





SD Access: Troubleshooting the fabric

Jon Balewicz, Technical Leader Engineering Michel Peters, Technical Leader Engineering

BRKARC-2020



Barcelona | January 27-31, 2020



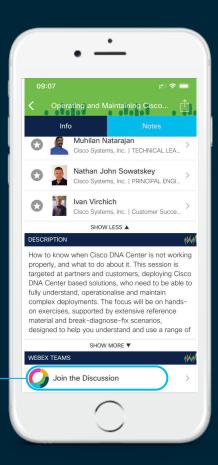
Cisco Webex Teams

Questions?

Use Cisco Webex Teams 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

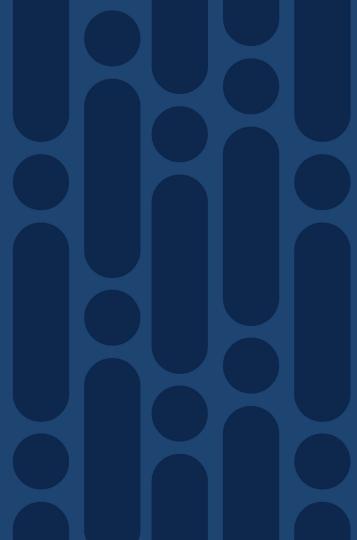


Agenda

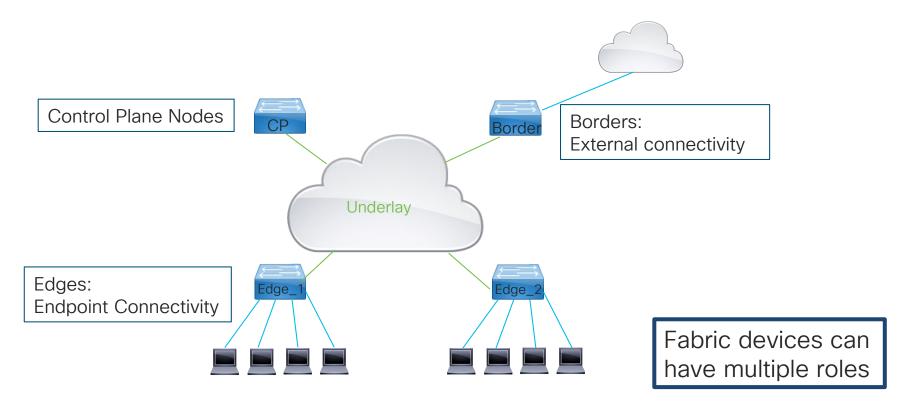
- Fabric
- Layer 3 forwarding
- Layer 2 forwarding
- Multicast Forwarding
- Security in the Fabric



Fabric

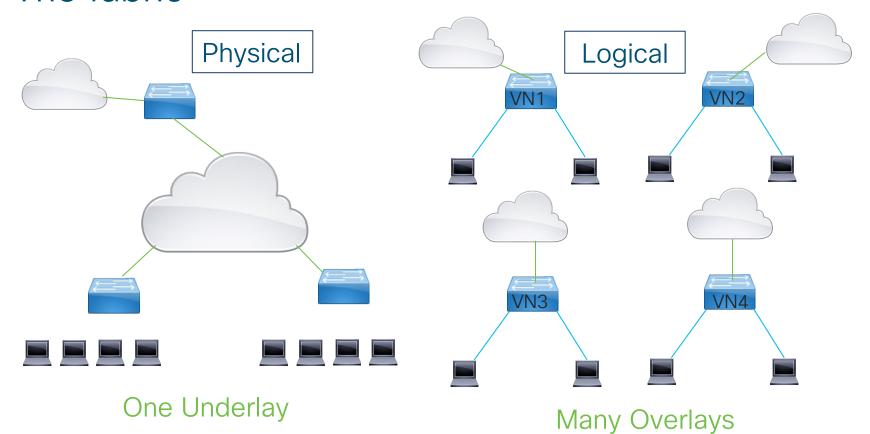


The basic fabric





The fabric



cisco Live!

SD Access Fabric Key Technologies

- Locator/ID Separation Protocol,
 Control plane protocol inside the fabric
- Cisco TrustSec,
 Assigning of Policy label to all packets and enforcing
- Authentication, Assigns endpoints using Dot1x/MAB with their respective authorization profiles and associated pools
- VXLAN,
 Used for encapsulating all Dataplane traffic trough the underlay to form
 the overlay networks



LISP Basic operation

- LISP is a routing architecture.
- LISP creates a level of indirection by using two spaces: "locators" (RLOC) and "endpoints" (EID)
- Advertise "locators" in core routing. Removes "hosts" from routing tables. Host prefixes moved to an alternative system database
- Routers in Underlay only need routing information to RLOC space, simplifies Underlay network
- To get path information to end hosts, routers query locator-end host map servers.
 Mapping analogous to DNS.
- Routers hold map-cache of locator-hosts.



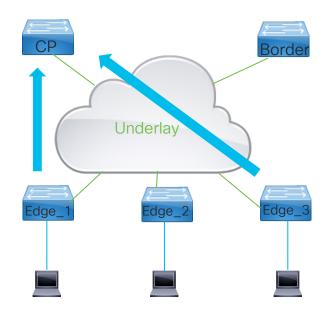
LISP Device	SD Access	Function
ETR (Egress Tunnel Router)& PETR (Proxy ETR)	Edge Device & Border node	Connects a LISP site to a LISP capable core network. Registers EID prefixes with Map Server (MS). Decapsulates LISP packets received from LISP core. PETR works on behalf of non-LISP domain and provides LISP-non-LISP connectivity.
ITR (Ingress Tunnel Router) & PITR (Proxy Ingress Tunnel Router)	Edge Device and Border node	Responsible for forwarding local traffic to external destinations. Resolves RLOC for a given destination by sending Map-request to Map Resolver. Encapsulates traffic and send to fabric. Typically, this is a Access Layer Switch. PITR works on behalf of non-LISP domain and provides LISP-non-LISP connectivity.
XTR (X Tunnel Router)	Edge Device	When both ITR and ETR functions are handled by one router, it is called XTR. This is typical in practice.
MR (Map Resolver)	Control Plane Node	Responds to Map-requests from ITR. Map-requests will be replied with a (Negative) Map-reply or forwarded to appropriate ETR
MS (Map Server)	Control Plane Node	Registers EID space upon receiving Map-register messages from ETR. Updates Map Resolver with EID and RLOC data.
MSMR (Map Server Map Resolver)	Control Plane Node	When a device acts as both Map Server and Map Resolver, it is called MS MR. This is typical in practice.
EID (Endpoint ID)	IP pools/End Points	Endpoint Identifier. IP addresses. Hidden from core network routing table. RLOC acts next-hop to reach EID space.
RLOC (Routing Locator)	Fabric Devices	Routing Locator. Exists in global routing tables. Authoritative to reach EID space.

LISP basic operation, registering with Map Server

RLOC	EID (mac address)
Edge_1	0050.5692.6d39
Edge_2	0050.5692.9735
Edge_3	70e4.22e5.c4f7

RLOC	EID (IPv4)
Edge_1	192.168.1.100
Edge_2	192.168.2.100
Edge_3	192.168.1.101

- Fabric devices learn the IPv4, IPv6 and Mac addresses of attached devices
- Fabric device register those with Map Server if they are in the defined EID Space
- Control Plane node keeps central database mapping all the EID to RLOC



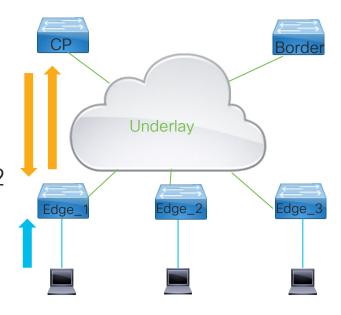


LISP basic operation, resolving

RLOC	EID (mac address)
Edge_1	0050.5692.6d39
Edge_2	0050.5692.9735
Edge_3	70e4.22e5.c4f7

RLOC	EID (IPv4)
Edge_1	192.168.1.100
Edge_2	192.168.2.100
Edge_3	192.168.1.101

- Endpoint 1 sends packet towards Endpoint 2
- Edge_1 initiates map request to CP node
- CP responds to Edge_2 with map-response containing RLOC information
- RLOC information added to map-cache to allow traffic forwarding to Endpoint 2

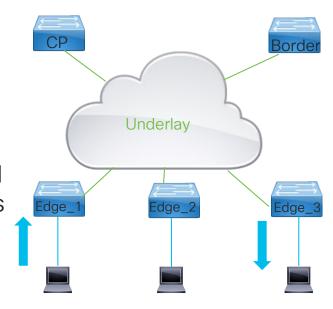


LISP basic operation, packet forwarding

RLOC	EID (mac address)
Edge_1	0050.5692.6d39
Edge_2	0050.5692.9735
Edge_3	70e4.22e5.c4f7

RLOC	EID (IPv4)
Edge_1	192.168.1.100
Edge_2	192.168.2.100
Edge_3	192.168.1.101

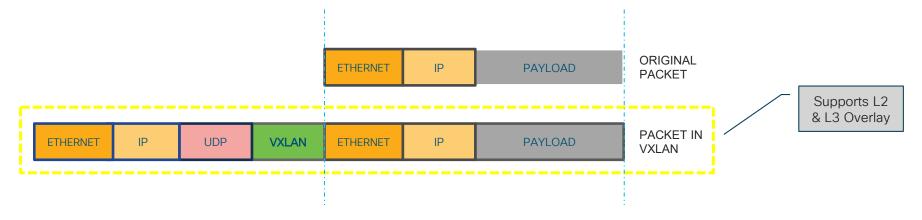
- Overlay traffic in SD Access is encapsulated in vxlan and send between RLOC addresses
- Loopback0 is typically used for RLOC
- Underlay Routing table provides reachability for RLOC's
- If reachability does not exist to RLOC traffic does not get forwarded





Data Plane

- In SD Access the entire packet is encapsulated
- VXLAN encapsulation used. Outer IP is RLOC
- VXLAN Network Identifier used for LISP instance ID
- Group Policy ID set to SGT



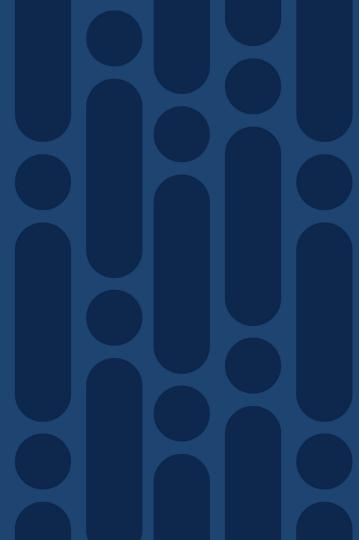


Packet Encapsulation

```
50.863252
                   192.168.0.1
                                        192,168,0,12
                                                             ICMP
                                                                       148 Echo (ping) request id=0x0b50, seq=1/256, ttl=63 (reply in 9)
    60.870066
                   192.168.0.1
                                        192.168.0.12
                                                             ICMP
                                                                       148 Echo (ping) request id=0x0b50, seq=2/512, ttl=63 (reply in 10)
     7 1 . 139082
                    10.255.1.14
                                        10.254.255.50
                                                             LISP
                                                                        82 Map-Request (RLOC-probe) for [4097] 192.168.0.1/32
                                                                        94 Map-Reply (RLOC-probe reply) for [4097] 192.168.0.1/32
     8 1 . 140831
                    10.255.1.22
                                        10.254.255.52
                                                             LTSP
    91.864089
                   192.168.0.12
                                                             ICMP
                                                                       148 Echo (ping) reply
                                                                                               id=0x0b50, seq=1/256, ttl=63 (request in 5)
                                        192.168.0.1
    101.864135
                                                             ICMP
                                                                       148 Echo (ping) reply
                                                                                               id=0x0b50, seq=2/512, ttl=63 (request in 6)
                   192.168.0.12
                                        192.168.0.1
                   192.168.0.1
                                                             ICMP
                                                                       148 Echo (ping) request id=0x0b50, seq=3/768, ttl=63 (reply in 12)
    11 1.869295
                                        192.168.0.12
    121.869346
                   192.168.0.12
                                        192.168.0.1
                                                             ICMP
                                                                       148 Echo (ping) reply
                                                                                               id=0x0b50, seq=3/768, ttl=63 (request in 11)
                                                                       148 Echo (ping) request id=0x0b50, seq=4/1024, ttl=63 (reply in 14)
   13 2.868296
                   192.168.0.1
                                        192.168.0.12
                                                             ICMP
                                                                       148 Echo (ping) reply
                                                                                               id=0x0b50, seq=4/1024, ttl=63 (request in 13)
                                                             ICMP
    14 2.868352
                   192.168.0.12
                                        192.168.0.1
Frame 5: 148 bytes on wire (1184 bits), 148 bytes captured (1184 bits) on interface 0
Ethernet II, Src: Cisco_9f:1d:40 (00:00:0c:9f:1d:40), Dst: Cisco e9:4c:7f (fc:99:47:e9:4c:7f)
Internet Protocol Version 4, Src: 10.255.1.22, Dst: 10.254.255.52
                                                                                    New Header
User Datagram Protocol, Src Port: 65359, Dst Port: 4789
Virtual eXtensible Local Area Network
                                                         SGT
 Flags: 0x8800, GBP Extension, VXLAN Network ID (VNI)
 Group Policy ID: 13
                                                                                   VXLAN Header
  VXLAN Network Identifier (VNI): 4097
                                                         LISP Instance ID
  Reserved: 0
Ethernet II, Src: Cisco 9f:00:00 (00:00:0c:9f:00:00), Dst: ba:25:cd:f4:ad:38 (ba:25:cd:f4:ad:38)
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.12
                                                                                   Encapsulated packet
Internet Control Message Protocol
```



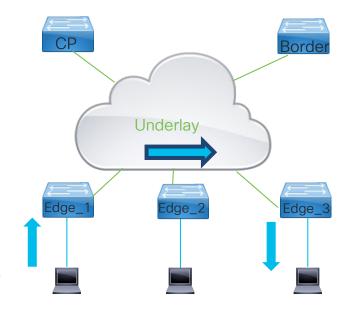
Layer 3 Forwarding



Layer 3 in the Fabric

RLOC	EID
Edge_1	192.168.1.100
Edge_2	192.168.2.100
Edge_3	192.168.1.101
Border	10.48.91.128/25

- All Edges use same IP for SVI (Anycast)
- Layer 3 LISP Instance ID's are in 4000 range
- Traffic forwarding
 - -> Outside Pool(other subnet): "Routed", Client sends to Anycast IP Mac Forwarding done based upon destination IP
 - -> Inside Pool (same subnet): "Bridged", Client sends to Mac of Endpoint Forwarding done based upon Mac Address

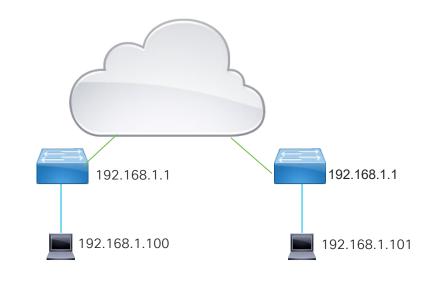




IP Anycast

- Every Edge Devices uses same Vlan, same IP address and same Mac Address
- Endpoints in the IP Pool(subnet) can be spread through the fabric.
- Default Gateway for Endpoint is set to Anycast IP

Edge 1#sh run int vlan 1024 interface Vlan1024 Edge 3#sh run int vlan 1024 mac-address 0000.0c91 interface Vlan1024 vrf forwarding CiscoLive mac-address 0000.0c9f.f45f ip address 192.168.2.1 vrf forwarding CiscoLive ip helper-address 10.48. ip address 192.168.2.1 255.255.255.0 no ip redirects ip helper-address 10.48.91.148 ip pim sparse-mode no ip redirects ip route-cache same-inte ip pim sparse-mode ip igmp version 3 ip route-cache same-interface ip igmp explicit-tracking ip igmp version 3 no lisp mobility liveness ip igmp explicit-tracking lisp mobility 192 168 2 no lisp mobility liveness test end lisp mobility 192 168 2 0-CiscoLive-IPV4 end





Locally Registered Endpoints

- VN's in SDA correlate to a VRF named as the VN
- Only endpoints belonging to EID space are added to the LISP database and registered with CP
- Interface LISP 0.<instance-id> part of VRF

```
Edge_1#sh ip vrf CiscoLive

Name Interfaces
CiscoLive Lo4100
v11022
LI0.4100
Tu2
v11024
```

```
Edge_1#sh run | s instance-id 4100
instance-id 4100
dynamic-eid 192_168_1_0-CiscoLive-IPV4
database-mapping 192.168.1.0/24 locator-set rloc_ab36f833-b546-4869-930f-578ba1cdf413
    !
dynamic-eid 192_168_2_0-CiscoLive-IPV4
database-mapping 192.168.2.0/24 locator-set rloc_ab36f833-b546-4869-930f-578ba1cdf413
    !
service ipv4
eid-table vrf CiscoLive
database-mapping 192.168.200.4/32 locator-set rloc_ab36f833-b546-4869-930f-578ba1cdf413
map-cache 0.0.0.0/0 map-request
```

Locally Registered Endpoints

```
Edge_1#sh ip arp vrf CiscoLive 192.168.1.100
Protocol Address Age (min) Hardware Addr Type Interface
Internet 192.168.1.100 3 0050.5692.6d39 ARPA Vlan1022
Edge_1#sh lisp instance-id 4100 ipv4 database 192.168.1.100/32
LISP ETR IPv4 Mapping Database for EID-table vrf CiscoLive (IID 4100), LSBs: 0x1
Entries total 2, no-route 0, inactive 0
192.168.1.100/32, dynamic-eid 192_168_1_0-CiscoLive-IPV4, inherited from default locator-set rloc_ab36f833-b546-4869-930f-578ba1cdf413
Locator Pri/Wgt Source State
172.31.255.109 10/10 cfg-intf site-self, reachable
```

- LISP Database registers only Learned Endpoints that are inside the EID Space
- Endpoints can be learned via ARP or DHCP Snooping
- Locator RLOC as advertised by Fabric Device registering the entry.
 RLOC IP address should be advertised in Underlay network as host route



Registration of Endpoints with Map Server (CP)

- IPv4/IPv6 Endpoints can be reached when learned by Edge and registered with CP
- Dynamic Endpoints learned via ARP and Device Tracking (DHCP/ARP)
- Once learned by Fabric Device it registered using LISP Reliable Transport with CP

```
Edge 1#sh lisp session
Sessions for VRF default, total: 2, established: 2
                                          Up/Down
Peer
                               State
                                                         In/Out
                                                                   Users
172.31.255.28:4342
                                          07:14:14
                                                        111/46
                               Uр
172.31.255.29:4342
                               Uρ
                                          07:14:14
                                                        111/46
Edge 1#sh lisp instance-id 4100 ipv4 statistics | sec Map-Register
 Map-Register records in/out:
                                                    0/28
   Map-Server AF disabled:
   Authentication failures:
Edge 1#sh lisp instance-id 4100 ipv4 statistics | sec Map-Requests
                                                    9/12
 Map-Requests in/out:
   Encapsulated Map-Requests in/out:
                                                    0/8
   RLOC-probe Map-Requests in/out:
                                                    9/4
   SMR-based Map-Requests in/out:
                                                    4/0
```



Control Plane Node (MSMR)

- Control Plane Node maintains table with all EID registrations
- Redundant Control Plane node do not synchronize each other.

```
CP 1#sh lisp site instance-id 4100
LISP Site Registration Information
* = Some locators are down or unreachable
# = Some registrations are sourced by reliable transport
Site Name
               Last
                         αU
                                Who Last
                                                      Inst
                                                               EID Prefix
               Register
                                Registered
                                                      ΙD
site uci
                                                      4100
                                                               0.0.0.0/0
               never
                         no
                                172.31.255.29:12616
                                                               10.48.91.128/25
               07:32:40
                         ves#
                                                      4100
                                                      4100
                                                               192.168.1.0/24
               never
                         nο
               00:03:39
                         ves#
                                172.31.255.109:13974 4100
                                                               192.168.1.100/32
                                                               192.168.1.101/32
               07:32:40
                         ves#
                                 172.31.255.111:43564 4100
                                                               192.168.2.0/24
                                                      4100
               never
                         no
               06:14:53
                         ves#
                                172.31.255.110:43692 4100
                                                               192.168.2.100/32
```



Control Plane Node (MSMR) details on EID

```
CP 1#sh lisp site 192.168.1.100/32 instance-id 4100
Requested EID-prefix:
 EID-prefix: 192.168.1.100/32 instance-id 4100
   First registered:
                        00:15:25
                                                              When registered on CP
                        00:15:25
   Last registered:
   Routing table tag:
   Origin:
                        Dynamic, more specific of 192.168.1.0/24
   Merge active:
                        No
                                                     Proxy Reply -> CP will respond
   Proxy reply:
                        Yes
   TTL:
                        1d00h
                                                     on behalf of registering
                        complete
   State:
   Registration errors:
                                                                      ETR Information
     Authentication failures:
     Allowed locators mismatch: 0
   ETR 172.31.255.109:13974, last registered 00:15:25, proxy-reply, map-notify
                            state complete, no security-capability
                            sourced by reliable transport
     Locator
              Local State
                                      Pri/Wqt Scope
                                                                   RLOC Information
     172.31.255.109 yes
                                       10/10 IPv4 none
                           up
```



Inactive clients

- Endpoints become inactive when no longer active on the network or roamed away to another fabric device
- Device Tracking sending regular ARP probes to ensure device reachability

```
Edge 1#sh lisp instance-id 4100 ipv4 database 192.168.1.100/32
LISP ETR IPv4 Mapping Database for EID-table vrf CiscoLive (IID 4100), LSBs: 0x1
Entries total 2, no-route 0, inactive 1
192.168.1.100/32, Inactive, expires: 23:58:48
Edge 1#sh lisp instance-id 4100 ipv4 away
LISP Away Table for router lisp 0 (CiscoLive) IID 4100
Prefix
                                        Producer
192.168.1.100/32
                                        local EID
Edge 1#sh lisp instance-id 4100 ipv4 smr
LISP SMR Table for router lisp 0 (CiscoLive) IID 4100
Prefix
                                        Producer
192.168.1.100/32
                                        away table
192.168.200.4/32
                                        local EID
```



Resolving Remote Destinations

Map Cache checked for Destination IP match.

→ Hit: traffic forwarded using cached information

→ No Hit: Map request is sent to the CP node(s)

- Responses from Control Plane Nodes are cached on fabric devices to build the map cache.
- Successful map-requests are cached with a TLL of 1 day
- Control plane node returns largest possible block containing requested EID when sending NMR.



Resolving Remote Destinations

```
Edge 2#sh lisp instance-id 4100 ipv4 map-cache
LISP IPv4 Mapping Cache for EID-table vrf CiscoLive (IID 4100), 8 entries
0.0.0.0/0, uptime: 1d03h, expires: never, via static-send-map-request
 Negative cache entry, action: send-map-request
8.0.0.0/7, uptime: 00:00:04, expires: 23:59:55, via map-reply, forward-native
 Encapsulating to proxy ETR
10.48.91.128/25, uptime: 00:00:16, expires: 23:59:44, via map-reply, complete
               Uptime State Pri/Wqt
 Locator
                                                Encap-IID
 172.31.255.29 00:00:16 up
                                   10/10
192.168.1.0/24, uptime: 1d03h, expires: never, via dynamic-EID, send-map-request
 Negative cache entry, action: send-map-request
192.168.1.100/32, uptime: 1d02h, expires: 03:39:23, via map-reply, complete
          Uptime State Pri/Wgt
 Locator
                                                Encap-IID
 172.31.255.109 20:20:36 up
                               10/10
192.168.2.0/24, uptime: 1d03h, expires: never, via dynamic-EID, send-map-request
 Negative cache entry, action: send-map-request
```

Map Cache shows EID range, source of cache entry and action to be taken.



LISP Remote forwarding on edge, more detail

- Routing table for VRF on edges show no Default Gateway or remote routes
- Entries in Database are inserted into routing table
- Remote entries in map-cache are not displayed or as Null routes

```
Edge 2#sh ip route vrf CiscoLive
Routing Table: CiscoLive
Gateway of last resort is not set
      192.168.2.0/24 is variably subnetted, 3 subnets, 2 masks
        192.168.2.0/24 is directly connected, Vlan1024
         192.168.2.1/32 is directly connected, Vlan1024
         192.168.2.100/32 [10/1] via 192.168.2.100, 2d18h, Vlan1024
     192.168.200.0/32 is subnetted, 1 subnets
         192.168.200.9 is directly connected, Loopback4100
CP 2#sh ip route vrf CiscoLive
Routing Table: CiscoLive
         192.168.1.0/24 [200/0], 3d20h, Null0
В
        192.168.1.1/32 is directly connected, Loopback1022
         192.168.1.100/32 [250/1], 2d14h, Null0
         192.168.1.101/32 [250/1], 2d22h, Null0
```

LISP Remote forwarding, more detail

```
Edge_2#sh ip cef vrf CiscoLive 192.168.1.100/32 detail

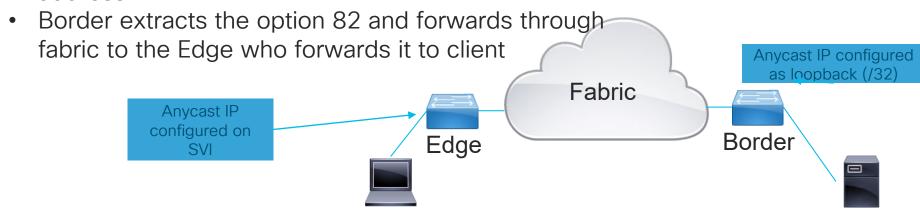
192.168.1.100/32, epoch 1, flags [subtree context, check lisp eligibility]
   SC owned, sourced: LISP remote EID - locator status bits 0x000000001
   LISP remote EID: 2 packets 1152 bytes fwd action encap, cfg as EID space, dynamic
EID need encap
   SC inherited: LISP cfg dyn-EID - LISP configured dynamic-EID
   LISP EID attributes: localEID No, c-dynEID Yes, d-dynEID No
   LISP source path list
   nexthop 172.31.255.109 LISP0.4100
2 IPL sources [no flags]
   nexthop 172.31.255.109 LISP0.4100
```

- CEF gives an accurate view of forwarding
- Next Hop egressing out of LISP interface is in Underlay network
- Using "internal" keyword provides even more detail
- Show ip cef <nexthop> gives egress interface information in underlay



DHCP in the fabric. Quick overview

- Host sends DHCP Discover
- DHCP Snooping inserts remote agent in option 82
- DHCP Relay forwards to DHCP server through fabric, setting giaddress to IP Anycast address
- DHCP Offer send by DHCP server to Anycast IP address.





Option 82 Agent Remote ID Decoding

```
AA BB CC CC CC DD EE EE EE

AA = Sub option, 03 = LISP (01 = mac address, 02 = string)

BB = length of option

CCCCCCC = LISP Instance ID

DD = Address Family IPv4 = 01 IPv6 -02

EEEEEEEE = Source locator
```

```
03 08 001002 01 c0a80106

03 Sub option lisp
08 Length of option
001002 = 4098 in decimals ->LISP Instance ID 4098
01= IPV4 locator
c0.a8.01.06 = 192.168.1.6 Source locator (Loopback 0 of xTR)
```



DHCP related debugs

- Debug ip dhcp snooping
 Enables showing detail with regards to DHCP snooping and the insertion of option 82 remote circuit
- Debug ip dhcp server
 Enables debug with regards to the relay function, insertion giaddress and relay functionality to the Server
- Debug dhcp detail
 Adds additional detail with regards to LISP in DHCP debugs



DHCP Debug - DHCP Snooping

```
Jan 27 18:23:14.889: DHCP SNOOPING: received new DHCP packet from input interface
(GigabitEthernet1/0/1)
Jan 27 18:23:14.890: DHCP SNOOPING: process new DHCP packet, message type: DHCPREQUEST,
input interface: Gi1/0/1, MAC da: ffff.ffff.ffff, MAC sa: 0050.5692.6d39, IP da:
255.255.255.255, IP sa: 0.0.0.0, DHCP ciaddr: 0.0.0.0, DHCP yiaddr: 0.0.0.0, DHCP
siaddr: 0.0.0.0, DHCP giaddr: 0.0.0.0, DHCP chaddr: 0050.5692.6d39, efp id: 0, vlan id:
1022
Jan 27 18:23:14.891: DHCP SNOOPING: add relay information option.
Jan 27 18:23:14.891: DHCP SNOOPING: Encoding opt82 CID in vlan-mod-port format
Jan 27 18:23:14.891: :VLAN case : VLAN ID 1022
Jan 27 18:23:14.891: VRF id is valid
Jan 27 18:23:14.891: LISP ID is valid, encoding RID in srloc format
Jan 27 18:23:14.892: DHCP SNOOPING: binary dump of relay info option, length: 22 data:
0x52 0x14 0x1 0x6 0x0 0x4 0x3 0xFE 0x1 0x1 0x2 0xA 0x3 0x8 0x0 0x10 0x4 0x1 0xAC 0x1F
OxFF Ox6D
Jan 27 18:23:14.893: DHCP SNOOPING: bridge packet get invalid mat entry:
FFFF.FFFF.FFFF, packet is flooded to ingress VLAN: (1022)
Jan 27 18:23:14.893: DHCP SNOOPING: bridge packet send packet to cpu port: Vlan1022.
```



DHCP Debug -DHCP Relay

DHCP Relay functionality sets GI address in DHCP packet and forwards

```
Jan 27 18:23:14.896: DHCPD: Finding a relay for client 0050.5692.6d39 on interface Vlan1022.

Jan 27 18:23:14.896: DHCPD: Locating relay for Subnet 192.168.1.1

Jan 27 18:23:14.896: DHCPD: there is no pool for 192.168.1.1.

Jan 27 18:23:14.896: DHCPD: Looking up binding using address 192.168.1.1

Jan 27 18:23:14.897: DHCPD: setting giaddr to 192.168.1.1.

Jan 27 18:23:14.897: DHCPD: BOOTREQUEST from 0050.5692.6d39 forwarded to 10.48.91.148.
```

Reply packet from DHCP server received by relay and forwarded

```
Jan 27 18:23:14.901: DHCPD: forwarding BOOTREPLY to client 0050.5692.6d39.
Jan 27 18:23:14.901: DHCPD: Option 125 not present in the msg.
Jan 27 18:23:14.902: DHCPD: src nbma addr as zero
Jan 27 18:23:14.902: DHCPD: ARP entry exists (192.168.1.100, 0050.5692.6d39).
Jan 27 18:23:14.902: DHCPD: egress Interfce Vlan1022
Jan 27 18:23:14.902: DHCPD: unicasting BOOTREPLY to client 0050.5692.6d39 (192.168.1.100).
```



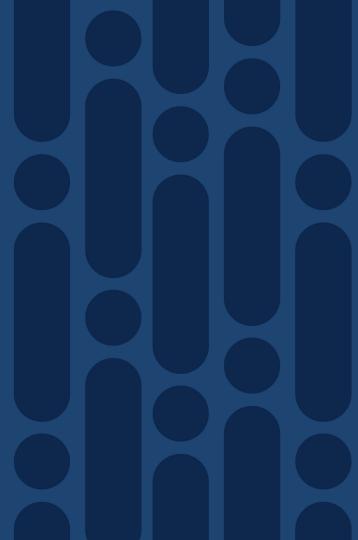
DHCP Debug -Snooping

```
Jan 27 18:23:14.903: DHCP SNOOPING: process new DHCP packet, message type: DHCPACK, input
interface: V11022, MAC da: 0050.5692.6d39, MAC sa: 0000.0c9f.f45d, IP da: 192.168.1.100, IP sa:
192.168.1.1, DHCP ciaddr: 0.0.0.0, DHCP yiaddr: 192.168.1.100, DHCP siaddr: 0.0.0.0, DHCP
giaddr: 192.168.1.1, DHCP chaddr: 0050.5692.6d39, efp id: 0, vlan id: 1022
Jan 27 18:23:14.904: DHCP SNOOPING: binary dump of option 82, length: 22 data:
0x52 0x14 0x1 0x6 0x0 0x4 0x3 0xFE 0x1 0x1 0x2 0xA 0x3 0x8 0x0 0x10 0x4 0x1 0xAC 0x1F 0xFF 0x6D
Jan 27 18:23:14.906: DHCP SNOOPING: binary dump of extracted circuit id, length: 8 data:
0x1 0x6 0x0 0x4 0x3 0xFE 0x1 0x1
Jan 27 18:23:14.907: DHCP SNOOPING: binary dump of extracted remote id, length: 12 data:
0x2 0xA 0x3 0x8 0x0 0x10 0x4 0x1 0xAC 0x1F 0xFF 0x6D
Jan 27 18:23:14.909: No entry found for mac(0050.5692.6d39) vlan(1022) GigabitEthernet1/0/1
Jan 27 18:23:14.909: host tracking not found for update add dynamic (192.168.1.100, 0.0.0.0,
0050.5692.6d39) vlan(1022)
Jan 27 18:23:14.909: DHCP SNOOPING: remove relay information option.
Jan 27 18:23:14.909: platform lookup dest vlan for input if: Vlan1022, is NOT tunnel,
if output: Vlan1022, if output->vlan id: 1022, pak->vlan id: 1022
Jan 27 18:23:14.910: DHCP SNOOPING: direct forward dhcp replyto output port:
GigabitEthernet1/0/1.
```



DHCP Snooping forwarding packet to Egress Interface

Layer 2 Forwarding



Layer 2 in the Fabric

EID (mac address)

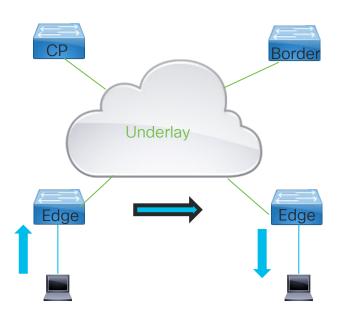
0050.5692.6d39

0050.5692.9735

70e4.22e5.c4f7

- Forwarding occurs inside an IP pool, based on Layer 2 Mac Addressing
- Complete Ethernet frame gets encapsulated in vxlan and transported through fabric
- All traffic inside an IP pool gets send via Layer 2 instances (8000 range)
- Mac Addresses are registered with CP node
- Edge Nodes resolve and cache remote mac addresses similar as done with Layer 3.
- Layer 2 Instances are associated with the Vlan corresponding to the SVI





Layer 2 Modes

- Layer 2 Extension mode on Cisco DNA Center adds Layer 2 transport through Fabric (Transports Known Unicast traffic) through Fabric
- Layer 2 Flooding mode allows flooding of selected traffic through the use of an underlay Mcast Group (broadcast-underlay) config present in config

```
Edge_3#sh run | s instance-id 8190
instance-id 8190
remote-rloc-probe on-route-change
service ethernet
eid-table vlan 1022
broadcast-underlay 239.0.0.3
database-mapping mac locator-set rloc_88
exit-service-ethernet
```

```
Edge_3#sh run int vlan 1022
interface Vlan1022
mac-address 0000.0c9f.f45d
vrf forwarding CiscoLive
ip address 192.168.1.1 255.255.255.0
ip helper-address 10.48.91.148
no lisp mobility liveness test
lisp mobility 192_168_1_0-CiscoLive-IPV4
```

Traffic inside Pool send via Layer 2, Traffic outside Layer 3 send via Layer 3



Layer 2 Mac Address Tables

- Local clients show as Dynamic or Static for Authenticated endpoints
- Remote Mac Addresses -> CP_LEARN and port Tu0
- Anycast IP with associated Mac learned on both clients
- ARP tables on clients hold mac address of remote, traffic from client to client is send to mac address of client

Edge_3	#sh mac add inc	1022 Type	
Vlan	Mac Address	Туре	Ports
1022 1022 1022 1022	0000.0c9f.f45d 58bf.eab6.4b75 70e4.22e5.c4f7 0050.5692.6d39	STATIC STATIC STATIC CP_LEARN	V11022 V11022 Gi1/0/1 Tu0

Local

```
guest@Client_3:~$ ip neig
192.168.1.100 dev eth0 lladdr 00:50:56:92:6d:39
192.168.1.1 dev eth0 lladdr 00:00:0c:9f:f4:5d
```

Remote

```
guest@Client_1:~$ ip neig
192.168.1.1 dev eth0 lladdr 00:00:0c:9f:f4:5d
192.168.1.101 dev eth0 lladdr 70:e4:22:e5:c4:f7
```



LISP Local registered mac addresses

- Layer 2 LISP use show lisp instance-id <instance> ethernet commands
- Similar to IP LISP maintains local entries in a database.
- All mac addresses part of Layer 2 EID space, all mac addresses can be learned and registered

```
Edge_3#sh lisp instance-id 8190 ethernet database
LISP ETR MAC Mapping Database for EID-table Vlan 1022 (IID 8190), LSBs: 0x1
Entries total 1, no-route 0, inactive 0

70e4.22e5.c4f7/48, dynamic-eid Auto-L2-group-8190, inherited from default locator-set rloc_88efd7b1-bb88-42d7-8a3f-68e1bfe94085
Locator Pri/Wgt Source State
172.31.255.111 10/10 cfg-intf site-self, reachable
```



Layer 2 Map-Cache

```
Edge 1#sh lisp instance-id 8190 ethernet map-cache detail
LISP MAC Mapping Cache for EID-table Vlan 1022 (IID 8190), 1 entries
70e4.22e5.c4f7/48, uptime: 04:09:04, expires: 19:50:55, via map-reply, complete
 Sources: map-reply
 State: complete, last modified: 04:09:04, map-source: 172.31.255.111
 Idle, Packets out: 0(0 bytes)
 Encapsulating dynamic-EID traffic
 Locator Uptime
                          State Pri/Wqt Encap-IID
 172.31.255.111 04:09:04 up 10/10
   Last up-down state change: 04:09:04, state change count: 1
                                    04:09:04, state change count: 1
   Last route reachability change:
   Last priority / weight change:
                                     never/never
   RLOC-probing loc-status algorithm:
     Last RLOC-probe sent:
                                     04:09:04 (rtt 3ms)
```

- Fabric Devices resolve RLOC when traffic send to unknown Destination mac-addresses using map-request.
- Similar to Layer 3 a map-cache is build for Layer 2 entries with result



Control Plane Node

- All Mac Addresses registered in Fabric on CP node as EID Prefix
- Show lisp instance-id <id> ethernet server uses Layer 2 instance-id or *

```
CP 1#sh lisp instance-id 8190 ethernet server
LISP Site Registration Information
* = Some locators are down or unreachable
# = Some registrations are sourced by reliable transport
Site Name
                                Who Last
                                                               EID Prefix
               Last
                         Uр
                                                      Inst
               Register
                                Registered
                                                      ΙD
                                                      8190
site uci
               never
                         nο
                                                               any-mac
                                                               0000.0c9f.f45d/48
               3d04h
                         ves#
                                172.31.255.19:2470
                                                      8190
               2d22h
                                                               0050.5692.6d39/48
                         ves#
                                172.31.255.109:13974 8190
               03:36:25
                         ves#
                                172.31.255.111:43564 8190
                                                               70e4.22e5.c4f7/48
               3d04h
                         ves#
                                172.31.255.19:2470
                                                      8190
                                                               fc99.47e9.4c7f/48
```



CP Node, Ethernet EID more detailed information

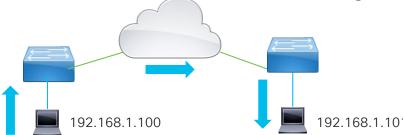
```
CP 1#sh lisp instance-id 8190 ethernet server 0050.5692.6d39
Requested EID-prefix:
 EID-prefix: 0050.5692.6d39/48 instance-id 8190
   First registered:
                         2d22h
                                                        Registration info
   Last registered:
                         2d22h
   Routing table tag:
   Origin:
                         Dynamic, more specific of any-mac
   Merge active:
                         No
   Proxy reply:
                         Yes
                                           CP responds to map-reply
   TTL:
                         1d00h
   State:
                         complete
   Registration errors:
     Authentication failures:
     Allowed locators mismatch: 0
   ETR 172.31.255.109:13974, last registered 2d22h, proxy-reply, map-notify
                             TTL 1d00h, sourced by reliable transport
     Locator
               Local State
                                       Pri/Wgt Scope
                                                               RLOC info
     172.31.255.109 yes
                                        10/10
                                                TPv4 none
                            up
```

Control Plane node detailed information on registered mac address



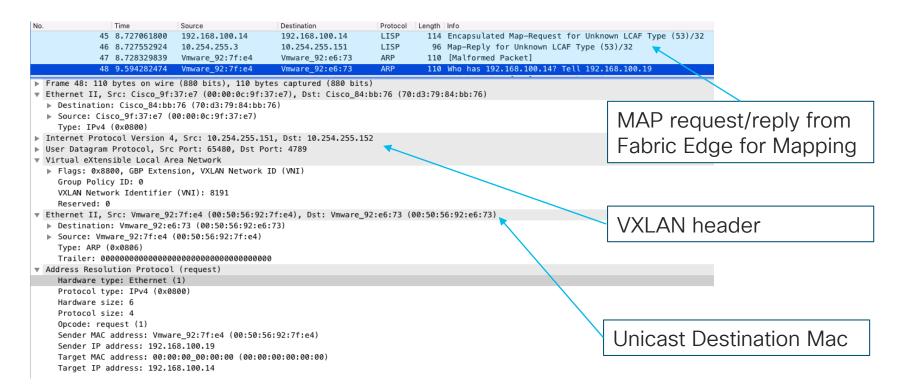
ARP in the Fabric

- ARP protocol relies on Layer 2 Broadcasts to resolve IP to Mac Address
- Layer 2 Broadcast domain (without Layer 2 flooding) constrained to just Fabric Edge
- Device Tracking enables ARP snooping, allowing rewriting of Destination Mac
- Fabric Edge register learned Address Resolution info with Control Plane node
- Fabric Edge's query Control Plane node for Address Resolution info to rewrite broadcast to Unicast Mac Address and send it through fabric as Unicast





ARP Captures





Device tracking

- Device tracking facilitates learning of End Points for Layer 2 Operation
- Learning happens for IPv4 and IPv6
- Probes used to verify/maintain reachability
- Remote entries shown via Interface Tu0, shorter aging time, no probing

```
Edge 1#show device-tracking database vlanid 1022
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
ARP 192.168.1.101
                     70e4.22e5.c4f7
                                                          1022 0005
                                           Tu0
                                                                     10s REACHABLE 297 s
   FE80::250:56FF:FE92:6D39 0050.5692.6d39 Gi1/0/1
                                                         1022 0005
                                                                     4mn REACHABLE 18 s try 0
DH4 192.168.1.100
                            0050.5692.6d39 Gi1/0/1
                                                          1022 0025
                                                                                     289 s try 0
                                                                       20s REACHABLE
   192.168.1.1
                            0000.0c9f.f45d V11022
                                                          1022 0100 5109mn REACHABLE
```



Local Mappings

- LISP maintains local database for Address Resolution
- Address Resolution is part of the Layer 2 Instance.
- Both IPv4 and Ipv6 Address are registered with Control Plane Node

```
Edge_1#sh lisp instance-id 8190 ethernet database address-resolution
LISP ETR Address Resolution for EID-table Vlan 1022 (IID 8190)
(*) -> entry being deleted
Hardware Address Host Address L3 InstID
0050.5692.6d39 FE80::250:56FF:FE92: 4100
192.168.1.100/32 4100
```

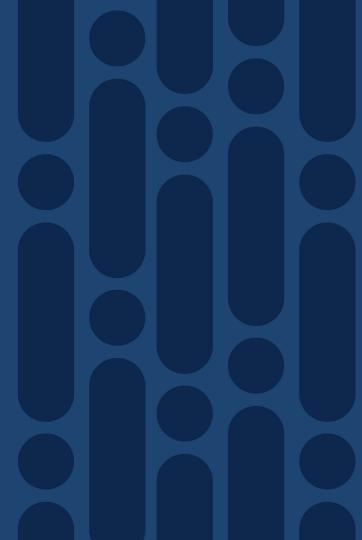


CP Address Resolution Mapping Info

- Control Plane Node maintains Address Resolution table for Layer 2 Instances
- Other Fabric Edges send mapping request to CP node when ARP entry is being received.
- CP Node responds to mapping queries from Fabric Edges



Multicast in the Fabric

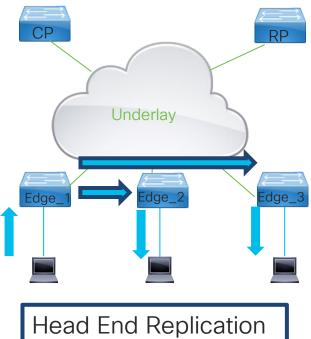


Multicasting in the Fabric

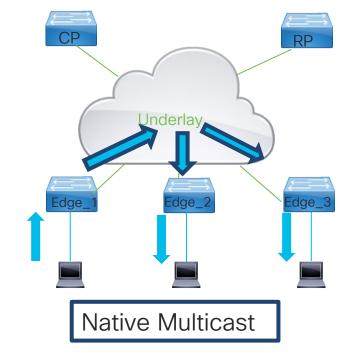
- Multicasting in SD Access modes:
 - Head End Replication (Default)
 - Native Multicast
- Head End Replication mode, multicast packets are replicated using Unicast to all fabric devices that joined the group.
- Native Multicast relies on underlay multicast topology using SSM groups Overlay Multicast groups are hashes to a range of groups in underlay network. Hashing collisions can occur
- Head End Replication can be enabled regardless of underlay multicast capable
- Native Multicast prevents Packet Duplication



Multicast Overview









RPF Resolution

Local

Remote

Edge_1#sh ip rpf vrf CiscoLive 192.168.1.100

RPF information for ? (192.168.1.100)

RPF interface: Vlan1022

RPF neighbor: ? (192.168.1.100) - directly connected

RPF route/mask: 192.168.1.100/32

RPF type: unicast (lisp)

Doing distance-preferred lookups across tables

RPF topology: ipv4 multicast base

Edge_1#sh ip rpf vrf CiscoLive 192.168.1.101

RPF information for ? (192.168.1.101)

RPF interface: LISP0.4100

RPF neighbor: ? (172.31.255.111) RPF route/mask: 192.168.1.101/32

RPF type: unicast ()

Doing distance-preferred lookups across tables

RPF topology: ipv4 multicast base

- RPF resolution for Sources reachable through the fabric:
 - RPF Interface LISP 0.<instance ID>
 - RPF Neighbor, RLOC IP address of Fabric Device source resides
- If RPF cannot be resolved multicast traffic will not flow



Head End Replication Mode, FHR

```
Edge_1#sh ip mroute vrf CiscoLive 239.100.100.100

IP Multicast Routing Table
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode
(*, 239.100.100.100), 02:29:39/stopped, RP 192.168.200.1, flags: SPF
Incoming interface: LISPO.4100, RPF nbr 172.31.255.28
Outgoing interface list: Null
(192.168.1.100, 239.100.100.100), 02:29:39/00:02:35, flags: FT
Incoming interface: Vlan1022, RPF nbr 0.0.0.0
Outgoing interface list:
LISPO.4100, 172.31.255.110, Forward/Sparse, 00:10:30/00:02:54
LISPO.4100, 172.31.255.111, Forward/Sparse, 01:09:35/00:02:46
```

- First Hop Router sending traffic through vxlan to both RLOCs with receivers
- All edge nodes join the *.G pointing to the RP RLOC IP address
- Traffic from Sender gets encapsulated into vxlan, similar to Unicast traffic



Head End Replication Mode, Egress Router

On receiver side the packet is de-encapsulated and send to the receiver

```
Edge 3#sh ip mroute vrf CiscoLive 239.100.100.100
(*, 239.100.100.100), 05:14:22/stopped, RP 192.168.200.1, flags: SJC
 Incoming interface: LISPO.4100, RPF nbr 172.31.255.28
 Outgoing interface list:
                                              RPF of (S,G) is RLOC of FHR
   Vlan1022, Forward/Sparse, 01:52:18/00:02:13
(192.168.1.100, 239.100.100.100), 01:29:05/00:02:09, flags: JT
 Incoming interface: LISPO.4100, RPF nbr 172.31.255.109
 Outgoing interface list:
   Vlan1022, Forward/Sparse, 01:29:05/00:02:13
                                              Ingress LISP Egress Vlan1022
Edge 3#sh ip igmp vrf CiscoLive groups
Group Address Interface
                             Uptime Expires Last
239.100.100.100 Vlan1022
                                      01:53:01 00:02:26 192.168.1.101
Edge 3#sh ip igmp snooping groups
                               Type Version Port List
Vlan Group
1022 239.100.100.100
                               igmp v3 Gi1/0/1
```

IGMP join on Gi 1/0/1 triggered the join.



Native Multicast – First Hop Router

```
Edge_1#sh ip mroute vrf CiscoLive 239.100.100.100 verbose

IP Multicast Routing Table
(*, 239.100.100.100), 23:32:06/stopped, RP 192.168.200.1, flags: SPF
    Incoming interface: LISP0.4100, RPF nbr 172.31.255.28, LISP:
[172.31.255.28, 232.0.3.1]
    Outgoing interface list: Null
(192.168.1.100, 239.100.100.100), 23:32:06/00:02:53, flags: FTp
    Incoming interface: Vlan1022, RPF nbr 0.0.0.0
Outgoing interface list:
    LISP0.4100, (172.31.255.109, 232.0.3.1), Forward/Sparse,

17:09:05/stopped, p
    172.31.255.111, 17:09:04/00:03:07
    172.31.255.110, 17:09:05/00:02:41
```



Native Multicast – First Hop Router

```
Edge_1#sh ip mfib 172.31.255.109 232.0.3.1
Forwarding Counts: Pkt Count/Pkts per second/Avg Pkt Size/Kbits per second
Other counts: Total/RPF failed/Other drops
I/O Item Counts: FS Pkt Count/PS Pkt Count
Default
(172.31.255.109,232.0.3.1) Flags: HW
SSM group overlay uses
SW Forwarding: 0/0/0/0, Other: 1/1/0
HW Forwarding: 61913/1/102/0, Other: 0/0/0
GigabitEthernet1/0/24 Flags: F NS
Pkts: 0/0

Egress port
```

- In underlay network the Overlay traffic is send encapsulated in vxlan
- Traffic is send as a multicast with source the RLOC of this fabric device



Native Multicast – Egress Router

```
Edge_2#sh ip mroute 232.0.3.1
IP Multicast Routing Table
(172.31.255.28, 232.0.3.1), 17:38:29/00:00:30, flags: sT
Incoming interface: GigabitEthernet2/0/47, RPF nbr 172.31.250.64
Outgoing interface list:
Nullo, Forward/Dense, 17:38:29/stopped
(172.31.255.109, 232.0.3.1), 17:38:29/00:00:30, flags: sT
Incoming interface: GigabitEthernet2/0/47, RPF nbr 172.31.250.64
Outgoing interface list:
Nullo, Forward/Dense, 17:38:29/stopped

S,G, sourced at RP
```

- Egress Interface showing Null, traffic is being De-encapsulated
- RPF neighbor for Underlay Multicast group is upstream router

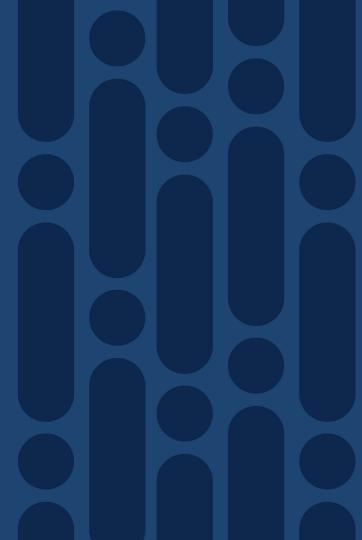


Native Multicast, Egress Router

- At the Egress Fabric Device traffic is de-encapsulated and send out
- RPF neighbor in Overlay is the RLOC of encapsulating device

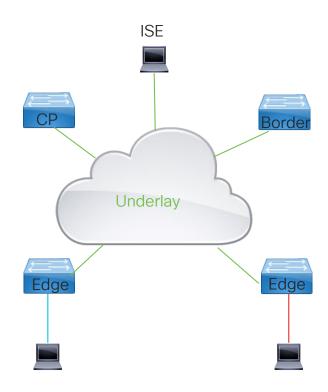


Security in the Fabric



Authentication in the Fabric

- Switch based authentication provides:
 - Access Control to Fabric
 - Assignment to VN/Pool
 - Policy Assignment to Endpoint
- ISE recommended, not mandatory
- Switches use 802.1x and Mac Address
 Bypass (MAB) to authenticate endpoints
- ISE can use profiling to determine type of endpoint



Authentication Profiles

- Default Profile per Fabric, applied to all Layer 2 Interfaces
 Can be overridden using host onboarding on Cisco DNA Center
- Order of Authentication and timers can be tuned on Cisco DNA Center.

Profiles:

- Closed Authentication, Most Secure
 Dot1x & MAB using Closed Authentication
- Open Authentication, Moderately Secure Dot1x & MAB using Open authentication
- Easy Connect, Moderately Secure
 Dot1x & MAB using open authentication and pre-auth ACL
- No Authentication, Unsecure



Access Session details

Edge 3#sh access-session interface gigabitEthernet 1/0/1 details GigabitEthernet1/0/1 Interface: 0x19558A98 IIF-ID: MAC Address: 70e4.22e5.c4f7 IPv6 Address: Unknown IPv4 Address: 192.168.1.101 User-Name: CLtestuser Device-type: Cisco-Device Status: Authorized Domain: DATA Oper host mode: multi-auth Oper control dir: both Session timeout: N/A Common Session ID: AC1FFA45000000107B7EA0EB Acct Session ID: 0x0000005 0x1d000006 Handle. Current Policy: PMAP DefaultWiredDot1xClosedAuth 1X MAB

- IPv4/IPv6 info from device tracking
- Username that authenticate
- Device-type from profiling
- Domain: Data or Voice
- Control Direction: in or both
- Policy: Applied policy on interface

- Server Policies, send from Radius
- Method: dot1x or mab and its state

Server Policies:

Vlan Group: Vlan: 1022

SGT Value: 200

Method status list:

Method Status Authc Success

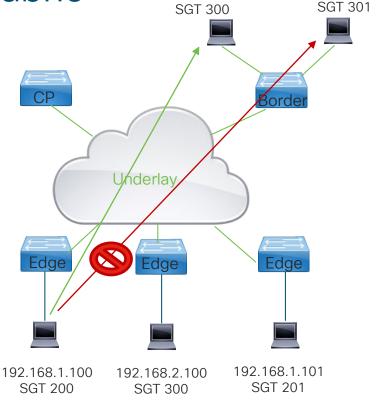


Security Policies inside the Fabric

SGT	Endpoint
201	192.168.1.100
300	192.168.1.101
200	192.168.1.3
300	10.48.91.151
301	10.48.91.251

SRC	DST	Action
200	300	Permit ssh Deny any
200	300	Deny ssh permit any

- Security based on Cisco TrustSec Solution
- Policy header inside vxlan header carries SGT
- Every endpoint assigned SGT Traffic policies enforced on egress not ingres
- Policies downloaded from ISE based on groups



10.48.91.151

10.48.91.251

Cisco TrustSec

- Every endpoint in the fabric gets assigned a Secure Group Tag
- Secure Group Tag transmitted in Policy Field in vxlan header of encapsulated frames
- Fabric devices download CTS environment data from ISE server.
- Fabric devices download permissions for all SGT on switch (Destination mappings only)
- Traffic being allowed/denied based upon SGT -> DGT mapping
- Traffic policy can be deny all, permit all, or SGACL
- Default action applied to all cells not populated.



CTS environment data

```
Edge 1#sh cts environment-data
                                   CTS environment data from ISE.
CTS Environment Data
                                   Crucial for Enforcement to occur
Current state = COMPLETE
Last status = Successful
Local Device SGT:
  SGT tag = 0-01:Unknown
Server List Info:
Installed list: CTSServerList1-0001, 1 server(s):
*Server: 10.48.91.222, port 1812, A-ID 25FCBAE325B2C0E4073058F860957868
         Status = ALIVE
         auto-test = TRUE, keywrap-enable = FALSE, idle-time = 60 mins, deadtime = 20 secs
Security Group Name Table:
                                                          Radius server used
   0-01:Unknown
   20-00: Phones
   200-01:CL Client 1
                                                          Groups known on ISE
   201-01:CL Client 2
Environment Data Lifetime = 86400 secs
Last update time = 16:26:35 UTC Wed Jan 8 2020
                                                              ISE can trigger
Env-data expires in 0:20:50:45 (dd:hr:mm:sec)
Env-data refreshes in 0:20:50:45 (dd:hr:mm:sec)
                                                              CoA to update
Cache data applied
                           = NONE
State Machine is running
```



CTS Enforcement

- All endpoints not assigned an SGT tag via Authentication or static configuration will belong to SGT 0 (unknown)
- SGT can be learned Locally on switch or via SXP sessions

```
Edge 1#sh cts role-based sgt-map vrf CiscoLive all
Active IPv4-SGT Bindings Information
IP Address
                         SGT
192.168.1.100
                        201
                                 LOCAL
CP 2#sh cts role-based sgt-map vrf CiscoLive all
Active IPv4-SGT Bindings Information
TP Address
                         SGT
                                 Source
10.48.91.151
                         300
                                 SXP
10.48.91.251
                         301
                                 SXP
```

Endpoint IP assigned SGT 201 via 802.1x

Border learned 2 mappings via SXP to ISE Server



CTS Policies

- Fabric Devices only Downloaded Policies it needs enforcing (egress enforcement) and is present on ISE
- All other traffic will hit a * * policy
- RBACL names are appended with a version,
 Ex: NoTelnet-00 is version 00 of RBACL name NoTelnet

```
CP 2#sh cts role-based permissions to 300
IPv4 Role-based permissions from group 200:CL Client 1 to group 300:CL Servers 1:
AllowSSHPING-00
IPv4 Role-based permissions from group 201:CL Client 2 to group 300:CL Servers 1:
allowping-00
CP 2#sh cts rbacl AllowSSHPING
                                        CP 2#sh cts rbacl allowping
CTS RBACL Policy
                                        CTS RBACL Policy
        = AllowSSHPING-00
                                                 = allowping-00
                                          name
 refcnt = 4
                                           refcnt = 4
   RBACL ACEs:
                                          RBACL ACEs:
   permit tcp dst eq 22
                                            permit icmp
   permit icmp
                                            deny tcp dst eq 22
    deny ip
                                            permit ip
```

Monitoring SGT traffic

- Counters are accumulative per device
- Traffic not hitting a more specific entry will hit * *
- Different Column for Software and Hardware enforcement

```
CP 2#sh cts role-based counters
Role-based TPv4 counters
                SW-Denied HW-Denied SW-Permitt HW-Permitt SW-Monitor HW-Monitor
From
                                      4965
                                                  312090
200
        300
201
       300
                           15
                                                  146
200
       301
201
        301
                                                  195
Edge 1#sh cts role-based counters
Role-based IPv4 counters
                SW-Denied HW-Denied SW-Permitt HW-Permitt SW-Monitor HW-Monitor
From
                                      13296
                                                  21927
200
        2.01
                                                  1.3
```



Complete your online session survey



- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live t-shirt.
- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Content Catalog on <u>ciscolive.com/emea</u>.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.com.



Continue your education





illilli CISCO

Thank you



cisco live!





You make possible