Nexus 7000 Configuration and Verification of LISP IGP Assist Extended Subnet Mode

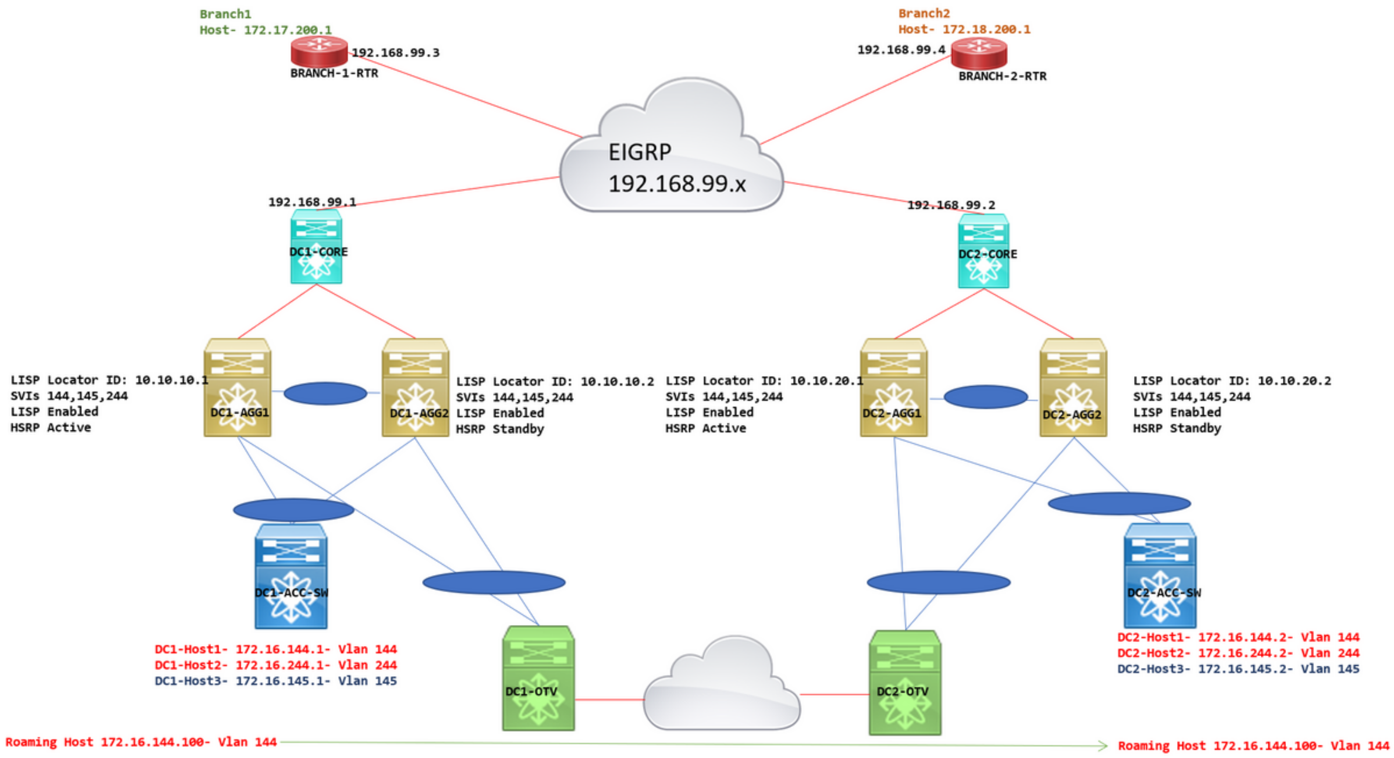
Contents

[Introduction]
[Topology]
[Details of the Topology]
[Components Used]
[Required Configurations on AGG switches]
[LISP Specific Configs on DC1-Agg1 and DC1-Agg2]
[LISP Specific Configs on DC2-Agg1 and DC2-Agg2]
[IGP Specific]
[Route-maps/Prefix-Lists]
[OTV VDC Configurations]
[Route population due to LISP configuration]
[Sequence of Events when Host comes online within a LISP enabled SVI]
[Map Notify Messages]
[Redistribution of LISP /32 routes to IGP]
[Path of the packet for Intra-vlan inter-DC]
[Path of the packet for Inter-vlan inter-DC(From Vlan 144 to Vlan 244)]
[Path of the packet for Inter-vlan inter-DC(From VRF-tenant-1 to VRF tenant-2)]
[Path of the Packet when a Branch-1 Host tries to reach a Silent Host which is present in DC2]
[Sequence of Events when a host moves(Roam) from DC1 to DC2]
[Useful Verification Commands]

Introduction

This Document is to explain how to deploy LISP IGP Assist Extended Subnet Mode(ESM) using a Nexus 7000

Topology



Details of the Topology

- DC1 and DC2 are two locations extended by OTV
- Vlans 144, 145 and 244 are configured on all Agg, Access layer and OTV Switches
- SVIs for these Vlans are configured on the Agg switches. SVIs 144 and 244 are in VRF tenant-1; SVI 145 is in VRF tenant-2.
- While deploying LISP IGP Assist, It is not a requirement that the SVIs have to be in VRFs; This example uses Multiple VRFs just to illustrate the Configuration changes that are required(under each relevant VRF context); All SVIs can be in the same VRF and still can use LISP IGP assist
- HSRP is configured in Vlans144, 145 and 244; FHRP Isolation is configured in this topology which would mean that in total 4 switches will be running HSRP and both sides will have active/standby pair. FHRP isolation is achieved by filtering HSRP Hello messages.
- DC1-agg1 and DC2-Agg2 are vPC pairs; Same applies for DC2-Agg1 and DC2-Agg2
- LISP configurations are applied under the SVIs 144, 145 and 244
- EIGRP neighborship is established from Agg to the Core switches per VRF. Sub-interfaces are running from Agg switches for each VRF up to the core switches and EIGRP Neighborship is formed over these sub-interfaces.
- Remote Routers(Branch) are also part of the same IGP Domain.
- When LISP IGP Assist is used, There is no LISP Encap/Decap and so the LISP routes will have to redistributed to the IGP(Here it is EIGRP). For this deployment model depicted in this document, the Branch Routers will not be having any LISP configuration.

Components Used

- Agg, Core switches are Nexus 7000 with SUP2E, F3/M3 running 8.2(4) NXOS Version
- Branch Routers are ASR1ks
- OTV Is Configured in another VDC on these Nexus 7000 switches; OTV and LISP have to be

on different VDCs. Sharing VDC is not an option.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, make sure that you understand the potential impact of any command.

Required Configurations on AGG switches

LISP Specific Configs on DC1-Agg1 and DC1-Agg2

Common Configuration on both DC1-Agg1 and DC1-Agg2

```
feature lisp
vrf context tenant-1                                                      # This example is
based on SVI 144 in VRF- tenant-1 and SVI 145 in VRF- tenant-2
  ip lisp etr                                                             # This is needed to
initialize LISP and only etr is needed on a IGP assist mode Environment
  lisp instance-id 2                                                      # Instance-ID should
be unique per VRF
  ip lisp locator-vrf default                                             # Locator Is
specified in Default VRF
  lisp dynamic-eid VLAN144                                                # Dynamic EID
definition for Vlan 144
    database-mapping 172.16.144.0/24 10.10.10.1 priority 50 weight 50     # Database-mapping
for 172.16.144.0/24 which is the Vlan 144; IP-> 10.10.10.1 is the Loopback100 IP address(which
is the unique IP on DC1-AGG1)
    database-mapping 172.16.144.0/24 10.10.10.2 priority 50 weight 50     # Database-mapping
for 172.16.144.0/24 which is the Vlan 144; IP-> 10.10.10.2 is the Loopback100 IP address(which
is the unique IP on DC1-AGG2)
    map-notify-group 239.254.254.254                                      # Multicast group
that will be used by LISP enabled switches to communicate about new EID learns or periodic EID
notification messages
    no route-export away-dyn-eid                                          # This is a hidden
command required to stop advertising any null0 /32 route for a remote host to the IGP
  lisp dynamic-eid VLAN244                                                # Dynamic EID
definition for Vlan 244
    database-mapping 172.16.244.0/24 10.10.10.1 priority 50 weight 50
    database-mapping 172.16.244.0/24 10.10.10.2 priority 50 weight 50
    map-notify-group 239.254.254.254
    no route-export away-dyn-eid

vrf context tenant-2
  ip lisp etr
  lisp instance-id 3
  ip lisp locator-vrf default
  lisp dynamic-eid VLAN145
    database-mapping 172.16.145.0/24 10.10.10.1 priority 50 weight 50
    database-mapping 172.16.145.0/24 10.10.10.2 priority 50 weight 50
    map-notify-group 239.254.254.254
    no route-export away-dyn-eid
```

Configuration on DC1-Agg1

```
interface Vlan144
  no shutdown
  vrf member tenant-1
  lisp mobility VLAN144
  lisp extended-subnet-mode                                              # SVI needs to be in
ESM Mode-Extended subnet mode
  ip address 172.16.144.250/24
```

```
  ip pim sparse-mode
  hsrp 144
    preempt
    priority 254
    ip 172.16.144.254

interface Vlan145
  no shutdown
  vrf member tenant-2
  lisp mobility VLAN145
  lisp extended-subnet-mode
  ip address 172.16.145.250/24
  ip pim sparse-mode
  hsrp 145
    preempt
    priority 254
    ip 172.16.145.254

interface Vlan244
  no shutdown
  vrf member tenant-1
  lisp mobility VLAN244
  lisp extended-subnet-mode
  ip address 172.16.244.250/24
  hsrp 244
    preempt
    priority 254
    ip 172.16.244.254

interface loopback100
  ip address 10.10.10.1/32
  ip router eigrp 100
  ip pim sparse-mode
```

Configuration on DC1-Agg2

```
interface Vlan144
  no shutdown
  vrf member tenant-1
  lisp mobility VLAN144
  lisp extended-subnet-mode
  ip address 172.16.144.251/24
  ip pim sparse-mode
  hsrp 144
    ip 172.16.144.254

interface Vlan145
  no shutdown
  vrf member tenant-2
  lisp mobility VLAN145
  lisp extended-subnet-mode
  ip address 172.16.145.251/24
  ip pim sparse-mode
  hsrp 145
    ip 172.16.145.254

interface Vlan244
  no shutdown
  vrf member tenant-1
  lisp mobility VLAN244
  lisp extended-subnet-mode
  no ip redirects
  ip address 172.16.244.251/24
  hsrp 244
```

```
    ip 172.16.244.254

interface loopback100
  ip address 10.10.10.2/32
  ip router eigrp 100
  ip pim sparse-mode
```

The database mapping has to be provided in such a way that in one side, both the DC1-Agg1 and DC1-Agg2 Loopback IP addresses are required to be specified; Within DC2-Agg1 and DC2-Agg2, a unique loopback will have to be created and put the same within the database-mapping.

In an IGP assist mode, if the configuration-> "ip lisp itr-etr" Is used, this will result in injecting /32 null0 host route for non LISP enabled Vlans; So the correct configuration is "ip lisp etr" for IGP assist mode.

LISP Specific Configs on DC2-Agg1 and DC2-Agg2

```
Common Configuration on both DC2-Agg1 and DC2-Agg2

feature lisp

vrf context tenant-1
  ip lisp etr
  lisp instance-id 2
  ip lisp locator-vrf default
  lisp dynamic-eid VLAN144
    database-mapping 172.16.144.0/24 10.10.20.1 priority 50 weight 50      # Note that the IP addresses used in DC2 Agg switches are 10.10.20.1 and 10.10.20.2(Which are Loopbacks Configured on DC2-Agg switches)
    database-mapping 172.16.144.0/24 10.10.20.2 priority 50 weight 50
    map-notify-group 239.254.254.254
    no route-export away-dyn-eid
  lisp dynamic-eid VLAN244
    database-mapping 172.16.244.0/24 10.10.20.1 priority 50 weight 50
    database-mapping 172.16.244.0/24 10.10.20.2 priority 50 weight 50
    map-notify-group 239.254.254.254
    no route-export away-dyn-eid
vrf context tenant-2
  ip lisp etr
  lisp instance-id 3
  ip lisp locator-vrf default
  lisp dynamic-eid VLAN145
    database-mapping 172.16.145.0/24 10.10.20.1 priority 50 weight 50
    database-mapping 172.16.145.0/24 10.10.20.2 priority 50 weight 50
    map-notify-group 239.254.254.254
    no route-export away-dyn-eid
```

Configuration on DC2-Agg1

```
interface Vlan144 no shutdown vrf member tenant-1 lisp mobility VLAN144 lisp extended-subnet-mode ip address 172.16.144.252/24 ip pim sparse-mode hsrp 144 preempt priority 254 ip 172.16.144.254 interface Vlan145 no shutdown vrf member tenant-2 lisp mobility VLAN145 lisp extended-subnet-mode ip address 172.16.145.252/24 ip pim sparse-mode hsrp 145 preempt priority 254 ip 172.16.145.254 interface Vlan244 no shutdown vrf member tenant-1 lisp mobility VLAN244 lisp extended-subnet-mode ip redirects ip address 172.16.244.252/24 hsrp 244 preempt priority 254 ip 172.16.244.254 interface loopback100 ip address 10.10.20.1/32 ip router eigrp 100 ip pim sparse-mode
```

Configuration on DC2-Agg2

```
interface Vlan144 no shutdown vrf member tenant-1 lisp mobility VLAN144 lisp extended-subnet-mode ip address 172.16.144.253/24 ip pim sparse-mode hsrp 144 ip 172.16.144.254 interface Vlan145 no shutdown vrf member tenant-2 lisp mobility VLAN145 lisp extended-subnet-mode ip address 172.16.145.253/24 ip pim sparse-mode hsrp 145 ip 172.16.145.254 interface Vlan244 no shutdown vrf member tenant-1 lisp mobility VLAN244 lisp extended-subnet-mode no ip redirects ip
```

```
address 172.16.244.253/24 hsrp 244 preempt ip 172.16.244.254 interface loopback100 ip address
10.10.20.2/32 ip router eigrp 100 ip pim sparse-mode
```

The difference between DC1 and DC2 Agg LISP configurations are the loopbacks defined in the "database mapping". In DC1 configuration, this will be defined with the loopbacks of DC1-Agg1 and DC1-Agg2 and for DC2, database mappings will be defined with the loopbacks which are in DC2-Agg1 and DC2-Agg2

Rest of the IGP/Route-maps/prefix-lists configurations shown below are going to be similar(The IP addresses assigned for the interfaces are indeed different)

IGP Specific

```
router eigrp 100
  address-family ipv4 unicast
  vrf tenant-1
    distance 90 245                                                   # External EIGRP
Routes have to have an AD which is higher than the default LISP AD(which is 240); Reason being,
if the redistributed route from dc1-agg1 comes back to dc1-agg2 via eigrp, default EIGRP
External is 170 which will override LISP route causing problems
    redistribute lisp route-map lisp-to-eigrp                         # This command is to
redistribute LISP /32 routes only to the IGP(EIGRP In this example)
    redistribute direct route-map direct                              # This is needed so
that the direct routes(/24 SVI routes in LISP) are redistributed to the IGP; This will be needed
if there is some device that is trying to communicate to a silent host in the LISP enabled Vlan
  vrf tenant-2
    distance 90 245
    redistribute lisp route-map lisp-to-eigrp
    redistribute direct route-map direct
```

LISP enabled AGG VDCs also will form IGP neighborship to the Core side

For this example, sub-interfaces which were part of each tenant VRFs were used to form the neighborship towards the core as shown below.

```
interface Ethernet3/6.111
  encapsulation dot1q 111
  vrf member tenant-1
  ip address 192.168.98.1/30
  ip router eigrp 100
  no shutdown


interface Ethernet3/6.212
  encapsulation dot1q 212
  vrf member tenant-2
  ip address 192.168.198.1/30
  ip router eigrp 100
  no shutdown
```

Route-maps/Prefix-Lists

```
ip prefix-list lisp-to-eigrp seq 5 permit 0.0.0.0/0 ge 32             # This is the prefix
list that is matching any /32 routes which are to be redistributed from LISP To IGP

route-map direct permit 10                                            # This is for the
Direct routes
```

```
route-map lisp-to-eigrp deny 10                                        # This is to prevent
any null0 routes from being redistributed to IGP from LISP
  match interface Null0
route-map lisp-to-eigrp permit 20                                      # This is to allow
redistribution of /32 host routes
  match ip address prefix-list lisp-to-eigrp
```

All the above Configurations are required on all AGG switches(DC1 and DC2). Keep in mind to provide unique IP addresses for the SVIs, Loopbacks, HSRP VIP will be the same for all SVIs

OTV VDC Configurations

HSRP Filtering

For IGP assist deployments, When extended by means of OTV or any other mechanism, the FHRP isolation has to be in place;

This is done by filtering FHRP Hello messages within the OTV VDC

In this example, N7k OTV is used and so below configurations were applied to filter the FHRP packets in OTV VDC.

```
ip access-list ALL_IPs
  10 permit ip any any
mac access-list ALL_MACs
  10 permit any any
ip access-list HSRP_IP
  10 permit udp any 224.0.0.2/32 eq 1985
  20 permit udp any 224.0.0.102/32 eq 1985
mac access-list HSRP_VMAC
  10 permit 0000.0c07.ac00 0000.0000.00ff any
  20 permit 0000.0c9f.f000 0000.0000.0fff any
arp access-list HSRP_VMAC_ARP
  10 deny ip any mac 0000.0c07.ac00 ffff.ffff.ff00
  20 deny ip any mac 0000.0c9f.f000 ffff.ffff.f000
  30 permit ip any mac any
vlan access-map HSRP_Localization 10
        match mac address HSRP_VMAC
        match ip address HSRP_IP
        action drop
vlan access-map HSRP_Localization 20
        match mac address ALL_MACs
        match ip address ALL_IPs
        action forward
vlan filter HSRP_Localization vlan-list 144-145
ip arp inspection filter HSRP_VMAC_ARP vlan 144-145

mac-list OTV_HSRP_VMAC_deny seq 10 deny 0000.0c07.ac00 ffff.ffff.ff00
mac-list OTV_HSRP_VMAC_deny seq 11 deny 0000.0c9f.f000 ffff.ffff.f000
mac-list OTV_HSRP_VMAC_deny seq 20 permit 0000.0000.0000 0000.0000.0000
route-map OTV_HSRP_filter permit 10
  match mac-list OTV_HSRP_VMAC_deny

otv-isis default
  vpn Overlay0
    redistribute filter route-map OTV_HSRP_filter
```

FHRP Filtering Configurations are required ONLY on OTV VDCs; If an ASR OTV deployment is used, the filtering mechanisms should be used as relevant and documented as per the ASR configuration guide.

OTV Suppress ARP

 # Disable ARP ND-cache feature on OTV VDCs

```
interface Overlay0
  no otv suppress-arp-nd >>>>>
```

Route population due to LISP configuration

```
DC1-AGG1# show ip route lisp vrf tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.0/25, ubest/mbest: 1/0
    *via Null0, [240/1], 07:22:30, lisp, dyn-eid
172.16.144.128/25, ubest/mbest: 1/0
    *via Null0, [240/1], 07:22:30, lisp, dyn-eid
```

When LISP is enabled on SVI 144, there will be two Null0 routes that are auto created; SVI 144 is a /24 subnet and so 1st null0 route would be from 172.16.144.0/25 and the second one will be 172.16.144.128/25 as shown above.

This is expected and by design; this is done in order to make sure that the packets sourced from undiscovered hosts trigger an RPF exception which will result in the packets to get punted to CPU and eventually will help in Host Detection(EID)

Sequence of Events when Host comes online within a LISP enabled SVI

Host detection on LISP enabled interfaces is based on the reception of L3 traffic from IP addresses within the range specified in the database-mapping configuration.

In order to facilitate detection of hosts, note that when LISP is enabled on an interface:
 # RPF exceptions are enabled on the interface, so that packets generated by unknown sources trigger the exception
 # LISP sourced Null0 routes are installed to ensure that unknown sources trigger the RPF exception

Since this solution relies on OTV for L2 extension between the two datacenters, ARP signaling cannot be directly used to detect IP hosts since in many cases is broadcasted to all the switches.

However ARP signals are used as an indication for LISP that an undetected host may be present. Since the host can reside at any side of the OTV bridge, LISP starts a localization mechanism

after learning a new IP-MAC binding.
The localization mechanism works as follows:
 # The switch learns a new IP-MAC binding (through GARP, RARP or an ARP request).
 # The switch that works as an active HSRP sends an echo request to the host but sourced from the HSRP VIP address
 # The host replies to the echo request, but following FHRP isolation in OTV, the echo reply is only received on the DC site where the host resides
 # Since the echo reply is an L3 packet, the host is then detected by LISP.

If an IP packet is received on any LISP enabled SVI, that itself will feed the LISP process informing that the End point is Local; there will not be any ICMP ECHO reqests send out to further confirm if the host is local or not. So it is critical to note that a Ping from DC2 Host to DC1-AGG SVI IP addresses will result in corruption of end point identification which also might result in ping loss or traffic blackhole as the host now is identified as a local EID in DC1 as opposed to DC2. So pings should not be originated from SVI IP addresses in a LISP environment as this may corrupt the routing table and will result in blackholing of traffic. The same issue will happen if the Hosts which are in LISP Enabled Vlan try to ping the SVI IP addresses; Pinging the VIP should be fine as the same is present and active on both Sides and site Local will catch the packet.

An example of routing table entry when a host is online in DC1 is as below;

```
DC1-AGG1# show ip route 172.16.144.1 vrf tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.1/32, ubest/mbest: 1/0, attached
    *via 172.16.144.1, Vlan144, [240/1], 3d05h, lisp, dyn-eid
     via 172.16.144.1, Vlan144, [250/0], 3d05h, am

DC1-AGG2# sh ip route 172.16.144.1 vr tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.1/32, ubest/mbest: 1/0, attached
    *via 172.16.144.1, Vlan144, [240/1], 3d05h, lisp, dyn-eid
     via 172.16.144.1, Vlan144, [250/0], 3d05h, am
```

As seen above, there are two routes; One by LISP process with the administrative Distance of 240 and other by AM-> Adjacency manager(populated by ARP process) which has AD of 250.

Both the Agg switches in DC1 will have the same entry.

also, LISP will list the same entry for the host in the dynamic EID table as shown below.

```
DC1-AGG1# show lisp dynamic-eid detail vrf tenant-1 | in 144.1, nex 1 172.16.144.1, Vlan144, uptime: 3d05h, last activity: 00:14:38 Discovered by: packet reception DC1-AGG2# show lisp dynamic-eid detail vrf tenant-1 | in 144.1, nex 1 172.16.144.1, Vlan144, uptime: 3d05h, last activity: 00:00:37 Discovered by: site-based Map-Notify
```

Discovery is different in both cases; DC1-AGG1 which is the HSRP active is recording the entry by means of "packet reception" which basically means that there was a packet that came in which

resulted in adding it as an EID

Once the Agg1 came to know about an EID, it sends out a multicast message from the Source IP-> Loopback100 IP address(defined under database mapping) to the group-> 239.254.254.254(configured above) and the vPC peer switch also receives it and populates the entry accordingly and considers this as a local EID due to the database mapping having both the IP addresses of dc1-agg1 and dc1-agg2. This same multicast packet would also traverse through the OTV to the remote sites; However, Remote sites would check the database mapping and since this packet is sourced from an IP address that is different from that of the "database mapping", it will not be considered as a local EID by the DC2 AGg switches.

Map Notify Messages

When a host is detected by LISP enabled SVI, a triggered "map-notify" message will be sent out to the multicast group that is defined under the corresponding dynamic EID configuration

Other than the triggered map-notify messages, there are periodic map-notify messages which are sent by the HSRP Active(Or FHRP active) switch in that vlan;

A PCAP of map notify message is as below.

```
> Frame 285: 122 bytes on wire (976 bits), 122 bytes captured (976 bits) on interface eth0, id 0
> Ethernet II, Src: de:ad:20:20:22:22 (de:ad:20:20:22:22), Dst: IPv4mcast_7e:fe:fe (01:00:5e:7e:fe:fe)
> Internet Protocol Version 4, Src: 10.10.20.2, Dst: 239.254.254.254
> User Datagram Protocol, Src Port: 4342, Dst Port: 4342
v Locator/ID Separation Protocol
    0100 .... .... .... .... .... = Type: Map-Notify (4)
    .... 0... .... .... .... .... = I bit (xTR-ID present): Not set
    .... .0.. .... .... .... .... = R bit (Built for an RTR): Not set
    .... ..00 0000 0000 0000 0000 = Reserved bits: 0x00000
    Record Count: 4
    Nonce: 0x0000000000000000
    Key ID: 0x0000
    Authentication Data Length: 0
    Authentication Data: <MISSING>
  > Mapping Record 1, EID Prefix: 172.16.144.2/32, TTL: -1610285056, Action: No-Action, Not Authoritative
  > Mapping Record 2, EID Prefix: 172.16.144.111/32, TTL: -1610285056, Action: No-Action, Not Authoritative
  > Mapping Record 3, EID Prefix: 172.16.144.252/32, TTL: -1610285056, Action: No-Action, Not Authoritative
  > Mapping Record 4, EID Prefix: 172.16.144.254/32, TTL: -1610285056, Action: No-Action, Not Authoritative
```

Redistribution of LISP /32 routes to IGP

This is they key for IGP assist mode; Any /32 LISP route would be redistributed to IGP; This is made possible by the "redistribute LISP" command that was applied under EIGRP.

Any /32 host route will be seen as an EIGRP external route after the redistribution. An adjustment of EIGRP Administrative distance was done to make it higher. This is to make sure that the LISP route stays in URIB as opposed to the incoming EIGRP External Route. eg; DC1-Agg1 and DC1-Agg2 are EIGRP Neighbors with DC1-core. A /32 route was injected by DC1-AGG1 to DC1-Core by means of redistribution. Now that the DC1-Core is EIGRP neighbor with DC1-Agg2, the same route may come back to DC1-Agg2 and has a chance of winning over LISP route(which has an AD of 240) if the EIGRP AD was 170; So, to avoid this, the EIGRP external route AD has been modified to 245.

The /32 route which was learnt by the DC1-Agg switches are redistributed to EIGRP and the DC1-core entry would look like below.

```
DC1-CORE# sh ip route 172.16.144.1
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.1/32, ubest/mbest: 2/0
    *via 192.168.98.1, Eth3/20.111, [170/51456], 00:00:01, eigrp-100, external
    *via 192.168.98.5, Eth3/22.112, [170/51456], 18:14:51, eigrp-100, external
```

The route is present in Global Routing table and no VRF is configured on the Core side.

And due to the "redistribute direct" that was configured on AGG Switches, the Core will also have a /24 ECMP route for the parent subnet as shown below. This will help to attract traffic for a silent host(For which there is no /32 route).

```
DC1-CORE# sh ip route 172.16.144.10                # Checking for a non existent Host
172.16.144.10
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.0/24, ubest/mbest: 2/0
    *via 192.168.98.1, Eth3/20.111, [170/51456], 00:02:13, eigrp-100, external
    *via 192.168.98.5, Eth3/22.112, [170/51456], 18:17:03, eigrp-100, external
```

Also a /24 ECMP route would be seen to both the DC1 and DC2 cores

```
Branch1-Router# sh ip route 172.16.144.10
Routing entry for 172.16.144.0/24
  Known via "eigrp 100", distance 170, metric 51712, type external
  Redistributing via eigrp 100
  Last update from 192.168.99.2 on GigabitEthernet0/0/1, 00:00:17 ago
  Routing Descriptor Blocks:
    192.168.99.2, from 192.168.99.2, 00:00:17 ago, via GigabitEthernet0/0/1      # 192.168.99.2
is DC2-Core
      Route metric is 51712, traffic share count is 1
      Total delay is 1020 microseconds, minimum bandwidth is 100000 Kbit
      Reliability 255/255, minimum MTU 1492 bytes
      Loading 1/255, Hops 2
  * 192.168.99.1, from 192.168.99.1, 00:00:17 ago, via GigabitEthernet0/0/1      # 192.168.99.1
is DC1-Core
      Route metric is 51712, traffic share count is 1
      Total delay is 1020 microseconds, minimum bandwidth is 100000 Kbit
      Reliability 255/255, minimum MTU 1492 bytes
      Loading 1/255, Hops 2
```

This route Would make sure that a Branch Host can reach to a silent host which lives in Either location.

Path of the packet for Intra-vlan inter-DC

When DC1-Host1 -> 172.16.144.1 tries to reach DC2-Host1-> 172.16.144.2, This is intra vlan inter Datacenter Traffic. DC1-Host 1 sends out an ARP Request which will traverse all the way through the OTV and reaches DC2-Host1

DC2-Host1 responds with an ARP Reply which comes back to the DC1-Host1

Subsequent ICMP packets are sent via the OTV

Path of the packet for Inter-vlan inter-DC(From Vlan 144 to Vlan 244)

When DC1-Host1-> 172.16.144.1 tries to reach DC2-Host2-> 172.16.244.2, the Packet will NOT be routed from vlan 144 to 244 in DC1; Rather, it follows a routed path from DC1-Agg to DC1-Core and then arrive at the DC2-Core and the final routing will be done by the DC2-Agg switches to the destination Vlan-244.

A traceroute from DC1-Host1 to DC2-Host2 is as below.

```
DC1-HOST# traceroute 172.16.244.2 vrf vlan144
traceroute to 172.16.244.2 (172.16.244.2), 30 hops max, 40 byte packets
 1  172.16.144.250 (172.16.144.250)  1.149 ms  0.841 ms  0.866 ms
# DC1-AGG1
 2  192.168.98.2 (192.168.98.2)  1.004 ms  0.67 ms  0.669 ms
# DC1-CORE
 3  192.168.99.2 (192.168.99.2)  0.756 ms  0.727 ms  0.714 ms
# DC2-CORE
 4  192.168.94.5 (192.168.94.5)  1.041 ms  0.937 ms 192.168.94.1 (192.168.94.1)  1.144 ms
# DC2-Agg1/DC2-Agg2
 5  172.16.244.2 (172.16.244.2)  2.314 ms *  2.046 ms
# DC2-Host2
```

Path of the packet for Inter-vlan inter-DC(From VRF-tenant-1 to VRF tenant-2)

This is going to follow the same as inter-vlan inter-DC communication from One vlan to another(previous example)

When DC1-host1-> 172.16.144.1 tries to reach DC2-Host3-> 172.16.145.2, This is inter-vlan inter-DC Traffic origingating in Vlan 144(VRF tenant-1) and destined to Vlan 145(VRF tenant-2). Unlike regular N7k OTV Deployments, this traffic will be treated slightly different. There will not be any inter-vlan routing happening on DC1 side; Rather this traffic will be routed and send up to the DC1-core and core will further route it through the IGP to the DC2-Core

For the sake of this document, inter-VRF leaking is done per site by the Core Switch. It could be any device(Like firewall); There are no changes from LISP Configuration perspective If Inter-VRF Leaking is present or not.

```
DC1-AGG1# sh ip route 172.16.145.2 vrf tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
```

```
'%<string>' in via output denotes VRF <string>

172.16.145.2/32, ubest/mbest: 1/0
    *via 192.168.98.2, Eth3/6.111, [245/51968], 00:00:46, eigrp-100, external
```

A Traceroute from DC1-Host1 to DC2-Host3 will reveal the same that its not inter-vlan routed, rather layer 3 routed through the Core. In short, the Inter-vlan traffic will not be using the OTV.

```
DC1-HOST# traceroute 172.16.145.2 vrf vlan144
traceroute to 172.16.145.2 (172.16.145.2), 30 hops max, 40 byte packets
 1  172.16.144.250 (172.16.144.250)  1.049 ms  0.811 ms  0.81 ms             # DC1-AGG1
 2  192.168.98.2 (192.168.98.2)  0.844 ms  0.692 ms  0.686 ms                # DC1-CORE
 3  192.168.99.2 (192.168.99.2)  0.814 ms  0.712 ms  0.735 ms                # DC2-CORE
 4  192.168.194.1 (192.168.194.1)  0.893 ms  0.759 ms 192.168.194.5 (192.168.194.5)  0.89 ms   # DC2-Agg1/DC2-Agg2
 5  172.16.145.2 (172.16.145.2)  1.288 ms *  1.98 ms                         # DC2-Host3
DC1-HOST#
```

Path of the Packet when a Branch-1 Host tries to reach a Silent Host which is present in DC2

Host in Branch-1-172.17.200.1 tries to reach DC2-Silent Host- 172.16.144.119. Since the host is silent, there will not be any /32 route present in DC2.

```
DC2-AGG1# show ip route 172.16.144.119 vr tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.0/25, ubest/mbest: 1/0
    *via Null0, [240/1], 20:48:29, lisp, dyn-eid

DC2-AGG2# show ip route 172.16.144.119 vr tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.0/25, ubest/mbest: 1/0
    *via Null0, [240/1], 20:48:13, lisp, dyn-eid
```

As per LISP design, the route 172.16.144.119 will be matching to the 172.16.144.0/25 null0 route.

When the Branch router receives a packet with the destination IP = 172.16.144.119, URIB has an ECMP /24 route to both DC1-core and DC2-core. Which essentially means the packet will be sent to one of the Core switches.

```
Branch1-Router# sh ip route 172.16.144.119
Routing entry for 172.16.144.0/24
  Known via "eigrp 100", distance 170, metric 51712, type external
```

```
  Redistributing via eigrp 100
  Last update from 192.168.99.2 on GigabitEthernet0/0/1, 00:08:54 ago
  Routing Descriptor Blocks:
    192.168.99.2, from 192.168.99.2, 00:08:54 ago, via GigabitEthernet0/0/1
      Route metric is 51712, traffic share count is 1
      Total delay is 1020 microseconds, minimum bandwidth is 100000 Kbit
      Reliability 255/255, minimum MTU 1492 bytes
      Loading 1/255, Hops 2
  * 192.168.99.1, from 192.168.99.1, 00:08:54 ago, via GigabitEthernet0/0/1
      Route metric is 51712, traffic share count is 1
      Total delay is 1020 microseconds, minimum bandwidth is 100000 Kbit
      Reliability 255/255, minimum MTU 1492 bytes
      Loading 1/255, Hops 2

Branch1-Router#sh ip cef exact-route 172.17.200.1 172.16.144.119 dest-port 1
172.17.200.1 -> 172.16.144.119 =>IP adj out of GigabitEthernet0/0/1, addr 192.168.99.1
```

The packet as per CEF is hashing to 192.168.99.1(which is DC1-Core)

DC1-Core has 2 ECMP paths; One towards the DC1-Agg1(HSRP Active) and the second, towards the DC1-Agg2(HSRP Standby). From Routing hash, the path selected is going to be the DC1-Agg2.

```
DC1-CORE# sh routing hash 172.17.200.1 172.16.144.119 1 1 Load-share parameters used for
software forwarding: load-share mode: address source-destination port source-destination
Universal-id seed: 0xfdba3ebe Hash for VRF "default" Hash Type is 1 Hashing to path
*192.168.98.5 Eth3/22.112
For route:
172.16.144.0/24, ubest/mbest: 2/0
    *via 192.168.98.1, Eth3/20.111, [170/51456], 00:19:57, eigrp-100, external
    *via 192.168.98.5, Eth3/22.112, [170/51456], 18:34:47, eigrp-100, external
```

```
DC1-CORE# sh cdp nei int e3/22
Capability Codes: R - Router, T - Trans-Bridge, B - Source-Route-Bridge
                  S - Switch, H - Host, I - IGMP, r - Repeater,
                  V - VoIP-Phone, D - Remotely-Managed-Device,
                  s - Supports-STP-Dispute

Device-ID          Local Intrfce  Hldtme Capability  Platform        Port ID
DC1-AGG2(JAF1534CHCJ)
                    Eth3/22        172    R S s     N7K-C7009        Eth3/7
```

Since the DC1-Agg2 doesn't have any entries in the URIB, the packet will be gleaned and sent to CPU which would force the DC1-Agg2 to generate an ARP request from the SVI IP address as shown below.

```
2020-02-18 15:09:05.673165 172.17.200.1 -> 172.16.144.119 ICMP 114 Echo (ping) request
id=0x0022, seq=0/0, ttl=254
2020-02-18 15:09:05.675041 de:ad:20:19:22:22 -> Broadcast ARP 60 Who has 172.16.144.119? Tell
172.16.144.251
```

This ARP request is a broadcast and it propagates in the entire Layer 2 domain which also includes the DC2 via the OTV Extension.

DC2-Silent Host now responds to the ARP request from DC1-Agg2

DC1-Agg2 receives this ARP reply from the silent host

```
2020-02-18 15:09:05.675797 64:12:25:97:46:41 -> de:ad:20:19:22:22 ARP 60 172.16.144.119 is at 
64:12:25:97:46:41
```

SInce the packet received was ARP(which serves as a hint to LISP), an ICMP ECHO Request is generated sourcing from the HSRP VIP-> 172.16.144.254 and destined to the silent host-> 172.16.144.119. The intention of sourcing the packet from the HSRP VIP Is to understand if the host is local or remote. If the Host is remote, then the FHRP Active is also present in the remote Datacenter which would catch the ICMP ECHO Reply packet from the Host and so this results in DC2-Agg2(which is the HSRP Active) to learn about this entry and LISP process will now make an EID Learn based on this IP Packet. The DC1-Agg2 which originally sourced the ICMP ECHO Request from the HSRP VIP never gets a response and so End point learning will never happen on DC1 side; Rather it will be on the DC2 side.

```
DC2-AGG2# show lisp dynamic-eid detail vrf tenant-1
LISP Dynamic EID Information for VRF "tenant-1"
Dynamic-EID name: VLAN144
  Database-mapping [2] EID-prefix: 172.16.144.0/24, LSBs: 0x00000003
    Locator: 10.10.20.1, priority: 50, weight: 50
             Uptime: 21:50:32, state: up
    Locator: 10.10.20.2, priority: 50, weight: 50
             Uptime: 21:50:13, state: up, local
  Registering more-specific dynamic-EIDs
  Registering routes: disabled
  Allowed-list filter: none applied
  Map-Server(s): none configured, use global Map-Server
  Site-based multicast Map-Notify group: 239.254.254.254
  Extended Subnet Mode configured on 1 interfaces
  Number of roaming dynamic-EIDs discovered: 3
  Last dynamic-EID discovered: 172.16.144.254, 00:01:10 ago
  Roaming dynamic-EIDs:
    172.16.144.2, Vlan144, uptime: 19:09:07, last activity: 00:05:21
      Discovered by: packet reception
    172.16.144.119, Vlan144, uptime: 00:05:55, last activity: 00:05:55
      Discovered by: packet reception
    172.16.144.252, Vlan144, uptime: 3d21h, last activity: 00:01:10
      Discovered by: packet reception

Secure-handoff pending for sources: none
```

Once the LISP process is aware of the EID on DC2-Agg2(HSRP Active), it will

a) Install a /32 locally

b) Redistribute the route to DC2-Core

c) Send a Site based notify as a multicast message in the Vlan(In this example, message will be destined to the group -> 239.254.254.254)

```
DC2-AGG1# show lisp dynamic-eid detail vrf tenant-1
LISP Dynamic EID Information for VRF "tenant-1"
Dynamic-EID name: VLAN144
  Database-mapping [2] EID-prefix: 172.16.144.0/24, LSBs: 0x00000003
    Locator: 10.10.20.1, priority: 50, weight: 50
             Uptime: 21:52:39, state: up, local
    Locator: 10.10.20.2, priority: 50, weight: 50
             Uptime: 21:52:08, state: up
```

```
  Registering more-specific dynamic-EIDs
  Registering routes: disabled
  Allowed-list filter: none applied
  Map-Server(s): none configured, use global Map-Server
  Site-based multicast Map-Notify group: 239.254.254.254
  Extended Subnet Mode configured on 1 interfaces
  Number of roaming dynamic-EIDs discovered: 4
  Last dynamic-EID discovered: 172.16.144.254, 00:03:07 ago
  Roaming dynamic-EIDs:
    172.16.144.2, Vlan144, uptime: 19:11:04, last activity: 00:00:21
      Discovered by: site-based Map-Notify
    172.16.144.110, Vlan144, uptime: 20:04:09, last activity: 20:04:09
      Discovered by: site-based Map-Notify
    172.16.144.119, Vlan144, uptime: 00:07:52, last activity: 00:00:21
      Discovered by: site-based Map-Notify
    172.16.144.252, Vlan144, uptime: 21:50:51, last activity: 00:00:21
      Discovered by: site-based Map-Notify

Secure-handoff pending for sources: none
```

At the end, the Branch-router1 will receive this /32 route which will result in the Branch router to send the traffic to the right DC2-core switch.

```
Branch1-Router# sh ip route 172.16.144.119
Routing entry for 172.16.144.119/32
  Known via "eigrp 100", distance 170, metric 51712, type external
  Redistributing via eigrp 100
  Last update from 192.168.99.2 on GigabitEthernet0/0/1, 00:06:25 ago
  Routing Descriptor Blocks:
  * 192.168.99.2, from 192.168.99.2, 00:06:25 ago, via GigabitEthernet0/0/1
      Route metric is 51712, traffic share count is 1
      Total delay is 1020 microseconds, minimum bandwidth is 100000 Kbit
      Reliability 255/255, minimum MTU 1492 bytes
      Loading 1/255, Hops 2
```

Sequence of Events when a host moves(Roam) from DC1 to DC2

Considering that L2 Extension is configured on this topology, a Host can move from DC1 to DC2.

Host-> 172.16.144.100 is in Vlan 144 and in DC1 initially.

The Route within DC1-Agg1 and DC1-Agg2 switches will be as following when the host is online in DC1

```
DC1-AGG1# sh ip route 172.16.144.100 vrf tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.100/32, ubest/mbest: 1/0, attached
    *via 172.16.144.100, Vlan144, [240/1], 00:05:03, lisp, dyn-eid
     via 172.16.144.100, Vlan144, [250/0], 00:05:05, am

DC1-AGG2# sh ip route 172.16.144.100 vrf tenant-1
```

```
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.100/32, ubest/mbest: 1/0, attached
    *via 172.16.144.100, Vlan144, [240/1], 00:08:05, lisp, dyn-eid
     via 172.16.144.100, Vlan144, [250/0], 00:08:07, am
```

 # A branch router will have the route pointing to the DC1-Core as below and a traceroute would point the DC1 Core/agg switches to reach the host which is in DC1

```
Branch1-Router#sh ip route 172.16.144.100
Routing entry for 172.16.144.100/32
  Known via "eigrp 100", distance 170, metric 51712, type external
  Redistributing via eigrp 100
  Last update from 192.168.99.1 on GigabitEthernet0/0/1, 00:00:06 ago
  Routing Descriptor Blocks:
  * 192.168.99.1, from 192.168.99.1, 00:00:06 ago, via GigabitEthernet0/0/1
      Route metric is 51712, traffic share count is 1
      Total delay is 1020 microseconds, minimum bandwidth is 100000 Kbit
      Reliability 255/255, minimum MTU 1492 bytes
      Loading 1/255, Hops 2

Branch1-Router#traceroute 172.16.144.100 source 172.17.200.1
Type escape sequence to abort.
Tracing the route to 172.16.144.100
VRF info: (vrf in name/id, vrf out name/id)
  1 192.168.99.1 1 msec 1 msec 0 msec                    # DC1-Core
  2 192.168.98.5 1 msec 1 msec                           # DC1-Agg2
    192.168.98.1 1 msec                                  # DC1-Agg1
  3 172.16.144.100 1 msec 0 msec 1 msec                  # DC1-Host
```

When the host Moves to DC2, it would send a GARP out in the Vlan 144. This would be seen at the DC2-Agg switches

```
2020-02-24 22:23:05.024902 Cisco_5a:4a:e7 -> Broadcast    ARP 60 Gratuitous ARP for
172.16.144.100 (Request)
```

As soon as a packet is received with the ARP/GARP/RARP, it triggers the localization mechanism to originate an ICMP Echo request out to the host sourced from the VIP

```
2020-02-24 22:23:05.026781 172.16.144.254 -> 172.16.144.100 ICMP 60 Echo (ping) request
id=0xac10, seq=0/0, ttl=128
```

Host-172.16.144.100 will now respond to the HSRP VIP

```
2020-02-24 22:23:07.035292 172.16.144.100 -> 172.16.144.254 ICMP 60 Echo (ping) reply
id=0xac10, seq=0/0, ttl=255
```

As soon as the IP Packet is received at the DC2-Agg1, this would result in LISP detecting the EID and making an entry in the Routing table for the Host and starts the redistribution process to the EIGRP

```
DC2-AGG1# sh ip route 172.16.144.100 vrf tenant-1
```

```
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.100/32, ubest/mbest: 1/0, attached
    *via 172.16.144.100, Vlan144, [240/1], 00:00:30, lisp, dyn-eid
     via 172.16.144.100, Vlan144, [250/0], 00:00:32, am
```

With the redistribution in place, the DC1-agg site(Which was the original Owner of this Host), would now see the Change in the RIB pointing to the EIGRP

```
DC1-AGG1#  sh ip route 172.16.144.100 vrf tenant-1
IP Route Table for VRF "tenant-1"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.144.100/32, ubest/mbest: 1/0
    *via 192.168.98.2, Eth3/6.111, [245/51968], 00:03:47, eigrp-100, external
```

A remote Branch Router will now see the route change and the traceroutes will reflect the path change to the DC2 core/Agg switches as shown below

```
Branch1-Router#sh ip route 172.16.144.100
Routing entry for 172.16.144.100/32
  Known via "eigrp 100", distance 170, metric 51712, type external
  Redistributing via eigrp 100
  Last update from 192.168.99.2 on GigabitEthernet0/0/1, 00:00:00 ago
  Routing Descriptor Blocks:
  * 192.168.99.2, from 192.168.99.2, 00:00:00 ago, via GigabitEthernet0/0/1
      Route metric is 51712, traffic share count is 1
      Total delay is 1020 microseconds, minimum bandwidth is 100000 Kbit
      Reliability 255/255, minimum MTU 1492 bytes
      Loading 1/255, Hops 2

Branch1-Router#traceroute 172.16.144.100 source 172.17.200.1
Type escape sequence to abort.
Tracing the route to 172.16.144.100
VRF info: (vrf in name/id, vrf out name/id)
  1 192.168.99.2 1 msec 0 msec 1 msec                   # DC2-Core
  2 192.168.94.1 1 msec 1 msec 1 msec                   # DC2-Agg1
  3 172.16.144.100 0 msec 0 msec 1 msec                 # Host-after move to DC2
```

Useful Verification Commands

show lisp dynamic-eid detail vrf <VRF Name>

Show ip route lisp vrf <VRF Name>

show lisp dynamic-eid summary vrf <VRF Name>