

為NXOS配置VXLAN vPC交換矩陣對等並排除故障

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[網路圖表](#)

[組態](#)

[TCAM配置](#)

[TCAM雕刻](#)

[vPC配置](#)

[VPC域](#)

[Keep-alive](#)

[虛擬對等鏈路的第3層介面](#)

[VPC對等鏈路](#)

[向上連結](#)

[SPINES配置](#)

[採用入口複製封裝的廣播、未知單播和組播流量](#)

[採用入口複製解除封裝的廣播、未知單播和組播流量](#)

[採用組播封裝的廣播、未知單播和組播流量](#)

[使用組播解封裝的廣播、未知單播和組播流量](#)

[驗證](#)

[相關資訊](#)

簡介

本文檔介紹如何為NXOS和BUM流量配置和驗證vPC交換矩陣對等。

必要條件

需求

思科建議瞭解以下主題：

- vPC (虛擬埠通道)
- 虛擬可擴充區域網路(VXLAN)

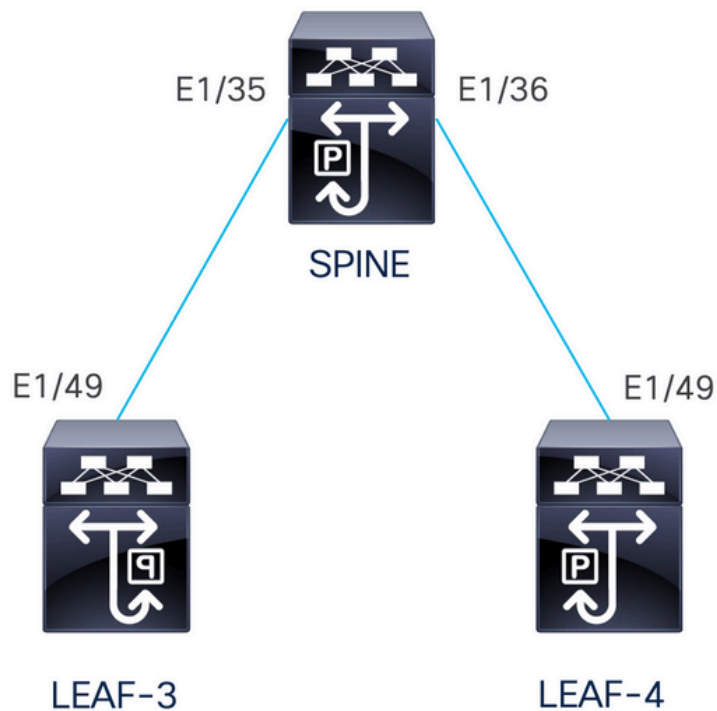
採用元件

本文中的資訊係根據以下軟體和硬體版本：

- 適用於枝葉交換機的N9K-C93240YC-FX2版本：10.3(3)
- N9K-C9336C-FX2（適用於主幹交換機）版本：10.3(3)

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

網路圖表



vPC交換矩陣對等提供增強型雙宿主接入解決方案，而不會浪費vPC對等鏈路的物理埠。此功能保留傳統vPC的所有特性。

在此部署中，我們將Leaf-3和Leaf-4配置為具有交換矩陣對等的vPC。

組態

TCAM配置

在配置之前，會對TCAM記憶體進行檢查：

```

LEAF-4(config-if)# sh hardware access-list tcam region
          NAT ACL[nat] size = 0
        Ingress PACL [ing-ifac1] size = 0
          VACL [vac1] size = 0
        Ingress RAACL [ing-racl] size = 2304
        Ingress L2 QOS [ing-l2-qos] size = 256
        Ingress L3/VLAN QOS [ing-l3-vlan-qos] size = 512
          Ingress SUP [ing-sup] size = 512
  
```

```

Ingress L2 SPAN filter [ing-l2-span-filter] size = 256
Ingress L3 SPAN filter [ing-l3-span-filter] size = 256
    Ingress FSTAT [ing-fstat] size = 0
        span [span] size = 512
    Egress RAcl [egr-racl] size = 1792
    Egress SUP [egr-sup] size = 256
    Ingress Redirect [ing-redirect] size = 0
    Egress L2 QOS [egr-l2-qos] size = 0
    Egress L3/VLAN QOS [egr-l3-vlan-qos] size = 0
    Ingress Netflow/Analytics [ing-netflow] size = 512
        Ingress NBM [ing-nbm] size = 0
        TCP NAT ACL[tcp-nat] size = 0
    Egress sup control plane[egr-copp] size = 0
    Ingress Flow Redirect [ing-flow-redirect] size = 0 <<<<<<<<<
Ingress PAcl IPv4 Lite [ing-ifacl-ipv4-lite] size = 0
Ingress PAcl IPv6 Lite [ing-ifacl-ipv6-lite] size = 0
    Ingress CNTACL [ing-cntacl] size = 0
    Egress CNTACL [egr-cntacl] size = 0
    MCAST NAT ACL[mcast-nat] size = 0
    Ingress DAcl [ing-dacl] size = 0
    Ingress PAcl Super Bridge [ing-pacl-sb] size = 0
    Ingress Storm Control [ing-storm-control] size = 0
    Ingress VAcl redirect [ing-vacl-nh] size = 0
    Egress PAcl [egr-ifacl] size = 0
    Egress Netflow [egr-netflow] size = 0

```

vPC交換矩陣對等要求應用TCAM區域劃分流重定向（即TCAM區域劃分）。TCAM雕刻需要在使用的功能之前儲存配置並重新載入交換機。

TCAM上的此空格是雙寬度，因此我們可以指定的最小值為512。

TCAM雕刻

在此場景中，ing-racl有足夠的空間來佔用512，並將那些512分配給ing-flow-redirect。

```

LEAF-4(config-if)# hardware access-list tcam region ing-racl 1792
Please save config and reload the system for the configuration to take effect

```

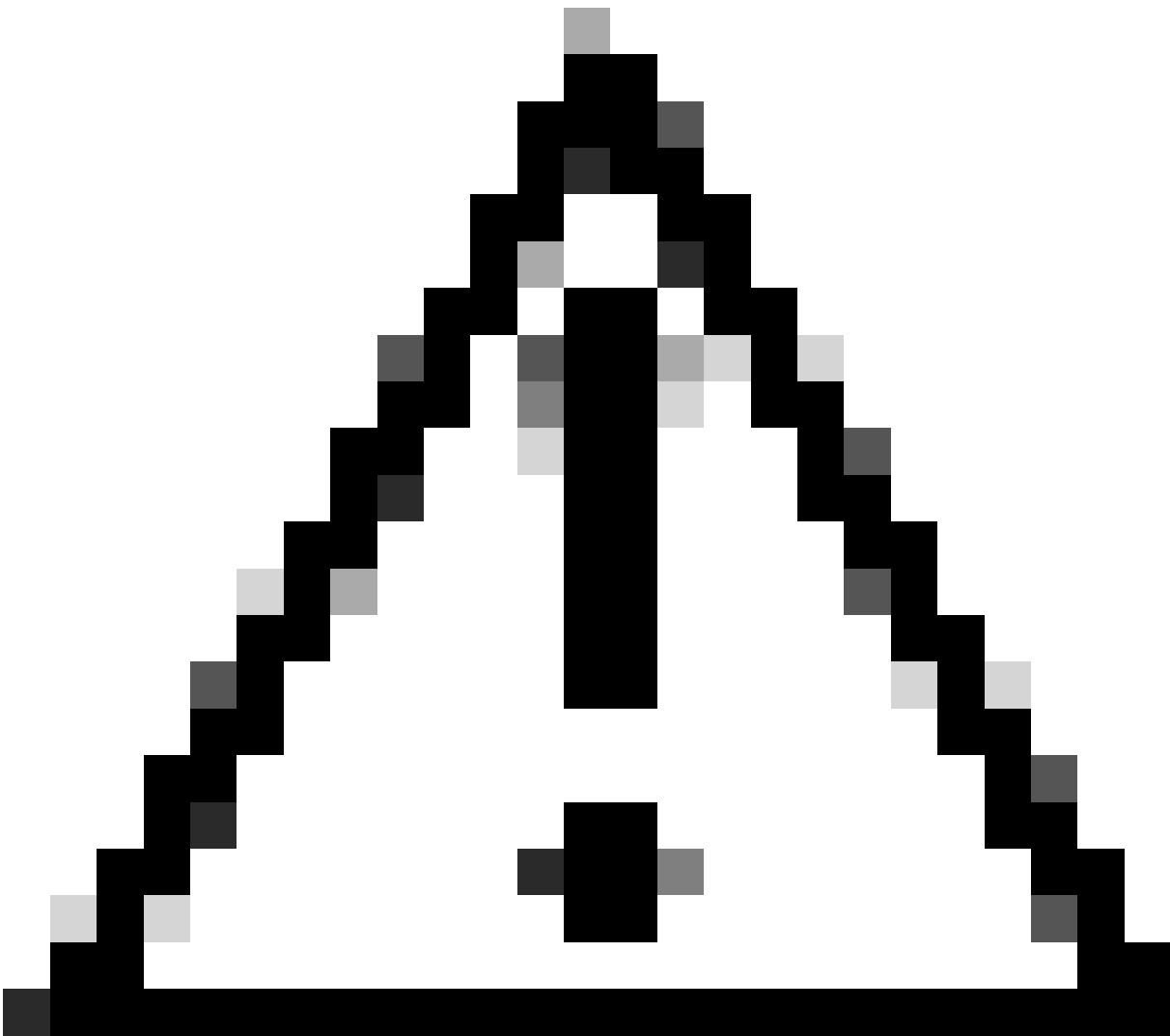
```

LEAF-4(config)# hardware access-list tcam region ing-flow-redirect 512
Please save config and reload the system for the configuration to take effect

```



```
Egress L3/VLAN QOS [egr-l3-vlan-qos] size = 0
Ingress Netflow/Analytics [ing-netflow] size = 512 <<<<<
    Ingress NBM [ing-nbm] size = 0
    TCP NAT ACL[tcp-nat] size = 0
    Egress sup control plane[egr-copp] size = 0
    Ingress Flow Redirect [ing-flow-redirect] size = 0
    Ingress PACL IPv4 Lite [ing-ifacl-ipv4-lite] size = 0
    Ingress PACL IPv6 Lite [ing-ifacl-ipv6-lite] size = 0
    Ingress CNTACL [ing-cntacl] size = 0
    Egress CNTACL [egr-cntacl] size = 0
    MCAST NAT ACL[mcast-nat] size = 0
    Ingress DAACL [ing-dacl] size = 0
    Ingress PACL Super Bridge [ing-pacl-sb] size = 0
    Ingress Storm Control [ing-storm-control] size = 0
    Ingress VAACL redirect [ing-vacl-nh] size = 0
    Egress PACL [egr-ifacl] size = 0
```



注意：請確保在TCAM上進行更改後重新載入裝置，否則，由於TCAM中未應用更改，VPC不會啟動。

vPC配置

VPC域

在VPC域中的LEAF-3和LEAF-4上，配置是為保持連線和虛擬對等鏈路指定IP地址

```
vpc domain 1
  peer-keepalive destination 192.168.1.1 source 192.168.1.2 vrf management
  virtual peer-link destination 10.10.10.2 source 10.10.10.1 dscp 56

interface port-channel1
  vpc peer-link
```

Keep-alive

vPC對等體之間的任何直接第3層鏈路必須僅用於對等保持活動。它必須位於僅用於保持連線的獨立VRF中。在此案例中，我們使用的是交換器的介面管理。

```
LEAF-3
interface mgmt0
  vrf member management
  ip address 192.168.1.1/24
```

```
LEAF-4
interface mgmt0
  vrf member management
  ip address 192.168.1.2/24
```

虛擬對等鏈路的第3層介面

用於虛擬對等鏈路的第3層介面不能與用於保持連線的第3層介面相同，您可以使用與底層相同的環回，也可以是Nexus上的專用環回

這裡，loopback0用於襯底，loopback2是用於虛擬對等鏈路的專用環回，而loopback1是關聯到介面NVE的介面。

```
LEAF-3
interface loopback0
  ip address 10.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

interface loopback1
  ip address 172.16.1.2/32
  ip address 172.16.1.1/32 secondary
  ip router ospf 1 area 0.0.0.0
```

```
ip pim sparse-mode

interface loopback2
ip address 10.10.10.2/32
ip router ospf 1 area 0.0.0.0
```

LEAF-4

```
interface loopback0
ip address 10.1.1.2/32
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode

interface loopback1
ip address 172.16.1.3/32
ip address 172.16.1.1/32 secondary
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode

interface loopback2
ip address 10.10.10.1/32
ip router ospf 1 area 0.0.0.0
```

VPC對等鏈路

即使我們不打算將物理介面分配給埠通道，對等鏈路也需要分配一個埠通道。

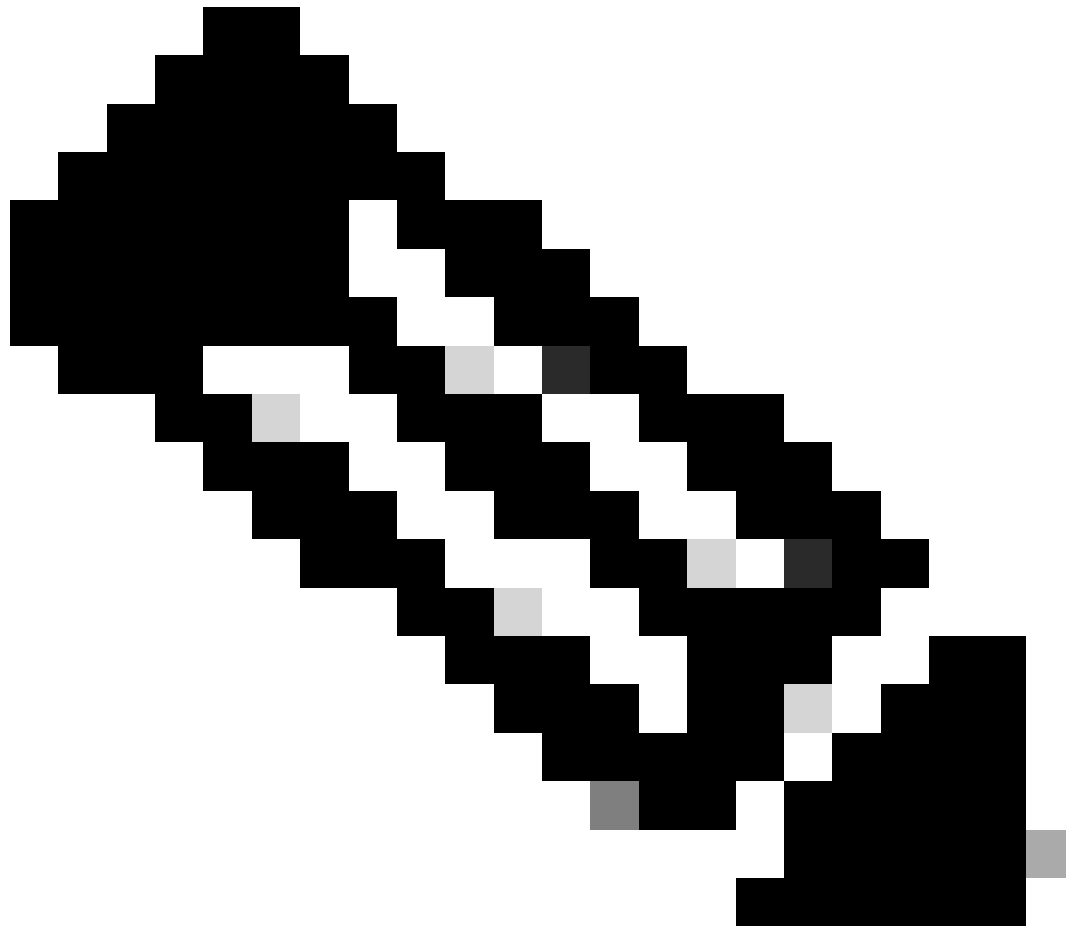
```
LEAF-3(config-if)# sh run interface port-channel 1 membership
```

```
interface port-channel1
switchport
switchport mode trunk
spanning-tree port type network
vpc peer-link
```

向上連結

配置的最後一部分是使用命令port-type fabric配置指向SPINE的兩個枝葉上的鏈路。

```
interface Ethernet1/49
port-type fabric <<<<<<<<<
medium p2p
ip unnumbered loopback0
ip router ospf 1 area 0.0.0.0
ip pim sparse-mode
no shutdown
```



注意：如果不配置埠型別交換矩陣，您將無法看到Nexus正在生成的保持連線

SPINES配置

建議在主幹上設定QoS以匹配VPC域上配置的DSCP值，因為vPC交換矩陣對等對等鏈路是透過傳輸網路建立的。

控制平面資訊CFS消息用於同步埠狀態資訊、VLAN資訊、VLAN到VNI對映、主機MAC地址和IGMP監聽組，這些消息透過交換矩陣傳輸。CFS消息使用相應的DSCP值進行標籤，必須在傳輸網路中保護該值。

```
class-map type qos match-all CFS
  match dscp 56
```

```
policy-map type qos CFS
  class CFS
    Set qos-group 7 <<< Depending on the platform it can be 4
```



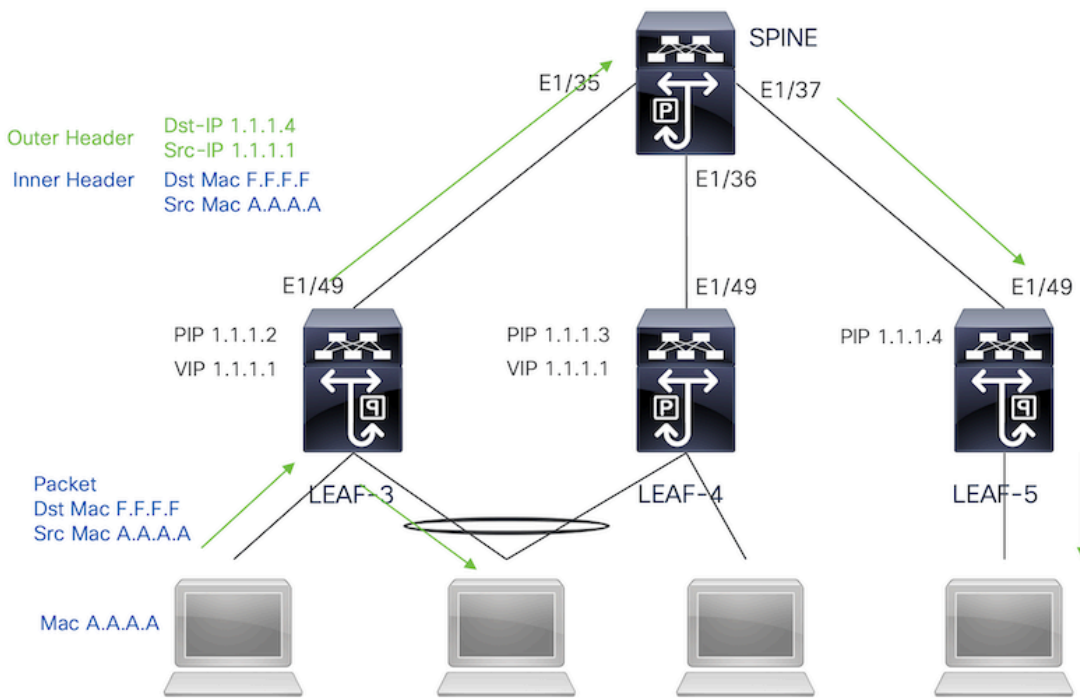
```
interface Ethernet 1/35-36
 service-policy type qos input CFS
```

採用入口複製封裝的廣播、未知單播和組播流量

當nexus收到需要廣播的資料包時，它會生成該資料包的2個副本。

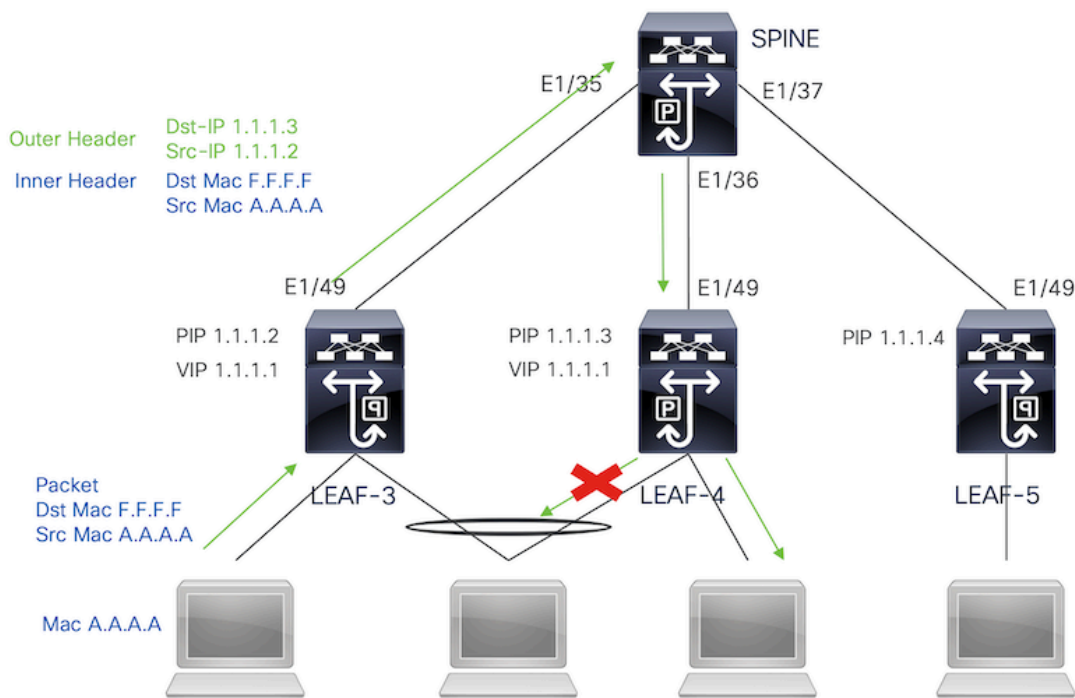
1. 到VNI泛洪清單中的所有遠端VTEP，包括本地接入埠
2. 到遠端VPC對等體

對於第一個副本，Nexus使用輔助IP地址的源IP和遠端VTEP的目標IP以及本地接入埠封裝流量。



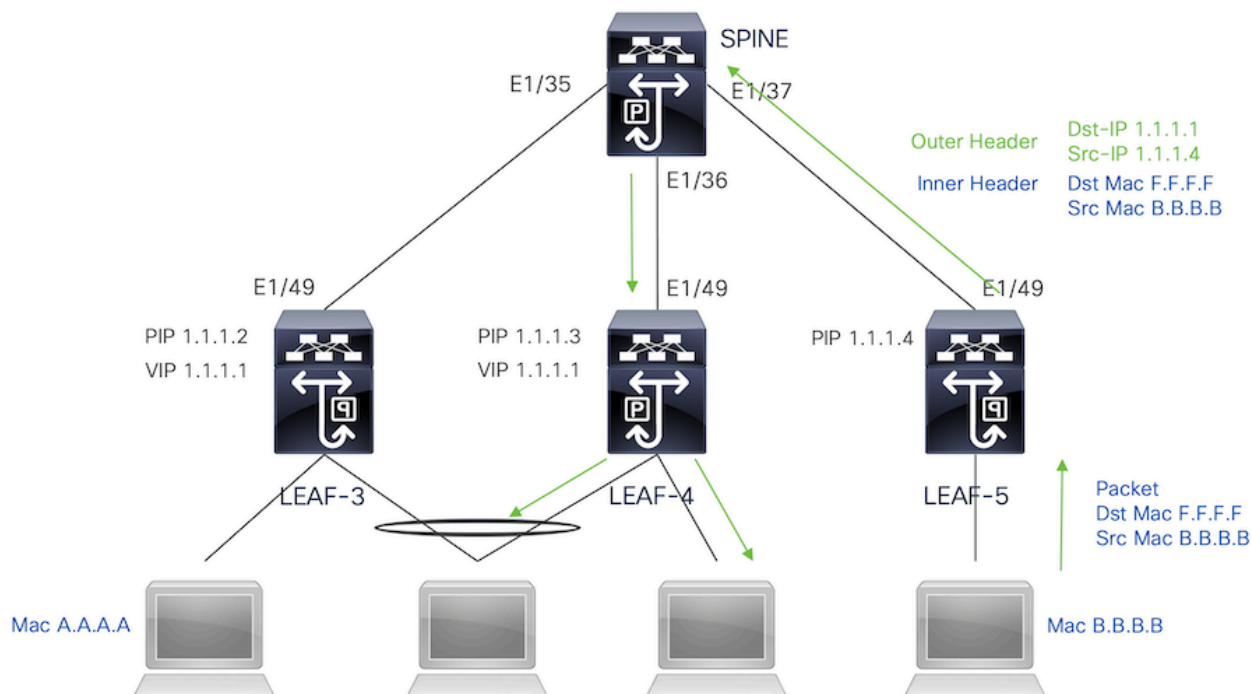
對於第二個副本，它將傳送到遠端VPC對等體，源IP將成為環回的主要IP，而目標IP是遠端VPC對等體的PIP。

從主幹收到資料包後，遠端VTEP將僅將資料包轉發到孤立埠。



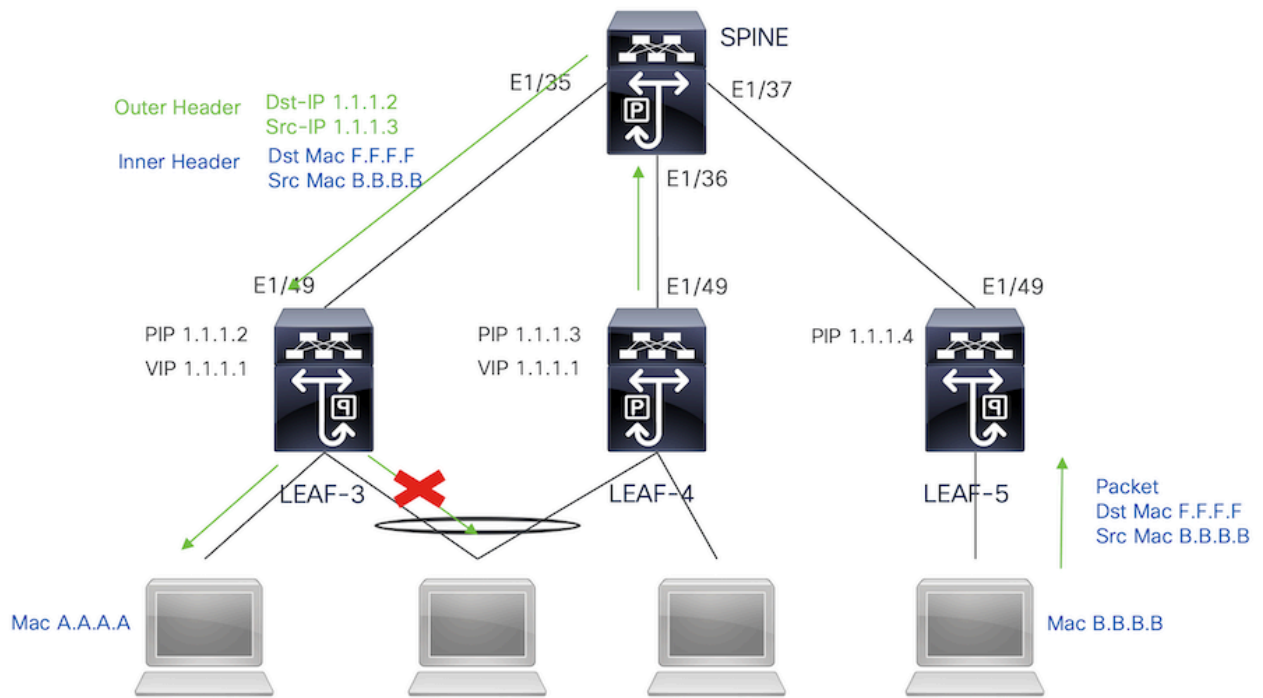
採用入口複製解除封裝的廣播、未知單播和組播流量

由於從另一個VTEP接收的BUM流量的目的地IP是VIP，因此流量會雜湊到其中一個VPC裝置，因此它會解除封裝封包並將其傳送至存取連線埠。



為了使流量到達遠端VPC對等體上連線的孤立埠，nexus會生成資料包的副本，並將它僅傳送到使用主IP地址作為源/目標IP的遠端VPC。

在遠端vpc對等裝置上收到流量後，nexus會解封流量，並僅將其轉發到孤立埠。

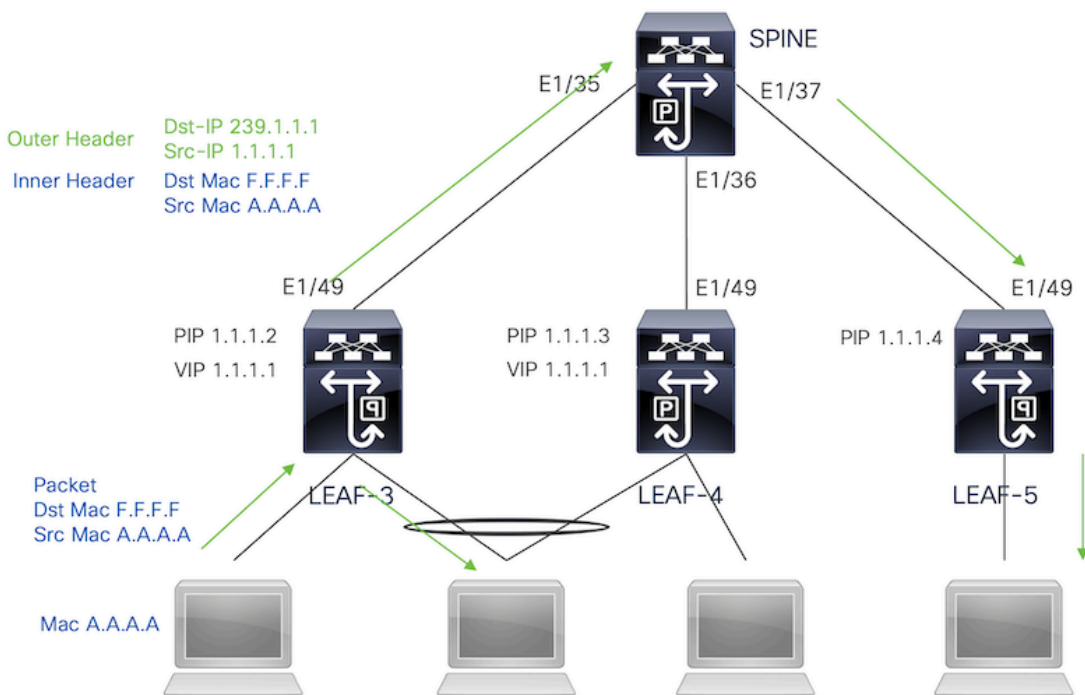


採用組播封裝的廣播、未知單播和組播流量

當nexus收到需要廣播的資料包時，它會生成該資料包的2個副本。

1. 資料包將傳送到組播S，G條目中的所有OIF，包括本地接入埠
2. 到遠端VPC對等體

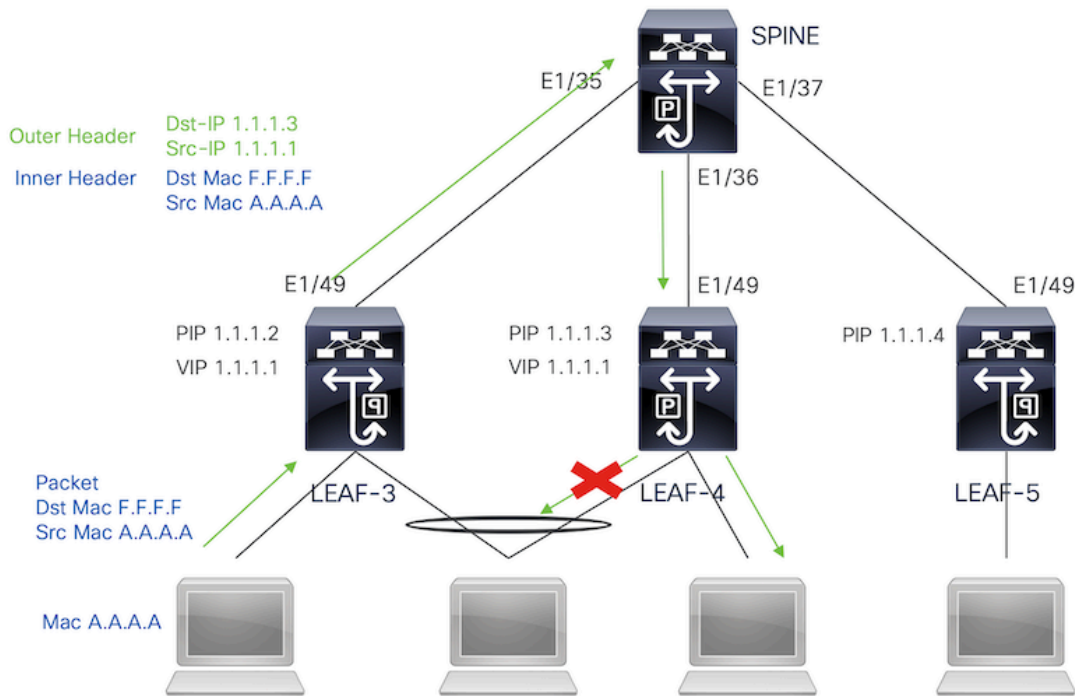
對於第一個副本，Nexus使用輔助IP地址的源IP和已配置組播組的目標IP封裝流量。



對於第二個副本，它將傳送到遠端VPC對等體，源IP將是環回的輔助埠，目標IP是遠端VPC對等體

的PIP。

從主幹收到封包後，遠端VTEP只會將封包轉送到孤立連線埠。

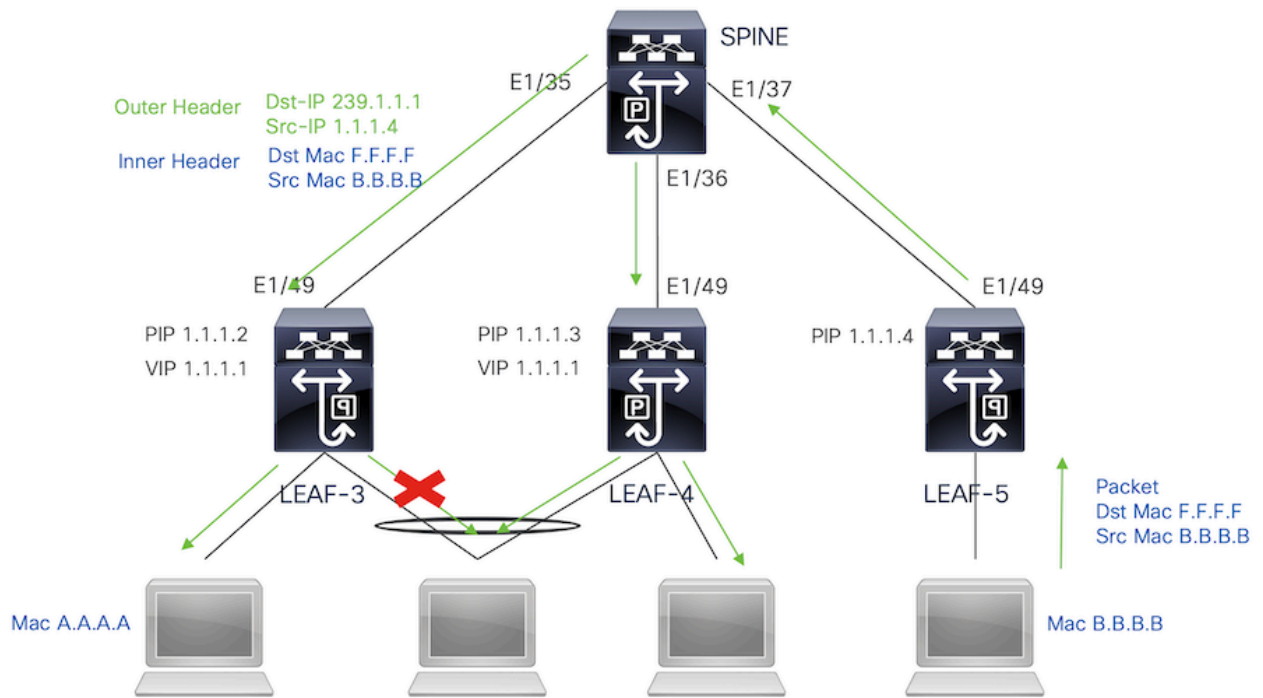


使用組播解封裝的廣播、未知單播和組播流量

對於解除封裝程式，封包將到達兩個VPC對等點。只有一台VPC裝置透過VPC埠通道轉發流量。這由命令中顯示的轉發器決定。

```
module-1# show forwarding internal vpc-df-hash
```

```
VPC DF: FORWARDER
```



驗證

要確保VPC已啟動，請運行以下命令：

驗證用於虛擬對等鏈路的IP地址的可達性。

```
LEAF-3# sh ip route 10.10.10.1
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>
```

```
10.10.10.1/32, ubest/mbest: 1/0
   *via 192.168.120.1, Eth1/49, [110/3], 01:15:01, ospf-1, intra
```

```
LEAF-3# ping 10.10.10.1
PING 10.10.10.1 (10.10.10.1): 56 data bytes
64 bytes from 10.10.10.1: icmp_seq=0 ttl=253 time=0.898 ms
64 bytes from 10.10.10.1: icmp_seq=1 ttl=253 time=0.505 ms
64 bytes from 10.10.10.1: icmp_seq=2 ttl=253 time=0.433 ms
64 bytes from 10.10.10.1: icmp_seq=3 ttl=253 time=0.465 ms
64 bytes from 10.10.10.1: icmp_seq=4 ttl=253 time=0.558 ms
```

```
LEAF-3(config-if)# show vpc brief
```

Legend:

(*) - local vPC is down, forwarding via vPC peer-link

```
vPC domain id          : 1
Peer status            : peer adjacency formed ok <<<<
```

```

vPC keep-alive status           : peer is alive <<<<
Configuration consistency status : success
Per-vlan consistency status     : success
Type-2 consistency status       : success
vPC role                         : secondary
Number of vPCs configured       : 0
Peer Gateway                     : Disabled
Dual-active excluded VLANs      : -
Graceful Consistency Check      : Enabled
Auto-recovery status            : Disabled
Delay-restore status             : Timer is off.(timeout = 30s)
Delay-restore SVI status        : Timer is off.(timeout = 10s)
Delay-restore Orphan-port status : Timer is off.(timeout = 0s)
Operational Layer3 Peer-router  : Disabled
Virtual-peerlink mode           : Enabled <<<<<<<

```

vPC Peer-link status

```

-----
id  Port  Status  Active vlans
--  ---  -
1   Po1   up      1,10,50,600-604,608,610-611,614-618,638-639,
                                662-663,701-704

```

要檢查VPC的角色，請運行命令：

```
LEAF-3(config-if)# sh vpc role
```

```

vPC Role status
-----
vPC role                : secondary <<<<
Dual Active Detection Status : 0
vPC system-mac          : 00:23:04:ee:be:01
vPC system-priority     : 32667
vPC local system-mac    : d0:e0:42:e2:09:6f
vPC local role-priority  : 32667
vPC local config role-priority : 32667
vPC peer system-mac     : 2c:4f:52:3f:46:df
vPC peer role-priority   : 32667
vPC peer config role-priority : 32667

```

在對等鏈路埠通道中允許的所有vlan必須對映到VNI，以防它們顯示為不一致

```

LEAF-3(config-if)# show vpc virtual-peerlink vlan consistency
Following vlans are inconsistent
1 608 610 611 614 615 616 617 618 638 639 701 702 703 704

```

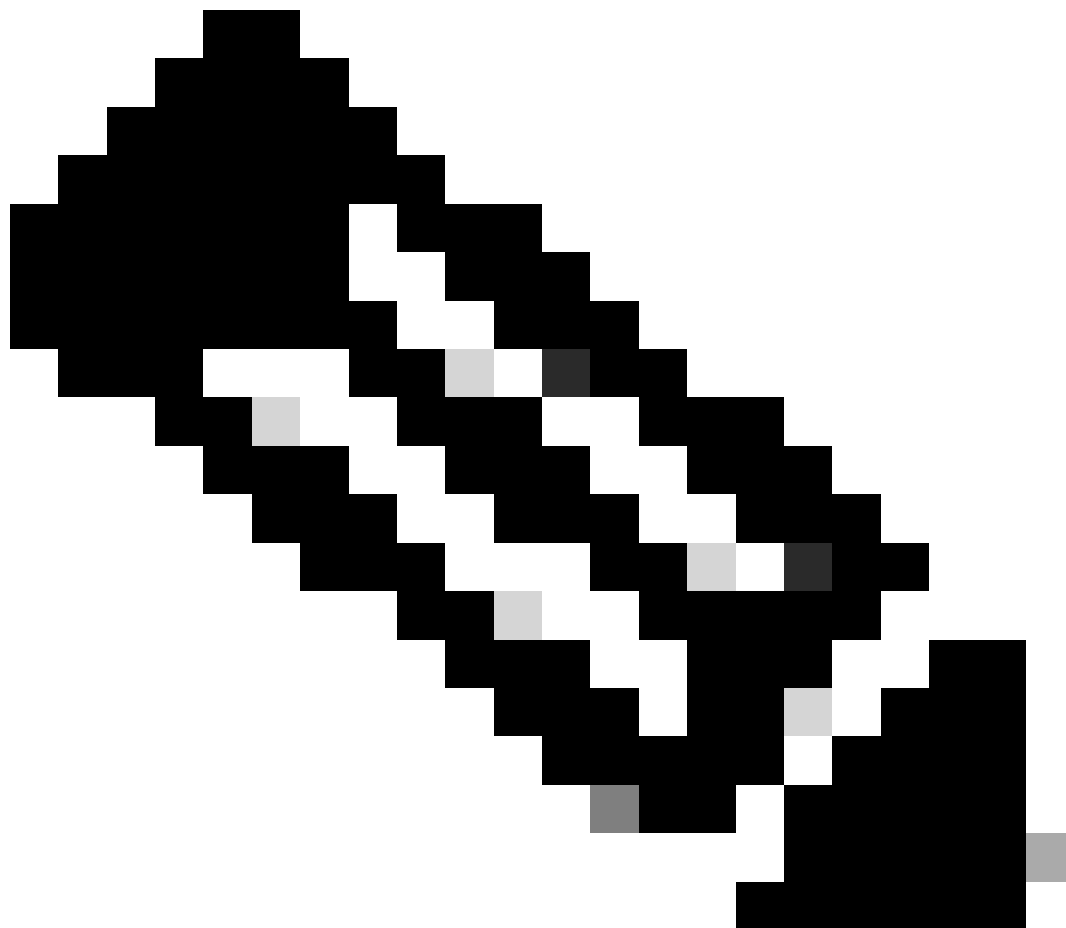
要確認上行鏈路上的配置已正確程式設計，請運行以下命令：

```
LEAF-3(config-if)# show vpc fabric-ports
```

Number of Fabric port : 1
Number of Fabric port active : 1

Fabric	Ports	State

Ethernet	1/49	UP



註：除非VPC處於運行狀態，否則NVE或與其相關聯的環回介面將會顯示。

相關資訊

- [思科技術支援與下載](#)

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。