# 使用vPC的高級VXLAN:L2VNI和L3VNI的配置與驗證

## 目錄



## 簡介

本文說明如何使用Nexus 9Kv交換機設定實驗，該實驗使用具有虛擬埠通道(vPC)的高級虛擬可擴展區域網(VXLAN)。

## 必要條件

### 需求

思科建議您瞭解以下主題：

- 瞭解路由和交換以及多重協定標籤交換(MPLS)技術
- 具備集結點(RP)和平台無關組播(PIM)等組播路由原則的經驗
- 瞭解邊界閘道通訊協定(BGP)位址系列指標(AFI)/後續位址系列指標(SAFI)

### 採用元件

本文件所述內容不限於特定軟體和硬體版本。

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。

## 背景資訊

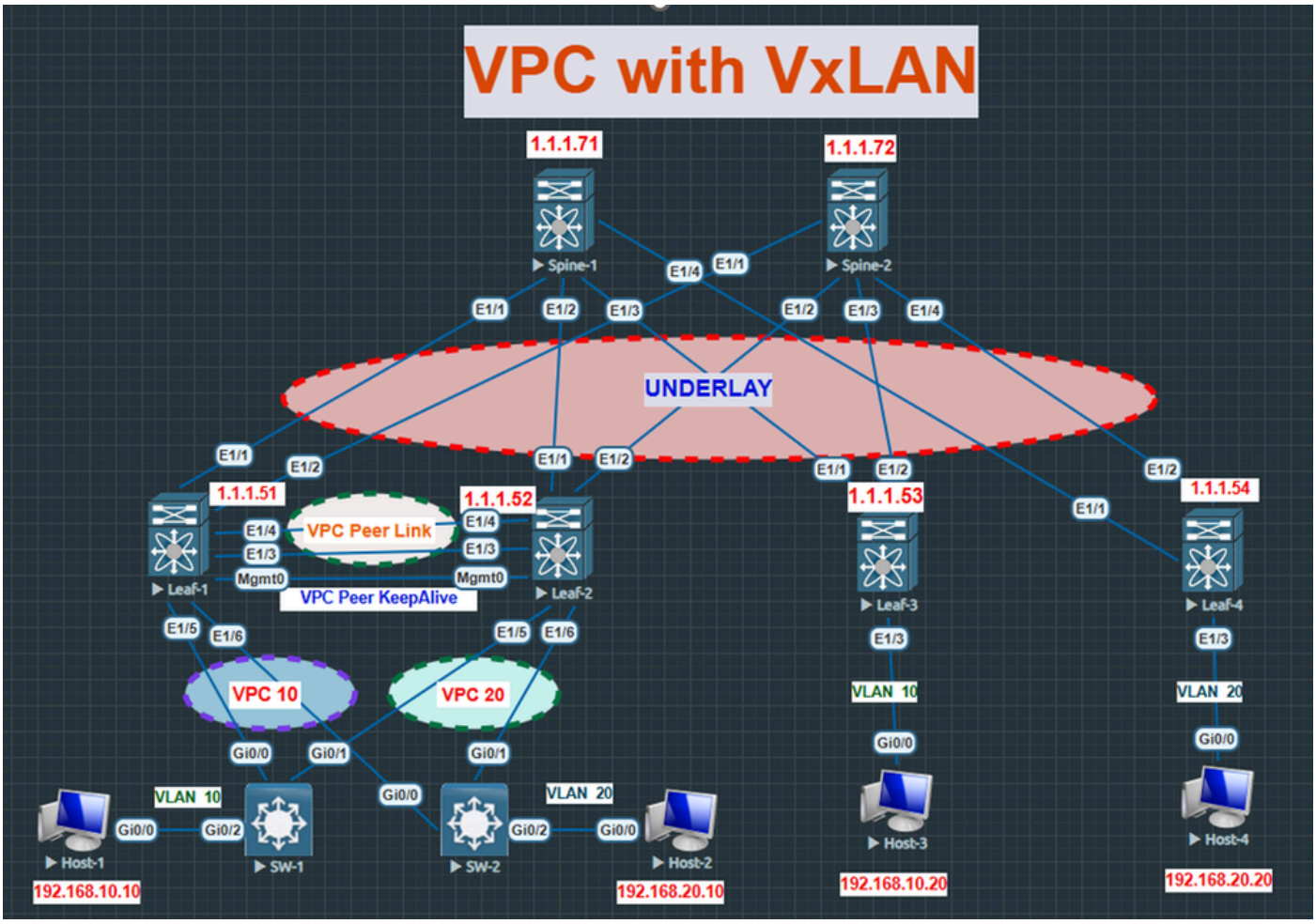
本文檔還提供有關部署實驗室以及驗證配置和操作的指導。

在本實驗中，枝葉和主幹都使用帶有Nexus 9000V交換機的EveNg。

| | |
|---|---|
| 虛擬通道端點(VTEP) | LEAF1、LEAF2、LEAF3、LEAF4 |
| vPC | LEAF1和LEAF2 |
| LEAF1主要和次要環回IP | Loopback0 - 1.1.1.51、Loopback1 - 10.1.1.100 |
| LEAF2主環回和輔助環回IP | Loopback0 - 1.1.1.52, Loopback1 - 10.1.1.100 |
| LEAF3環回IP | 1.1.1.53 |
| LEAF4環回IP | 1.1.1.54 |
| SPINE1環回和任播RP | Loopback0 - 1.1.1.71、Loopback1 - 10.1.2.10（任播RP） |
| SPINE2環回和任播RP | Loopback0 - 1.1.1.72, Loopback1 - 10.1.2.10（任播RP） |
| 主機1 | 192.168.10.10(0000.000.aaaa)(VLAN 10) |
| 主機2 | 192.168.20.10(0000.000.bbb)(VLAN 20) |
| 主機3 | 192.168.10.20(0000.000.cccc)(VLAN 10) |
| 主機4 | 192.168.20.20(0000.000.ddd)(VLAN 20) |
| VLAN 10 | L2VNI支100010 |
| VLAN 20 | L2VNI支100020 |
| VLAN 500 | L3VNI裝50000 |

# 設定

## 網路圖表




## 組態

- 底層和PIM鄰居已經建立。

枝葉交換機：

```
feature ospf

router ospf UNDERLAY
  log-adjacency-changes

interface loopback0
  ip router ospf UNDERLAY area 0.0.0.0

interface Ethernet1/1
  ip ospf cost 4
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0

interface Ethernet1/2
  ip ospf cost 4
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
```

在枝葉交換機上啟用開放最短路徑優先(OSPF)

```
feature pim

ip pim rp-address 10.1.2.10 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8

vrf context TENANT1
  ip pim ssm range 232.0.0.0/8

interface Vlan10
  ip pim sparse-mode

interface Vlan20
  ip pim sparse-mode

interface loopback0
  ip pim sparse-mode

interface Ethernet1/1
  ip pim sparse-mode

interface Ethernet1/2
  ip pim sparse-mode
```

在枝葉交換機上啟用PIM

```
LEAF-1# show ip ospf neighbors
 OSPF Process ID UNDERLAY VRF default
 Total number of neighbors: 2
 Neighbor ID     Pri State            Up Time  Address         Interface
 1.1.1.71          1 FULL/ -          04:32:03 192.168.11.1    Eth1/1
 1.1.1.72          1 FULL/ -          04:17:47 192.168.21.2    Eth1/2
LEAF-1# sh ip pim neighbor
PIM Neighbor Status for VRF "default"
Neighbor        Interface            Uptime    Expires   DR       Bidir-  BFD    ECMP Redirect
                                                         Priority Capable State     Capable
192.168.11.1    Ethernet1/1          04:32:14  00:01:30  1        yes     n/a     no
192.168.21.2    Ethernet1/2          04:17:58  00:01:44  1        yes     n/a     no
LEAF-1#
```

OSPF鄰居

主幹交換機：

```
feature pim

ip pim rp-address 10.1.2.10 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8
ip pim anycast-rp 10.1.2.10 1.1.1.71
ip pim anycast-rp 10.1.2.10 1.1.1.72
```

在主幹交換機上啟用PIM

- 底層和PIM鄰居已經建立。
- 對於整個組播組224.0.0.0/4，兩個主幹交換機將是相同的任播RP。
- 枝葉交換機和主幹交換機之間的介面上的最大傳輸單元(MTU)設定為9000/9216。

首先，在Leaf1和Leaf2之間設定vPC。

步驟1. vPC功能和域啟用。

- 啟用功能vPC和連結彙總控制通訊協定(LACP)。
- 配置vPC域。
- 管理0介面用作對等保持連線鏈路，而Eth1/3和Eth1/4將是vPC對等鏈路（埠通道1）的一部分。
- 確保將peer-switch命令配置為與降序交換機共用一個公用MAC地址。

```
feature lacp
feature vpc
```

在枝葉交換機上啟用功能

```
LEAF-1# sh run vpc

!Command: show running-config vpc
!Running configuration last done at: Sat Dec 28 07:17:18 2024
!Time: Sat Dec 28 07:39:48 2024

version 7.0(3)I7(9) Bios:version
feature vpc

vpc domain 1
  peer-switch
  role priority 100
  peer-keepalive destination 192.168.0.52
  peer-gateway

interface port-channel1
  vpc peer-link
```

在枝葉交換機1上啟用vPC

```
LEAF-2# sh run vpc

!Command: show running-config vpc
!Running configuration last done at: Sat Dec 28 07:17:14 2024
!Time: Sat Dec 28 07:40:20 2024

version 7.0(3)I7(9) Bios:version
feature vpc

vpc domain 1
  peer-switch
  role priority 200
  peer-keepalive destination 192.168.0.51
  peer-gateway

interface port-channel1
  vpc peer-link
```

在枝葉交換機2上啟用vPC

步驟2.埠成員分配。

- 將埠成員分配給通道組，並將它們包括在vPC中。在本例中，使用了兩個vPC。vPC 20和vPC 10。

```
LEAF-1# sh run int port-channel 10, port-channel 20 membership

!Command: show running-config interface port-channel10, port-channel20 membership
!Running configuration last done at: Sat Dec 28 07:17:18 2024
!Time: Sat Dec 28 07:42:44 2024

version 7.0(3)I7(9) Bios:version

interface port-channel10
  switchport mode trunk
  vpc 10

interface Ethernet1/5

  switchport mode trunk
  channel-group 10 mode active

interface port-channel20
  switchport mode trunk
  vpc 20

interface Ethernet1/6

  switchport mode trunk
  channel-group 20 mode active

LEAF-1#
```

在枝葉交換機1上分配埠通道

```
LEAF-2# sh run int port-channel 10, port-channel 20 membership

!Command: show running-config interface port-channel10, port-channel20 membership
!Running configuration last done at: Sat Dec 28 07:17:14 2024
!Time: Sat Dec 28 07:43:16 2024

version 7.0(3)I7(9) Bios:version

interface port-channel10
  switchport mode trunk
  vpc 10

interface Ethernet1/5

  switchport mode trunk
  channel-group 10 mode active

interface port-channel20
  switchport mode trunk
  vpc 20

interface Ethernet1/6

  switchport mode trunk
  channel-group 20 mode active

LEAF-2#
```

在枝葉交換機2上分配埠通道

- 此處建立了vPC，對等體開始交換keepalive消息以驗證可用性。

```
LEAF-1# show vpc
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                          : primary
Number of vPCs configured         : 2
Peer Gateway                      : Enabled
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)
Operational Layer3 Peer-router    : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id    Port   Status Active vlans
--    ----   ------ --------------------------------------------------
1     Po1    up     1,10,20,500

vPC status
----------------------------------------------------------------------------
Id    Port          Status Consistency Reason                Active vlans
--    ------------  ------ ----------- ------                ---------------
10    Po10          up     success     success               1,10,20,500

20    Po20          up     success     success               1,10,20,500


Please check "show vpc consistency-parameters vpc <vpc-num>" for the
consistency reason of down vpc and for type-2 consistency reasons for
any vpc.

LEAF-1#
```

枝葉交換機1上的vPC狀態

```
LEAF-2# sh vpc
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         : 2
Peer Gateway                      : Enabled
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)
Operational Layer3 Peer-router    : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id     Port   Status Active vlans
--     ----   ------ -----------------------------------------------
1      Po1    up     1,10,20,500

vPC status
----------------------------------------------------------------------------
Id     Port          Status Consistency Reason                Active vlans
--     ------------  ------ ----------- ------                ---------------
10     Po10          up     success     success               1,10,20,500

20     Po20          up     success     success               1,10,20,500


Please check "show vpc consistency-parameters vpc <vpc-num>" for the
consistency reason of down vpc and for type-2 consistency reasons for
any vpc.

LEAF-2#
```

枝葉交換機2上的vPC狀態

- VLAN 10、20、500已配置並透過vPC成員埠和vPC對等鏈路。

步驟3.配置輔助IP地址。

- 當vPC包含在VXLAN交換矩陣中時，兩個vPC VTEP對等體都開始使用虛擬IP(VIP)地址作為源地址，而不是其物理IP地址(PIP)。 這也表示當BGP乙太網路VPN(EVPN)預設通告路由型別2（MAC/IP通告）和5（IP首碼路由）時，VIP會用作下一個躍點。本例中的Loopback 0介面設定為使用兩個IP地址：10.1.1.100/32(VIP)作為輔助IP，1.1.1.51/32(PIP)作為主IP。
- 這裡，一個公用IP地址配置為loopback 0介面下的次要IP地址。

```
LEAF-1# sh run int lo

!Command: show running-config interface loopback0
!Running configuration last done at: Sat Dec 28 07:51:58 2024
!Time: Sat Dec 28 07:55:26 2024

version 7.0(3)I7(9) Bios:version

interface loopback0
  ip address 1.1.1.51/32
  ip address 10.1.1.100/32 secondary
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

LEAF-1#
```

枝葉交換機1上的輔助IP

```
LEAF-2# sh run int lo

!Command: show running-config interface loopback0
!Running configuration last done at: Sat Dec 28 07:52:05 2024
!Time: Sat Dec 28 07:55:37 2024

version 7.0(3)I7(9) Bios:version

interface loopback0
  ip address 1.1.1.52/32
  ip address 10.1.1.100/32 secondary
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode

LEAF-2#
```

枝葉交換機2上的輔助IP

步驟4.啟用VXLAN和相關功能。

- 網路虛擬化(nV)覆蓋 — 啟用VXLAN
- Feature nV overlay EVPN — 啟用EVPN控制平面
- 功能交換矩陣轉發 — 啟用主機移動管理器
- 功能虛擬網路(VN) — 網段 — 基於VLAN — 啟用基於VLAN的VXLAN

```
LEAF-1# sh run | sec "feature|nv over"
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
LEAF-1#
```

枝葉交換機上的功能

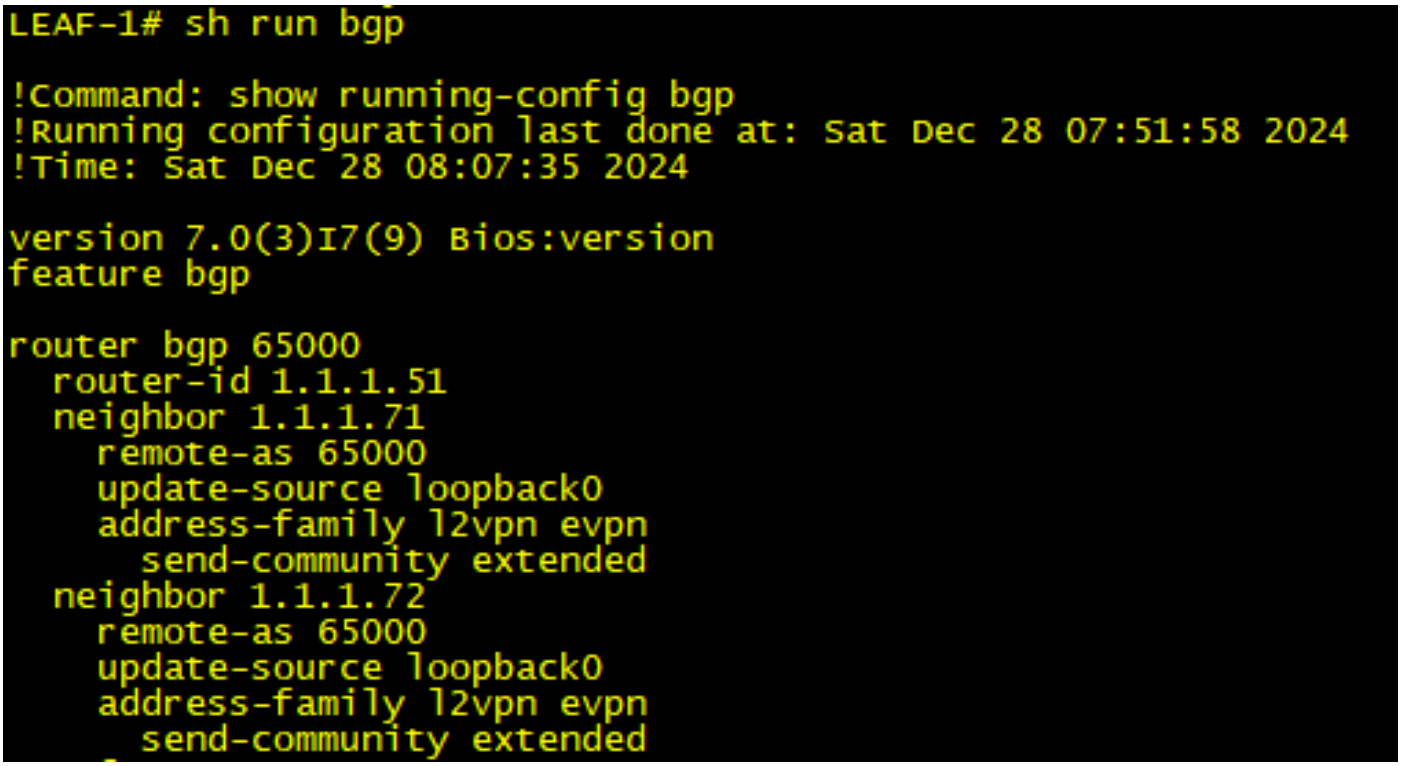
```
SPINE-1# sh run | sec "feature|nv over"
nv overlay evpn
feature ospf
feature bgp
feature pim
feature nv overlay
SPINE-1#
```

主幹交換機上的功能

- 由於主幹不需要瞭解客戶端的VLAN資訊，因此無需啟用VN網段和交換矩陣功能。

步驟5.啟用BGP鄰居關係。

- 必須啟用枝葉和主幹交換機之間的BGP。脊柱將在實驗中用作路由反射器。
- 雖然可以選擇設定路由反射器(RR)，但為了可擴充性，思科建議使用RR。




在枝葉交換機上啟用BGP

```
SPINE-1# sh run bgp

!Command: show running-config bgp
!Running configuration last done at: Sat Dec 28 07:16:33 2024
!Time: Sat Dec 28 08:08:21 2024

version 7.0(3)I7(9) Bios:version
feature bgp

router bgp 65000
  router-id 1.1.1.71
  neighbor 1.1.1.51
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 1.1.1.52
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 1.1.1.53
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client
  neighbor 1.1.1.54
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
      route-reflector-client


SPINE-1#
```

在脊柱交換機上啟用BGP

```
LEAF-1# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 1.1.1.51, local AS number 65000
BGP table version is 62, L2VPN EVPN config peers 2, capable peers 2
10 network entries and 13 paths using 2228 bytes of memory
BGP attribute entries [10/1600], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [4/16]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
1.1.1.71        4 65000     146     121       62    0    0 01:45:52 3
1.1.1.72        4 65000     141     114       62    0    0 01:39:12 3
LEAF-1#
```

枝葉交換機上的BGP狀態

```
SPINE-1# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 1.1.1.71, local AS number 65000
BGP table version is 98, L2VPN EVPN config peers 4, capable peers 4
9 network entries and 9 paths using 2124 bytes of memory
BGP attribute entries [7/1120], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
1.1.1.51        4 65000     147     124       98    0    0 01:46:29 2
1.1.1.52        4 65000     147     124       98    0    0 01:46:30 2
1.1.1.53        4 65000     128     155       98    0    0 02:01:15 1
1.1.1.54        4 65000     191     225       98    0    0 03:03:08 2
SPINE-1#
```

主幹交換器上的BGP狀態

步驟6.在枝葉交換機上啟用VRF情景。VRF可分隔客戶流量，並促進兩個不同的L2VNI之間通過L3VNI的通訊。

- 在VRF TENANT1下50000配L3VNI路由。

```
vrf context TENANT1
  vni 50000
  ip pim ssm range 232.0.0.0/8
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
```

L3VNI分配

步驟7.網路虛擬介面(NVE)、VXLAN識別碼(VNI)和VLAN設定。

- 設定NVE介面，使用Loopback 0作為源。定義每個VNI的多點傳送群組(其中將傳送第2層廣播、未知單點傳播和多點傳送(BUM)流量)，然後將VNI 100010和100020 ID附加到NVE介面。VXLAN標頭包含VNI用來識別其所屬的VXLAN區段的資訊。
- L3VNI實50000已連結到VRF例項(當將其傳送到主幹交換機時，VNI 50000已連線在VRF表中)。
- host-reachability protocol BGP命令在VXLAN隧道中啟用EVPN地址系列，這表示通過控制平面而非資料平面中的BGP協定獲知MAC地址和IP地址。
- 在NVE介面下配置suppress-arp。
- 將第2層和第3層VLAN連線到相關VNI。

抑制位址解析通訊協定(ARP):

多重通訊協定(MP)-BGP EVPN控制平面提供了一種稱為ARP抑制的改進，以減少ARP要求產生的廣播流量所帶來的網路泛濫。每個VNI的VTEP都會為已知IP主機保留一個ARP抑制快取表，並在為該VNI啟用ARP抑制時，在VNI網段中保留與它們對應的MAC地址。其本地VTEP會攔截ARP請求

，並在ARP抑制快取表中查詢ARP解析的IP地址，只要VNI中的終端主機提交針對另一個終端主機IP地址的ARP請求。如果本地VTEP發現匹配項，則代表遠端終端主機傳送ARP響應。然後，ARP響應向本地主機提供遠端主機的MAC地址。如果本地VTEP的ARP抑製表中沒有ARP解析的IP地址，則ARP請求會泛洪到VNI中的其他VTEP。對於向靜默網路主機發出的第一個ARP請求，可能會發生此ARP泛洪。

```
LEAF-1# sh run interface nve 1

!Command: show running-config interface nve1
!Running configuration last done at: Sat Dec 28 07:51:58 2024
!Time: Sat Dec 28 08:44:44 2024

version 7.0(3)I7(9) Bios:version

interface nve1
  no shutdown
  host-reachability protocol bgp
  source-interface loopback0
  member vni 50000 associate-vrf
  member vni 100010
    suppress-arp
    mcast-group 239.0.0.10
  member vni 100020
    suppress-arp
    mcast-group 239.0.0.20

LEAF-1#
```

NVE介面

```
LEAF-1# sh run vlan

!Command: show running-config vlan
!Running configuration last done at: Sat Dec 28 07:51:58 2024
!Time: Sat Dec 28 08:46:44 2024

version 7.0(3)I7(9) Bios:version
vlan 1,10,20,500
vlan 10
  vn-segment 100010
vlan 20
  vn-segment 100020
vlan 500
  vn-segment 50000


LEAF-1#
```

VLAN到VN網段對映

- 通過向Spine傳送PIM加入消息，NVE介面將在啟動後立即分別加入組播組239.0.0.10和239.0.0.20。

- 您還可以在映像中看到其他(S， G)表(1.1.1.54,239.0.0.20)和(10.1.1.100, 239.0.0.10/239.0.0.20)，這些表已在不同枝葉交換機的骨幹中註冊。

```
LEAF-1# sh ip mroute summary
IP Multicast Routing Table for VRF "default"
Route Statistics unavailable - only liveness detected

Total number of routes: 7
Total number of (*,G) routes: 2
Total number of (S,G) routes: 4
Total number of (*,G-prefix) routes: 1
Group count: 2, rough average sources per group: 2.0

Group: 232.0.0.0/8, Source count: 0
Source          packets      bytes          aps   pps       bit-rate     oifs
(*,G)           0            0              0     0         0.000   bps  0

Group: 239.0.0.10/32, Source count: 2
Source          packets      bytes          aps   pps       bit-rate     oifs
(*,G)           1            100            100   0         0.000   bps  1
1.1.1.53        48           4644           96    0         78.267  bps  1
10.1.1.100      1124         113514         100   0         131.467 bps  1

Group: 239.0.0.20/32, Source count: 2
Source          packets      bytes          aps   pps       bit-rate     oifs
(*,G)           1            100            100   0         0.000   bps  1
1.1.1.54        51           4944           96    0         63.200  bps  1
10.1.1.100      1116         112729         101   0         70.667  bps  1
LEAF-1#
```

Mroute表

步驟8.啟用EVPN例項。

- 為BGP下的EVPN和VRF啟用EVPN例項以及地址系列。

```
LEAF-1# sh run bgp

!Command: show running-config bgp
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 09:43:07 2024

version 7.0(3)I7(9) Bios:version
feature bgp

router bgp 65000
  router-id 1.1.1.51
  neighbor 1.1.1.71
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
  neighbor 1.1.1.72
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community extended
  vrf TENANT1
    address-family ipv4 unicast
      redistribute direct route-map REDIST
evpn
  vni 100010 l2
    rd auto
    route-target import auto
    route-target export auto
  vni 100020 l2
    rd auto
    route-target import auto
    route-target export auto
vrf context TENANT1
```

EVPN例項

- route-map REDIST的唯一用途是允許所有內容。
- 使用redistribute direct命令，將連線的VRF感知路由提升為MP-BGP（第5類路由）。
- 上面顯示的EVPN配置與BGP用於通告MAC路由（型別2路由）的network語句相同。

步驟9.為VRF下的終端主機配置每個VLAN的交換機虛擬介面(SVI)。

- 在每個枝葉交換機上，SVI配置為本地配置的VLAN，一個SVI配置為L3VNI VLAN，以實現對稱路由資訊庫(RIB)。

對稱肋：

- 當終端主機將資料包傳送到不同的網路並接收到枝葉交換機時，它將首先被處理為L2VNI，然後使用VRF將其置於L3VNI並傳送到遠端枝葉。
- 遠端枝葉首先使用路由接收VRF表中的資料包，然後橋接到L2VNI並將其傳送到終端主機。
- 通過這種方式，實現了對稱路由(B-R-R-B)。

```
LEAF-1# sh run interface vlan 10,vlan 20,vlan 500

!Command: show running-config interface Vlan10, Vlan20, Vlan500
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 10:00:26 2024

version 7.0(3)I7(9) Bios:version

interface Vlan10
  no shutdown
  mtu 9216
  vrf member TENANT1
  no ip redirects
  ip address 192.168.10.254/24
  no ipv6 redirects
  ip pim sparse-mode
  fabric forwarding mode anycast-gateway

interface Vlan20
  no shutdown
  mtu 9216
  vrf member TENANT1
  no ip redirects
  ip address 192.168.20.254/24
  no ipv6 redirects
  ip pim sparse-mode
  fabric forwarding mode anycast-gateway

interface Vlan500
  no shutdown
  vrf member TENANT1
  no ip redirects
  ip forward
  no ipv6 redirects

LEAF-1#
```

VLAN介面

- VLAN 500下的IP forward命令用於為所有VXLAN啟用第3層轉發。無需配置IP地址，因為它只處理從L2VNI表到L3VNI表的資料包。

```
LEAF-1# show bgp vpnv4 unicast vrf TENANT1
BGP routing table information for VRF default, address family VPNv4 Unicast
BGP table version is 15, Local Router ID is 1.1.1.51
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 1.1.1.51:3    (VRF TENANT1)
*>r192.168.10.0/24    0.0.0.0                  0        100      32768 ?
*>i192.168.10.20/32   1.1.1.53                 0        100          0 i
*>r192.168.20.0/24    0.0.0.0                  0        100      32768 ?
*>i192.168.20.20/32   1.1.1.54                 0        100          0 i

LEAF-1#
```

學習VRF TENANT1的BGP VPNv4路由

- 每個VLAN的IP地址對於所有枝葉交換機上的所有SVI都是通用的。這稱為任播IP，它用於移動性管理，在該管理中，終端可以無縫地與其他主機通訊，而不會發生任何中斷。

步驟10.為終端主機啟用交換矩陣轉發任播網關MAC。

- 它確保連線到交換矩陣的裝置的無縫第3層網關冗餘和最佳化轉發。
- 任播網關MAC地址是一個全域性一致的MAC地址，用於交換矩陣中的所有第3層網關。
- 此概念與第一躍點備援通訊協定(FHRP)中使用的概念相同，其中每個群組都獲發一個虛擬MAC。

```
LEAF-1# show running-config fabric forwarding

!Command: show running-config fabric forwarding
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 10:08:08 2024

version 7.0(3)I7(9) Bios:version
nv overlay evpn
feature fabric forwarding

fabric forwarding anycast-gateway-mac 0000.1234.5678


interface Vlan10
  fabric forwarding mode anycast-gateway

interface Vlan20
  fabric forwarding mode anycast-gateway

LEAF-1#
```

啟用交換矩陣轉發

步驟11.為成員埠啟用接入/中繼VLAN。

vPC交換機：

```
LEAF-1# sh run int po10 membership

!Command: show running-config interface port-channel10 membership
!Running configuration last done at: Sat Dec 28 09:22:19 2024
!Time: Sat Dec 28 10:13:19 2024

version 7.0(3)I7(9) Bios:version

interface port-channel10
  switchport mode trunk
  vpc 10

interface Ethernet1/5

  switchport mode trunk
  channel-group 10 mode active

LEAF-1#
```

啟用到vPC成員介面的中繼埠

非vPC交換機：

```
LEAF-3# show running-config interface e1/3

!Command: show running-config interface Ethernet1/3
!Running configuration last done at: Sat Dec 28 09:28:18 2024
!Time: Sat Dec 28 10:14:42 2024

version 7.0(3)I7(9) Bios:version

interface Ethernet1/3
  switchport access vlan 10
  spanning-tree port type edge

LEAF-3#
```

啟用連線到非vPC成員介面的中繼埠

## 驗證

- 檢查ARP和MAC地址表。

```
LEAF-1# sh ip arp vrf TENANT1

Flags: * - Adjacencies learnt on non-active FHRP router
       + - Adjacencies synced via CFSoE
       # - Adjacencies Throttled for Glean
       CP - Added via L2RIB, Control plane Adjacencies
       PS - Added via L2RIB, Peer Sync
       RO - Re-Originated Peer Sync Entry
       D - Static Adjacencies attached to down interface

IP ARP Table for context TENANT1
Total number of entries: 2
Address         Age       MAC Address     Interface       Flags
192.168.20.10   00:00:36  0000.0000.bbbb  Vlan20
192.168.10.10   00:04:19  0000.0000.aaaa  Vlan10
LEAF-1#  sh ip arp suppression-cache deta

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   00:04:33 0000.0000.aaaa   10 port-channel10      L
192.168.10.20   00:55:53 0000.0000.cccc   10 (null)              R        1.1.1.53
192.168.20.10   00:00:50 0000.0000.bbbb   20 port-channel20      L
192.168.20.20   03:26:04 0000.0000.dddd   20 (null)              R        1.1.1.54
LEAF-1#
```

枝葉交換機1上的ARP和MAC表

```
LEAF-2# show ip arp vrf TENANT1

Flags: * - Adjacencies learnt on non-active FHRP router
       + - Adjacencies synced via CFSoE
       # - Adjacencies Throttled for Glean
       CP - Added via L2RIB, Control plane Adjacencies
       PS - Added via L2RIB, Peer Sync
       RO - Re-Originated Peer Sync Entry
       D - Static Adjacencies attached to down interface

IP ARP Table for context TENANT1
Total number of entries: 2
Address         Age       MAC Address     Interface       Flags
192.168.20.10   00:01:28  0000.0000.bbbb  vlan20          +
192.168.10.10   00:00:11  0000.0000.aaaa  vlan10          +
LEAF-2#
```

枝葉交換機2上的ARP和MAC表

- 兩個對等點都會維護ARP專案。
- 檢查網路虛擬介面(NVI)狀態。

vPC交換機：

```
LEAF-1# show nve peers
Interface Peer-IP          State LearnType Uptime   Router-Mac
--------- ---------------  ----- --------- -------- -----------------
nve1      1.1.1.53         Up    CP        01:09:04 5000.0003.0007
nve1      1.1.1.54         Up    CP        03:39:16 5000.0004.0007

LEAF-1# show nve vni
Codes: CP - Control Plane        DP - Data Plane
       UC - Unconfigured         SA - Suppress ARP
       SU - Suppress Unknown Unicast
       Xconn - Crossconnect
       MS-IR - Multisite Ingress Replication

Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      50000    n/a               Up    CP   L3 [TENANT1]
nve1      100010   239.0.0.10        Up    CP   L2 [10]            SA
nve1      100020   239.0.0.20        Up    CP   L2 [20]            SA

LEAF-1#
```

vPC交換機上的NVE對等體

非vPC交換機：

```
LEAF-3# show nve peers
Interface Peer-IP          State LearnType Uptime   Router-Mac
--------- ---------------  ----- --------- -------- -----------------
nve1      1.1.1.54         Up    CP        01:14:00 5000.0004.0007
nve1      10.1.1.100       Up    CP        01:14:16 5000.0001.0007

LEAF-3#
```

非vPC交換機上的NVE對等體

- 在這裡，您會發現對等IP是10.1.1.100而不是主環回IP地址，因此返回的資料包將針對該IP路由到任何vPC交換機。
- 檢查BGP EVPN路由。

```
LEAF-1# show l2route evpn mac-ip all
Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
(Dup):Duplicate (Spl):Split (Rcv):Recv(D):Del Pending (S):Stale (C):Clear
(Ps):Peer Sync (Ro):Re-Originated
Topology    Mac Address    Prod   Flags         Seq No     Host IP         Next-Hops
----------- -------------- ------ ---------- ---------- --------------- ---------------
10          0000.0000.aaaa HMM    --            0          192.168.10.10   Local
10          0000.0000.cccc BGP    --            0          192.168.10.20   1.1.1.53
20          0000.0000.bbbb HMM    --            0          192.168.20.10   Local
20          0000.0000.dddd BGP    --            0          192.168.20.20   1.1.1.54
LEAF-1#
```

BGP l2route EVPN MAC-IP

```
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

Topology    Mac Address    Prod   Flags         Seq No     Next-Hops
----------- -------------- ------ ------------ ---------- ----------------
10          0000.0000.aaaa Local  L,            0          Po10
10          0000.0000.cccc BGP    Spl           0          1.1.1.53
20          0000.0000.bbbb Local  L,            0          Po20
20          0000.0000.dddd BGP    SplRcv        0          1.1.1.54
500         5000.0003.0007 VXLAN  Rmac          0          1.1.1.53
500         5000.0004.0007 VXLAN  Rmac          0          1.1.1.54
LEAF-1#
```

BGP l2route EVPN MAC

```
LEAF-1# show bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 1.1.1.51, local AS number 65000
BGP table version is 134, L2VPN EVPN config peers 2, capable peers 2
12 network entries and 15 paths using 2568 bytes of memory
BGP attribute entries [12/1920], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [4/16]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
1.1.1.71        4 65000     312      263      134    0    0 03:46:01 3
1.1.1.72        4 65000     307      256      134    0    0 03:39:21 3
LEAF-1#
```

BGP EVPN摘要

```
LEAF-1# show bgp l2vpn evpn
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 146, Local Router ID is 1.1.1.51
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network            Next Hop            Metric     LocPrf     Weight Path
Route Distinguisher: 1.1.1.51:32777    (L2VNI 100010)
*>l[2]:[0]:[0]:[48]:[0000.0000.aaaa]:[0]:[0.0.0.0]/216
                      10.1.1.100                             100      32768 i
*>l[2]:[0]:[0]:[48]:[0000.0000.aaaa]:[32]:[192.168.10.10]/272
                      10.1.1.100                             100      32768 i
*>i[2]:[0]:[0]:[48]:[0000.0000.cccc]:[32]:[192.168.10.20]/272
                      1.1.1.53                               100          0 i

Route Distinguisher: 1.1.1.51:32787    (L2VNI 100020)
*>l[2]:[0]:[0]:[48]:[0000.0000.bbbb]:[0]:[0.0.0.0]/216
                      10.1.1.100                             100      32768 i
*>i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[0]:[0.0.0.0]/216
                      1.1.1.54                               100          0 i
*>l[2]:[0]:[0]:[48]:[0000.0000.bbbb]:[32]:[192.168.20.10]/272
                      10.1.1.100                             100      32768 i
*>i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[32]:[192.168.20.20]/272
                      1.1.1.54                               100          0 i

Route Distinguisher: 1.1.1.53:32777
*>i[2]:[0]:[0]:[48]:[0000.0000.cccc]:[32]:[192.168.10.20]/272
                      1.1.1.53                               100          0 i
* i                   1.1.1.53                               100          0 i

Route Distinguisher: 1.1.1.54:32787
* i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[0]:[0.0.0.0]/216
                      1.1.1.54                               100          0 i
*>i                   1.1.1.54                               100          0 i
* i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[32]:[192.168.20.20]/272
                      1.1.1.54                               100          0 i
*>i                   1.1.1.54                               100          0 i

Route Distinguisher: 1.1.1.51:3    (L3VNI 50000)
*>i[2]:[0]:[0]:[48]:[0000.0000.cccc]:[32]:[192.168.10.20]/272
                      1.1.1.53                               100          0 i
*>i[2]:[0]:[0]:[48]:[0000.0000.dddd]:[32]:[192.168.20.20]/272
                      1.1.1.54                               100          0 i

LEAF-1#
```

BGP EVPN路由

- 常見的問題是枝葉交換機如何獲取遠端主機的MAC條目。此過程由無償ARP實現。啟用網路埠後，它會立即傳送ARP請求以驗證IP地址的唯一性。然後，每台枝葉交換機記錄MAC地址並將其包含在BGP更新資料包中。這允許其他枝葉交換機相應地更新其各自的MAC地址表。但是可能出現終端主機不生成無償ARP（靜默主機）的情況，在這種情況下，ARP請求將廣播到枝葉，並且由於它是廣播請求，枝葉交換機將生成針對特定VNI的組播請求到相應的組。在本例中為239.0.0.10和239.0.0.20。

- 允許在同一VNI中從Host-1 ping Host-3，並檢視捕獲。

```
HOST-1#ping 192.168.10.20 rep 2
Type escape sequence to abort.
Sending 2, 100-byte ICMP Echos to 192.168.10.20, timeout is 2 seconds:
!!
Success rate is 100 percent (2/2), round-trip min/avg/max = 11/11/12 ms
HOST-1#
```

從HOST-1 ping HOST-3

透過VXLAN傳輸的網際網路控制訊息通訊協定(ICMP)封包：

```
> Frame 213: 164 bytes on wire (1312 bits), 164 bytes captured (1312 bits) on interface -, id 0
> Ethernet II, Src: 50:00:00:06:00:07 (50:00:00:06:00:07), Dst: 50:00:00:03:00:07 (50:00:00:03:00:07)
> Internet Protocol Version 4, Src: 10.1.1.100, Dst: 1.1.1.53
v User Datagram Protocol, Src Port: 50413, Dst Port: 4789
    Source Port: 50413
    Destination Port: 4789
    Length: 130
  > Checksum: 0x0000 [zero-value ignored]
    [Stream index: 24]
    [Stream Packet Number: 1]
  > [Timestamps]
    UDP payload (122 bytes)
v Virtual eXtensible Local Area Network
  > Flags: 0x0800, VXLAN Network ID (VNI)
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 100010
    Reserved: 0
> Ethernet II, Src: 00:00:00_00:aa:aa (00:00:00:00:aa:aa), Dst: 00:00:00_00:cc:cc (00:00:00:00:cc:cc)
v Internet Protocol Version 4, Src: 192.168.10.10, Dst: 192.168.10.20
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 100
    Identification: 0x0000 (0)
  > 000. .... = Flags: 0x0
    ...0 0000 0000 0000 = Fragment Offset: 0
    Time to Live: 255
    Protocol: ICMP (1)
    Header Checksum: 0x262a [validation disabled]
    [Header checksum status: Unverified]
    Source Address: 192.168.10.10
    Destination Address: 192.168.10.20
    [Stream index: 11]
> Internet Control Message Protocol
```

顯示ICMP請求資料包通過L2VNI資料包傳輸的Wireshark捕10010

- 您可以看到，來源IP是10.1.1.100，且連線埠4789作為UDP目的地。
- 由於這是VNI內通訊，VLAN 10將使用VNI 100010，而VLAN 20將使用VNI 1000。
- 允許使用不同的VNI從Host-1 ping Host-4，並檢視捕獲。

```
HOST-1#ping 192.168.20.20
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.20.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 11/13/21 ms
HOST-1#
```

從HOST-1 ping HOST-4

使用VXLAN的ICMP封包：

```
> Frame 27: 164 bytes on wire (1312 bits), 164 bytes captured (1312 bits) on interface -, id 0
> Ethernet II, Src: 50:00:00:05:00:07 (50:00:00:05:00:07), Dst: 50:00:00:04:00:07 (50:00:00:04:00:07)
> Internet Protocol Version 4, Src: 10.1.1.100, Dst: 1.1.1.54
v User Datagram Protocol, Src Port: 54712, Dst Port: 4789
    Source Port: 54712
    Destination Port: 4789
    Length: 130
  > Checksum: 0x0000 [zero-value ignored]
    [Stream index: 3]
    [Stream Packet Number: 1]
  > [Timestamps]
    UDP payload (122 bytes)
v Virtual eXtensible Local Area Network
  > Flags: 0x0800, VXLAN Network ID (VNI)
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 50000
    Reserved: 0
> Ethernet II, Src: 50:00:00:01:00:07 (50:00:00:01:00:07), Dst: 50:00:00:04:00:07 (50:00:00:04:00:07)
> Internet Protocol Version 4, Src: 192.168.10.10, Dst: 192.168.20.20
> Internet Control Message Protocol
```

顯示ICMP請求資料包通過L3VNI資料包傳輸的Wireshark捕50000

- 由於這是VNI間通訊，因此將使用L50000VNI通訊協定。
- 檢查終端主機的ARP表。

```
HOST-1#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.10.10           -   0000.0000.aaaa  ARPA   GigabitEthernet0/0
Internet  192.168.10.20          18   0000.0000.cccc  ARPA   GigabitEthernet0/0
Internet  192.168.10.254          3   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-1#
```

主機1 ARP條目

```
HOST-2#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.20.10           -   0000.0000.bbbb  ARPA   GigabitEthernet0/0
Internet  192.168.20.20          44   0000.0000.dddd  ARPA   GigabitEthernet0/0
Internet  192.168.20.254          4   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-2#
```

主機2 ARP條目

```
HOST-3#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.10.10         103   0000.0000.aaaa  ARPA   GigabitEthernet0/0
Internet  192.168.10.20           -   0000.0000.cccc  ARPA   GigabitEthernet0/0
Internet  192.168.10.254         10   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-3#
```

主機3 ARP條目

```
HOST-4#sh ip arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.20.10          43   0000.0000.bbbb  ARPA   GigabitEthernet0/0
Internet  192.168.20.20           -   0000.0000.dddd  ARPA   GigabitEthernet0/0
Internet  192.168.20.254          6   0000.1234.5678  ARPA   GigabitEthernet0/0
HOST-4#
```

主機4 ARP條目

```
HOST-4#tclsh
HOST-4(tcl)#set ip_list {192.168.10.10 192.168.10.20 192.168.20.10 192.168.20.20}
192.168.10.10 192.168.10.20 192.168.20.10 192.168.20.20
HOST-4(tcl)#foreach ip $ip_list {
HOST-4(tcl)#foreach ip $ip_list {
+>            puts "Pinging $ip rep 50 size 1500"
+>            set result [exec ping $ip]
+>            puts $result
+>        }
Pinging 192.168.10.10 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.10.10, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 12/14/16 ms
Pinging 192.168.10.20 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.10.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 10/12/15 ms
Pinging 192.168.20.10 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.20.10, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 8/11/17 ms
Pinging 192.168.20.20 rep 50 size 1500

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.20.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/3 ms

HOST-4(tcl)#
```

從HOST-4 ping所有其他終端主機

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）。