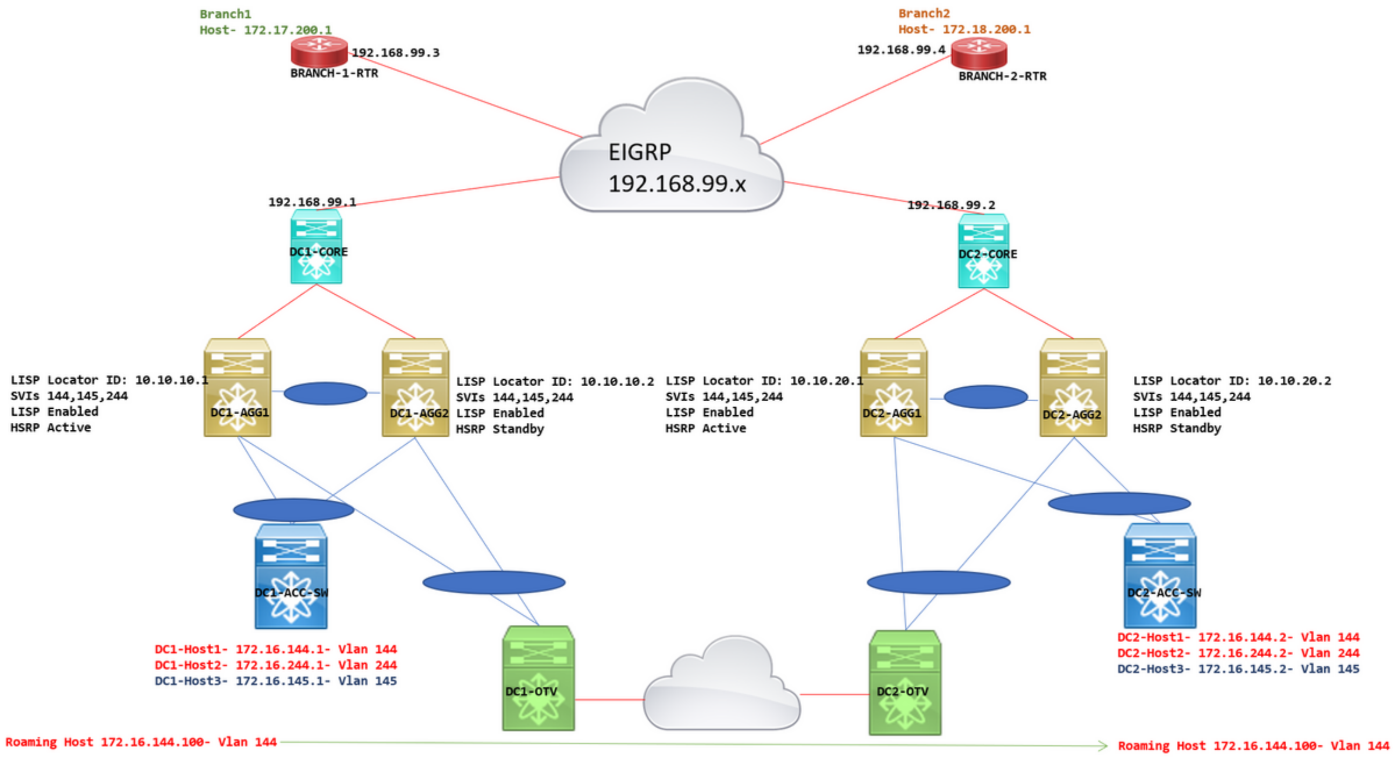
Nexus 7000配置和验证LISP IGP辅助扩展子网模式

目录

简介

本文档将介绍如何使用Nexus 7000部署LISP IGP辅助扩展子网模式(ESM)

拓扑



拓扑详细信息

- DC1和DC2是OTV扩展的两个位置
- VLAN 144、145和244在所有Agg、接入层和OTV交换机上配置
- 这些VLAN的SVI在Agg交换机上配置。SVI 144和244位于VRF租户-1中；SVI 145在VRF租户2中。
- 在部署LISP IGP Assist时，SVI不必位于VRF中；本示例使用多个VRF只是为了说明所需的配置更改（在每个相关VRF环境下）；所有SVI都可以位于同一VRF中，并且仍然可以使用LISP IGP辅助
- HSRP在Vlan144、145和244中配置；此拓扑中配置了FHRP隔离，这意味着总共4台交换机将运行HSRP，并且两端都有主用/备用对。FHRP隔离通过过滤HSRP Hello消息来实现。
- DC1-agg1和DC2-Agg2是vPC对；DC2-Agg1和DC2-Agg2也适用
- LISP配置在SVI 144、145和244下应用
- 每个VRF从Agg建立到核心交换机的EIGRP邻居关系。子接口从每个VRF的Agg交换机运行到核心交换机，并且EIGRP邻居关系形成在这些子接口上。
- 远程路由器（分支）也属于同一IGP域。
- 使用LISP IGP Assist时，没有LISP封装/解封，因此LISP路由必须重分发到IGP（此处为EIGRP）。 对于本文档中描述的此部署模式，分支路由器将没有任何LISP配置。

使用的组件

- Nexus 7000是核心交换机，带SUP2E，F3/M3运行8.2(4)NXOS版本
- 分支路由器是ASR1k
- OTV在这些Nexus 7000交换机的另一个VDC中配置；OTV和LISP必须位于不同的VDC上。共享VDC不是选项。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您使用的是真实网络，请确保您已经了解所有命令的潜在影响。

AGG交换机上所需的配置

DC1-Agg1和DC1-Agg2上的LISP特定配置

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
```

#数据库映射的提供方式必须是，在一端，DC1-Agg1和DC1-Agg2环回IP地址都需要指定；在DC2-Agg1和DC2-Agg2中，必须创建一个唯一环回，并将其放在数据库映射中。

#在IGP辅助模式下，如果使用配置 — > "ip lisp itr-etr"，则会为未启用LISP的VLAN注入/32 null0主机路由；因此，IGP辅助模式的正确配置是“ip lisp etr”。

DC2-Agg1和DC2-Agg2上的LISP特定配置

```
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
```

DC1和DC2 Agg LISP配置之间的区别是“数据库映射”中定义的环回。 在DC1配置中，这将使用DC1-Agg1和DC1-Agg2的环回定义，对于DC2，数据库映射将使用DC2-Agg1和DC2-Agg2中的环回定义

#下面显示的其余IGP/路由映射/前缀列表配置将相似（为接口分配的IP地址确实不同）

IGP特定

```
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
```

#启用LISP的AGG VDC还将与核心端形成IGP邻居关系

#在本示例中，作为每个租户VRF一部分的子接口用于形成与核心的邻居关系，如下所示。

```
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
```

路由映射/前缀列表

```
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
```

#所有AGG交换机（DC1和DC2）都需要上述所有配置。 请记住，为SVI、环回、HSRP VIP提供唯一的IP地址对于所有SVI都是相同的

OTV VDC配置

HSRP过滤

#对于IGP辅助部署，当通过OTV或任何其他机制进行扩展时，必须进行FHRP隔离；

#这通过过滤OTV VDC中的FHRP Hello消息来完成

#在本示例中，使用N7k OTV，因此应用了以下配置来过滤OTV VDC中的FHRP数据包。

```
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
```

#仅OTV VDC上需要FHRP过滤配置；如果使用ASR OTV部署，则过滤机制应作为相关机制使用，并根据ASR配置指南进行记录。

OTV抑制ARP

#在OTV VDC上禁用ARP ND缓存功能

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

由于LISP配置导致的路由填充

```
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
```

#在SVI 144上启用LISP后，将自动创建两条Null0路由；SVI 144是/24子网，因此第1条null0路由来自172.16.144.0/25，第二条路由来自172.16.144.128/25，如上所示。

#这是预期的，而且是按设计的；这样做是为了确保源自未发现主机的数据包触发RPF异常，该异常将导致数据包被传送到CPU，并最终有助于主机检测(EID)

当主机在启用LISP的SVI内联机时的事件顺序

#启用LISP的接口上的主机检测基于从数据库映射配置中指定的范围内的IP地址接收的L3流量。

为便于检测主机，请注意，当在接口上启用LISP时：
#接口上启用了RPF异常，因此未知源生成的数据包会触发异常
#源于Null0的LISP路由已安装，以确保未知源触发RPF异常

由于此解决方案依赖OTV在两个数据中心之间进行L2扩展，因此ARP信令不能直接用于检测IP主机，因为在许多情况下，ARP信令会广播到所有交换机。

但是，ARP信号用作LISP指示可能存在未检测到的主机。由于主机可驻留在OTV网桥的任一端，因此LISP在学习新的IP-MAC绑定后启动本地化机制。
定位机制的工作原理如下：
#交换机获知新的IP-MAC绑定（通过GARP、RARP或ARP请求）。
#作为活动HSRP工作的交换机向主机发送回应请求，但源自HSRP VIP地址
#主机响应回应请求，但在OTV中进行FHRP隔离后，仅在主机所在的DC站点上收到回应应答
#由于回应应答是L3数据包，因此LISP会检测到主机。

#如果在任何启用LISP的SVI上收到IP数据包，则它本身将向LISP进程提供通知终端为本地的信息；不会发送任何ICMP ECHO请求，以进一步确认主机是否是本地主机。因此，务必注意，从DC2主机对DC1-AGG SVI IP地址执行Ping操作会导致端点标识损坏，这也可能导致ping丢失或流量黑洞，因为主机现在被标识为DC1中的本地EID，而不是DC2。因此，ping不应从LISP中的SVI IP地址发起环境，因为这可能损坏路由表，并导致流量黑洞。如果LISP已启用VLAN的主机尝试ping SVI IP地址，则会出现同样的问题；对VIP执行ping操作应该正常，因为两端存在相同且处于活动状态，并且本地站点将捕获数据包。

主机在DC1上线时的路由表条目示例如下：

```
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
```

#如上所示，有两条路由；AD为250的AM->邻接管理器（由ARP进程填充）使用管理距离为240的LISP进程。

DC1中的两台Agg交换机将具有相同的条目。

#此外，LISP将在动态EID表中列出主机的相同条目，如下所示。

```
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
```

#发现在这两种情况下都不同；HSRP活动的DC1-AGG1通过“数据包接收”记录条目，这基本上意味着传入的数据包导致将其添加为EID

Agg1得知EID后，会从源IP-> Loopback100 IP地址（在数据库映射下定义）向组 — > 239.254.254.254（在上面配置）发送组播消息，vPC对等交换机也会收到该消息并相应地填充条目由于数据库映射同时具有dc1-agg1和dc1-agg2的IP地址，因此将其视为本地EID。此相同的组播数据包也会通过OTV传输到远程站点；但是，远程站点将检查数据库映射，由于此数据包来自与“数据库映射”不同的IP地址，因此DC2 AGg交换机不会将其视为本地EID。

映射通知消息

#当启用LISP的SVI检测到主机时，触发的“映射通知”消息将发送到在相应动态EID配置下定义的组播组

#除触发的映射通知消息外，该VLAN中有由HSRP活动（或FHRP活动）交换机发送的定期映射通知消息；

#映射通知消息的PCAP如下所示。

```
> 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
```

将LISP /32路由重分发到IGP

#这是IGP辅助模式的关键；任何/32 LISP路由都将重分发到IGP;这可通过在EIGRP下应用的“redistribute LISP”命令实现。

#在重分发后，任何/32主机路由都将被视为EIGRP外部路由。调整EIGRP管理距离以使其更高。这是为了确保LISP路由与传入EIGRP外部路由保持在URIB中。例如：DC1-Agg1和DC1-Agg2是具有DC1核心的EIGRP邻居。DC1-AGG1通过重分发将/32路由注入DC1-Core。既然DC1-Core是与DC1-Agg2的EIGRP邻居，那么如果EIGRP AD是170，则同一路由可能返回到DC1-Agg2，并且有机会赢取LISP路由（AD为240）；因此，为避免这种情况，EIGRP外部路由AD已修改为245。

DC1-Agg交换机获知的/32路由将重分发到EIGRP，DC1-core条目如下所示。

```
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
```

#该路由存在于全局路由表中，且核心端未配置VRF。

#由于在AGG交换机上配置了“redistribute direct”，因此核心层还将为父子网提供/24 ECMP路由，如下所示。这将有助于吸引静默主机（没有/32路由）的流量。

```
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
```

#此外，DC1和DC2核心都会看到/24 ECMP路由

```
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
```

#此路由可确保分支主机可以连接到位于任一位置的静默主机。

DC内VLAN间数据包的路径

#当DC1-Host1 -> 172.16.144.1尝试到达DC2-Host1-> 172.16.144.2时，这是数据中心间VLAN流量。DC1-Host 1发出ARP请求，该请求将一路穿过OTV并到达DC2-Host1

DC2-Host1以返回DC1-Host1的ARP应答作出响应

#后续ICMP数据包通过OTV发送

DC间VLAN间的数据包路径（从VLAN 144到VLAN 244）

#当DC1-Host1-> 172.16.144.1尝试到达DC2-Host2-> 172.16.244.2时，数据包不会从DC1的vlan 144路由到244;相反，它遵循从DC1-Agg到DC1-Core的路由路径，然后到达DC2-Core，最终路由将由DC2-Agg交换机完成到目的Vlan-244。

#从DC1-Host1到DC2-Host2的traceroute如下所示。

```
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
```

DC间VLAN间数据包的路径（从VRF租户–1到VRF租户–2）

#这将遵循与从一个VLAN到另一个VLAN的DC间通信相同的方法（上例）

#当DC1-host1-> 172.16.144.1尝试到达DC2-Host3-> 172.16.145.2时，这是发往Vlan 144（VRF租户–1）的DC间流量，发往Vlan 145（VRF租户–2）。 与常规N7k OTV部署不同，此流量的处理方式略有不同。DC1端不会发生任何VLAN间路由；相反，此流量将被路由并发送到DC1核心，而核心层将通过IGP进一步将其路由到DC2核心层

#为了本文档，核心交换机按站点执行VRF间泄漏。它可以是任何设备（如防火墙）；如果VRF间泄漏存在或不存在，则从LISP配置角度看没有更改。

```
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
```

#从DC1-Host1到DC2-Host3的Traceroute将显示与其未通过VLAN间路由的路由相同，而是通过核心路由的第3层。简而言之，VLAN间流量将不使用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#
```

当Branch-1主机尝试到达DC2中存在的静默主机时数据包的路径

Branch-1-172.17.200.1中的主机尝试到达DC2-Silent Host- 172.16.144.119。由于主机处于静默状态，因此DC2中不会存在任何/32路由。

```
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
```

#根据LISP设计，路由172.16.144.119将与172.16.144.0/25 null0路由匹配。

#当Branch路由器收到目的IP = 172.16.144.119的数据包时，URIB具有到DC1核和DC2核的ECMP /24路由。这实质上意味着数据包将发送到其中一台核心层交换机。

```
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
```

#根据CEF，数据包散列到192.168.99.1（即DC1-Core）

DC1-Core有2条ECMP路径；一个指向DC1-Agg1（HSRP活动），另一个指向DC1-Agg2（HSRP备用）。 从路由散列中，选择的路径将是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
```

#由于DC1-Agg2在URIB中没有任何条目，因此数据包将被收集并发送到CPU，这将强制DC1-Agg2从SVI IP地址生成ARP请求，如下所示。

```
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
```

#此ARP请求是广播，它在整个第2层域内传播，该域还包括通过OTV扩展的DC2。

DC2-Silent Host现在响应来自DC1-Agg2的ARP请求

DC1-Agg2从静默主机收到此ARP应答

```
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
```

#由于收到的数据包是ARP（用作LISP的提示），因此会从HSRP VIP-> 172.16.144.254生成ICMP ECHO请求，该请求发往静默主机> 172.16.144.119。从HSRP VIP获取数据包的目的是了解主机是本地还是远程。如果主机是远程的，则FHRP活动也存在于远程数据中心中，该数据中心将从主机捕获ICMP ECHO应答数据包，因此这会导致DC2-Agg2（即HSRP活动）了解此条目，LISP进程现在将基于此IP数据包创建EID学习。最初从HSRP VIP发出ICMP ECHO请求的DC1-Agg2从未收到响应，因此DC1端永远不会发生终端学习；而是在DC2侧。

```
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
```

#一旦LISP进程知道DC2-Agg2（HSRP活动）上的EID，它将

a)在本地安装/32

b)重分布到DC2-Core的路由

c)在Vlan中以组播消息的形式发送基于站点的通知（在本例中，消息将发往组 — > 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
```

#最后，Branch-router1将收到此/32路由，这将导致Branch路由器将流量发送到正确的DC2核心交换机。

```
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
```

主机从DC1移动（漫游）到DC2时的事件顺序

#考虑到此拓扑上配置了L2扩展，主机可以从DC1移动到DC2。

Host-> 172.16.144.100最初在VLAN 144和DC1中。

#当主机在DC1上线时，DC1-Agg1和DC1-Agg2交换机内的路由将如下所示

```
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
```

#分支路由器将具有指向DC1-Core的路由，如下所示，而traceroute将指向DC1 Core/agg交换机以到达DC1中的主机

```
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
```

#当主机移动到DC2时，它会在Vlan 144中发送GARP。在DC2-Agg交换机上会看到此情况

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

#一旦收到包含ARP/GARP/RARP的数据包，就会触发本地化机制，向源自VIP的主机发出ICMP回应请求

```
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
```

#主机172.16.144.100现在将响应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
```

#一旦在DC2-Agg1处收到IP数据包，这将导致LISP检测EID并在主机的路由表中输入条目，并启动到EIGRP的重分发过程

```
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
```

#当重分发到位后，DC1-agg站点（该站点是此主机的原始所有者）现在将看到指向EIGRP的RIB中的更改

```
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
```

#远程分支路由器现在将看到路由更改，traceroutes将反映到DC2核心/Agg交换机的路径更改，如下所示

```
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
```

有用的验证命令

show lisp dynamic-eid detail vrf <VRF Name>

Show ip route lisp vrf <VRF Name>

show lisp dynamic-eid summary vrf <VRF Name>