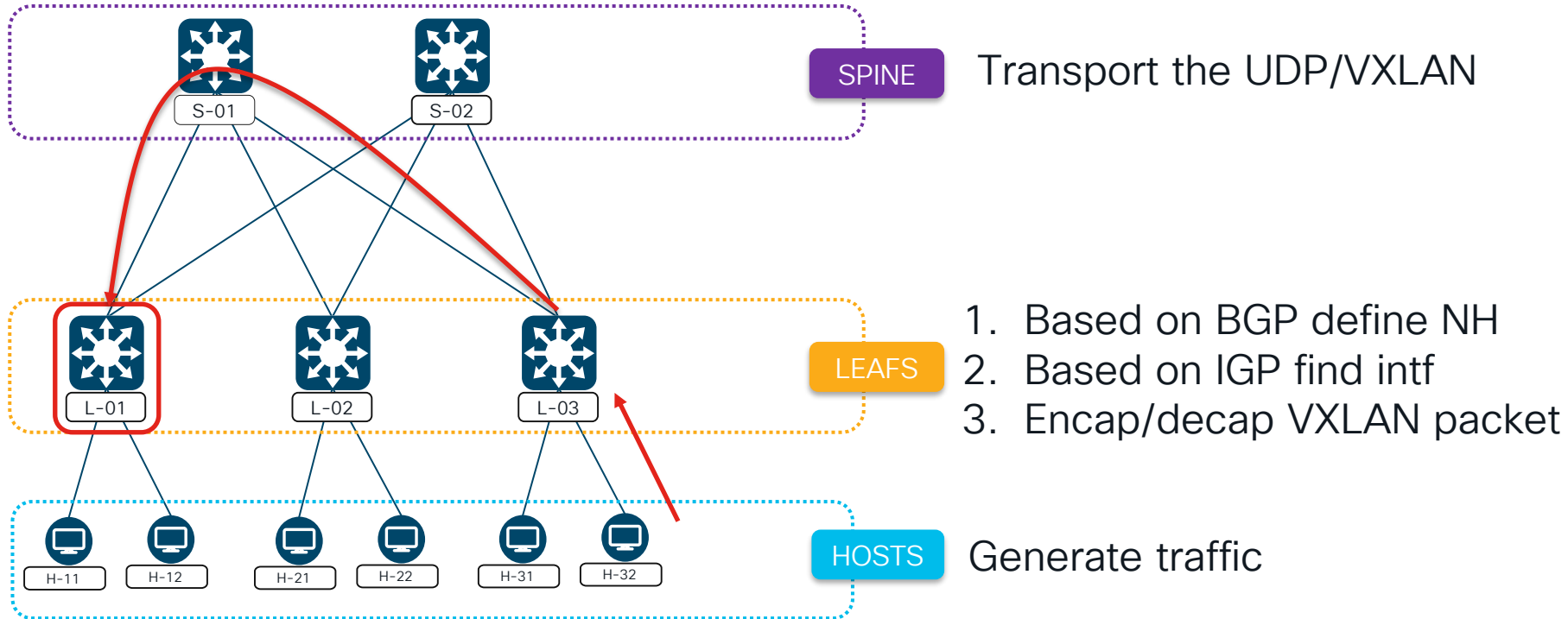
# EVPN Signaling Underlay



Underlay signaling:

- Leaf/Spine Loopback reachability
- IGP – OSPF/ISIS
- BGP

# EVPN Essence



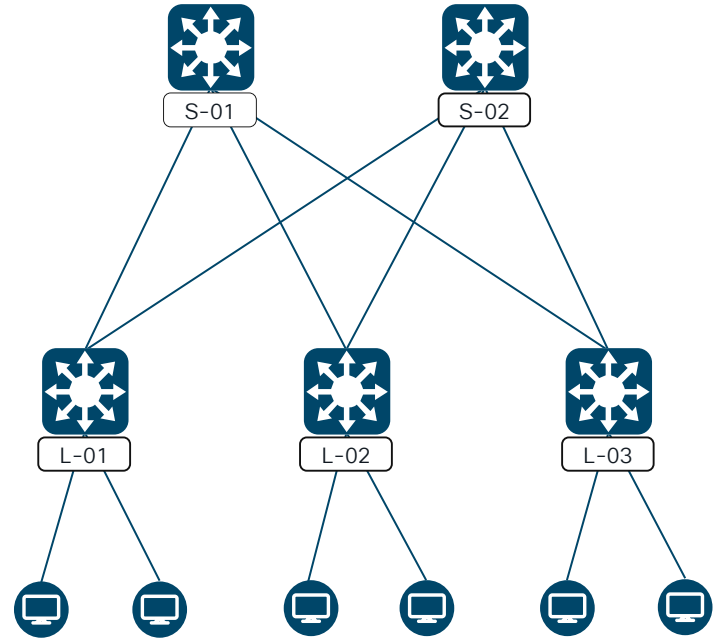
# Underlay



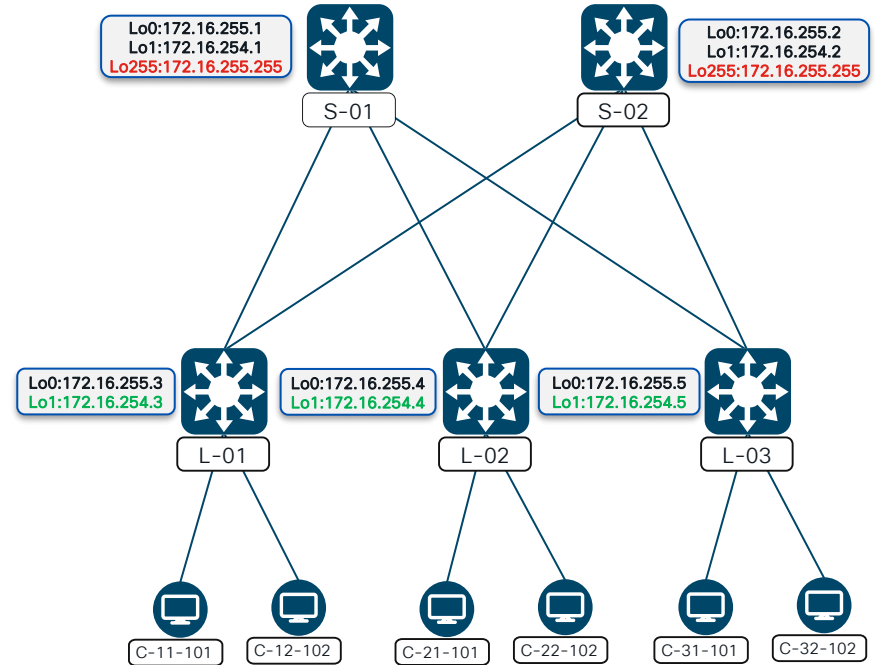
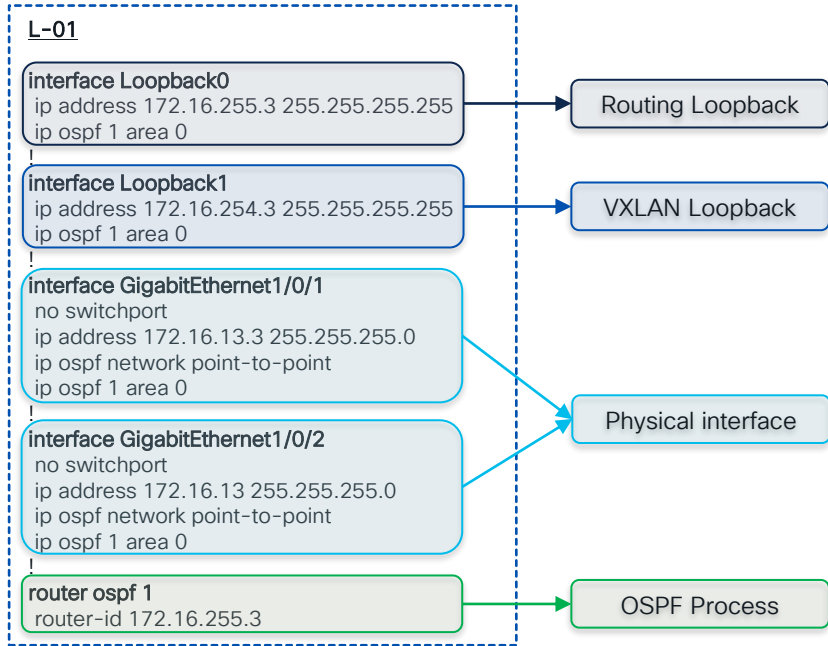
# Underlay unicast

- Catalyst 9000, IOS-XE
- Underlay: OSPF
- AS 65001, iBGP
- S-01, S-02: RRs, Anycast RP

Vlan	VNI	Notes
101	10101	Mcast 225.0.0.101
102	10102	Ingress-replication
901	50901	SVI for L3VNI VRF “green”

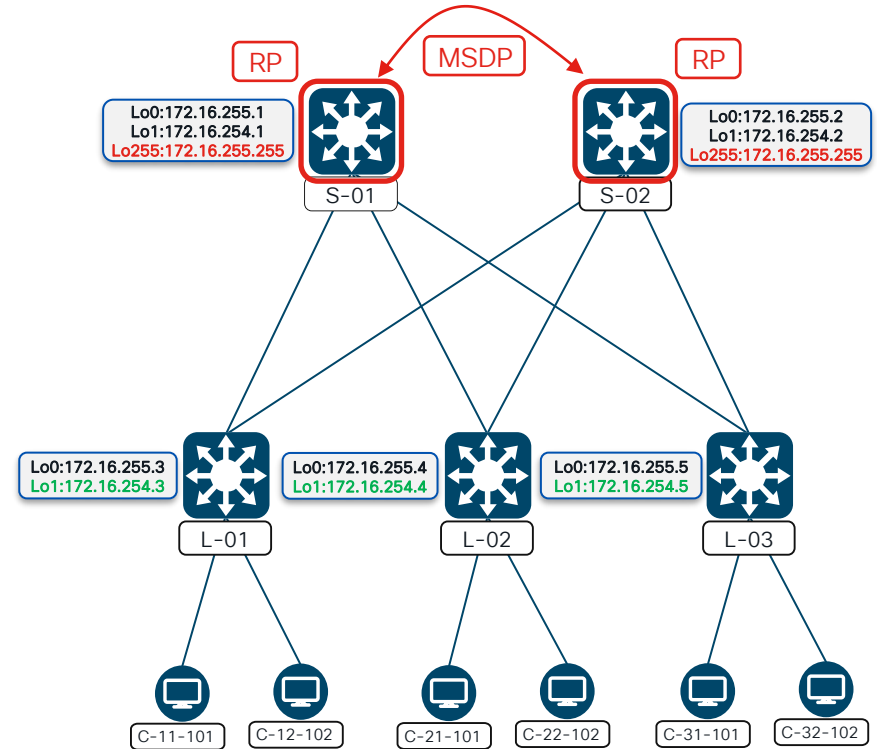


# Underlay unicast configuration

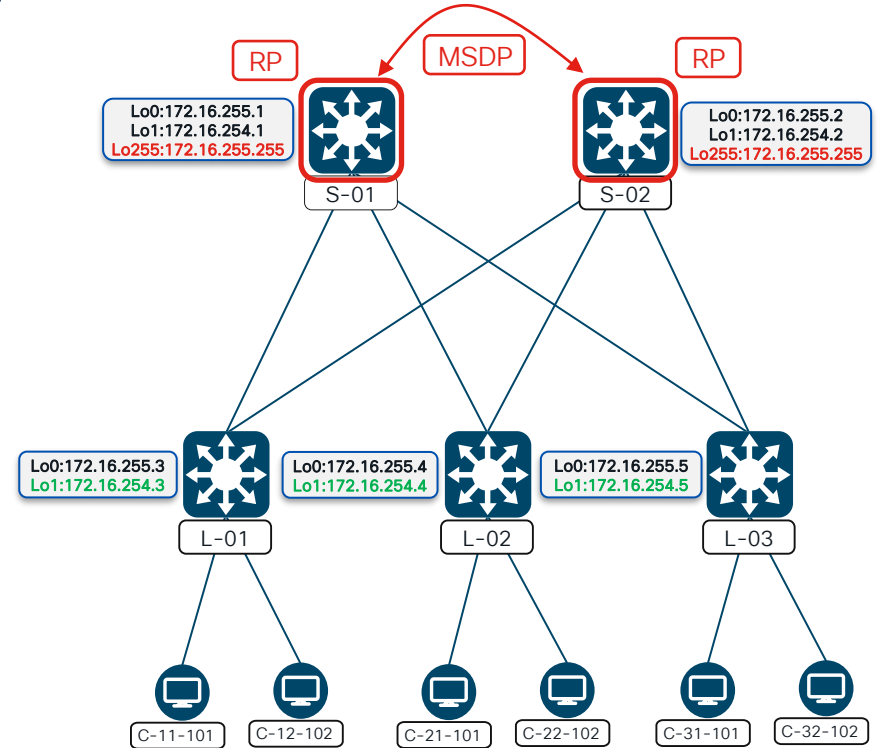
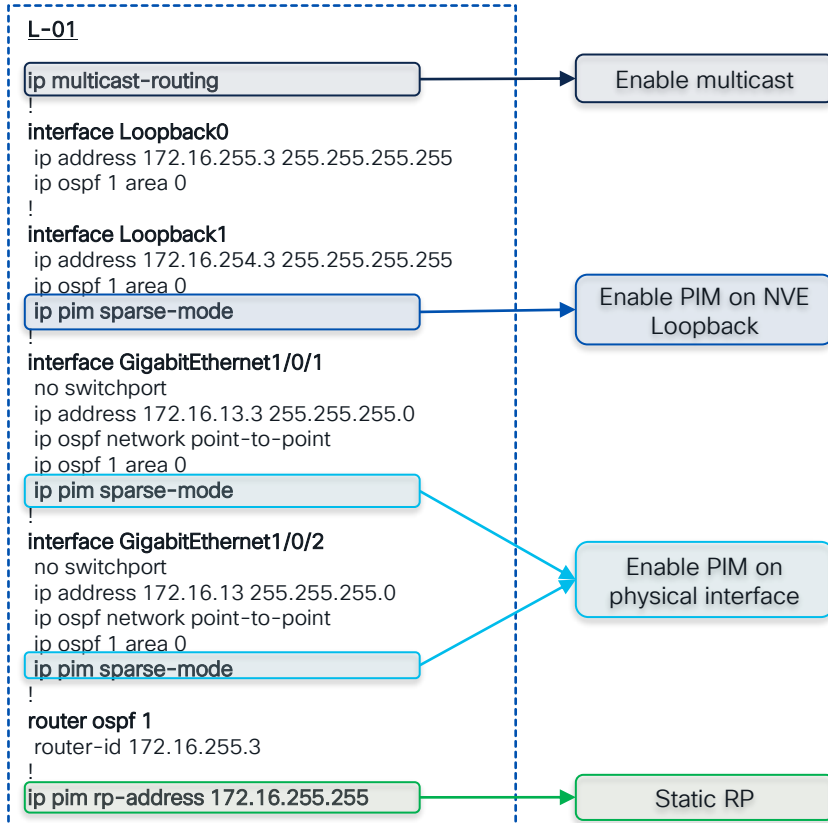


# Underlay multicast

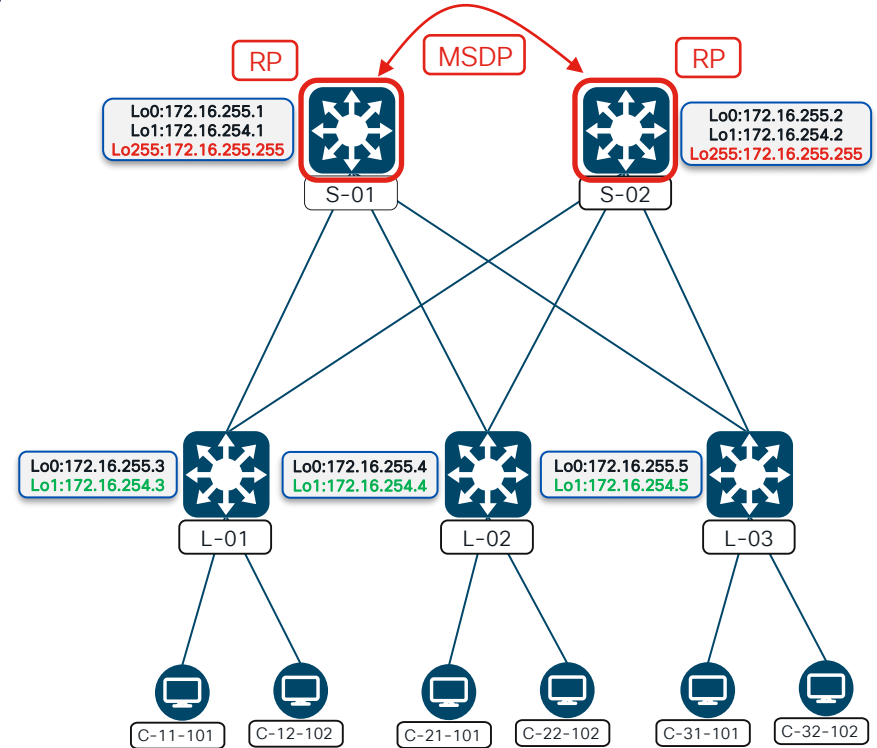
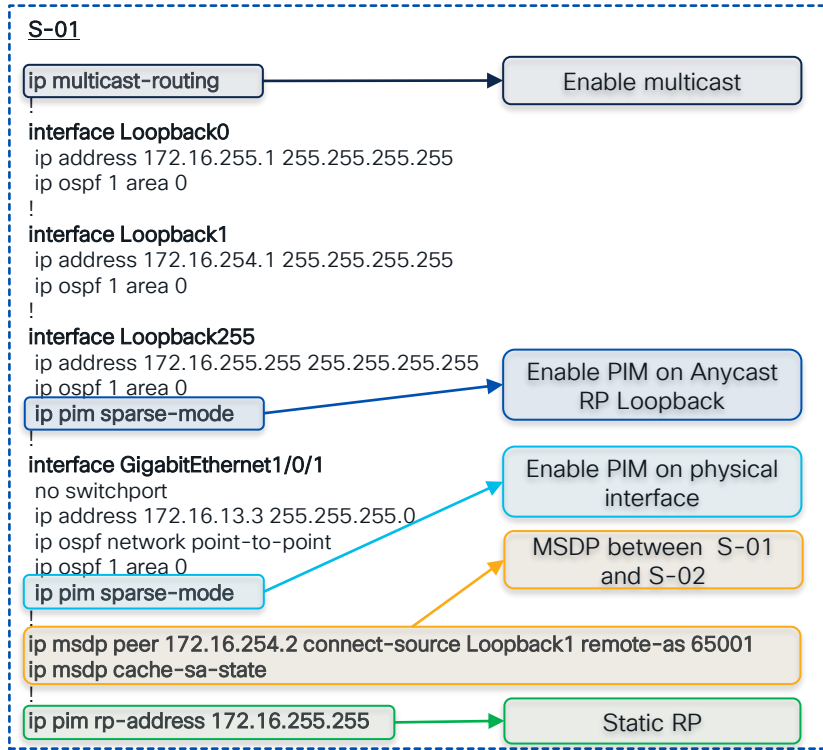
- Used for BUM replication
- PIM enabled:
  - Physical interfaces to the Core
  - Lo1 (NVE Loopback)
  - Anycast RP Loopback Lo255
- MSDP is configured between S-01 and S-02
- Static RP on all devices



# Underlay multicast configuration

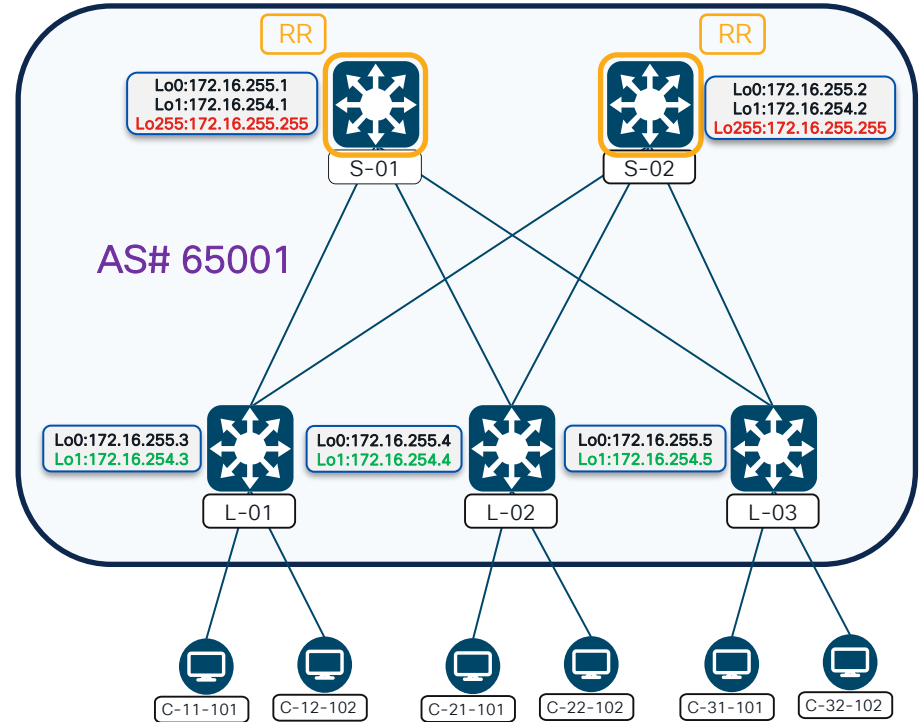


# Underlay multicast configuration

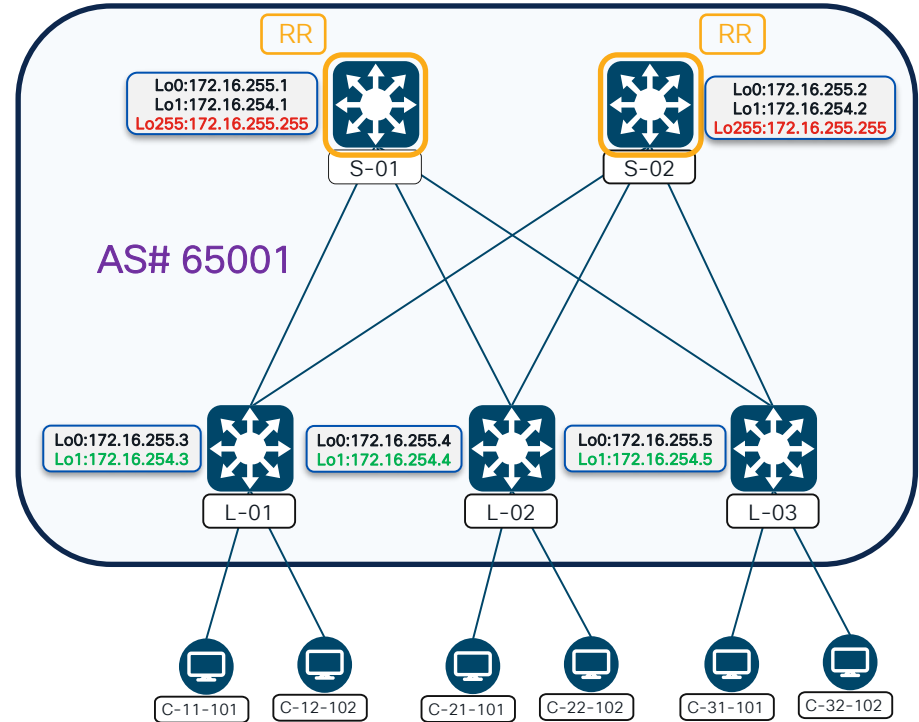
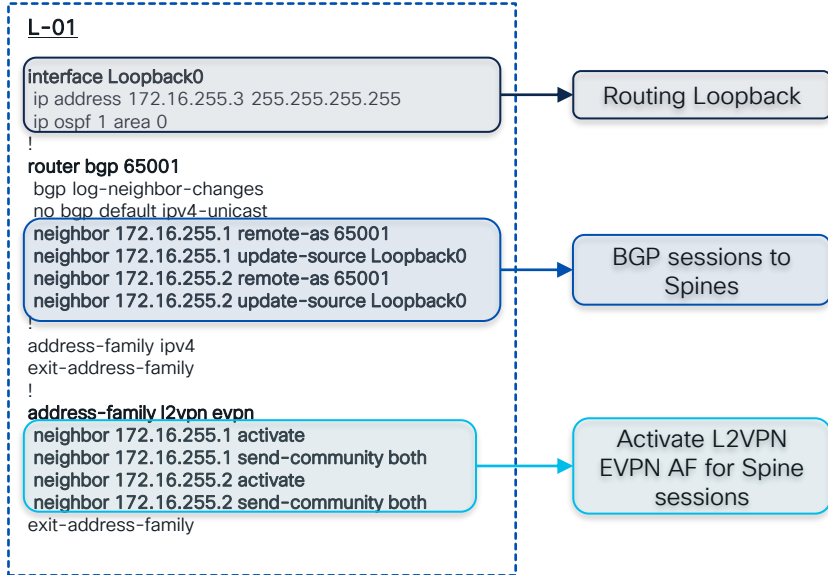


# Underlay BGP

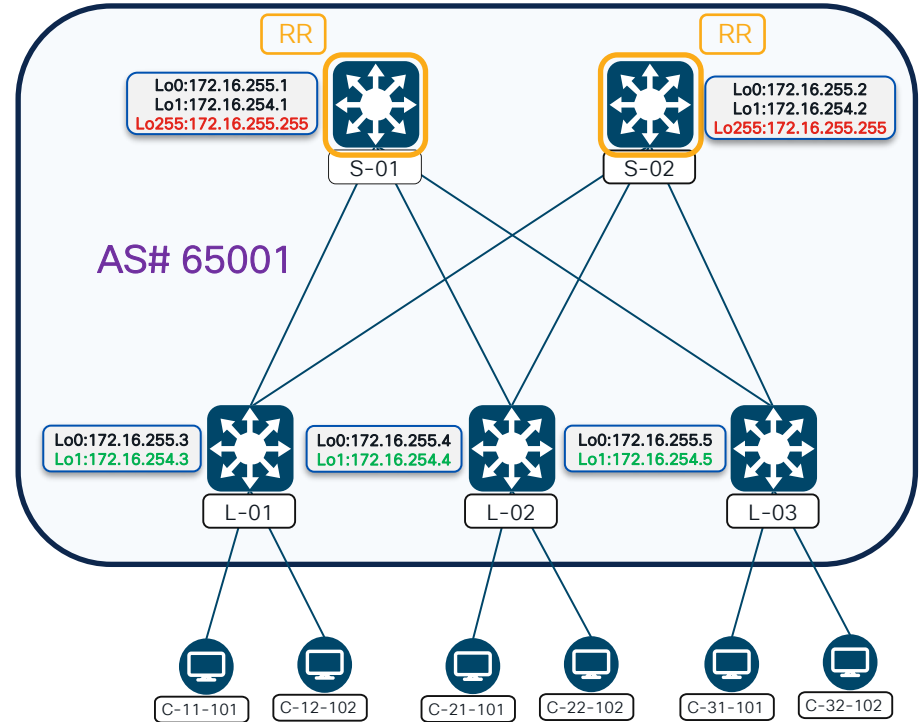
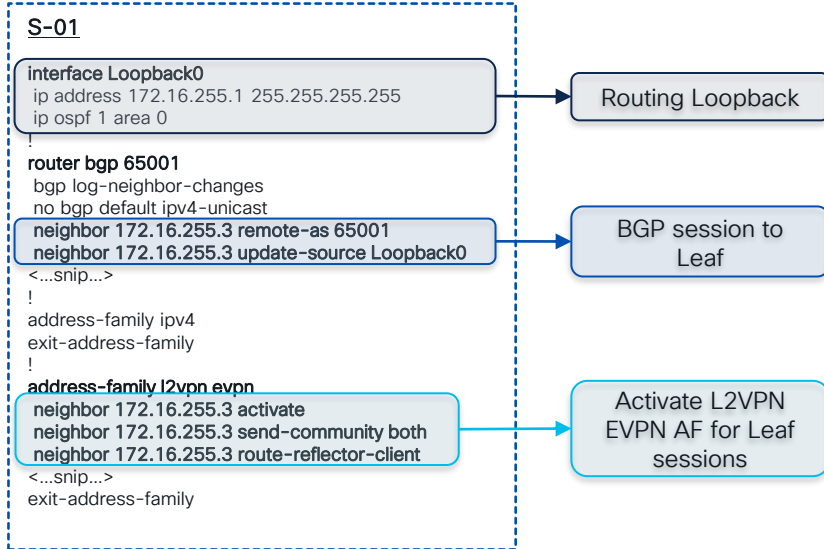
- AS 65001, iBGP
- S-01, S-02: RRs



# Underlay BGP Configuration - Leaf



# Underlay BGP Configuration - Spine



# Bridging



# Bridging

- BUM
  - Broadcast (ARP, GARP, ect)
  - Unknown unicast
  - L2 multicast
  
- Known unicast

# BUM handling



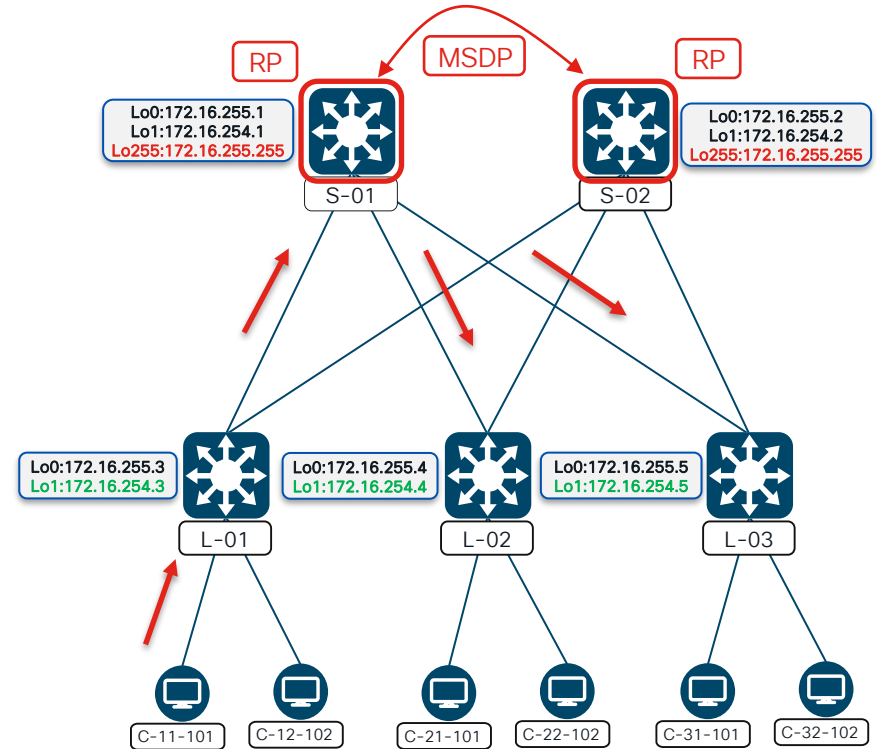
# BUM handling - Replication

Replication methods:

- Multicast replication (“static” in configuration, multicast in nature)
- Ingress replication (multiple unicast copies)

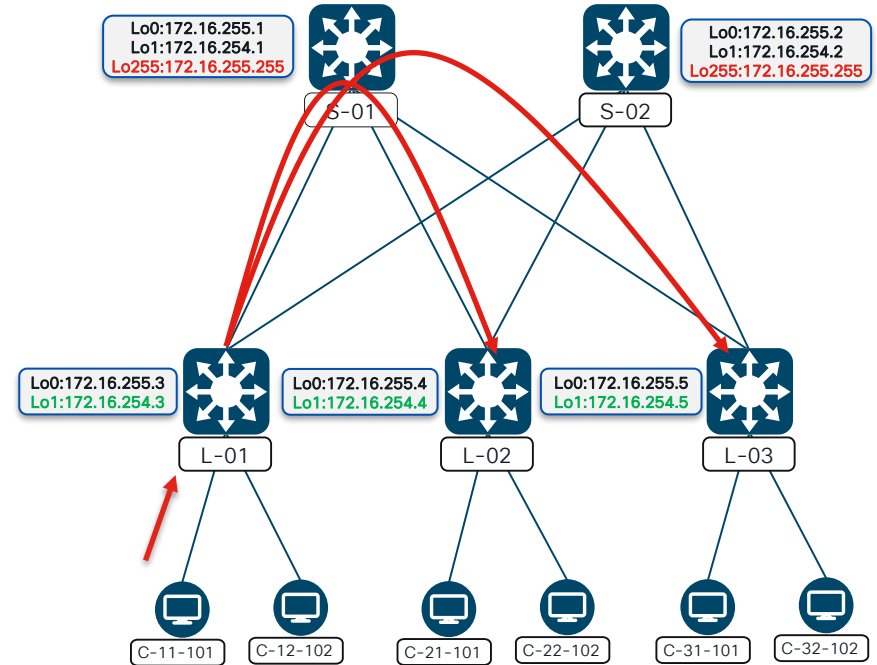
# Multicast Replication

- Pros:
  - Efficient bandwidth utilization
  - Scalable solution
- Cons:
  - Multicast configuration

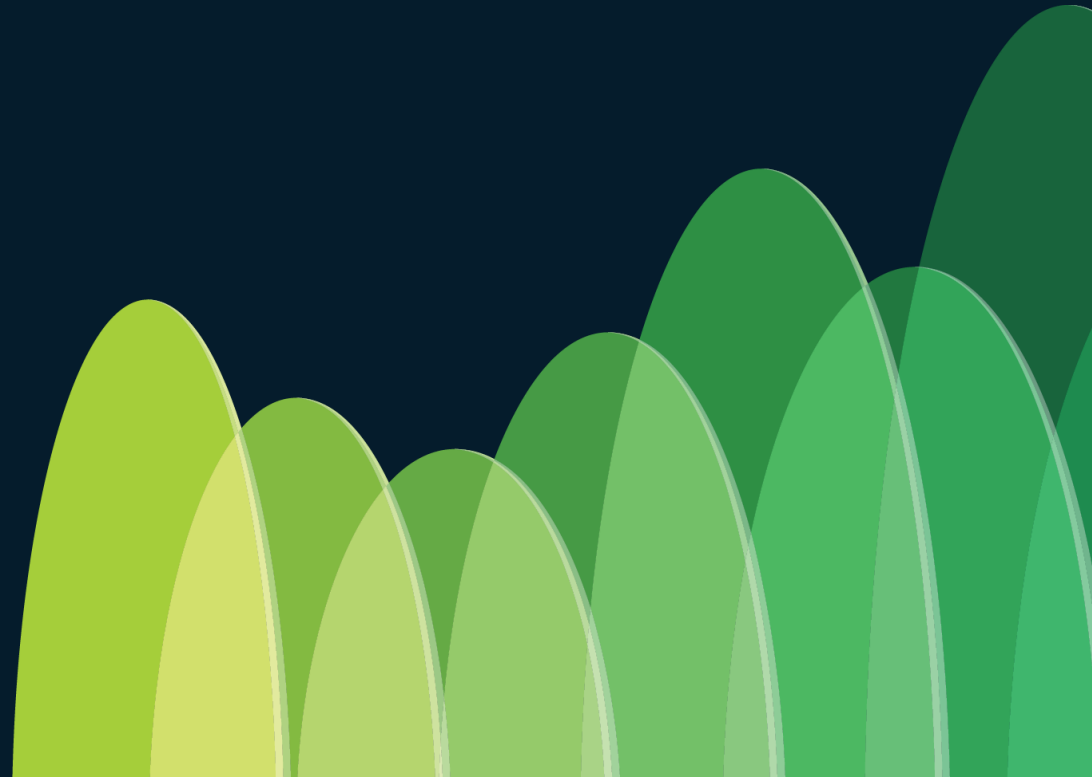


# Ingress Replication

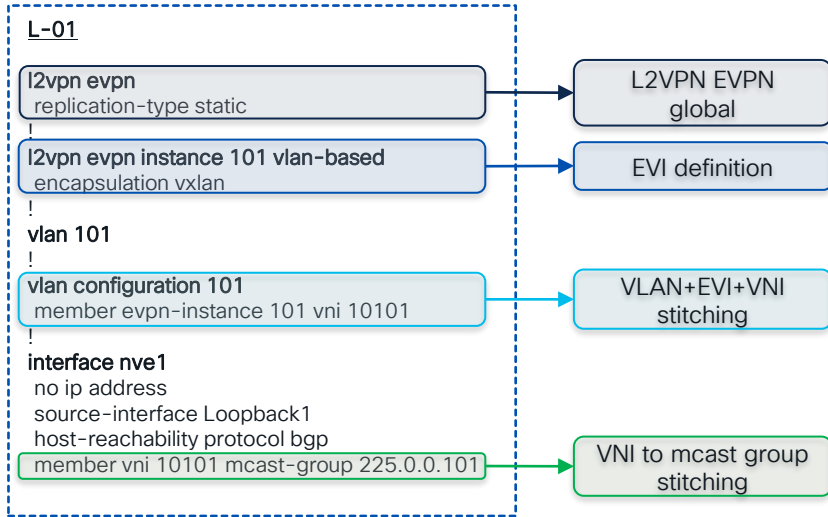
- Pros:
  - Easy to configure
  - No need multicast in the CORE
- Cons:
  - Non-optimal bandwidth utilization



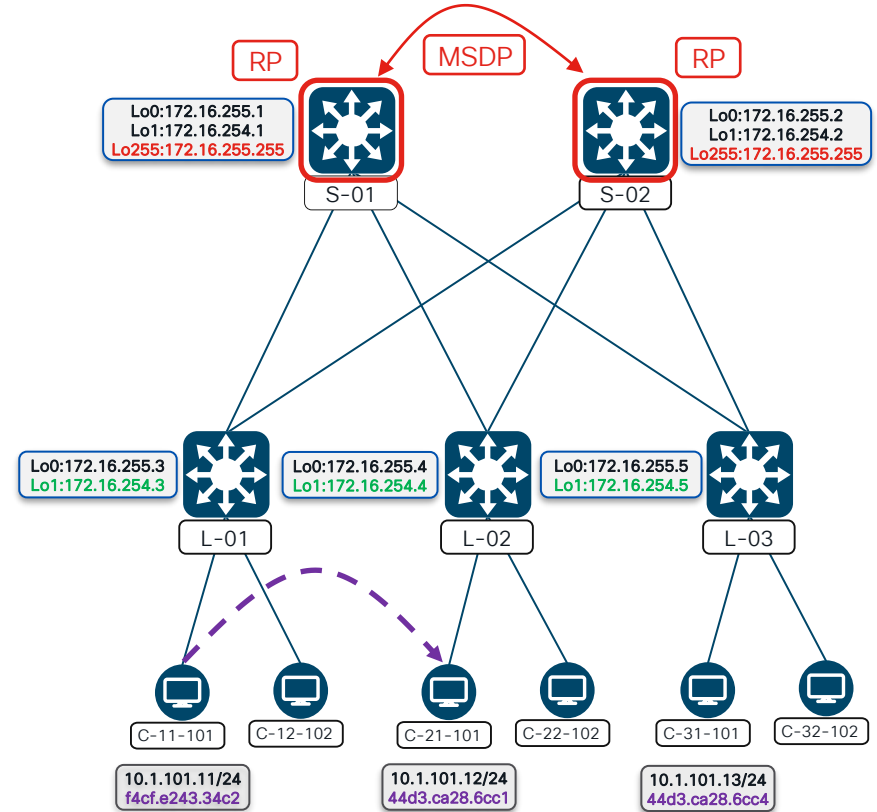
# BUM Multicast Replication



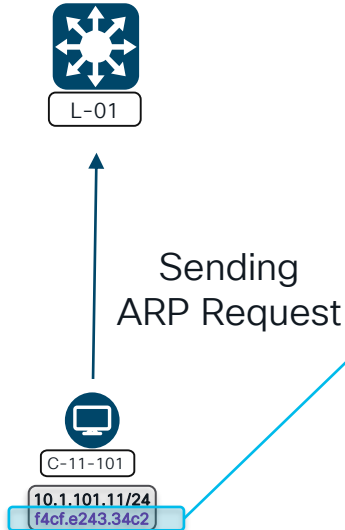
# BUM Multicast Replication



C-11-101#ping 10.1.101.12



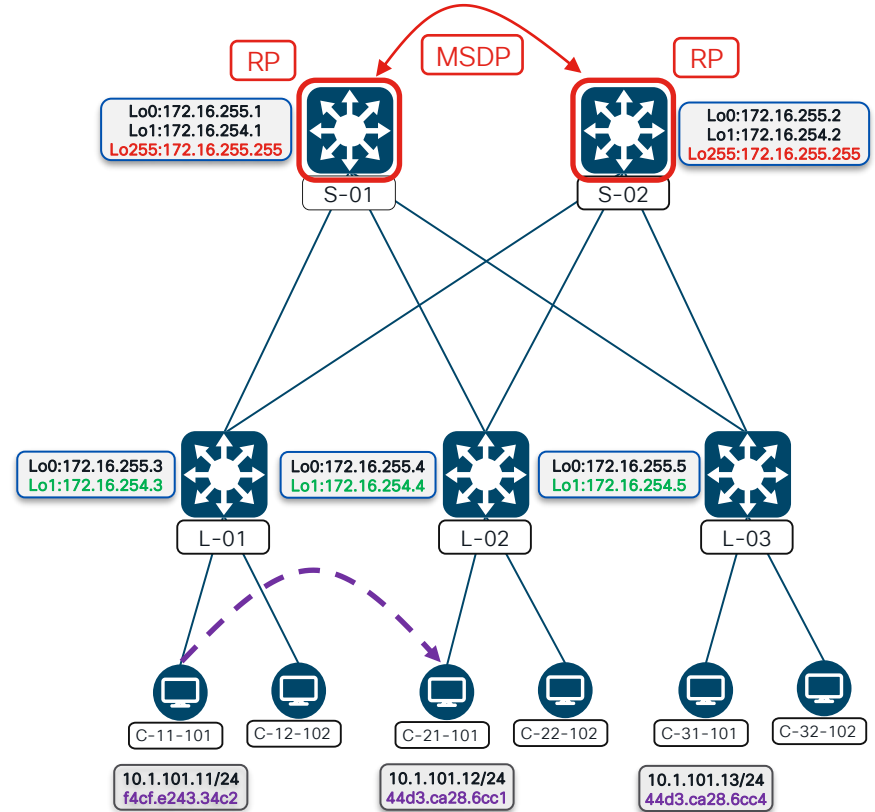
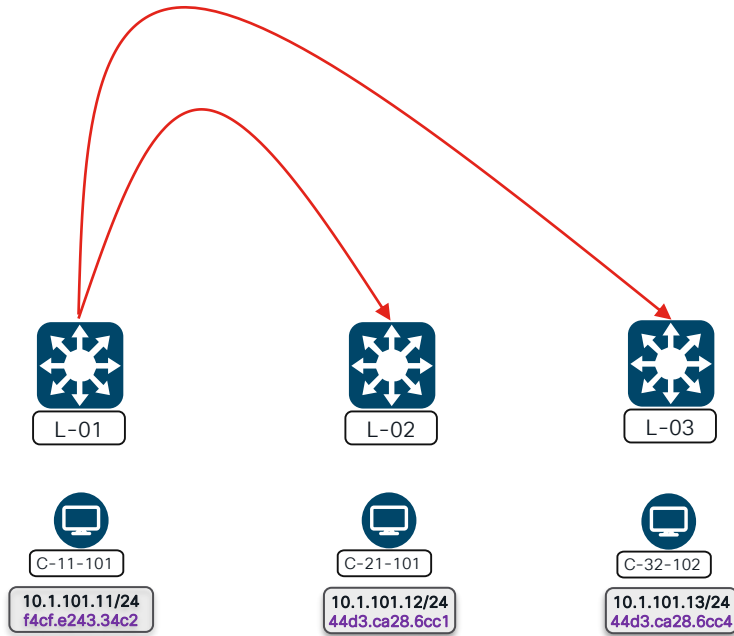
# BUM Multicast Replication



```
L-01#sh mon cap TAC buffer display-filter "arp" det
Starting the packet display ..... Press Ctrl + Shift + 6 to exit

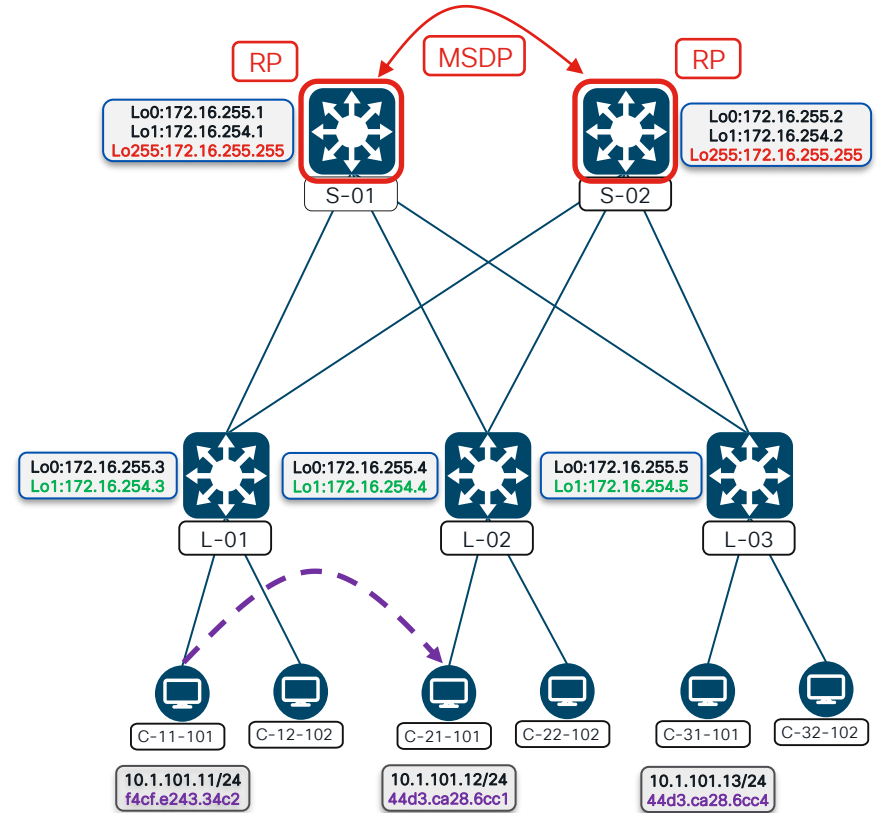
Frame 12: 60 bytes on wire (480 bits), 60 bytes captured (480 bits) on interface
/tmp/epc_ws/wif_to_ts_pipe, id 0
[Protocols in frame: eth:ethertype:arp]
Ethernet II, Src: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2), Dst: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)
  Destination: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)
    Address: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)
      ....1. .... = LG bit: Locally administered address (this is NOT the factory
default)
      ....1. .... = IG bit: Group address (multicast/broadcast)
    Source: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2)
      Address: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2)
        ....0. .... = LG bit: Globally unique address (factory default)
        ....0. .... = IG bit: Individual address (unicast)
    Type: ARP (0x0806)
    Padding: 0000000000000000000000000000000000000000000000000000000000000000
Address Resolution Protocol (request)
  Hardware type: Ethernet (1)
  Protocol type: IPv4 (0x0800)
  Hardware size: 6
  Protocol size: 4
  Opcode: request (1)
  Sender MAC address: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2)
  Sender IP address: 10.1.101.11
  Target MAC address: 00:00:00:00:00:00 (00:00:00:00:00:00)
  Target IP address: 10.1.101.12
```

# BUM Multicast Replication



# BUM Multicast Replication

- Every VTEP is the SRC and DST
- (\*,G) entries are signaled when NVE interface is up
- (S,G) entries are signaled when BUM traffic is present
- ASM (not SSM) for BUM



# BUM Multicast Replication (\*,G)

Incoming interface for (\*,G) is pointing to RP

```
L-01#sh ip rpf 172.16.255.255
RPF information for ? (172.16.255.255)
  RPF interface: GigabitEthernet1/0/2
  RPF neighbor: ? (172.16.23.2)
  RPF route/mask: 172.16.255.255/32
  RPF type: unicast (ospf 1)
  Doing distance-preferred lookups across tables
  RPF topology: ipv4 multicast base, originated from
  ipv4 unicast base
```

```
Leaf-01#sh ip mro 225.0.0.101 | b \ (
(*, 225.0.0.101), 4d00h/stopped, RP 172.16.255.255, flags: SJCx
Incoming interface: GigabitEthernet1/0/2, RPF nbr 172.16.23.2
Outgoing interface list:
Tunnel0, Forward/Sparse-Dense, 4d00h/00:01:22
```

Outgoing interface is Tunnel0  
VXLAN Tunnel

```
L-01#sh interface Tunnel0
Tunnel0 is up, line protocol is up
Hardware is Tunnel
<...snip...>
Tunnel protocol/transport MUDP/IP
TEID 0x0, sequencing disabled
Checksumming of packets disabled
source port:4789, destination_port:0
<...snip...>
```

# BUM Multicast Replication (S,G)

```
Leaf-01#sh ip mro 225.0.0.101 | b \\  
(*, 225.0.0.101), 4d00h/stopped, RP 172.16.255.255, flags: SJCFx  
  Incoming interface: GigabitEthernet1/0/2, RPF nbr 172.16.23.2  
  Outgoing interface list:  
    Tunnel0, Forward/Sparse-Dense, 4d00h/00:01:45  
  
(172.16.254.3, 225.0.0.101), 00:00:04/00:03:27, flags: FTx  
  Incoming interface: Loopback1, RPF nbr 0.0.0.0  
  Outgoing interface list:  
    GigabitEthernet1/0/2, Forward/Sparse, 00:00:04/00:03:25
```

Incoming interface for (S,G) on L-01 is NVE Loopback1

```
interface nve1  
no ip address  
source-interface Loopback1  
host-reachability protocol bgp  
member vni 10101 mcast-group 225.0.0.101
```

Outgoing interface for (S,G) on L-01 points to Spine

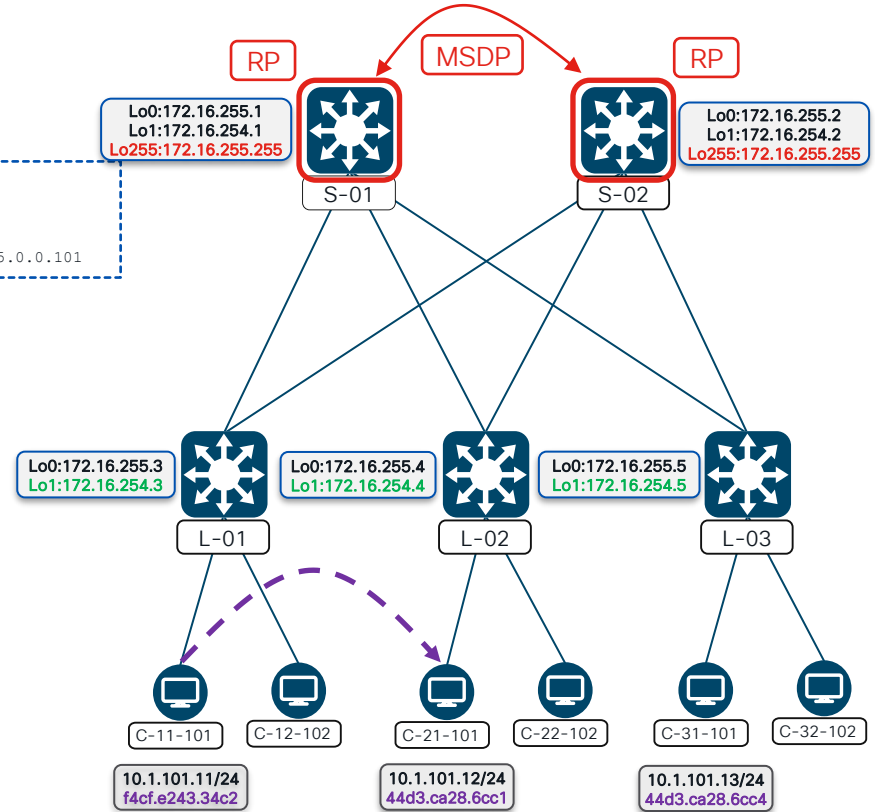
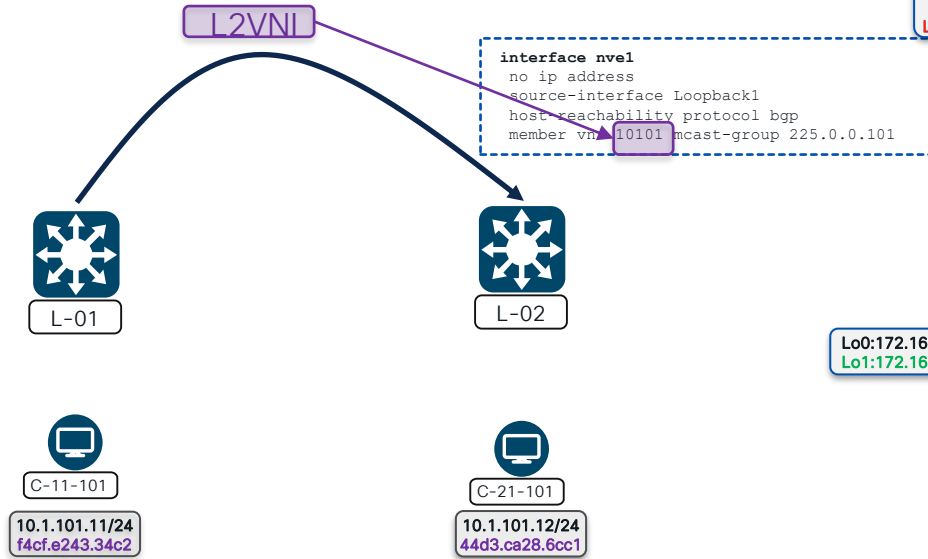
```
Leaf-02#sh ip mro 225.0.0.101 | b \\  
(*, 225.0.0.101), 5d00h/stopped, RP 172.16.255.255, flags: SJCx  
  Incoming interface: GigabitEthernet1/0/2, RPF nbr 172.16.24.2  
  Outgoing interface list:  
    Tunnel1, Forward/Sparse-Dense, 5d00h/00:00:56  
  
(172.16.254.3, 225.0.0.101), 00:00:29/00:02:30, flags: FTx  
  Incoming interface: GigabitEthernet1/0/2, RPF nbr 172.16.24.2  
  Outgoing interface list:  
    Tunnel0, Forward/Sparse-Dense, 00:00:29/00:02:30
```

Incoming interface for (S,G) on L-02 points to Spine

Outgoing interface for (S,G) on L-02 is VXLAN tunnel

# BUM Multicast Replication

## Sending ARP Request



# BUM Multicast Replication

```
Leaf-02#sh mon cap TAC buffer display-filter "arp" det
Starting the packet display ..... Press Ctrl + Shift + 6 to exit
```

```
Frame 4: 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface 0
Ethernet II, Src: 7c:21:0d:92:b2:d6 (7c:21:0d:92:b2:d6), Dst: 7c:21:0d:bd:95:64 (7c:21:0d:bd:95:64)
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.4
User Datagram Protocol, Src Port: 65418, Dst Port: 4789
  Source Port: 65418
  Destination Port: 4789
  Length: 76
```

## Virtual eXtensible Local Area Network

VXLAN Network Identifier (VNI): 10101

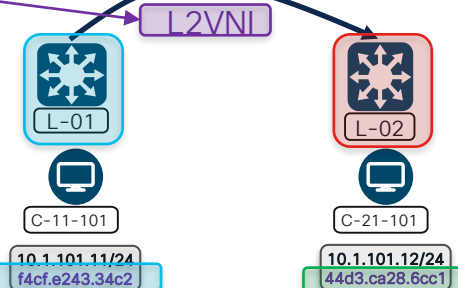
Reserved: 0

```
Ethernet II, Src: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2), Dst: 44:d3:ca:28:6c:c1 (44:d3:ca:28:6c:c1)
Address Resolution Protocol (request)
```

```
Hardware type: Ethernet (1)
Protocol type: IPv4 (0x0800)
Hardware size: 6
Protocol size: 4
Opcode: request (1)
Sender MAC address: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2)
Sender IP address: 10.1.101.11
Target MAC address: 00:00:00:00:00:00 (00:00:00:00:00:00)
Target IP address: 10.1.101.12
```

UNICAST???

Why not  
ff:ff:ff:ff:ff:ff ???



# ARP Suppression

## draft-ietf-bess-evpn-proxy-arp-nd-16

As a high level example, the operation of the **EVPN Proxy-ARP/ND** function in the network of Figure 1 is described below. In this example we assume IP1, IP2 and IP3 are IPv4 addresses:

1. Proxy-ARP/ND is enabled in BD1 of PE1, PE2 and PE3.
2. The PEs start adding dynamic, static and EVPN-learned entries to their Proxy tables:
  - A. PE3 adds IP1->M1 and IP2->M2 based on the EVPN routes received from PE1 and PE2. Those entries were previously learned as dynamic entries in PE1 and PE2 respectively, and advertised in BGP EVPN.
  - B. PE3 adds IP4->M4 as dynamic. This entry is learned by snooping the corresponding ARP messages sent by CE4.
  - C. An operator also provisions the static entry IP3->M3.
3. When CE3 sends an ARP Request asking for the MAC address of IP1, PE3 will:
  - A. Intercept the ARP Request and perform a Proxy-ARP lookup for IP1.
  - B. If the lookup is successful (as in Figure 1), PE3 will send an ARP Reply with IP1->M1. The ARP Request will not be flooded to the EVPN network or any other local CEs.
  - C. If the lookup is not successful, PE3 will flood the ARP Request in the EVPN network and the other local CEs.

### 3.4. Unicast-forward Sub-Function

As discussed in [Section 3.3](#), in some cases the operator may want to **'unicast-forward' certain ARP-Request and NS messages** as opposed to reply to them. The implementation of a 'unicast-forward' function is RECOMMENDED. This option can be enabled with one of the following parameters:

- a. unicast-forward always
- b. unicast-forward unknown-options

If 'unicast-forward always' is enabled, the PE will perform a Proxy-ARP/ND table lookup and in case of a hit, the PE will forward the packet to the owner of the MAC found in the Proxy-ARP/ND table. This is irrespective of the options carried in the ARP/ND packet. This option provides total transparency in the BD and yet reduces the amount of flooding significantly.

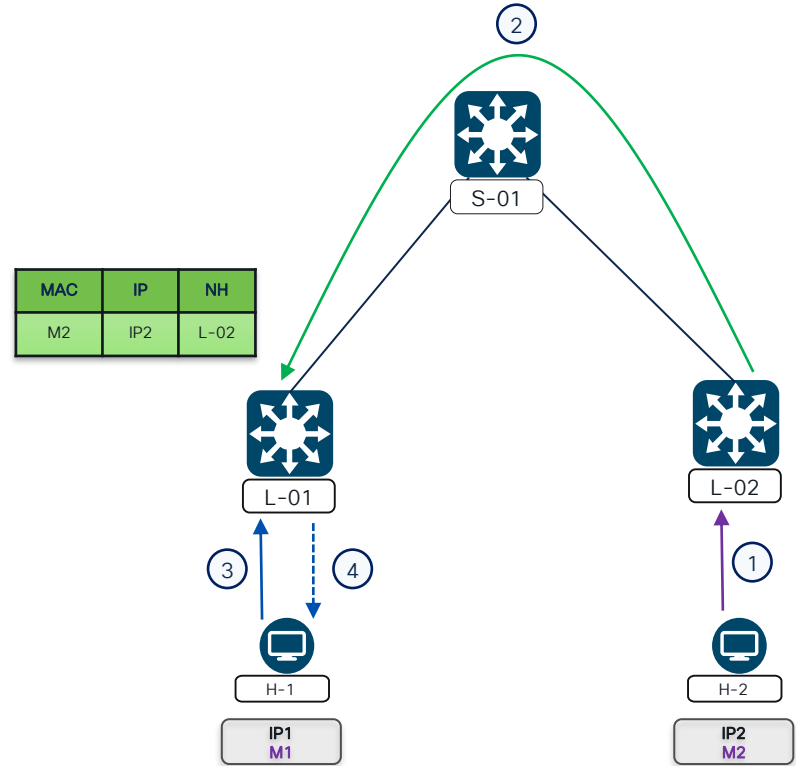
If 'unicast-forward unknown-options' is enabled, upon a successful Proxy-ARP/ND lookup, the PE will perform a 'unicast-forward' action only if the ARP-Request or NS messages carry unknown options, as explained in [Section 3.3](#). The 'unicast-forward unknown-options' configuration allows the support of new applications using ARP/ND in the BD while still reducing the flooding.

Irrespective of the enabled option, if there is no successful Proxy-ARP/ND lookup, the unknown ARP-Request/NS will be flooded in the context of the BD, as per [Section 3.6](#).

# ARP Suppression (Proxy-ARP)

1. Host H2 is active and learned by Fabric
2. BGP Update is advertised
3. H-1 sends ARP Requests
4. L-01 sends ARP Reply instead of H-2

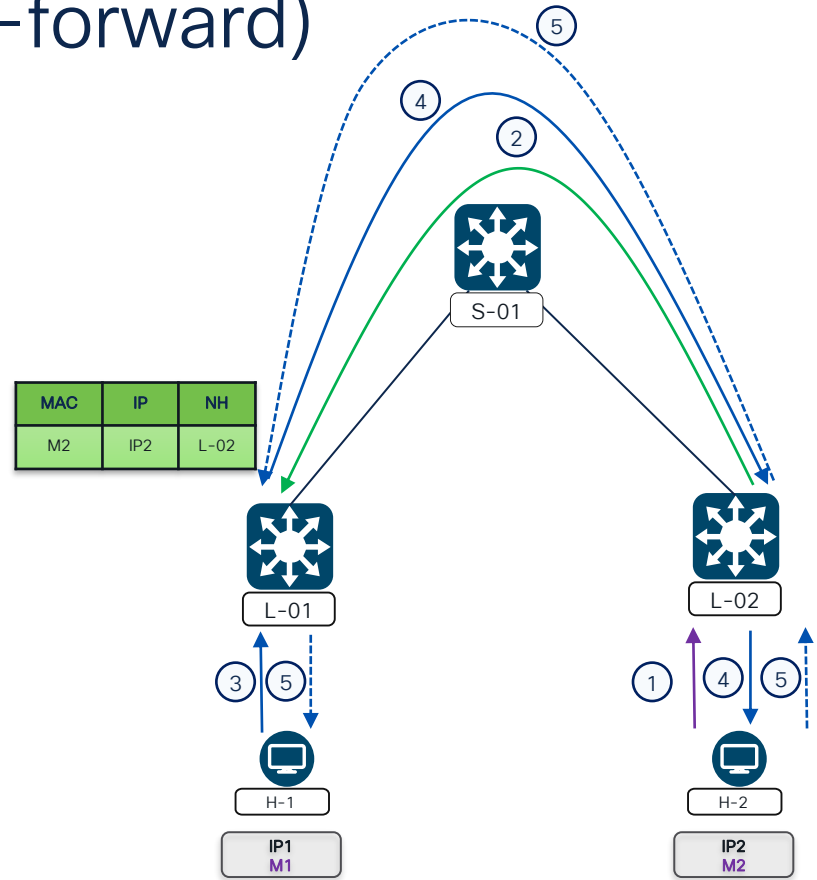
*If H-2 is unknown - flood*



# ARP Suppression (unicast-forward)

1. Host H2 is active and learned by Fabric
2. BGP Update is advertised
3. H-1 sends ARP Requests
4. L-01 converts broadcast ARP to unicast ARP and sends to L-02
5. H-2 sends unicast ARP Reply

*If H-2 is unknown - flood*



# ARP Suppression

\* ARP/ND suppression by default starting from 17.1.1

```
L-01#sh l2vpn evpn summary | i ARP
ARP/ND Flooding Suppression: Enabled
```

To disable ARP/ND suppression globally

```
L-01#conf t
Enter configuration commands, one per line. End with CNTL/Z.
L-01(config)#
L-01(config)#l2vpn evpn
L-01(config-evpn)#flooding-suppression address-resolution disable
L-01(config-evpn)#end

L-01#sh l2vpn evpn summ | i ARP
ARP/ND Flooding Suppression: Disabled
```

Or per EVI

```
L-01#conf t
Enter configuration commands, one per line. End with CNTL/Z.
L-01(config)#l2vpn evpn instance 101 vlan-based
L-01(config-evpn-evi)#flooding-suppression address-resolution disable
L-01(config-evpn-evi)#end
```

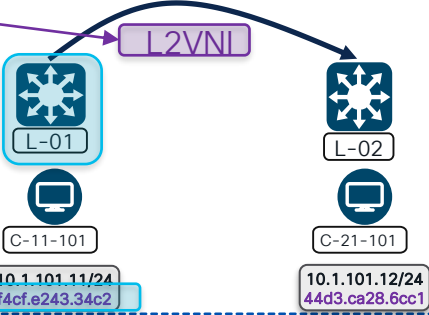
# ARP Suppression (unknown host)

```
L-02#sh mon cap TAC buff display-filter "arp" det
Starting the packet display ..... Press Ctrl + Shift + 6 to exit

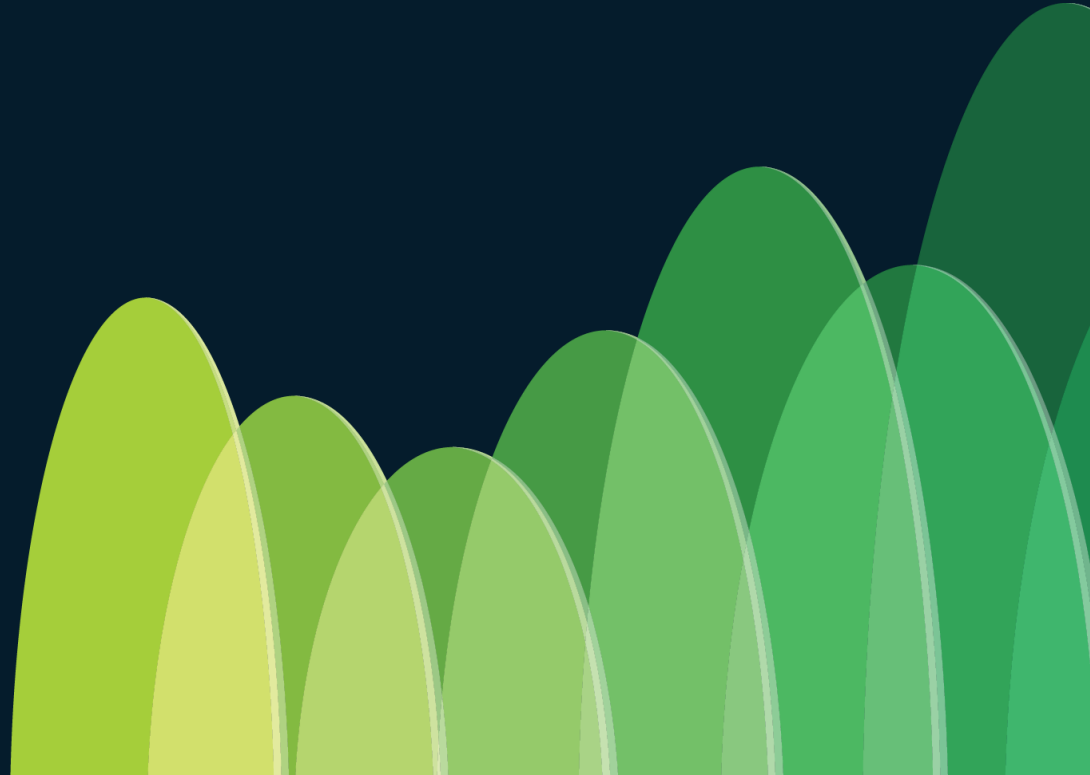
Frame 7: 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface 0
Ethernet II, Src: 7c:21:0d:bd:2c:d6 (7c:21:0d:bd:2c:d6), Dst: 01:00:5e:00:00:65 (01:00:5e:00:00:65)
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 225.0.0.101
User Datagram Protocol, Src Port: 65418, Dst Port: 4789
    Source Port: 65418
    Destination Port: 4789
    Length: 76
Virtual eXtensible Local Area Network
    VXLAN Network Identifier (VNI): 10101
    Reserved: 0
Ethernet II, Src: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2), Dst: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)
Address Resolution Protocol (request)
    Hardware type: Ethernet (1)
    Protocol type: IPv4 (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: request (1)
    Sender MAC address: f4:cf:e2:43:34:c2 (f4:cf:e2:43:34:c2)
    Sender IP address: 10.1.101.11
    Target MAC address: 00:00:00:00:00:00 (00:00:00:00:00:00)
    Target IP address: 10.1.101.254
```

MULTICAST

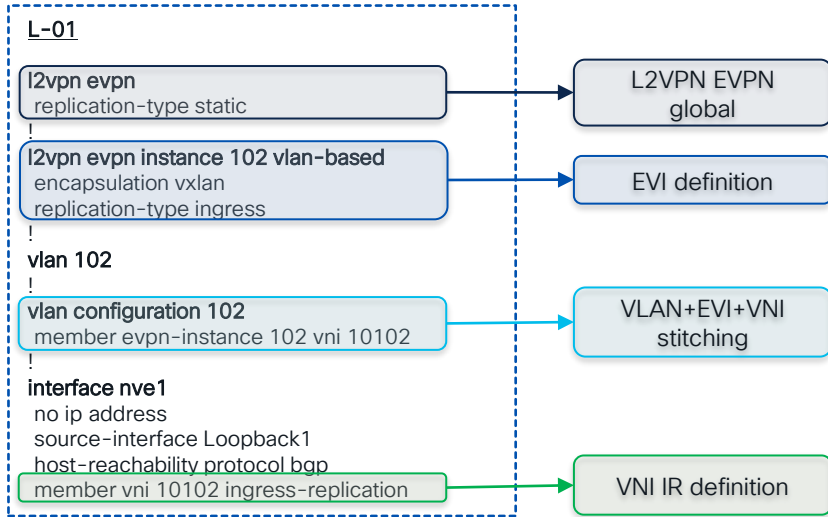
Broadcast



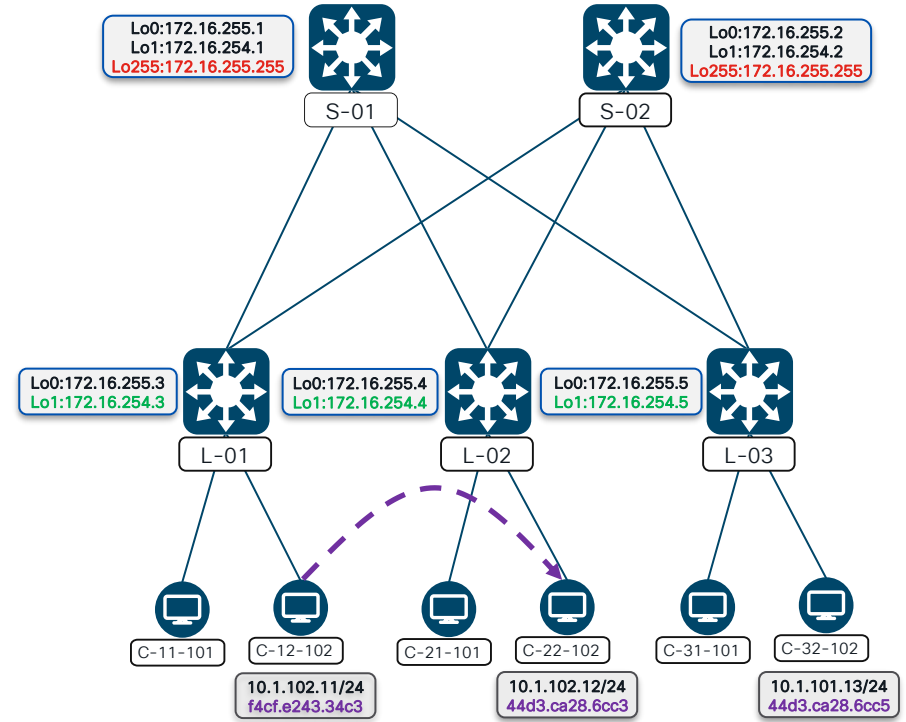
# BUM Ingress Replication



# BUM Ingress Replication

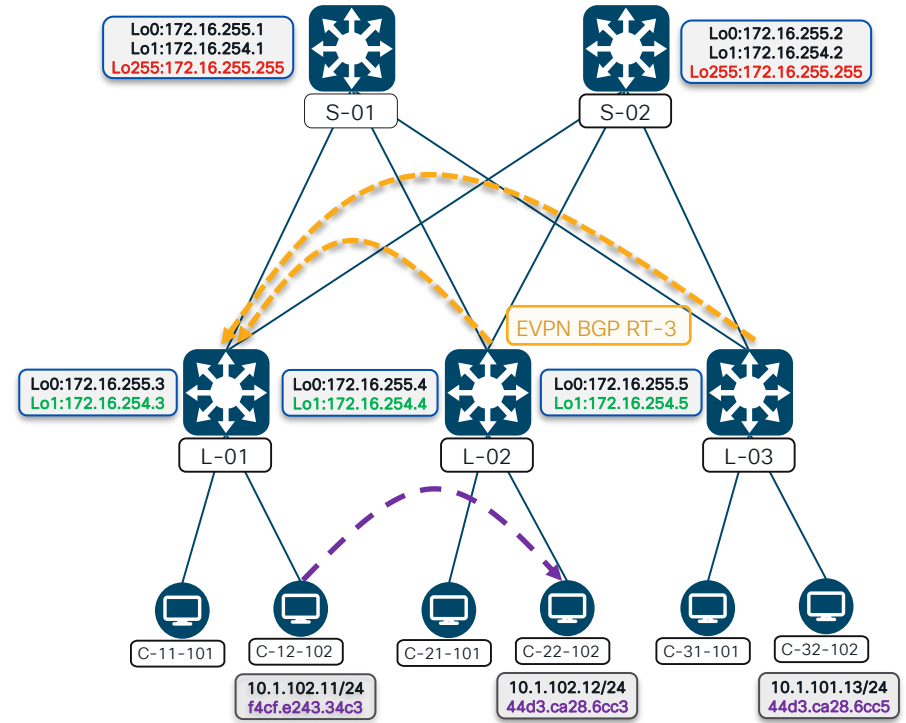


C-12-102#ping 10.1.102.12



# BUM Ingress Replication

- No Multicast in the Core
- Unicast replication in the Core
- BGP EVPN RT-3 are used for building flood list



# BGP RD/RT in EVPN

- EVPN – Ethernet Virtual Private Network
- EVI or Ethernet Virtual Instance – Ethernet VRF

```
L-01#sh l2vpn evpn summary | i Router ID
```

```
Router ID: 172.16.254.3
```

```
L-01#show l2vpn evpn evi 102 detail
```

```
EVPN instance: 102 (VLAN Based)
RD: 172.16.254.3:102 (auto)
Import-RTs: 65001:102
Export-RTs: 65001:102
<...skip...>
```

```
L-01#show run | i router bgp
```

```
router bgp 65001
```

```
L-01
```

```
l2vpn evpn
 replication-type ingress
 !
 l2vpn evpn instance 102 vlan-based
 encapsulation vxlan
 !
 vlan configuration 102
 member evpn-instance 102 vni 10102
```

- RD - <L2VPN RID>:<EVI>
- RT - <AS NUMBER>:<EVI>

# BUM Ingress Replication RT-3

```
L-01#sh bgp l2vpn evpn
```

```
<...snip...>
```

```
Route Distinguisher: 172.16.254.3:102
```

```
*> [3][172.16.254.3:102][0][32][172.16.254.3]/17
      ::
      32768 ?
*>i [3][172.16.254.3:102][0][32][172.16.254.4]/17
      172.16.254.4 0 100 0 ?
*>i [3][172.16.254.3:102][0][32][172.16.254.5]/17
      172.16.254.5 0 100 0 ?
```

RT-3 from  
L-01/02/03

```
<...snip...>
```

```
L-01#sh bgp l2vpn evpn detail [3][172.16.254.3:102][0][32][172.16.254.4]/17
```

```
BGP routing table entry for [3][172.16.254.3:102][0][32][172.16.254.4]/17, version 7
```

```
Paths: (1 available, best #1, table evi_102)
```

```
Flag: 0x100
```

```
Not advertised to any peer
```

```
Refresh Epoch 4
```

```
Local, imported path from [3][172.16.254.4:102][0][32][172.16.254.4]/17 (global)
```

```
172.16.254.4 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
```

```
Origin incomplete, metric 0, localpref 100, valid, internal, best
```

```
Extended Community: RT:65001:102, ENCAP:0 → L2 RT
```

```
Originator: 172.16.255.4, Cluster list: 172.16.255.1
```

```
PMSI Attribute: Flags:0x0, Tunnel type:IR, length 4, vni:10102 tunnel identifier: < Tunnel Endpoint: 172.16.254.4 >
```

```
rx pathid: 0, tx pathid: 0x0
```

```
Updated on Dec 14 2022 15:58:04 UTC
```

IR tunnel : 6

L2 VNI tied to EVI 102

VTEP IP

# BUM Ingress Replication

```
L-01#sh l2route evpn imet
```

EVI	ETAG	Prod	Router IP Addr	Type	Label	Tunnel ID	Multicast	Proxy
101	0	Static	225.0.0.101	0	10101	225.0.0.101		No
102	0	BGP	172.16.254.4	6	10102	172.16.254.4		No
102	0	BGP	172.16.254.5	6	10102	172.16.254.5		No
102	0	L2VPN	172.16.254.3	6	10102	172.16.254.3		No

```
L-01#sh l2route evpn imet topology 102
```

EVI	ETAG	Prod	Router IP Addr	Type	Label	Tunnel ID	Multicast	Proxy
102	0	BGP	172.16.254.4	6	10102	172.16.254.4		No
102	0	BGP	172.16.254.5	6	10102	172.16.254.5		No
102	0	L2VPN	172.16.254.3	6	10102	172.16.254.3		No

PMSI Tunnel types:

0 - No tunnel information

present

<...snip...>

6 - Ingress Replication

7 - mLDP MP2MP LSP

# BUM Ingress Replication - Known Host

```
L-02#show mon cap TAC buff detail
Starting the packet display ..... Press Ctrl + Shift + 6 to exit
```

```
Frame 11: 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface /tmp/epc_ws/wif_to_ts_pipe, id 0
```

```
[Protocols in frame: eth:ethertype:ip:udp:vxlan:eth:ethertype:arp]
```

```
Ethernet II, Src: 7c:21:0d:bd:2c:d6 (7c:21:0d:bd:2c:d6), Dst: 7c:21:0d:bd:95:56 (7c:21:0d:bd:95:56)
```

```
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.4
```

```
User Datagram Protocol, Src Port: 65357, Dst Port: 4789
```

```
Source Port: 65357
```

```
Destination Port: 4789
```

```
Length: 76
```

```
Virtual eXtensible Local Area Network
```

```
VXLAN Network Identifier (VNI): 10102
```

```
Reserved: 0
```

```
Ethernet II, Src: f4:cf:e2:43:34:c3 (f4:cf:e2:43:34:c3), Dst: 44:d3:ca:28:6c:c3 (44:d3:ca:28:6c:c3)
```

```
Address Resolution Protocol (request)
```

```
Hardware type: Ethernet (1)
```

```
Protocol type: IPv4 (0x0800)
```

```
Hardware size: 6
```

```
Protocol size: 4
```

```
Opcode: request (1)
```

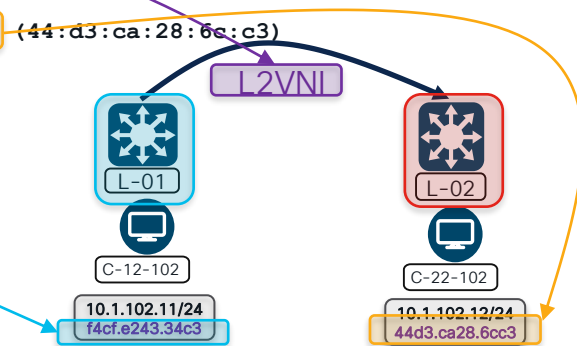
```
Sender MAC address: f4:cf:e2:43:34:c3 (f4:cf:e2:43:34:c3)
```

```
Sender IP address: 10.1.102.11
```

```
Target MAC address: 00:00:00:00:00:00 (00:00:00:00:00:00)
```

```
Target IP address: 10.1.102.12
```

Unicast



# BUM Ingress Replication - Unknown Host

```
L-02#show mon cap TAC buff detail
Starting the packet display ..... Press Ctrl + Shift + 6 to exit

Frame 1: 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface /tmp/
[Protocols in frame: eth:ethertype:ip:udp:vxlan:eth:ethertype:arp]
Ethernet II, Src: 7c:21:0d:bd:2c:d6 (7c:21:0d:bd:2c:d6), Dst: 7c:21:0d:bd:95:56 (7c:21:0d:bd:95:56)
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.4
User Datagram Protocol, Src Port: 65357, Dst Port: 4789
    Source Port: 65357
    Destination Port: 4789
    Length: 76
Virtual eXtensible Local Area Network
    VXLAN Network Identifier (VNI): 10102
    Reserved: 0
Ethernet II, Src: f4:cf:e2:43:34:c3 (f4:cf:e2:43:34:c3), Dst: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)
Address Resolution Protocol (request)
    Hardware type: Ethernet (1)
    Protocol type: IPv4 (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: request (1)
    Sender MAC address: f4:cf:e2:43:34:c3 (f4:cf:e2:43:34:c3)
    Sender IP address: 10.1.102.11
    Target MAC address: 00:00:00:00:00:00 (00:00:00:00:00:00)
    Target IP address: 10.1.102.12
```

Broadcast

VXLAN Network Identifier (VNI): 10102

L-01

C-12-102

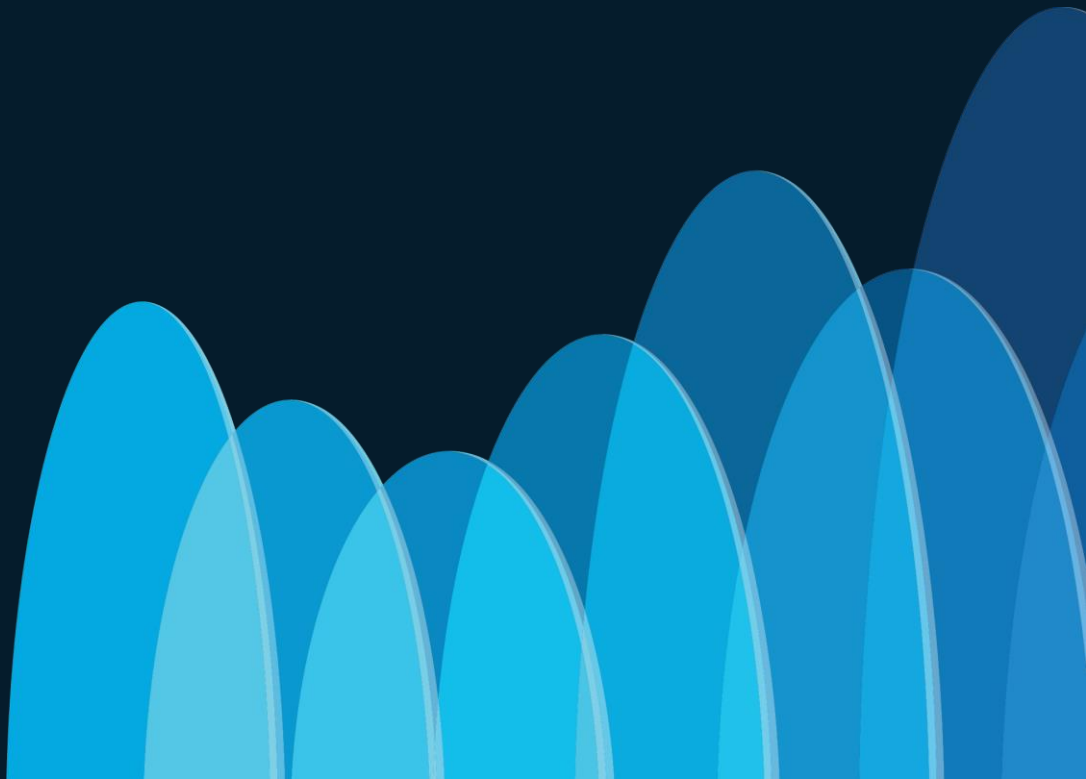
10.1.102.11/24  
f4cf.e243.34c3

L-02

C-22-102

10.1.102.12/24  
44d3.ca28.6cc3

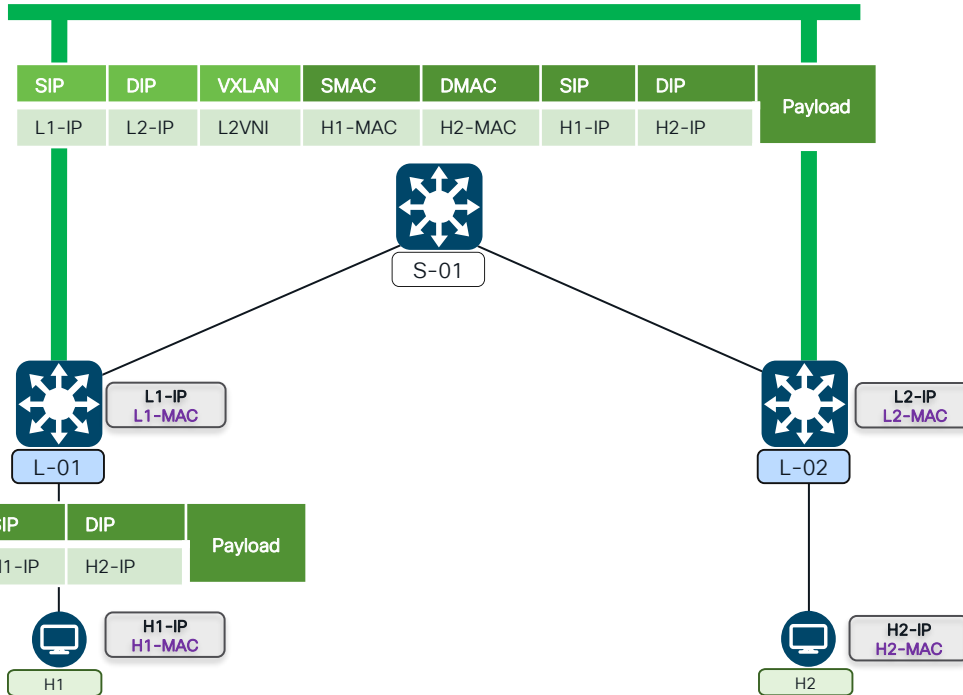
# L2 Unicast



# Bridging

Bridge

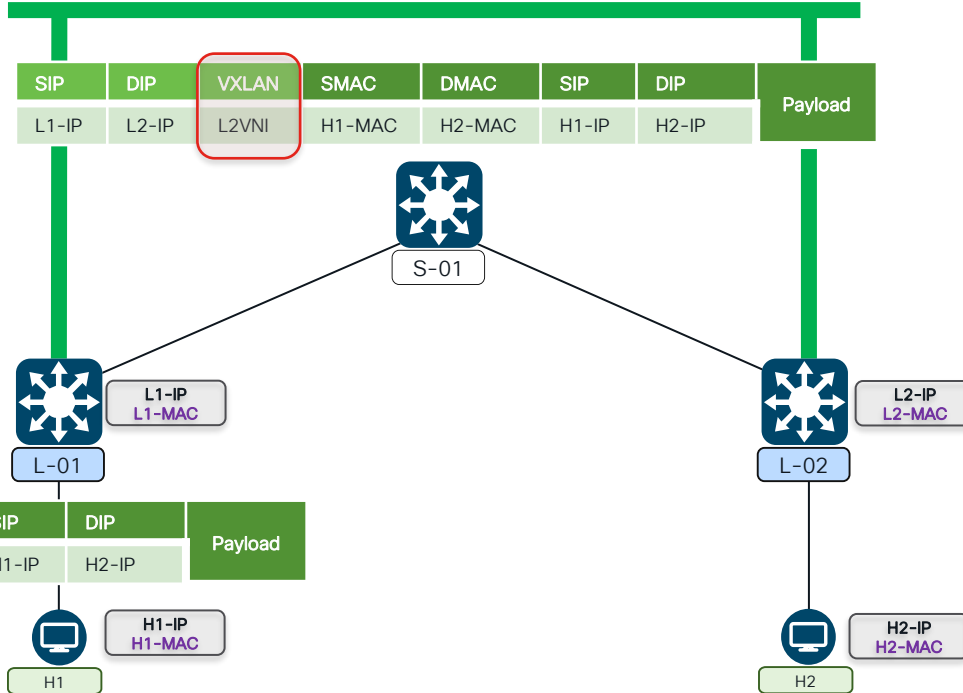
L2VNI



# Bridging

Bridge

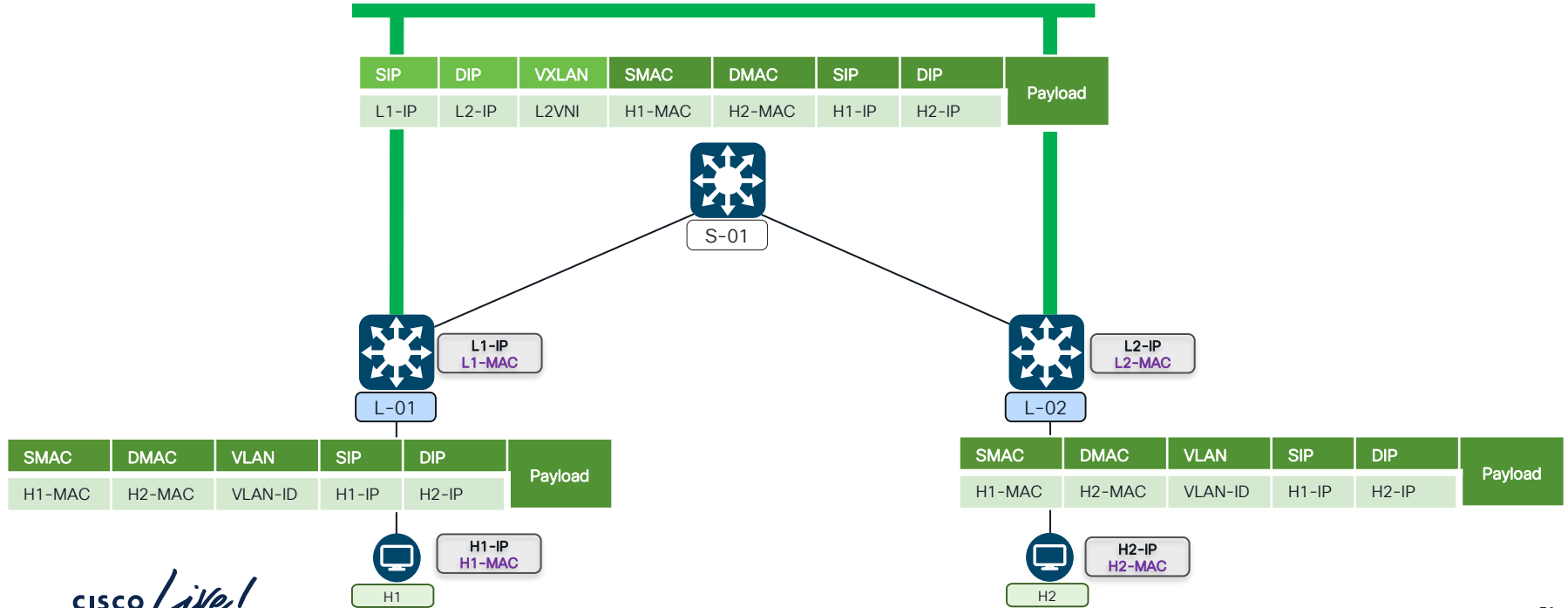
L2VNI



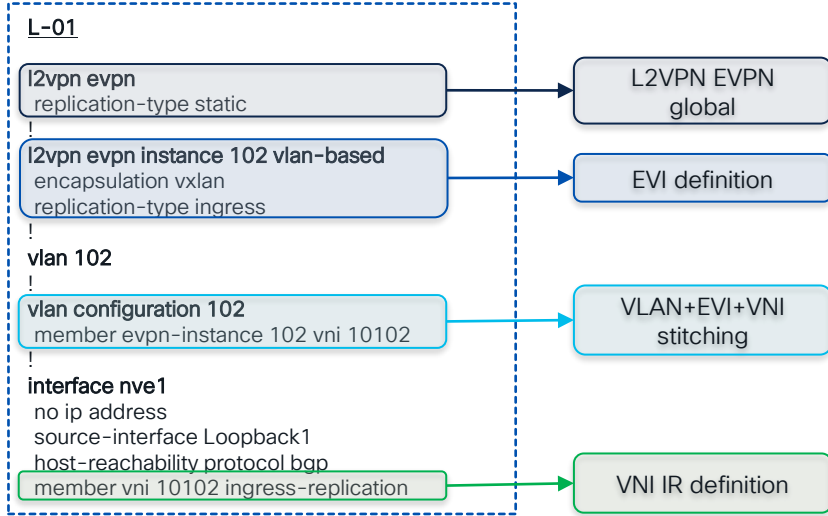
# Bridging

Bridge

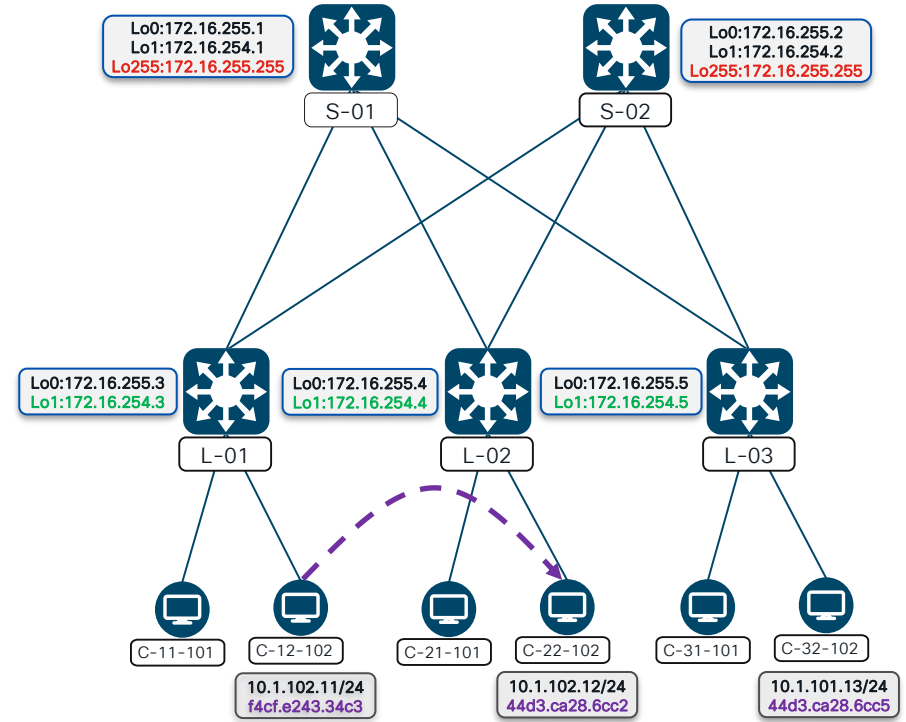
L2VNI



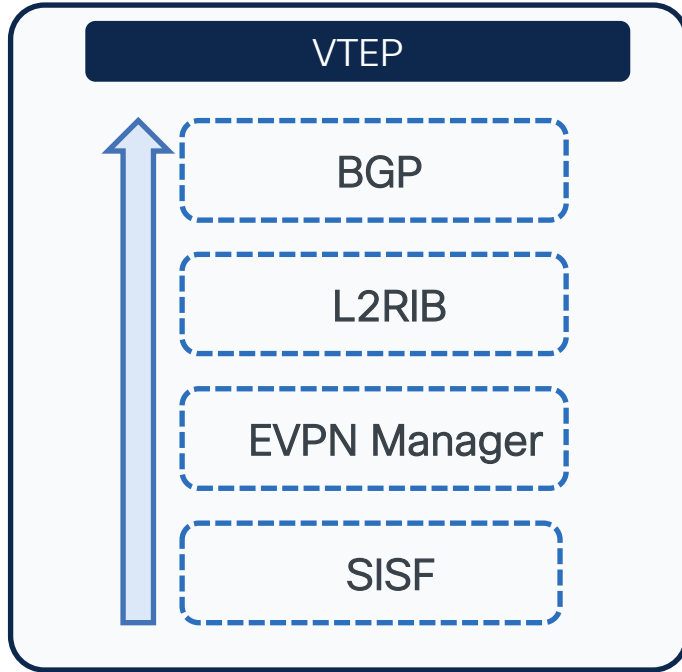
# L2 Unicast Flow



C-12-102#ping 10.1.102.12



# MAC/IP Learning process



Border Gateway Protocol

SW component L2 Routing Information Base

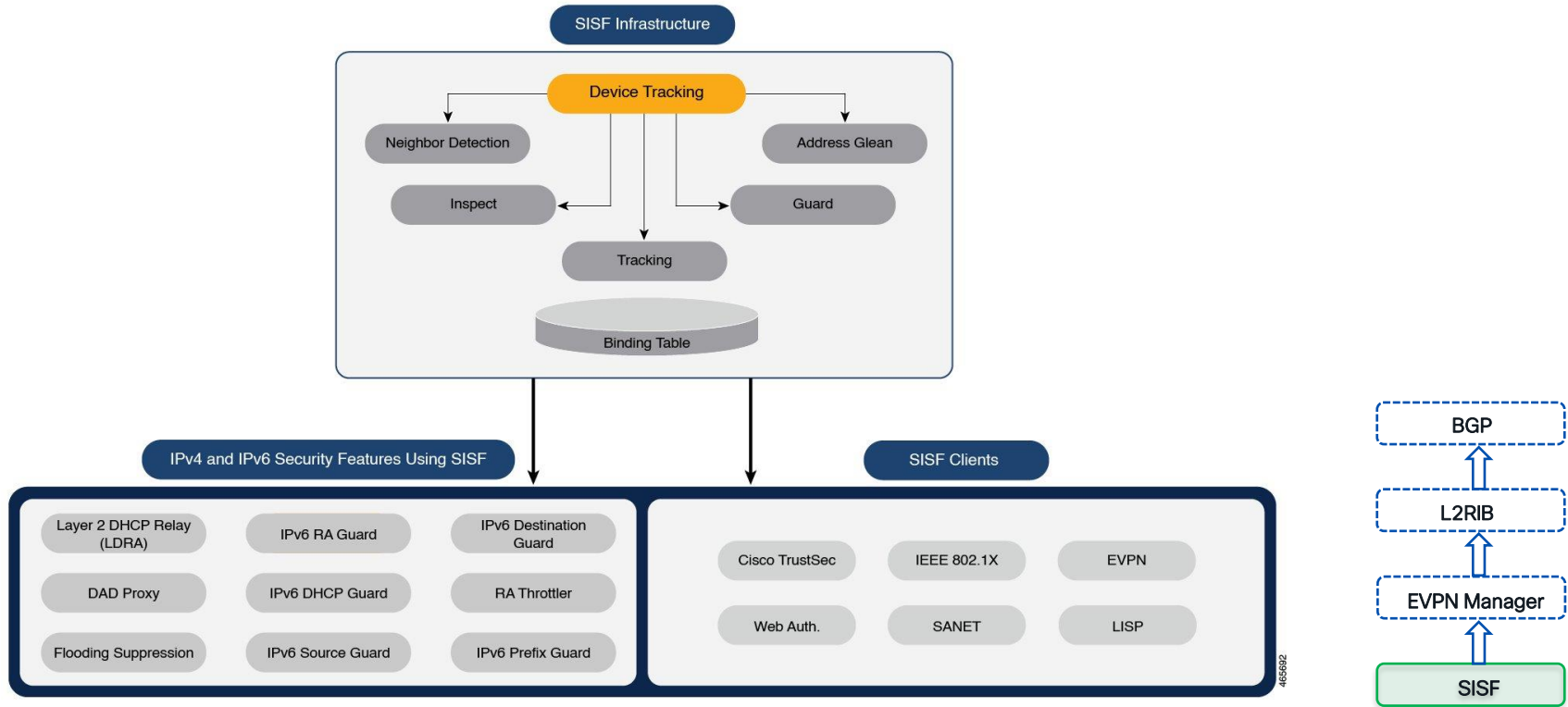
SW component EVPN Manager

Switch Integrated Security Features (SISF)



Host

# SISF (Switch Integrated Security Features)



# SISF table

```
L-02#show device-tracking database mac 44d3.ca28.6cc2 interface gi1/0/11
```

```
Codes: L - Local, S - Static, ND - Neighbor Discovery, ARP - Address Resolution Protocol, DH4 - IPv4 DHCP, DH6 - IPv6 DHCP, PKT - Other Packet, API - API created
```

```
Preflevel flags (prlvl):
```

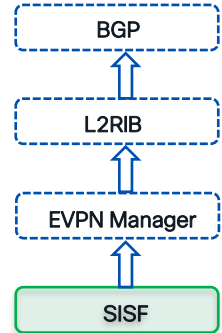
```
0001:MAC and LLA match      0002:Orig trunk          0004:Orig access
0008:Orig trusted trunk    0010:Orig trusted access 0020:DHCP assigned
0040:Cga authenticated     0080:Cert authenticated  0100:Statically assigned
```

Network Layer Address	Link Layer Address	Interface	vlan	prlvl	age	state	Time left
macDB has 1 entries for mac 44d3.ca28.6cc2,vlan 102, 1 dynamic							
ARP 10.1.102.12	44d3.ca28.6cc2	Gi1/0/11	102	0005	22s	REACHABLE	282 s try 0

```
# debug device-tracking
```

```
# debug device-tracking parser interface gi1/0/1
```

*SISF – initial point for learning*



# EVPN Manager

```
L-02#sh l2vpn evpn mac evi 102
```

MAC Address	EVI	VLAN	ESI	Ether Tag	Next Hop(s)
44d3.ca28.6cc2	102	102	0000.0000.0000.0000.0000	0	Gi1/0/11:102
44d3.ca28.6cc5	102	102	0000.0000.0000.0000.0000	0	172.16.254.5
f4cf.e243.34c3	102	102	0000.0000.0000.0000.0000	0	172.16.254.3

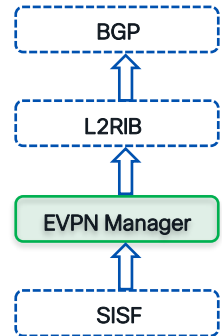
```
L-02#sh l2vpn evpn mac ip evi 102
```

IP Address	EVI	VLAN	MAC Address	Next Hop(s)
10.1.102.11	102	102	f4cf.e243.34c3	172.16.254.3
10.1.102.12	102	102	44d3.ca28.6cc2	Gi1/0/11:102
10.1.102.13	102	102	44d3.ca28.6cc5	172.16.254.5

```
# conf t
(config)# monitor event-trace evpn event include detail
(config)# monitor event-trace evpn event size 1000000
(config)# end

# show monitor event-trace evpn event all
# show monitor event-trace evpn error all

# debug l2vpn evpn event
# debug l2vpn evpn event detail
```



# L2RIB (L2 Routing Information Base)

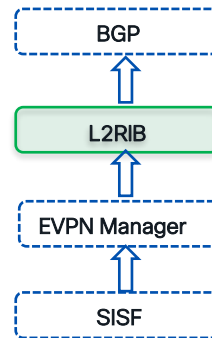
```
L-02#show l2route evpn mac topology 102
```

EVI	ETag	Prod	Mac Address	Next Hop(s)	Seq Number
102	0	BGP	10b3.d56a.8fcd	V:10102 172.16.254.3	0
102	0	L2VPN	44d3.ca28.6cc2	Gi1/0/11:102	0
102	0	BGP	44d3.ca28.6cc5	V:10102 172.16.254.5	0
102	0	BGP	7c21.0dbd.274d	V:10102 172.16.254.5	0
102	0	BGP	f4cf.e243.34c3	V:10102 172.16.254.3	0

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

# show monitor event-trace l2rib event all
# show monitor event-trace l2rib error all

# debug l2rib event
# debug l2rib event detail
```



# BGP (Border Gateway Protocol)

```
L-02#sh bgp l2vpn evpn 10.1.102.12/32
```

```
BGP routing table entry for [2][172.16.254.4:102][0][48][44D3CA286CC2][32][10.1.102.12]/24, version 15167
```

```
Paths: (1 available, best #1, table evi_102)
```

```
  Advertised to update-groups:
```

```
    1
```

```
  Refresh Epoch 1
```

```
  Local
```

```
    0.0.0.0 (via default) from 0.0.0.0 (172.16.255.4)
```

```
      Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
```

```
      EVPN ESI: 00000000000000000000, Label1 10102
```

```
      Extended Community: RT:65001:102 ENCAP:8
```

```
  Local irb vxlan vtep:
```

```
    vrf:not found, 13-vni:0
```

```
    local router mac:0000.0000.0000
```

```
    core-irb interface:(not found)
```

```
    vtep-ip:172.16.254.4
```

```
  rx pathid: 0, tx pathid: 0x0
```

```
  Updated on Jan 3 2025 10:23:30 UTC
```

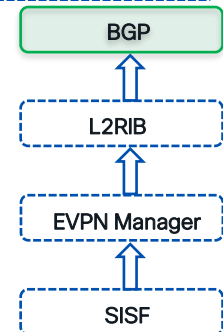
```
# debug bgp l2vpn evpn events
```

```
# debug bgp l2vpn evpn updates
```

```
# debug bgp l2vpn evpn evi context detail
```

```
# debug bgp l2vpn evpn evi event detail
```

```
# debug bgp l2vpn evpn nve detail
```



# EVPN Type-2 route (MAC/IP in detail)

```
L-02#show bgp l2vpn evpn detail [2][172.16.254.3:102][0][48][F4CFE24334C3][0][*]/20
```

```
BGP routing table entry for [2][172.16.254.3:102][0][48][F4CFE24334C3][0][*]/20, version 15173
```

```
Paths: (2 available, best #2, table EVPN-BGP-Table)
```

```
Not advertised to any peer
```

```
Refresh Epoch 44
```

```
Local
```

NH

```
172.16.254.3 (metric 3) (via default) from 172.16.255.2 (172.16.255.2)
```

```
Origin incomplete, metric 0, localpref 100, valid, internal
```

```
EVPN ESI: 00000000000000000000, Label1 10102
```

L2 VNI

```
Extended Community: RT:65001:102-ENCAP:8
```

L2 RT

```
Originator: 172.16.255.3, Cluster list: 172.16.255.2
```

```
rx pathid: 0, tx pathid: 0
```

```
Updated on Jan 2 2025 11:58:15 UTC
```

MAC address

IP address

```
L-02#sh bgp l2vpn evpn detail [2][172.16.254.3:102][0][48][F4CFE24334C3][32][10.1.102.11]/24
```

```
BGP routing table entry for [2][172.16.254.3:102][0][48][F4CFE24334C3][32][10.1.102.11]/24, version 15248
```

```
Paths: (2 available, best #1, table EVPN-BGP-Table)
```

```
Not advertised to any peer
```

```
Refresh Epoch 45
```

```
Local
```

NH

```
172.16.254.3 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
```

```
Origin incomplete, metric 0, localpref 100, valid, internal, best
```

```
EVPN ESI: 00000000000000000000, Label1 10102
```

L2 VNI

```
Extended Community: RT:65001:102-ENCAP:8
```

L2 RT

```
Originator: 172.16.255.3, Cluster list: 172.16.255.1
```

```
rx pathid: 0, tx pathid: 0x0
```

```
Updated on Jan 3 2025 12:37:28 UTC
```

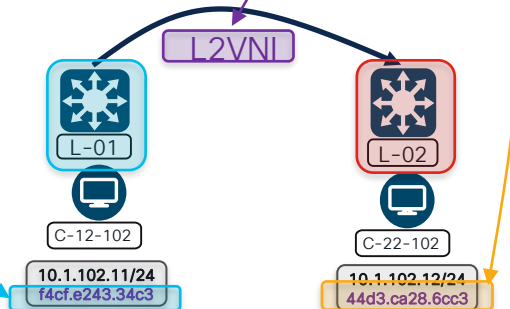
MAC address

IP address

# Unicast Packet Capture

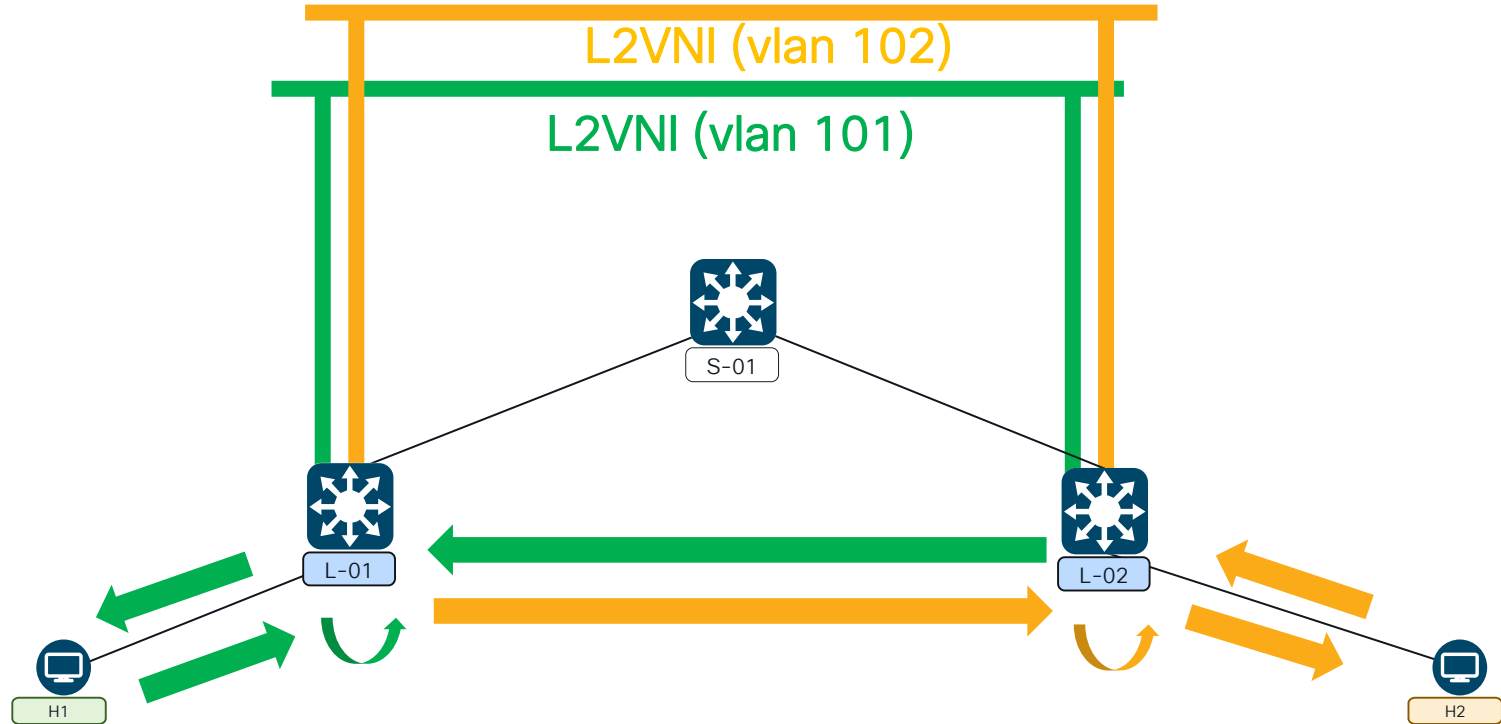
```
L-02#sh mon cap TAC buff display-filter "icmp" det
Starting the packet display ..... Press Ctrl + Shift + 6 to exit

[Protocols in frame: eth:ethertype:ip:udp:vxlan:eth:ethertype:ip:icmp:data]
Ethernet II, Src: 7c:21:0d:92:b2:d6 (7c:21:0d:92:b2:d6), Dst: 7c:21:0d:bd:95:64 (7c:21:0d:bd:95:64)
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.4
User Datagram Protocol, Src Port: 65357, Dst Port: 4789
Virtual eXtensible Local Area Network
  VXLAN Network Identifier (VNI): 10102
  Reserved: 0
Ethernet II, Src: f4:cf:e2:43:34:c3 (f4:cf:e2:43:34:c3), Dst: 44:d3:ca:28:6c:c3 (44:d3:ca:28:6c:c3)
Internet Protocol Version 4, Src: 10.1.102.11, Dst: 10.1.102.12
Internet Control Message Protocol
  Type: 8 (Echo (ping) request)
<...snip...>
```

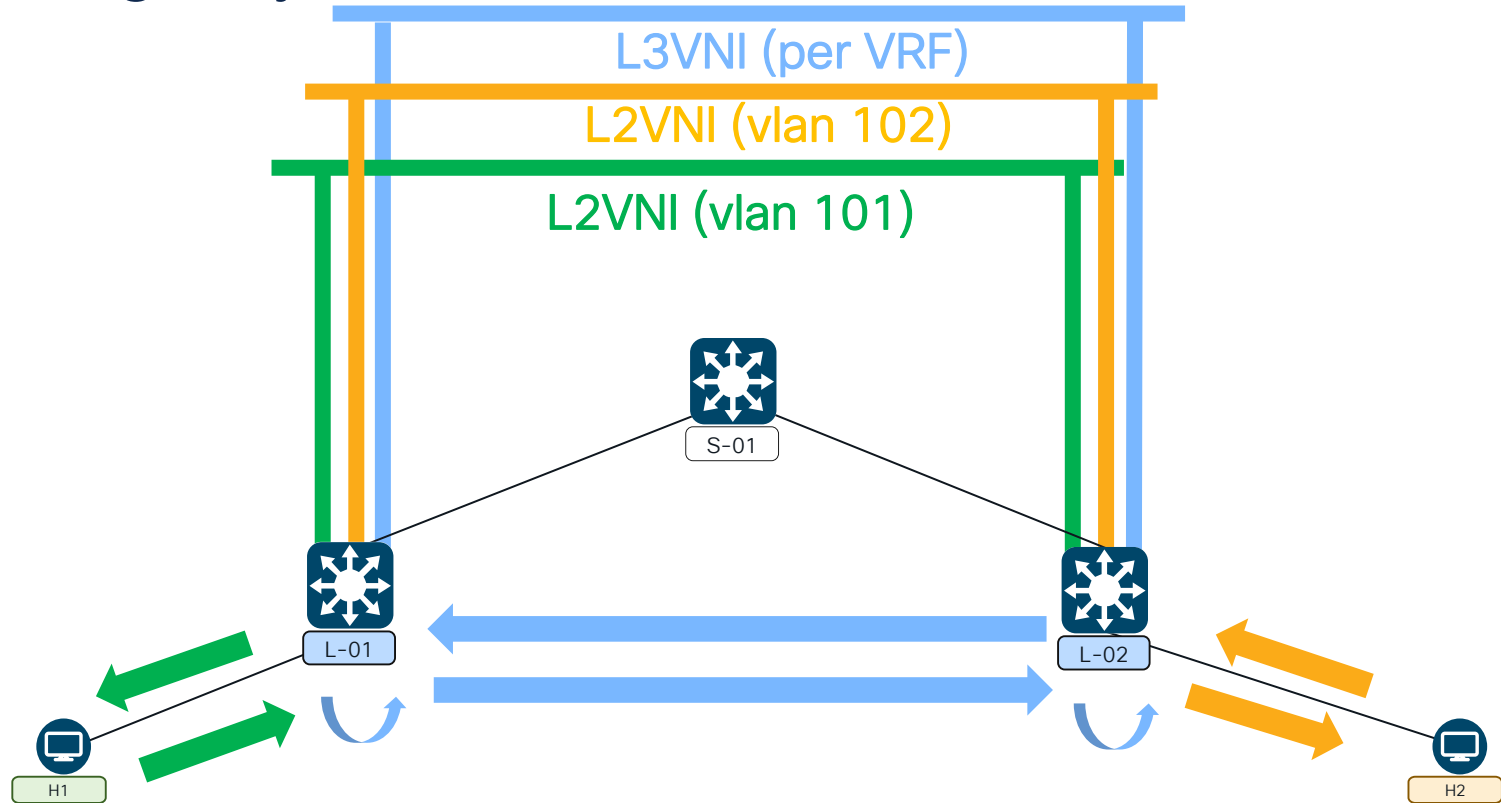


# Routing

# Routing Symmetric IRB



# Routing Asymmetric IRB



# Routing of the packet

Routing

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L2-IP	L3VNI	L1-MAC	L2-MAC	H1-IP	H2-IP	



S-01



L-01

L1-IP  
L1-MAC



L-02

L2-IP  
L2-MAC

SMAC	DMAC	VLAN	SIP	DIP	Payload
H1-MAC	GW MAC	VLAN-ID	H1-IP	H2-IP	



H1

H1-IP  
H1-MAC



H2

H2-IP  
H2-MAC

# Routing of the packet

Routing

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L2-IP	L3VNI	L1-MAC	L2-MAC	H1-IP	H2-IP	



S-01



L-01

L1-IP  
L1-MAC



L-02

L2-IP  
L2-MAC

SMAC	DMAC	VLAN	SIP	DIP	Payload
H1-MAC	GW MAC	VLAN-ID	H1-IP	H2-IP	



H1

H1-IP  
H1-MAC



H2

H2-IP  
H2-MAC

# Routing of the packet

Routing

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L2-IP	L3VNI	L1-MAC	L2-MAC	H1-IP	H2-IP	



L1-IP  
L1-MAC



L2-IP  
L2-MAC

SMAC	DMAC	VLAN	SIP	DIP	Payload
H1-MAC	GW MAC	VLAN-ID	H1-IP	H2-IP	



H1-IP  
H1-MAC

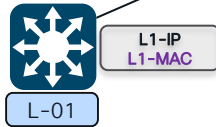


H2-IP  
H2-MAC

# Routing of the packet

Routing

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L2-IP	L3VNI	L1-MAC	L2-MAC	H1-IP	H2-IP	



SMAC	DMAC	VLAN	SIP	DIP	Payload
H1-MAC	GW MAC	VLAN-ID	H1-IP	H2-IP	



# Routing of the packet

Routing

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L2-IP	L3VNI	L1-MAC	L2-MAC	H1-IP	H2-IP	



S-01



L-01

L1-IP  
L1-MAC



L-02

L2-IP  
L2-MAC

SMAC	DMAC	VLAN	SIP	DIP	Payload
H1-MAC	GW MAC	VLAN-ID	H1-IP	H2-IP	



H1

H1-IP  
H1-MAC



H2

H2-IP  
H2-MAC

# Routing of the packet

Routing

SIP	DIP	VXLAN	SMAC	DMAC	SIP	DIP	Payload
L1-IP	L2-IP	L3VNI	L1-MAC	L2-MAC	H1-IP	H2-IP	



S-01



L-01

L1-IP  
L1-MAC



L-02

L2-IP  
L2-MAC

SMAC	DMAC	VLAN	SIP	DIP	Payload
H1-MAC	GW MAC	VLAN-ID	H1-IP	H2-IP	



H1

H1-IP  
H1-MAC

SMAC	DMAC	VLAN	SIP	DIP	Payload
GW MAC	H2-MAC	VLAN-ID	H1-IP	H2-IP	



H2

H2-IP  
H2-MAC

# L3 Unicast Flow

L-01

```
vrf definition green
description green VRF definition
rd 1:1
!
address-family ipv4
route-target export 1:1
route-target import 1:1
route-target export 1:1 stitching
route-target import 1:1 stitching
exit-address-family
!
address-family ipv6
route-target export 1:1
route-target import 1:1
route-target export 1:1 stitching
route-target import 1:1 stitching
exit-address-family
!
```

router bgp 65001

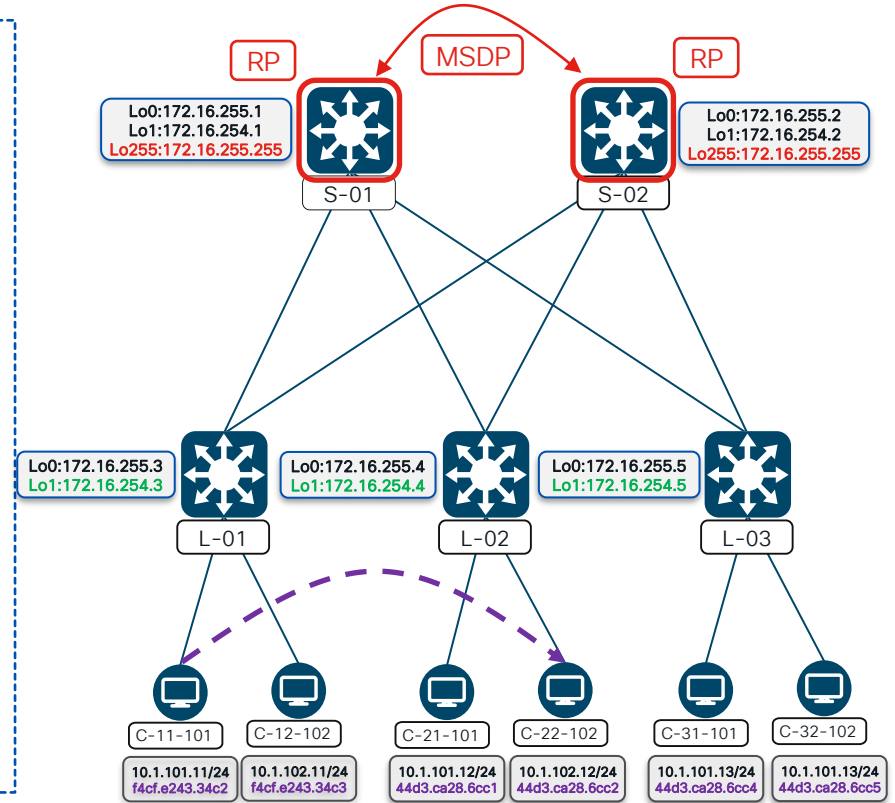
<...snip...>

```
address-family ipv4 vrf green
advertise l2vpn evpn
redistribute static
redistribute connected
exit-address-family
!
address-family ipv6 vrf green
<...snip...>
exit-address-family
<...continue...>
```

L-01

```
<...continue>
!
vlan configuration 101
member evpn-instance 101 vni 10101
vlan configuration 102
member evpn-instance 102 vni 10102
vlan configuration 901
member vni 50901
!
interface Vlan101
vrf forwarding green
ip address 10.1.101.1 255.255.255.0
!
interface Vlan102
vrf forwarding green
ip address 10.1.102.1 255.255.255.0
ipv6 address 2001:102::1/64
!
interface Vlan901
vrf forwarding green
ip unnumbered Loopback1
ipv6 enable
no autostate
!
interface nve1
no ip address
source-interface Loopback1
host-reachability protocol bgp
member vni 10101 mcast-group 225.0.0.101
member vni 10102 ingress-replication
member vni 50901 vrf green
```

C-11-101#ping 10.1.102.12



# EVPN Type-2 MAC/IP for Routing

```
L-02#sh bgp l2vpn evpn detail [2][172.16.254.3:102][0][48][F4CFE24334C3][32][10.1.102.11]/24
BGP routing table entry for [2][172.16.254.3:102][0][48][F4CFE24334C3][32][10.1.102.11]/24, version 15459
Paths: (2 available, best #2, table EVPN-BGP-Table)
  Not advertised to any peer
  Refresh Epoch 46
  Local
  [172.16.254.3] (metric 3) (via default) from 172.16.255.2 (172.16.255.2)
  Origin incomplete, metric 0, localpref 100, valid, internal
  EVPN ESI: 00000000000000000000, Label1 10102, Label2 50901
  Extended Community: RT:1:1 RT:65001:102 ENCAP:8
  Router MAC:10B3.D56A.8FC8
  Originator: 172.16.255.3, Cluster list: 172.16.255.2
  rx pathid: 0, tx pathid: 0
  Updated on Jan 3 2025 16:44:31 UTC
  Refresh Epoch 46
```

MAC address

IP address

NH

L2 VNI

L3 VNI

L2 RT

L3 RT

RMAC

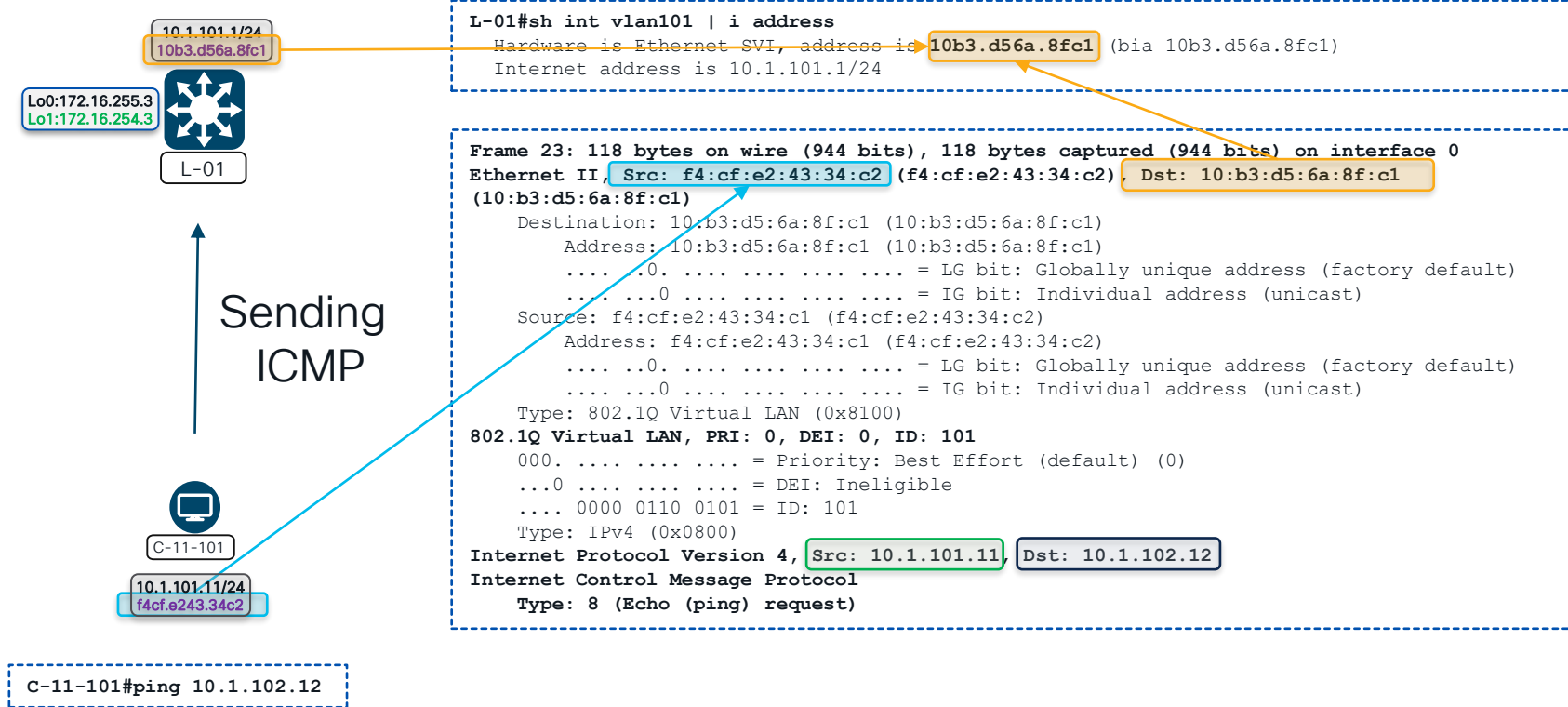
# EVPN Type-5 IP-only route

```
L-01#sh bgp l2vpn evpn detail [5][1:1][0][32][10.1.255.51]/17
BGP routing table entry for [5][1:1][0][32][10.1.255.51]/17, version 16261
Paths: (2 available, best #2, table EVPN-BGP-Table)
Not advertised to any peer
Refresh Epoch 1
Local
  172.16.254.5 (metric 3) (via default) from 172.16.255.2 (172.16.255.2)
  Origin incomplete, metric 0, localpref 100, valid, internal
  EVPN ESI: 00000000000000000000, Gateway Address: 0.0.0.0, VNI Label 50901, MPLS VPN Label 0
  Extended Community: RT:1:1 ENCAP:8 Router MAC:7C21.0DBD.2748
  Originator: 172.16.255.5, Cluster list: 172.16.255.2
  rx pathid: 0, tx pathid: 0
  Updated on Jan 3 2025 17:11:17 UTC
```

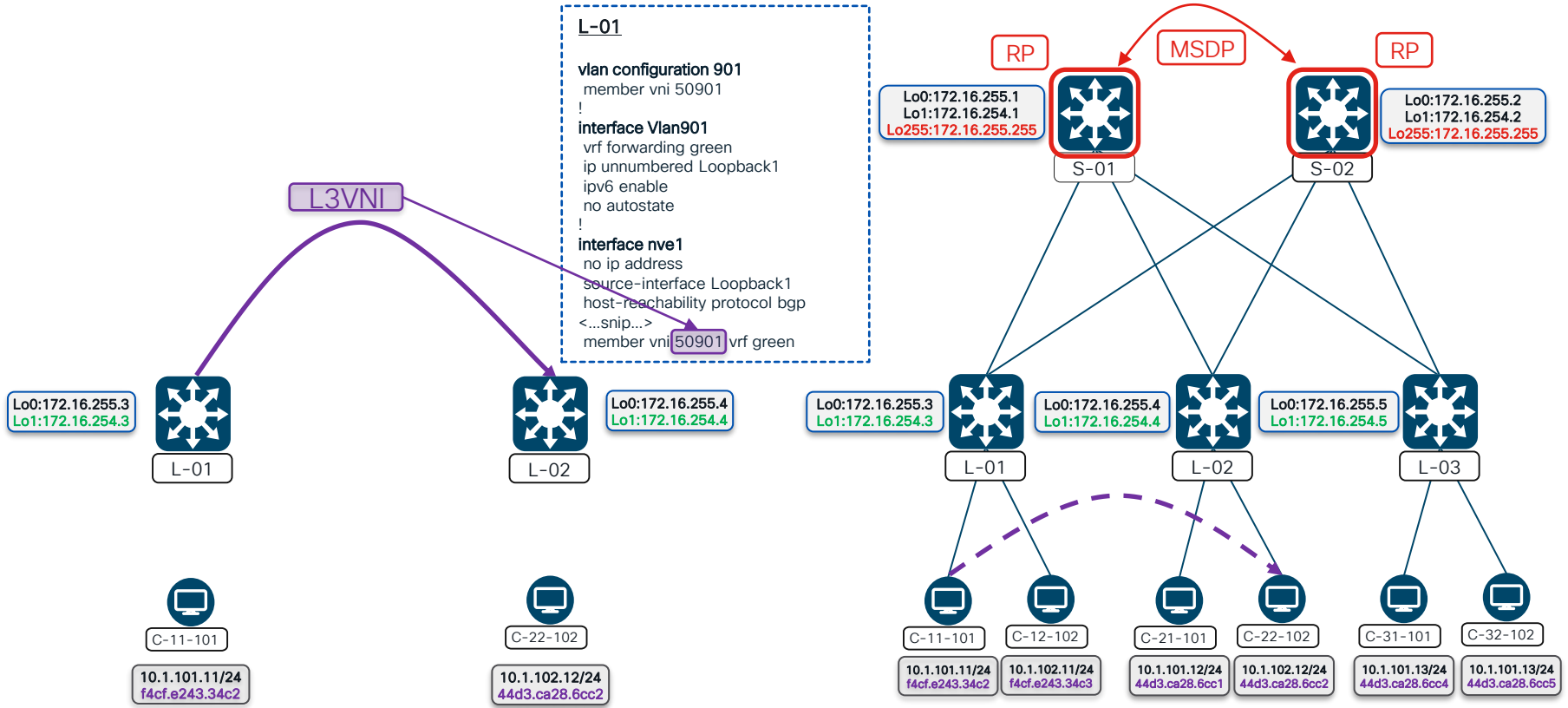
Diagram illustrating the EVPN Type-5 IP-only route configuration and its components:

- NH** (Next Hop): 172.16.254.5
- IP length** and **IP address**: [32][10.1.255.51]/17
- L3 VNI** (VNI Label): 50901
- L3 RT** (RT): 1:1
- RMAC** (Router MAC): 7C21.0DBD.2748

# Host to GW



# VTEP to VTEP



# VTEP to VTEP Packet Capture

```
Frame 9: 164 bytes on wire (1312 bits), 164 bytes captured (1312 bits) on interface 0
Ethernet II, Src: 7c:21:0d:92:b2:d6 (7c:21:0d:92:b2:d6), Dst: 7c:21:0d:bd:95:64 (7c:21:0d:bd:95:64)
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.4
User Datagram Protocol, Src Port: 65418, Dst Port: 4789
```

## Virtual eXtensible Local Area Network

```
Flags: 0x0800, VXLAN Network ID (VNI)
0... .. = GBP Extension: Not defined
... .. .0... .. = Don't Learn: False
... 1... .. = VXLAN Network ID (VNI): True
... .. .0... .. = Policy Applied: False
.000 .000 0.00 .000 = Reserved(R): 0x0000
```

Group Policy ID: 0

VXLAN Network Identifier (VNI): 50901

Reserved: 0

```
Ethernet II, Src: 10:b3:d5:6a:8f:c8 (10:b3:d5:6a:8f:c8), Dst: 7c:21:0d:bd:95:48 (7c:21:0d:bd:95:48)
```

```
Destination: 7c:21:0d:bd:95:48 (7c:21:0d:bd:95:48)
Address: 7c:21:0d:bd:95:48 (7c:21:0d:bd:95:48)
... ..0... .. = LG bit: Globally unique address (factory default)
... ..0... .. = IG bit: Individual address (unicast)
```

```
Source: 10:b3:d5:6a:8f:c8 (10:b3:d5:6a:8f:c8)
Address: 10:b3:d5:6a:8f:c8 (10:b3:d5:6a:8f:c8)
... ..0... .. = LG bit: Globally unique address (factory default)
... ..0... .. = IG bit: Individual address (unicast)
```

Type: IPv4 (0x0800)

```
Internet Protocol Version 4, Src: 10.1.101.11, Dst: 10.1.102.12
```

<...snip...>

```
Source: 10.1.101.11
Destination: 10.1.102.12
```

```
Internet Control Message Protocol
```

<...snip...>

```
L-01#sh int vlan 901 | i address
```

```
Hardware is Ethernet SVI, address is 10b3.d56a.8fc8 (bia 10b3.d56a.8fc8)
Interface is unnumbered. Using address of Loopback1 (172.16.254.3)
```

```
L-02#sh int vlan 901 | i address
```

```
Hardware is Ethernet SVI, address is 7c21.0dbd.9548 (bia 7c21.0dbd.9548)
Interface is unnumbered. Using address of Loopback1 (172.16.254.4)
```

L3VNI

Lo0:172.16.255.3  
Lo1:172.16.254.3



Lo0:172.16.255.4  
Lo1:172.16.254.4



# RIB verification

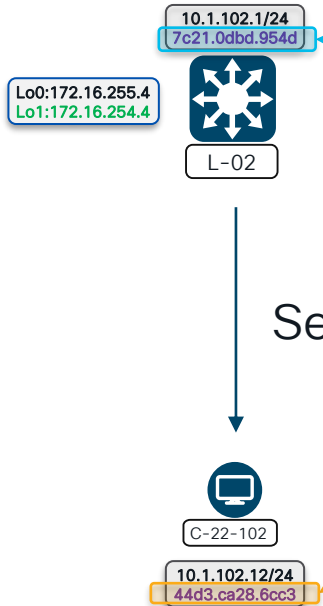
```
L-01#sh ip ro vrf green 10.1.102.12
```

```
Routing Table: green
Routing entry for 10.1.102.12/32
  Known via "bgp 65001", distance 200, metric 0, type internal
  Last update from 172.16.254.4 on Vlan901, 03:51:54 ago
  Routing Descriptor Blocks:
  * 172.16.254.4 (default), from 172.16.255.1, 03:51:54 ago, via Vlan901
    opaque_ptr 0x7F69318490D8
    Route metric is 0, traffic share count is 1
    AS Hops 0
    MPLS label: none
```

```
L-01#sh ip cef vrf green 10.1.101.12/32 internal
```

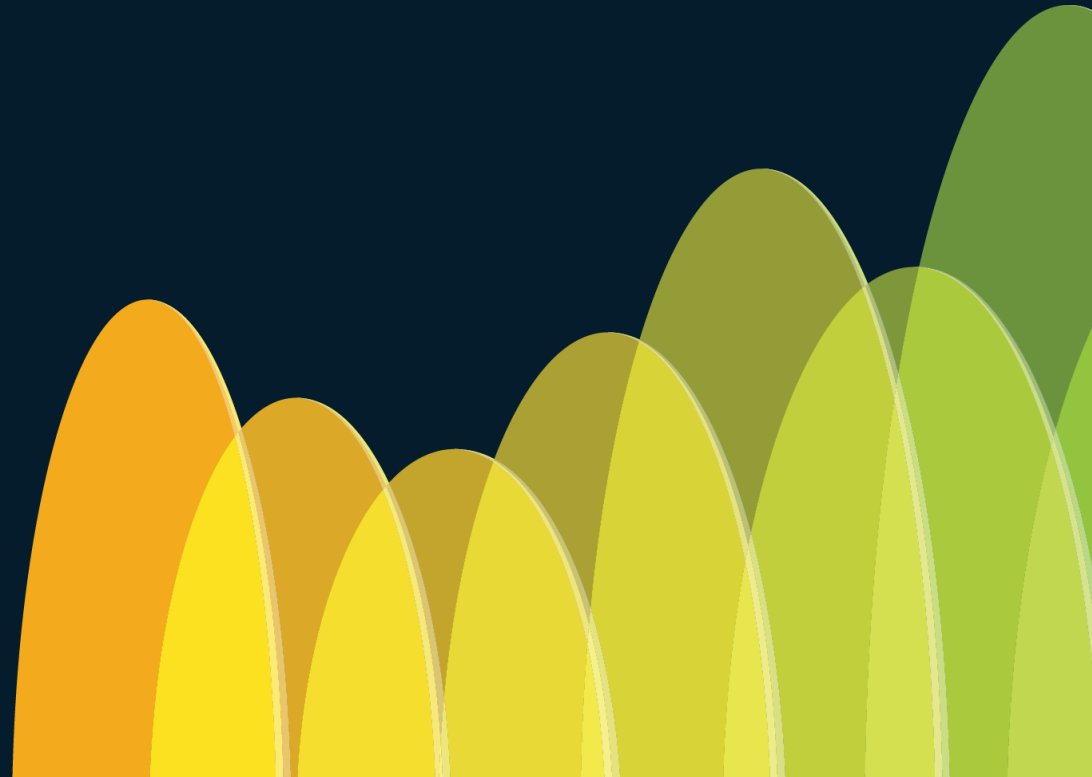
```
10.1.101.12/32, epoch 0, flags [rnolbl, rlbls], RIB[B], refcnt 6, per-destination sharing
sources: RIB
feature space:
  IPRM: 0x00018000
  Broker: linked, distributed at 3rd priority
ifnums:
  Vlan901(59): 172.16.254.4
path list 77955B26DBF0, 3 locks, per-destination, flags 0x249 [shble, rif, hwc, bgp]
  path 77955ACAA470, share 1/1, type attached nexthop, for IPv4
    nexthop 172.16.254.4 Vlan901, IP adj out of Vlan901, addr 172.16.254.4 77955B1BBD90
output chain:
  IP adj out of Vlan901, addr 172.16.254.4 77955B1BBD90
```

# VTEP to Host



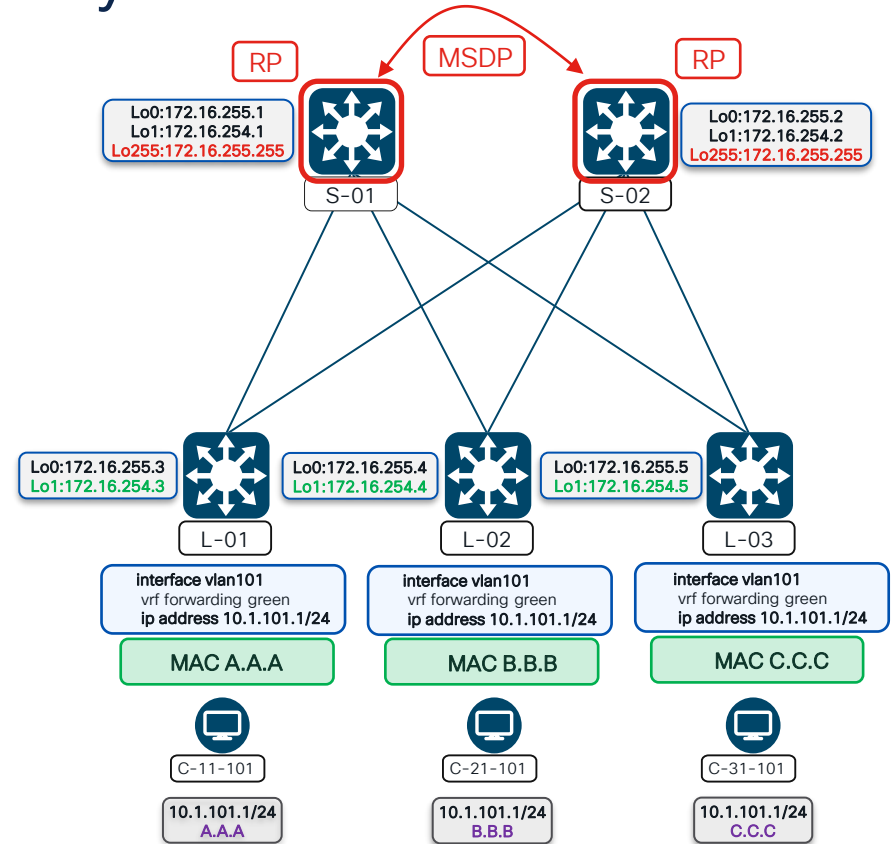
```
Frame 29: 114 bytes on wire (912 bits), 114 bytes captured (912 bits) on interface /tmp/epc_ws/wif_to_ts_pipe, id 0
[Protocols in frame: eth:ethertype:ip:icmp:data]
Ethernet II, Src: 7c:21:0d:bd:95:4d (7c:21:0d:bd:95:4d), Dst: 44:d3:ca:28:6c:c3 (44:d3:ca:28:6c:c3)
  Destination: 44:d3:ca:28:6c:c2 (44:d3:ca:28:6c:c2)
    Address: 44:d3:ca:28:6c:c2 (44:d3:ca:28:6c:c2)
      ....00. .... = LG bit: Globally unique address (factory default)
      ....00. .... = IG bit: Individual address (unicast)
  Source: 7c:21:0d:bd:95:4d (7c:21:0d:bd:95:4d)
    Address: 7c:21:0d:bd:95:4d (7c:21:0d:bd:95:4d)
      ....00. .... = LG bit: Globally unique address (factory default)
      ....00. .... = IG bit: Individual address (unicast)
  Type: IPv4 (0x0800)
Internet Protocol Version 4, Src: 10.1.101.11, Dst: 10.1.102.12
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
  Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    0000 00.. = Differentiated Services Codepoint: Default (0)
    .... 0000 = Explicit Congestion Notification: Not ECN-Capable Transport (0)
  Total Length: 100
  Identification: 0x62ef (25327)
  Flags: 0x00
    0... .... = Reserved bit: Not set
    .0... .... = Don't fragment: Not set
    ..0... .... = More fragments: Not set
  Fragment Offset: 0
  Time to Live: 252
  Protocol: ICMP (1)
  Header Checksum: 0x7c90 [validation disabled]
  [Header checksum status: Unverified]
  Source Address: 10.1.101.11
  Destination Address: 10.1.102.12
Internet Control Message Protocol
  Type: 8 (Echo (ping) request)
<...snip...>
```

# Distributed Anycast Gateway



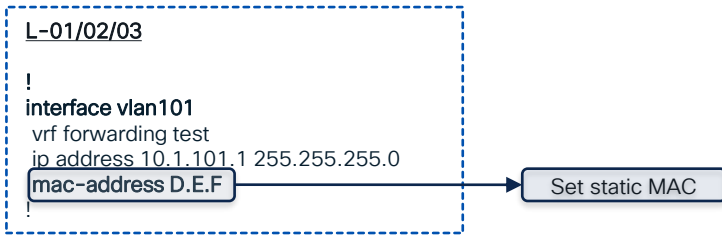
# Distributed Anycast Gateway

- Every VTEP has the same IP on SVI (like default GW)
- Every SVI by default has its own HW address
- Different GW ARP for clients connected to different VTEPs
- Deployment options:
  - Set static MAC for SVI
  - Use aliasing

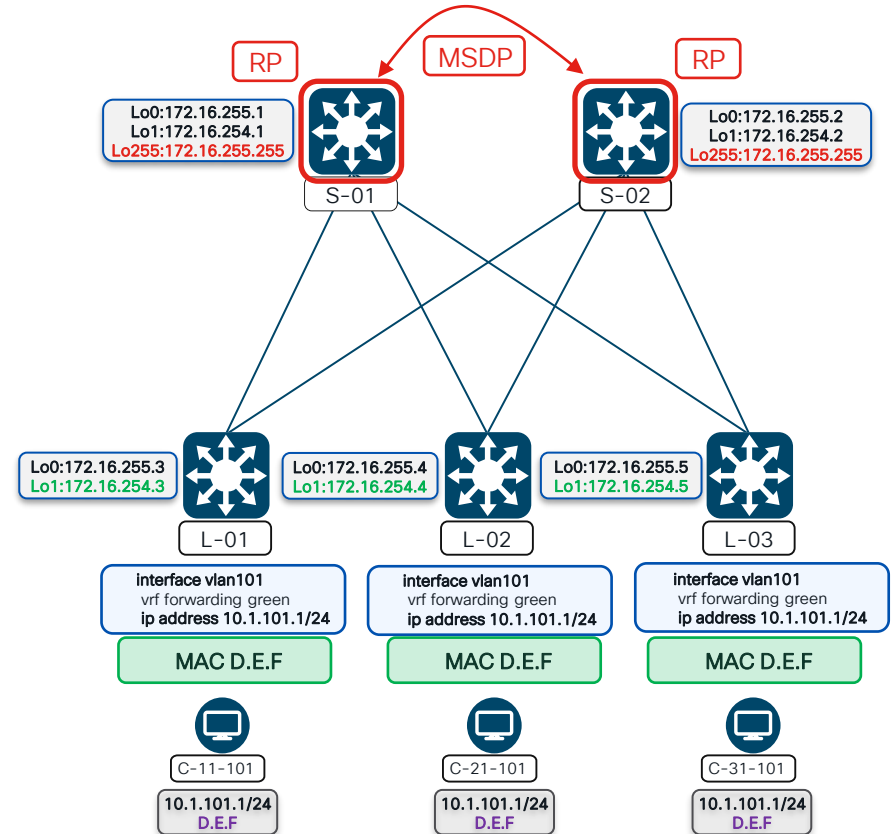


# DAG (static MAC)

- Set the same static MAC on SVI on all VTEPs

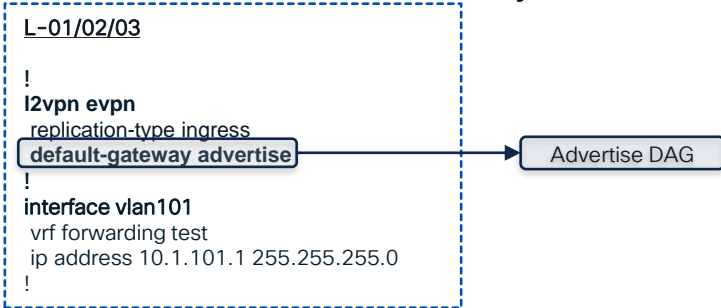


- ARP on clients for the default GW connected to the different VTEPs will be the same

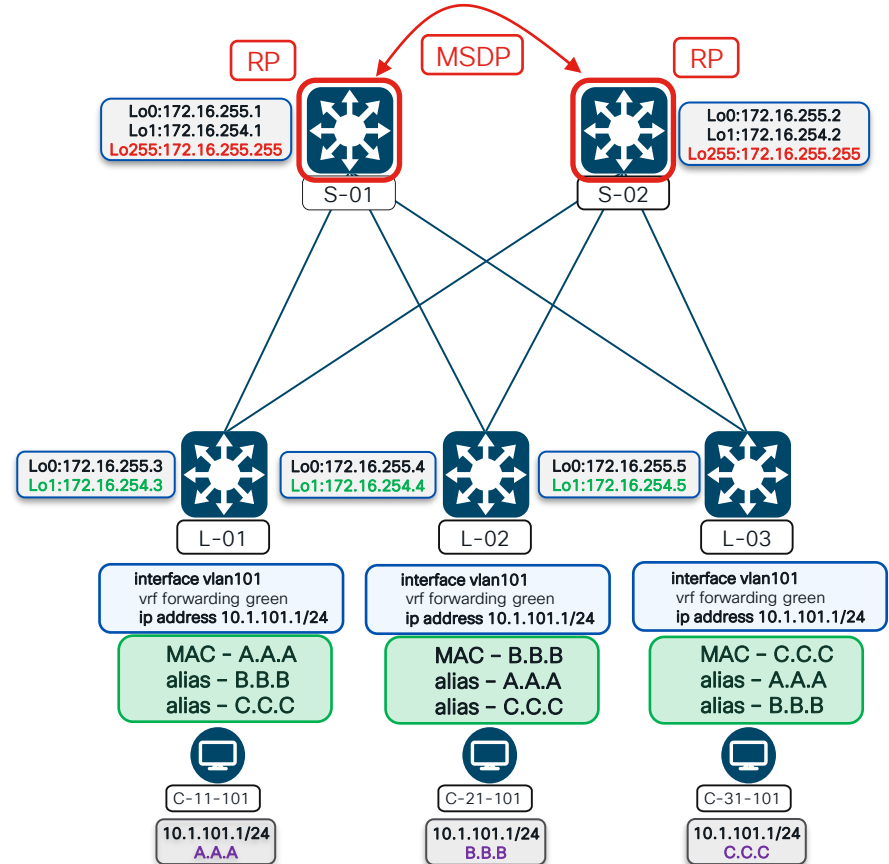


# DAG (aliasing MAC)

- Advertise BGP Type-2 route for DAG with DEF GW community



- As a result, remote MACs for DAGs will be installed in HW and traffic with DST MACs of DAGs on remote VTEPs will be processed



# DAG (aliasing MAC)

```
L-01#sh int vlan101 | i address
```

```
Hardware is Ethernet SVI, address is 10b3.d56a.8fc1 (bia 10b3.d56a.8fc1)
```

SVI MAC

```
L-01#sh bgp l2vpn evpn detail [2][172.16.254.3:101][0][48][10B3D56A8FC1][32][10.1.101.1]/24
```

```
BGP routing table entry for [2][172.16.254.3:101][0][48][10B3D56A8FC1][32][10.1.101.1]/24, version 16240
```

```
Paths: (1 available, best #1, table evi_101)
```

SVI MAC

```
Advertised to update-groups:
```

```
1
```

```
Refresh Epoch 1
```

```
Local
```

```
0.0.0.0 (via default) from 0.0.0.0 (172.16.255.3)
```

```
Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
```

```
EVPN ESI: 00000000000000000000, Label1 10101, Label2 50901
```

```
Extended Community: RT:1:1 RT:65001:101 ENCAP:8 EVPN DEF GW:0:0
```

```
Router MAC:10B3.D56A.8FC8
```

Alias  
Community

```
Local irb vxlan vtep:
```

```
vrf:green, l3-vni:50901
```

```
local router mac:10B3.D56A.8FC8
```

```
core-irb interface:Vlan901
```

```
vtep-ip:172.16.254.3
```

```
rx pathid: 0, tx pathid: 0x0
```

```
Updated on Jan 3 2025 16:42:27 UTC
```

# DAG (aliasing MAC)

```
L-01#sh bgp l2vpn evpn detail | i 10.1.101.1
```

```
BGP routing table entry for [2][172.16.254.3:101][0][48][10B3D56A8FC1][32][10.1.101.1]/24, version 195  
BGP routing table entry for [2][172.16.254.4:101][0][48][7C210DBD9541][32][10.1.101.1]/24, version 236  
BGP routing table entry for [2][172.16.254.5:101][0][48][7C210DBD2741][32][10.1.101.1]/24, version 175
```

```
L-01#sh int vlan101 | i address
```

```
Hardware is Ethernet SVI, address is 10b3.d56a.8fc1 (bia 10b3.d56a.8fc1)  
Internet address is 10.1.101.1/24
```

```
L-02#sh int vlan101 | i address
```

```
Hardware is Ethernet SVI, address is 7c21.0dbd.9541 (bia 7c21.0dbd.9541)  
Internet address is 10.1.101.1/24
```

```
L-03#sh int vlan101 | i address
```

```
Hardware is Ethernet SVI, address is 7c21.0dbd.2741 (bia 7c21.0dbd.2741)  
Internet address is 10.1.101.1/24
```

## Check in EVPN Manager:

```
L-01#sh l2vpn evpn default-gateway vlan 101
```

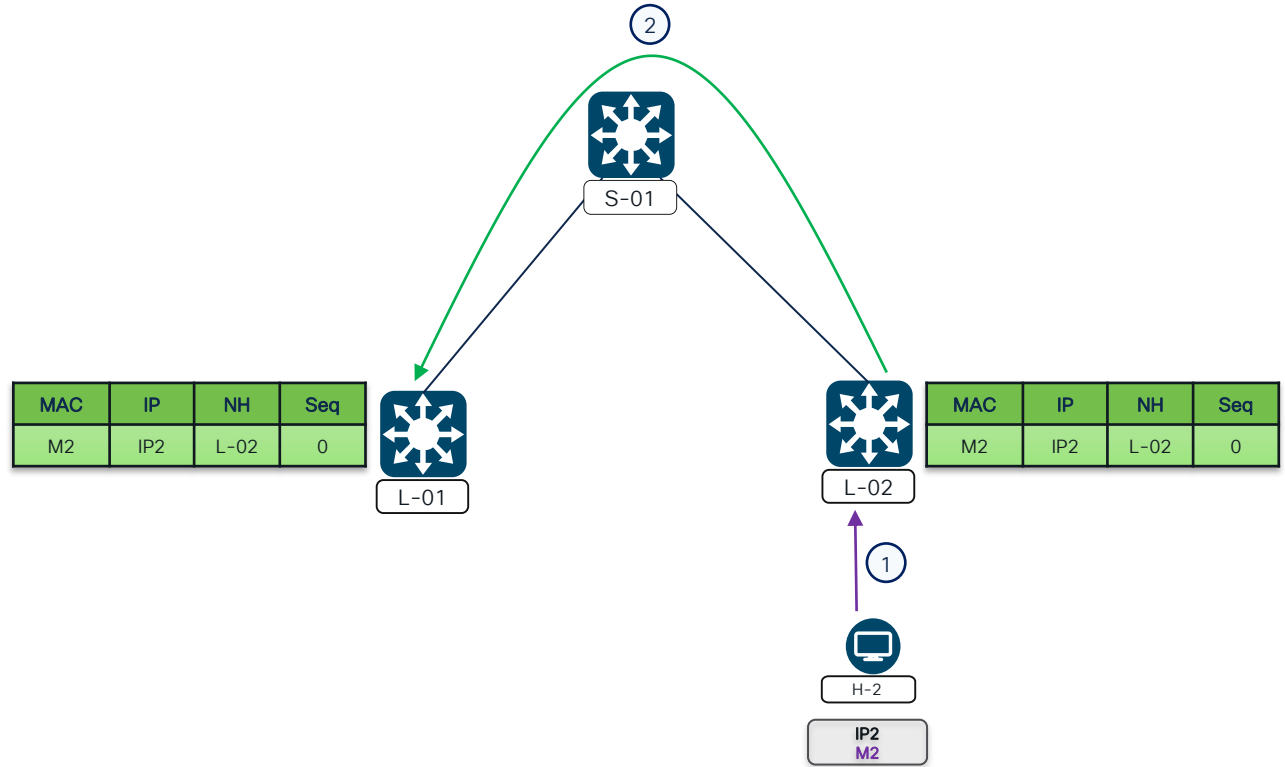
Valid	Default Gateway Address	EVI	VLAN	MAC Address	Source
Y	10.1.101.1	101	101	10b3.d56a.8fc1	Vl101
Y	10.1.101.1	101	101	7c21.0dbd.9541	172.16.254.4
Y	10.1.101.1	101	101	7c21.0dbd.2741	172.16.254.5

# Mobility and loop prevention



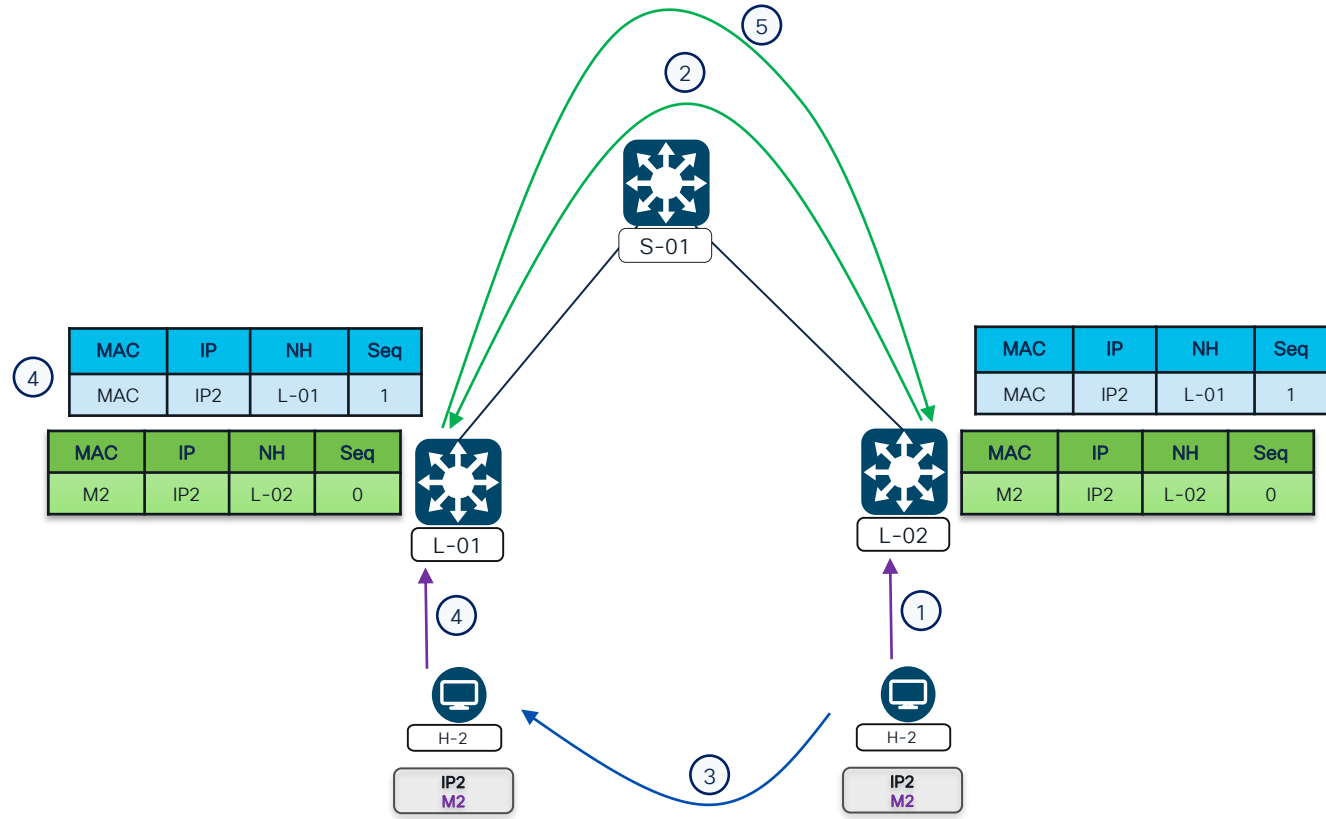
# Mobility

1. Host H2 is active and learned by Fabric
2. BGP Update is advertised



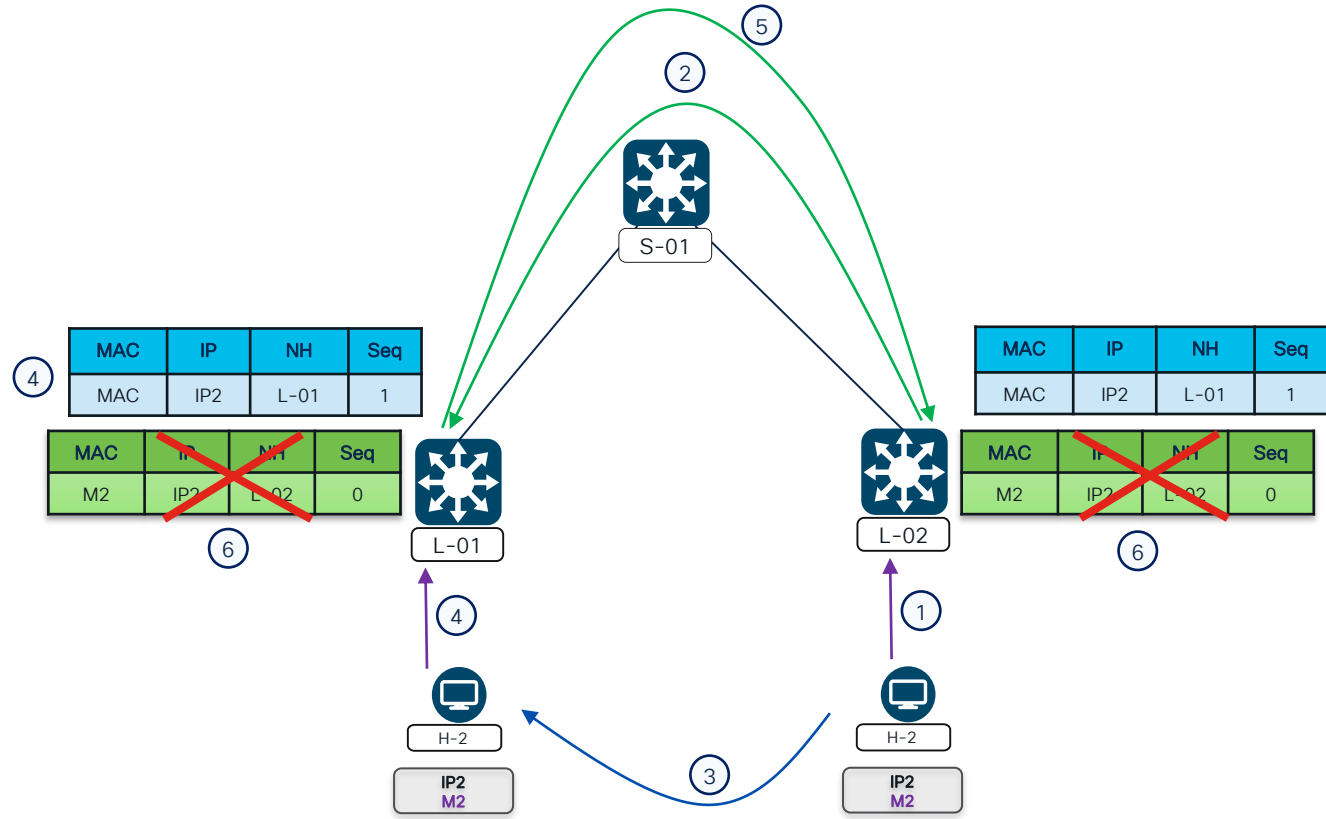
# Mobility

1. Host H2 is active and learned by Fabric
2. BGP Update is advertised
3. Host H2 moved to L-01
4. Host H2 is active and learned by L-02 w/ Seq#1
5. BGP Update is advertised



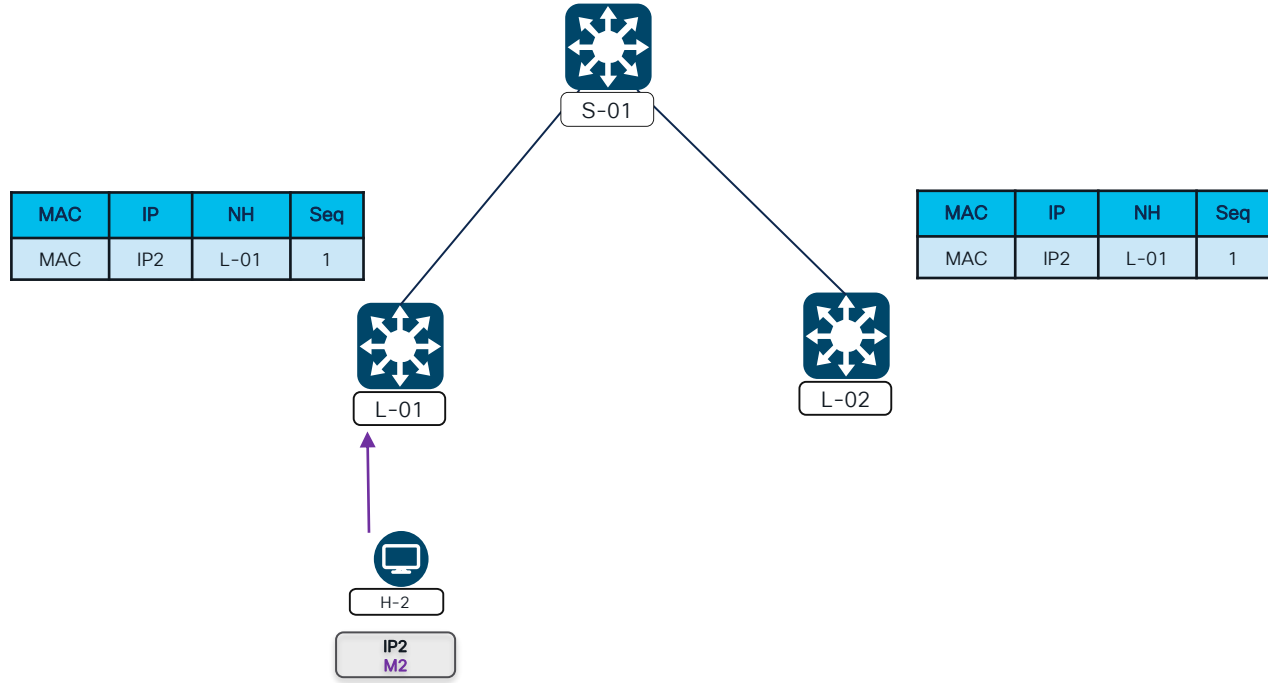
# Mobility

1. Host H2 is active and learned by Fabric
2. BGP Update is advertised
3. Host H2 moved to L-01
4. Host H2 is active and learned by L-02 w/ Seq#1
5. BGP Update is advertised
6. L-02 compare two routes and withdraw the route with lowest Seq#0



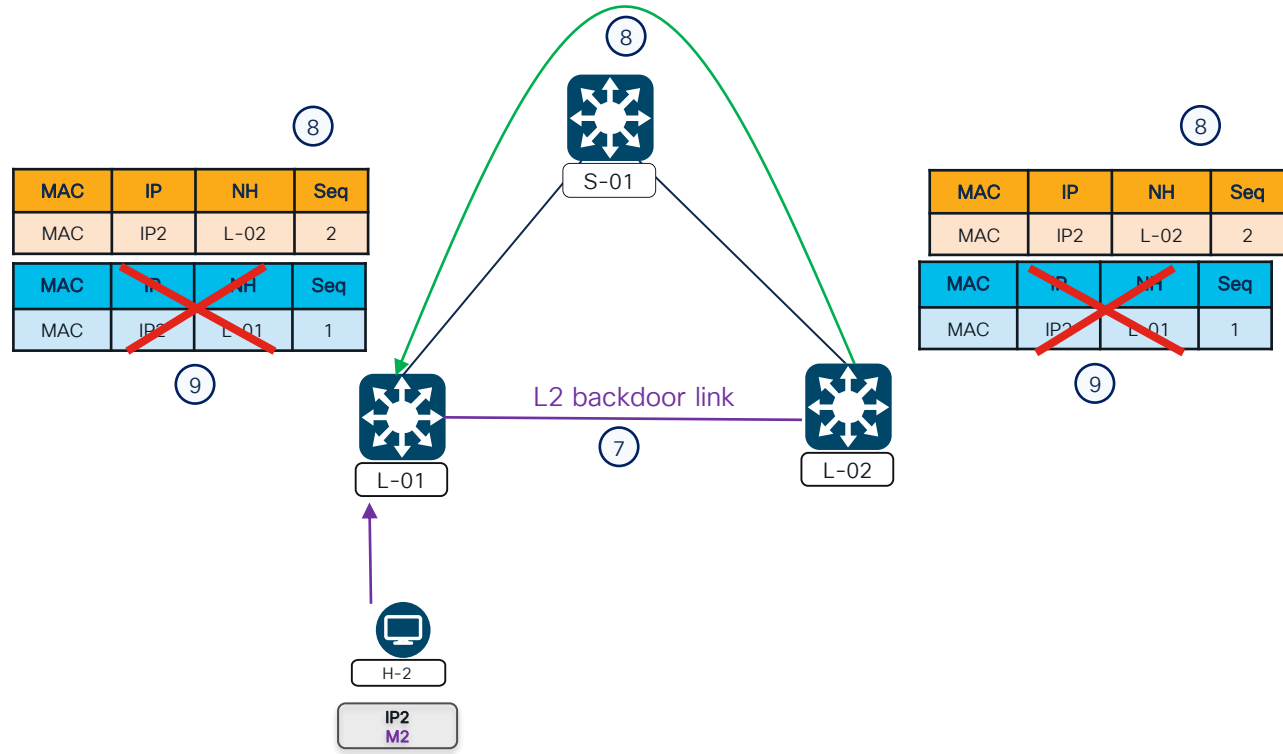
# Mobility (final stage)

1. Host H2 is active and learned by Fabric
2. BGP Update is advertised
3. Host H2 moved to L-01
4. Host H2 is active and learned by L-02 w/ Seq#1
5. BGP Update is advertised
6. L-02 compare two routes and withdraw the route with lowest Seq#0



# Loop prevention

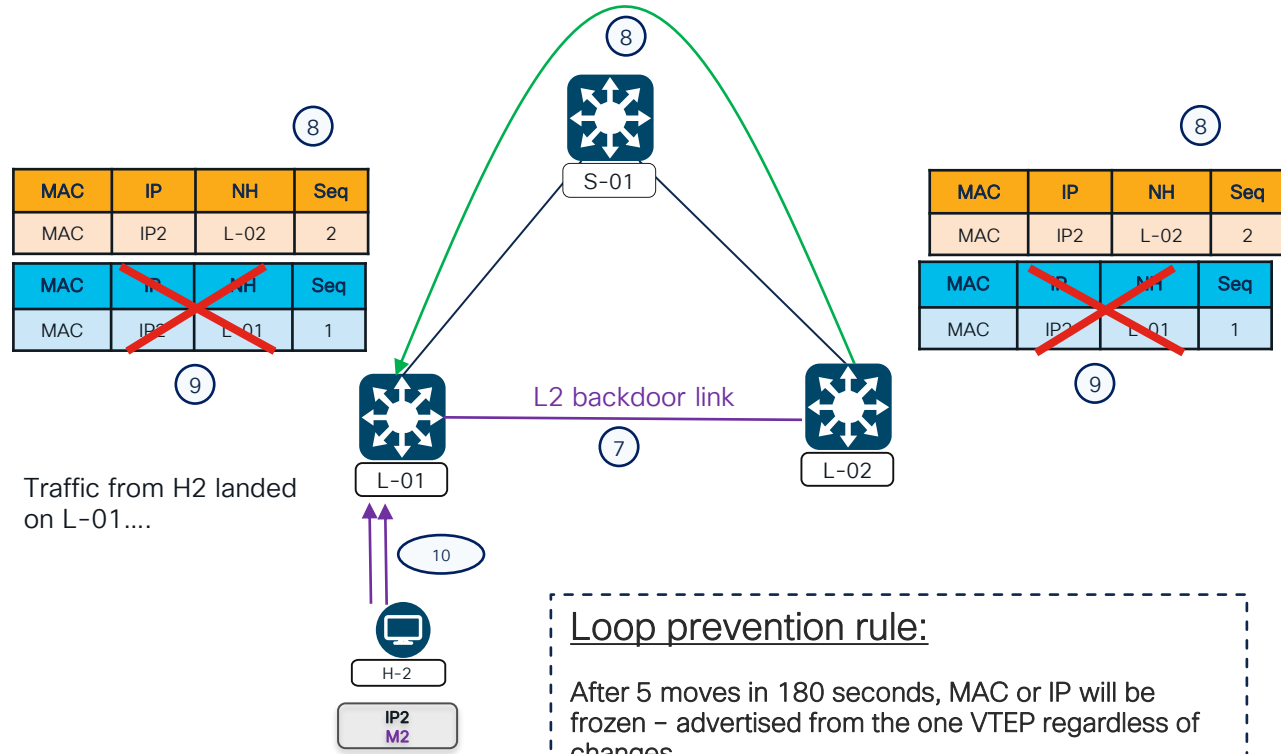
- Host H2 is active and learned by Fabric
- BGP Update is advertised
- Host H2 moved to L-01
- Host H2 is active and learned by L-02 w/ Seq#1
- BGP Update is advertised
- L-02 compare two routes and withdraw the route with lowest Seq#0
- Add L2 link b/w Leafs
- Traffic from H2 landed on L-02 and BGP route w/ Seq#2 is advertised
- L-01 compares two routes and withdraw the route with lowest Seq#1



# Loop prevention

- Host H2 is active and learned by Fabric
- BGP Update is advertised
- Host H2 moved to L-01
- Host H2 is active and learned by L-02 w/ Seq#1
- BGP Update is advertised
- L-02 compare two routes and withdraw the route with lowest Seq#0
- Add L2 link b/w Leafs
- Traffic from H2 landed on L-02 and BGP route w/ Seq#2 is advertised
- L-01 compares two routes and withdraw the route with lowest Seq#1

- Traffic from H2 landed on L-01....



Loop prevention rule:  
 After 5 moves in 180 seconds, MAC or IP will be frozen – advertised from the one VTEP regardless of changes

# Sequence Number

```
L-02#sh bgp l2vpn evpn 10.1.101.11/32
```

```
<...snip...>
```

```
BGP routing table entry for [2][172.16.254.4:101][0][48][F4CFE24334C2][32][10.1.101.11]/24, version 15628
```

```
Paths: (1 available, best #1, table evi_101)
```

```
Not advertised to any peer
```

```
Refresh Epoch 46
```

```
Local, imported path from [2][172.16.254.3:101][0][48][F4CFE24334C2][32][10.1.101.11]/24 (global)
```

```
172.16.254.3 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
```

```
Origin incomplete, metric 0, localpref 100, valid, internal, best
```

```
EVPN ESI: 00000000000000000000, Label1 10101, Label2 50901
```

```
Extended Community: RT:1:1 RT:65001:101 ENCAP:8 EVPN MAC:Seq-4
```

```
Router MAC:10B3.D56A.8FC8
```

```
Originator: 172.16.255.3, Cluster list: 172.16.255.1
```

```
rx pathid: 0, tx pathid: 0x0
```

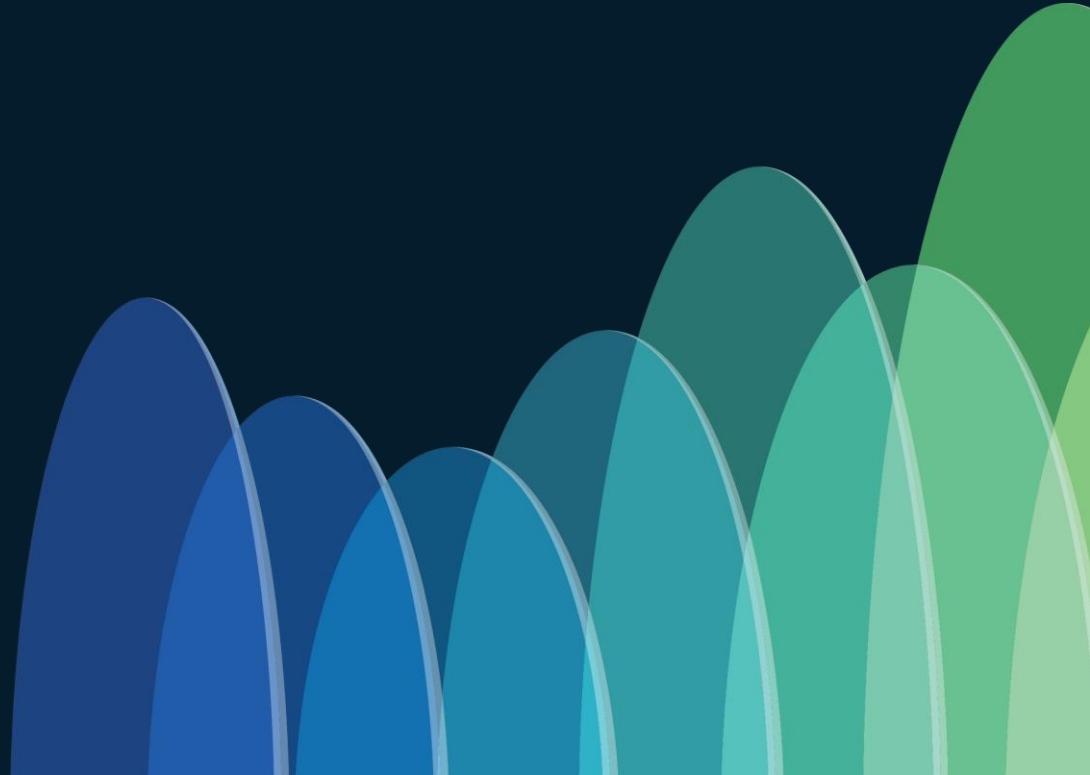
```
Updated on Jan 10 2025 20:15:42 UTC
```

EVPN MAC:Seq-4

Seq#

*If Seq# == 0, doesn't present in the output. Only non zero sequence number is printed in the output.*

# Conclusions



# Conclusions

- BUM – Broadcast, Unknown Unicast, Multicast (bridging)
  - Multicast Replication (aka static)
  - Ingress Replication (RT-3 routes)
  - ARP suppression (unicast-forwarding)
- Bridging Known Unicast (RT-2 routes)
- Routing
  - One L3VNI per VRF
  - RT-2 based (between Fabric hosts) or RT-5 based (to external network)
- DAG (static MAC and MAC aliasing)
- Mobility (Seq#)

# Webex App

## Questions?

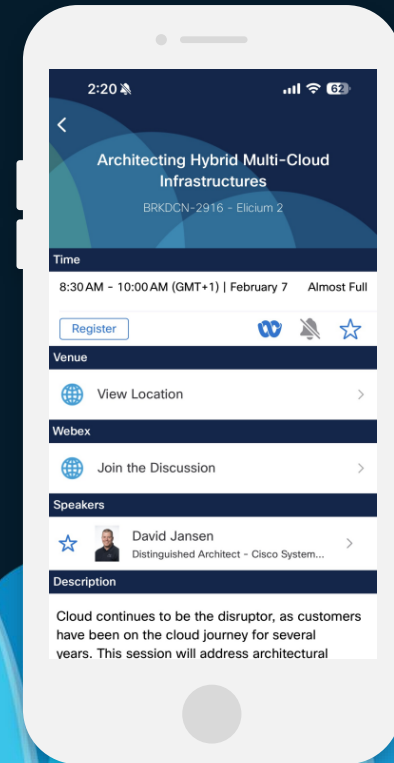
Use the Webex app to chat with the speaker after the session

## How

- 1 Find this session in the Cisco Events mobile app
- 2 Click “Join the Discussion”
- 3 Install 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 February 28, 2025.

**CISCO** *Live!*



# Fill Out Your Session Surveys



Participants who fill out a minimum of 4 session surveys and the overall event survey will get a unique Cisco Live t-shirt.

(from 11:30 on Thursday, while supplies last)



All surveys can be taken in the Cisco Events mobile app or by logging in to the Session Catalog and clicking the 'Participant Dashboard'



Content Catalog

# 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 [ciscolive.com/on-demand](https://ciscolive.com/on-demand). Sessions from this event will be available from March 3.

Contact me at: [dvishchu@cisco.com](mailto:dvishchu@cisco.com)

# Action Steps

- Check the presentation
- Go to your lab
- Repeat the steps
- If there are questions:
  - Webex space
  - TAC case



Thank you

CISCO *Live!*

CISCO *Live!*

GO BEYOND

A series of overlapping, rounded, teardrop-shaped abstract forms in various shades of blue, ranging from light to dark, positioned on the right side of the image.