配置 VXLAN

目录

简介

先决条件

- 要求
- 使用的组件

背景信息

- 术语
- 什么是VXLAN
- 为什么选择VXLANBGP EVPN
- BUM复制选项

配置

- 网络图
- 编址
- 配置
  - 主干–1
  - 主干–2
  - 枝叶–1(vPC VTEP)
  - 枝叶–2(vPC VTEP)
  - 枝叶–3（独立VTEP）
  - 可选：入口复制而不是组播

验证

- 底层
- 重叠控制平面
- NVE和VNI
- MAC/IP学习
- L3 VNI路由
- vPC
- 数据包格式

故障排除

相关信息

简介

本文档介绍如何在主干 — 枝叶交换矩阵中的Cisco Nexus 9000系列交换机上使用BGP EVPN控制平面配置和验证VXLAN。

先决条件

要求

Cisco 建议您了解以下主题：

- 虚拟端口通道(vPC)概念
- IP路由（OSPF或IS-IS）和BGP，包括路由反射器
- 组播路由概念，例如协议无关组播(PIM)稀疏模式和交汇点(RP)

使用的组件

本文档中的信息基于以下软件和硬件版本：

- Network Advantage许可证
- 思科NX-OS版本10.4(x)
- 2个Cisco Nexus N9K-C9332D-GX2B作为主干（BGP路由反射器和PIM任播RP）
- 3个Cisco Nexus N9K-C93180YC-FX3作为枝叶(VTEP);枝叶1和枝叶2是一对vPC

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。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。



注意：本文档中的配置适用于独立NX-OS模式下的Nexus 9000云扩展平台（ — EX、-FX、-FX2、-FX3、-GX、-GX2及更高版本）。它不适用于ACI模式下的Nexus 9000交换机。

背景信息

术语

期限	描述
VXLAN	虚拟可扩展LAN:在具有MAC-in-UDP封装

	（UDP端口4789）的路由IP底层上提供第2层和第3层重叠服务。
VNI	VXLAN网络标识符:24位网段ID。L2 VNI映射到VLAN（网桥域）。 L3 VNI映射到VRF。
VTEP	VXLAN隧道终端:封装和解封VXLAN流量的设备。
NVE	网络虚拟化边缘接口（接口nve1）：发生封装和解封的逻辑接口。
BGP EVPN	在VTEP之间通告MAC、MAC/IP（路由类型2）、IP前缀（路由类型5）和VTEP成员（路由类型3）的控制平面。
AnyCast网关	在每个枝叶上使用相同的SVI IP和MAC地址，并且默认网关始终是本地枝叶。
BUM	广播、未知单播和组播流量。

什么是VXLAN

- VXLAN在任何路由IP底层之上重叠第2层和第3层网段。
- 它使用MAC-in-UDP封装，为原始帧增加50字节的开销。
- 与传统VLAN的4094相比，它最多支持1600万个网段。
- 底层对所有链路使用等价多路径(ECMP)，枝叶和主干之间没有生成树协议。

为什么选择VXLAN BGP EVPN

本文档的早期版本使用flood-and-learn，其中VTEP从数据平面泛洪获取远程MAC地址。BGP EVPN用控制平面代替数据平面学习，是推荐的部署模式：

- VTEP通过BGP相互发现并学习远程MAC和IP地址，从而减少泛洪。
- ARP抑制允许本地VTEP应答已知主机的ARP请求。
- 分布式任播网关和每个VRF的第3层VNI在第一跳枝叶的子网之间提供最佳路由。

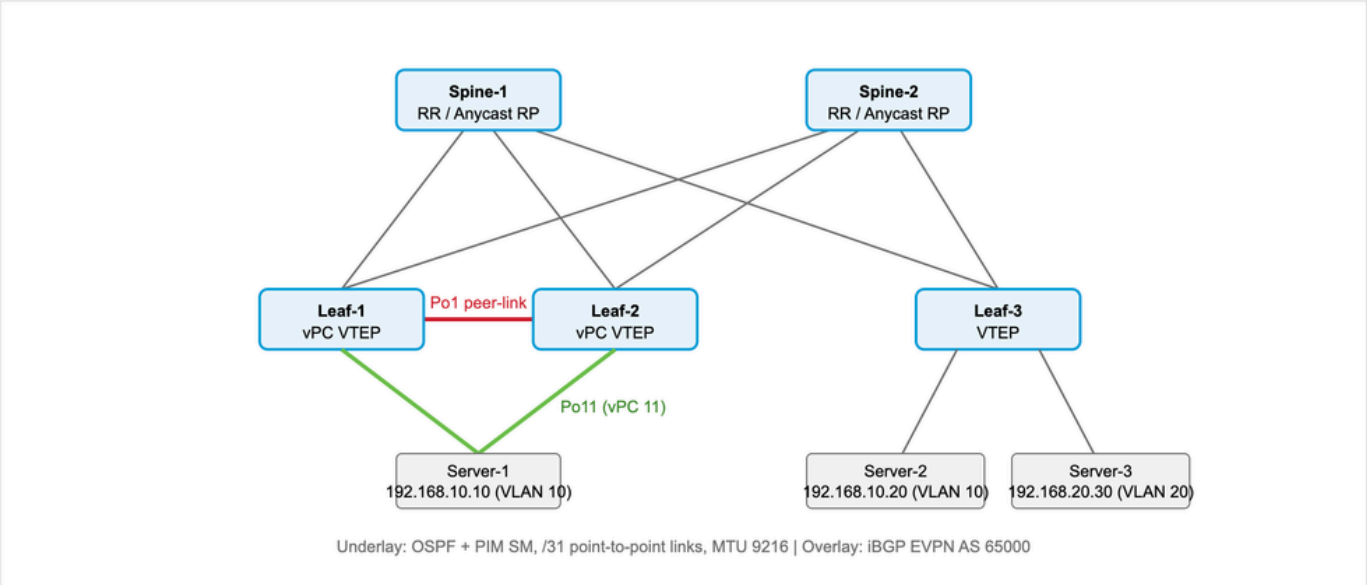
- 支持多租户、多宿主（vPC或EVPN ESI）和多站点扩展。

BUM复制选项

模式	配置	衬底要求
组播（用于本文档）	vni下的mcast-group <group>	所有底层链路和环回上的PIM稀疏模式，以及RP（主干上的任播RP很常见）
入口复制	vni下的入口复制协议bgp	仅单播。入口VTEP向BGP路由类型3为VNI获取的每个远程VTEP发送副本。

配置

网络图



编址

设备	Loopback0（路由、BGP、PIM）	Loopback1（NVE源）	上行链路（Eth1/49到Spine-1,Eth1/50到Spine-2）
主干–1	10.0.0.1	-	Eth1/1-3到Leaf-1/2/3:10.1.1.0、.2、.4/31
主干–2	10.0.0.2	-	Eth1/1-3到Leaf-1/2/3:10.1.2.0、.2、.4/31
枝叶1	10.0.0.11	10.0.1.11，辅助10.0.1.112	10.1.1.1/31，10.1.2.1/31
枝叶–2	10.0.0.12	10.0.1.12，辅助10.0.1.112	10.1.1.3/31，10.1.2.3/31
枝叶–3	10.0.0.13	10.0.1.13	10.1.1.5/31，10.1.2.5/31

覆盖对象	价值
任播RP	10.254.254.1（两个主干上的环回254），组239.1.1.0/24
VLAN 10 / L2 VNI	10010，组播组239.1.1.10，网关192.168.10.1/24
VLAN 20 / L2 VNI	10020，组播组239.1.1.20，网关192.168.20.1/24
VRF租户 — A/L3 VNI	50000(VLAN 500)
任播网关MAC	2020.0000.00aa
vPC备份底层SVI	VLAN 3900、10.255.255.0/31(Leaf-1)和10.255.255.1/31(Leaf-2)



注意：为NVE源(Loopback1)、路由和BGP对等(Loopback0)使用单独的环回。 在vPC VTEP上，NVE源环回具有两个对等体上相同的辅助IP。远程VTEP使用此共享地址

（vPC任播VTEP）作为单个VTEP到达vPC对。

配置

这些配置仅显示与底层和VXLAN相关的部分。根据要求配置管理、AAA和其他功能。

主干–1

```
nv overlay evpn
feature ospf
feature bgp
feature pim
feature nv overlay

ip pim rp-address 10.254.254.1 group-list 239.1.1.0/24
ip pim anycast-rp 10.254.254.1 10.0.0.1
ip pim anycast-rp 10.254.254.1 10.0.0.2

interface loopback0
  description Routing / BGP / Anycast-RP member
  ip address 10.0.0.1/32
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

interface loopback254
  description Anycast-RP
  ip address 10.254.254.1/32
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

interface Ethernet1/1
  description to Leaf-1 Eth1/49
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.1.0/31
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet1/2
  description to Leaf-2 Eth1/49
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.1.2/31
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```

```
interface Ethernet1/3
  description to Leaf-3 Eth1/49
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.1.4/31
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown

router ospf UNDERLAY
  router-id 10.0.0.1

router bgp 65000
  router-id 10.0.0.1
  address-family l2vpn evpn
  template peer VTEP
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
      route-reflector-client
  neighbor 10.0.0.11
    inherit peer VTEP
  neighbor 10.0.0.12
    inherit peer VTEP
  neighbor 10.0.0.13
    inherit peer VTEP
```

主干–2

```
interface loopback0
  ip address 10.0.0.2/32
interface Ethernet1/1
  ip address 10.1.2.0/31
interface Ethernet1/2
  ip address 10.1.2.2/31
interface Ethernet1/3
  ip address 10.1.2.4/31
router ospf UNDERLAY
  router-id 10.0.0.2
router bgp 65000
  router-id 10.0.0.2
```

枝叶–1(vPC VTEP)

```
nv overlay evpn
feature ospf
```

```
feature bgp
feature pim
feature fabric forwarding
feature interface-vlan
feature vn-segment-vlan-based
feature lacp
feature vpc
feature nv overlay

fabric forwarding anycast-gateway-mac 2020.0000.00aa
ip pim rp-address 10.254.254.1 group-list 239.1.1.0/24

vlan 10
  vn-segment 10010
vlan 20
  vn-segment 10020
vlan 500
  name L3VNI-Tenant-A
  vn-segment 50000
vlan 3900
  name vPC-backup-underlay

system nve infra-vlans 3900

vrf context Tenant-A
  vni 50000
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn

vpc domain 10
  peer-switch
  role priority 10
  peer-keepalive destination 172.16.0.2 source 172.16.0.1 vrf management
  delay restore 150
  peer-gateway
  ip arp synchronize
  ipv6 nd synchronize

interface port-channel1
  description vPC peer-link
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10,20,500,3900
  spanning-tree port type network
  vpc peer-link

interface port-channel11
  description Server-1
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10,20
  spanning-tree port type edge trunk
  mtu 9216
  vpc 11

interface Vlan10
  no shutdown
  vrf member Tenant-A
  no ip redirects
  ip address 192.168.10.1/24
```

```
  fabric forwarding mode anycast-gateway

interface Vlan20
  no shutdown
  vrf member Tenant-A
  no ip redirects
  ip address 192.168.20.1/24
  fabric forwarding mode anycast-gateway

interface Vlan500
  description L3VNI Tenant-A
  no shutdown
  vrf member Tenant-A
  no ip redirects
  ip forward

interface Vlan3900
  description vPC backup underlay
  no shutdown
  mtu 9216
  no ip redirects
  ip address 10.255.255.0/31
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

interface nve1
  no shutdown
  host-reachability protocol bgp
  advertise virtual-rmac
  source-interface loopback1
  member vni 10010
    suppress-arp
    mcast-group 239.1.1.10
  member vni 10020
    suppress-arp
    mcast-group 239.1.1.20
  member vni 50000 associate-vrf

interface Ethernet1/1
  description Server-1
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10,20
  mtu 9216
  channel-group 11 mode active
  no shutdown

interface Ethernet1/53-54
  description vPC peer-link to Leaf-2
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10,20,500,3900
  channel-group 1 mode active
  no shutdown

interface Ethernet1/49
  description to Spine-1 Eth1/1
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.1.1/31
```

```
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet1/50
  description to Spine-2 Eth1/1
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.2.1/31
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface loopback0
  description Routing / BGP
  ip address 10.0.0.11/32
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

interface loopback1
  description NVE source
  ip address 10.0.1.11/32
  ip address 10.0.1.112/32 secondary
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

router ospf UNDERLAY
  router-id 10.0.0.11

router bgp 65000
  router-id 10.0.0.11
  address-family l2vpn evpn
    advertise-pip
  template peer SPINE-RR
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
  neighbor 10.0.0.1
    inherit peer SPINE-RR
  neighbor 10.0.0.2
    inherit peer SPINE-RR
  vrf Tenant-A
    address-family ipv4 unicast
      advertise l2vpn evpn

evpn
  vni 10010 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 10020 l2
    rd auto
    route-target import auto
    route-target export auto
```

注意：在云扩展vPC VTEP上，备份基础SVI(VLAN 3900)必须列在系统四个基础vlan中。此VLAN必须在默认VRF中，不能有vn网段，并且必须在对等链路上允许。当一个vPC对等体丢失其所有上行链路时，它通过孤立端口的对等链路传输VXLAN流量。请勿配置ID为512的独立VLAN，例如10和522。

从vPC对等体通告-pip和通告虚拟rmac，后者使用主IP(PIP)和自己的路由器MAC通告类型5（前缀）路由和孤立主机路由。如果没有这些信息，只可将流向一个对等体后面的主机的流量发送到无法到达该对等体的对等体。

ARP抑制(suppress-arp)仅支持VNI，其中VTEP是该VNI的分布式任播网关。

枝叶–2(vPC VTEP)

```
vpc domain 10
  role priority 20
  peer-keepalive destination 172.16.0.1 source 172.16.0.2 vrf management

interface Vlan3900
  ip address 10.255.255.1/31

interface Ethernet1/49
  ip address 10.1.1.3/31
interface Ethernet1/50
  ip address 10.1.2.3/31

interface loopback0
  ip address 10.0.0.12/32
interface loopback1
  ip address 10.0.1.12/32
  ip address 10.0.1.112/32 secondary

router ospf UNDERLAY
  router-id 10.0.0.12
router bgp 65000
  router-id 10.0.0.12
```

警告：两个vPC对等体上的NVE配置（VNI、组播组、源接口、辅助IP）必须相同。不匹配是vPC类型1一致性故障，并且NVE接口或VNI被挂起。通过运行show vpc consistency-parameters global命令进行检查。

枝叶–3（独立VTEP）

```
nv overlay evpn
feature ospf
feature bgp
feature pim
feature fabric forwarding
feature interface-vlan
feature vn-segment-vlan-based
feature nv overlay

fabric forwarding anycast-gateway-mac 2020.0000.00aa
ip pim rp-address 10.254.254.1 group-list 239.1.1.0/24

vlan 10
  vn-segment 10010
vlan 20
  vn-segment 10020
vlan 500
  name L3VNI-Tenant-A
  vn-segment 50000

vrf context Tenant-A
  vni 50000
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn

interface Vlan10
  no shutdown
  vrf member Tenant-A
  no ip redirects
  ip address 192.168.10.1/24
  fabric forwarding mode anycast-gateway

interface Vlan20
  no shutdown
  vrf member Tenant-A
  no ip redirects
  ip address 192.168.20.1/24
  fabric forwarding mode anycast-gateway

interface Vlan500
  no shutdown
  vrf member Tenant-A
  no ip redirects
  ip forward

interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback1
  member vni 10010
    suppress-arp
    mcast-group 239.1.1.10
  member vni 10020
    suppress-arp
    mcast-group 239.1.1.20
  member vni 50000 associate-vrf

interface Ethernet1/1
  description Server-2
  switchport
```

```
  switchport access vlan 10
  spanning-tree port type edge
  no shutdown

interface Ethernet1/2
  description Server-3
  switchport
  switchport access vlan 20
  spanning-tree port type edge
  no shutdown

interface Ethernet1/49
  description to Spine-1 Eth1/3
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.1.5/31
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface Ethernet1/50
  description to Spine-2 Eth1/3
  no switchport
  mtu 9216
  medium p2p
  ip address 10.1.2.5/31
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown

interface loopback0
  ip address 10.0.0.13/32
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

interface loopback1
  description NVE source
  ip address 10.0.1.13/32
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

router ospf UNDERLAY
  router-id 10.0.0.13

router bgp 65000
  router-id 10.0.0.13
  template peer SPINE-RR
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
  neighbor 10.0.0.1
    inherit peer SPINE-RR
  neighbor 10.0.0.2
    inherit peer SPINE-RR
  vrf Tenant-A
    address-family ipv4 unicast
      advertise l2vpn evpn
```

```
evpn
  vni 10010 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 10020 l2
    rd auto
    route-target import auto
    route-target export auto
```

可选：入口复制而不是组播

如果底层不运行组播，请在每个VTEP上将mcast-group替换为基于BGP的入口复制。L2 VNI不需要PIM和RP配置。

```
interface nve1
  member vni 10010
    suppress-arp
    ingress-replication protocol bgp
  member vni 10020
    suppress-arp
    ingress-replication protocol bgp
```

验证

使用本部分可确认配置能否正常运行。

思科 CLI 分析器（仅适用于注册客户）支持某些 show 命令。使用Cisco CLI分析器查看show命令输出的分析。

区域	命令
底层	show ip ospf neighbors、show ip route 10.0.1.13、show ip pim neighbor、show ip pim rp
重叠控制平面	show bgp l2vpn evpn summary， show bgp l2vpn evpn

NVE	show nve interface nve1 detail、show nve peers、show nve vni
MAC/IP学习	show mac address-table vlan 10、show l2route evpn mac all、show l2route evpn mac-ip all、show ip arp suppression-cache detail
L3 VNI	show ip route vrf Tenant-A、show bgp l2vpn evpn route-type 2、show bgp l2vpn evpn route-type 5
组播复制	show ip mroute 239.1.1.10
vPC	show vpc、show vpc consistency-parameters global、show system nve infra-vlan

这些输出来自服务器–1、服务器–2和服务器–3之间发送流量后的稳定状态。

底层

```
Leaf-1# show ip pim rp
PIM RP Status Information for VRF "default"
BSR disabled
Auto-RP disabled
BSR RP Candidate policy: None
BSR RP policy: None
Auto-RP Announce policy: None
Auto-RP Discovery policy: None

Anycast-RP 10.254.254.1 members:
  10.0.0.1  10.0.0.2

RP: 10.254.254.1, (0),
 uptime: 1d01h   priority: 255,
 RP-source: (local),
 group ranges:
 239.1.1.0/24
```

重叠控制平面

```
Leaf-1# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.0.0.11, local AS number 65000
BGP table version is 58, L2VPN EVPN config peers 2, capable peers 2

Neighbor        V    AS    MsgRcvd    MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.0.1        4 65000       1520       1488       58    0    0 1d01h    9
10.0.0.2        4 65000       1519       1490       58    0    0 1d01h    9
```

NVE和VNI

```
Leaf-1# show nve interface nve1 detail
Interface: nve1, State: Up, encapsulation: VXLAN
 VPC Capability: VPC-VIP-Only [notified]
 Local Router MAC: 5254.0011.1b08
 Host Learning Mode: Control-Plane
 Source-Interface: loopback1 (primary: 10.0.1.11, secondary: 10.0.1.112)
 Source Interface State: Up
 Virtual RMAC Advertisement: Yes
 NVE Flags:
 Interface Handle: 0x49000001
 Source Interface hold-down-time: 180
 Source Interface hold-up-time: 30
 Remaining hold-down time: 0 seconds
 Virtual Router MAC: 0200.0a00.0170

Leaf-1# show nve peers
Interface Peer-IP                                 State LearnType Uptime   Router-Mac
--------- --------------------------------------  ----- --------- -------- -----------------
nve1      10.0.1.13                               Up    CP        1d01h    5254.0013.1b08

Leaf-1# show nve vni
Codes: CP - Control Plane        DP - Data Plane
       UC - Unconfigured         SA - Suppress ARP
       S-ND - Suppress ND
       SU - Suppress Unknown Unicast
       Xconn - Crossconnect
       MS-IR - Multisite Ingress Replication
       HYB - Hybrid IRB mode

Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      10010    239.1.1.10        Up    CP   L2 [10]            SA
nve1      10020    239.1.1.20        Up    CP   L2 [20]            SA
nve1      50000    n/a               Up    CP   L3 [Tenant-A]
```



注意：使用BGP EVPN时，当远程VTEP通告本地配置的VNI的路由(LearnType CP)时，会出现NVE对等体。 Leaf-3将vPC对视为一个对等体，即AnyCast VTEP 10.0.1.112。通过advertise-pip，Leaf-3还将vPC主IP 10.0.1.11和10.0.1.12视为类型5和孤立路由。

MAC/IP学习

```
Leaf-1# show l2route evpn mac all
Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv (AD):Auto-Delete (D):Del Pending
(S):Stale (C):Clear, (Ps):Peer Sync (O):Re-Originated (Nho):NH-Override
(Pf):Permanently-Frozen, (Orp): Orphan

Topology    Mac Address    Prod   Flags         Seq No     Next-Hops
----------- -------------- ------ ------------- ---------- ----------------
10          0050.5601.0a0a Local  L,            0          Po11
10          0050.5603.0a14 BGP    SplRcv        0          10.0.1.13 (Label: 10010)
20          0050.5603.1e1e BGP    SplRcv        0          10.0.1.13 (Label: 10020)

Leaf-1# show ip arp suppression-cache detail
Flags: + - Adjacencies synced via CFSoE
       L - Local Adjacency
       R - Remote Adjacency
       L2 - Learnt over L2 interface
       PS - Added via L2RIB, Peer Sync
       RO - Dervied from L2RIB Peer Sync Entry

Ip Address      Age      Mac Address    Vlan Physical-ifindex    Flags    Remote Vtep Addrs

192.168.10.10   00:02:11 0050.5601.0a0a   10 port-channel11       L
192.168.10.20   00:05:42 0050.5603.0a14   10 (null)               R        10.0.1.13
```

L3 VNI路由

```
Leaf-1# show ip route vrf Tenant-A 192.168.20.30
IP Route Table for VRF "Tenant-A"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%

     ' in via output denotes VRF


192.168.20.30/32, ubest/mbest: 1/0
    *via 10.0.1.13%default, [200/0], 00:10:21, bgp-65000, internal, tag 65000, segid: 50000 tunnelid: 0
```

---

vPC

```
Leaf-1# show vpc consistency-parameters global | include Nve|Src|Sec|Host
Nve1 Adm St, Src Adm St,    1     Up, Up, 10.0.1.112, CP,    Up, Up, 10.0.1.112, CP,
Sec IP, Host Reach, VMAC          TRUE, Disabled,               TRUE, Disabled,
Adv, SA,mcast l2, mcast           0.0.0.0, 0.0.0.0, Disabled,   0.0.0.0, 0.0.0.0, Disabled,
Nve Vni Configuration       1     10010-10020                   10010-10020

Leaf-1# show system nve infra-vlans
Currently active infra-vlans: 3900
```

数据包格式

所有VXLAN数据包均源自入口VTEP的NVE源环回。对于vPC端口上的主机，这是共享辅助IP(10.0.1.112)。 已知单播流量流向远程VTEP的NVE IP。BUM流量进入VNI的多播组，或者使用入口复制，单播副本进入每个远程VTEP。

报头	示例：Server-1到Server-2(VLAN 10、L2 VNI)	示例：Server-1到Server-3（路由，L3 VNI）
外部以太网	枝叶–1上行链路MAC到主干MAC	枝叶–1上行链路MAC到主干MAC
外部IP	10.0.1.112 到 10.0.1.13	10.0.1.112到10.0.1.13（如果流量来自孤立端口且使用advertise-pip，则为10.0.1.11到10.0.1.13）
外部UDP	源端口是内部流的哈希（ECMP的熵）；目的端口4789	相同
VXLAN	标志0x08（I标志），VNI 10010	标志0x08,VNI 50000
内部以太网	Server-1 MAC到Server-2 MAC	Leaf-1路由器MAC到Leaf-3路由器MAC
内部IP	192.168.10.10 到	192.168.10.10到

	192.168.10.20	192.168.20.30（TTL递减）

注意：VXLAN向原始帧添加50个字节。将底层MTU（所有主干和枝叶上行链路以及vPC备份SVI）设置为比最大主机MTU至少多50字节。建议使用9216。

故障排除

症状	检查	共同原因
NVE对等体不在show nve peers中	show bgp l2vpn evpn summary、show ip route <remote NVE IP>、show nve vni	BGP EVPN会话关闭；未在底层中通告的远程NVE环回；两个VTEP上均未配置VNI;缺少发送社区扩展
vPC VTEP上的NVE接口关闭或VNI挂起	show nve interface nve1 detail，show vpc consistency-parameters global	对等体之间的NVE配置不匹配；重新加载后，源环回仍然处于抑制计时器中；辅助IP缺失或不同
同一VNI中的主机无法通信，或ARP失败	show l2route evpn mac all、show ip arp suppression-cache detail、show ip mroute <group>	未收到RT-2（路由目标或VNI不匹配）；BUM未复制（NVE环回上没有PIM邻居、没有RP、PIM丢失）；suppress-arp on a VNI without an anycast gateway SVI
VNI之间的路由失败	show ip route vrf Tenant-A、show nve vni(L3 VNI state)、show bgp l2vpn evpn route-type 5	L3 VNI VLAN或SVI丢失或关闭；vrf情景下缺少vni;通告BGP VRF下缺少l2vpn evpn;route-target both auto evpn missing
流量仅对孤立端口失败，或者在一个vPC对等体失去其上行链路后失败	show system nve infra-vlan， show ip ospf neighbors vrf default	备份底层SVI丢失，不在系统无基础设施VLAN中，或不允许在对等链路上
大数据包被丢弃，小数据包工	show interface	对于VXLAN开销来说，底层

作	<uplink>（TX错误），ping NVE环回之间设置了DF位	MTU太小

注意：这些项目不支持VXLAN上行链路：

- 第3层子接口
- 未在系统nve infra-vlan中配置的SVI

NVE源环回必须在默认VRF中。

相关信息

- Cisco Nexus 9000系列NX-OS VXLAN配置指南，版本10.5(x)
- 通过最佳实践了解和配置 Nexus 9000 vPC
- 技术支持和文档 - Cisco Systems

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。