

# VXLAN資料平面學習實驗室（泛洪和學習機制）

## 目錄

---

- [簡介](#)
- [必要條件](#)
  - [需求](#)
  - [採用元件](#)
- [背景資訊](#)
  - [使用的術語](#)
- [設定](#)
  - [網路圖表](#)
  - [組態](#)

---

## 簡介

本文檔介紹如何使用VXLAN和Flood and Learn方法設定Nexus 9Kv交換機的CML。

## 必要條件

### 需求

思科建議您瞭解以下主題：

- 瞭解路由和交換
- 多點傳送路由概念，例如集結點(RP)和平台無關多點傳送(PIM)

### 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，枝葉和主幹均採用帶Nexus 9000V交換機的思科建模實驗室(CML)。

|     |                     |                         |
|-----|---------------------|-------------------------|
| 枝葉1 | Loopback0 - 1.1.1.1 | Loopback1 - 10.10.10.10 |
|-----|---------------------|-------------------------|

|      |                     |                                |
|------|---------------------|--------------------------------|
| 枝葉2  | Loopback0 - 2.2.2.2 | Loopback1 - 20.20.20.20        |
| 枝葉3  | Loopback0 - 3.3.3.3 | Loopback1 - 30.30.30.30        |
| 骨幹1  | Loopback0 - 4.4.4.4 | Loopback1 - 60.60.60.60 — 任播RP |
| 骨幹2  | Loopback0 - 5.5.5.5 | Loopback1 - 60.60.60.60 — 任播RP |
| 案頭子網 | 192.168.100.0/24    |                                |

## 使用的術語

虛擬可擴充區域網路(VXLAN)通道端點(VTEP) — 將MAC流量封裝到IP流量中，並將MAC流量路由到其他VTEP。

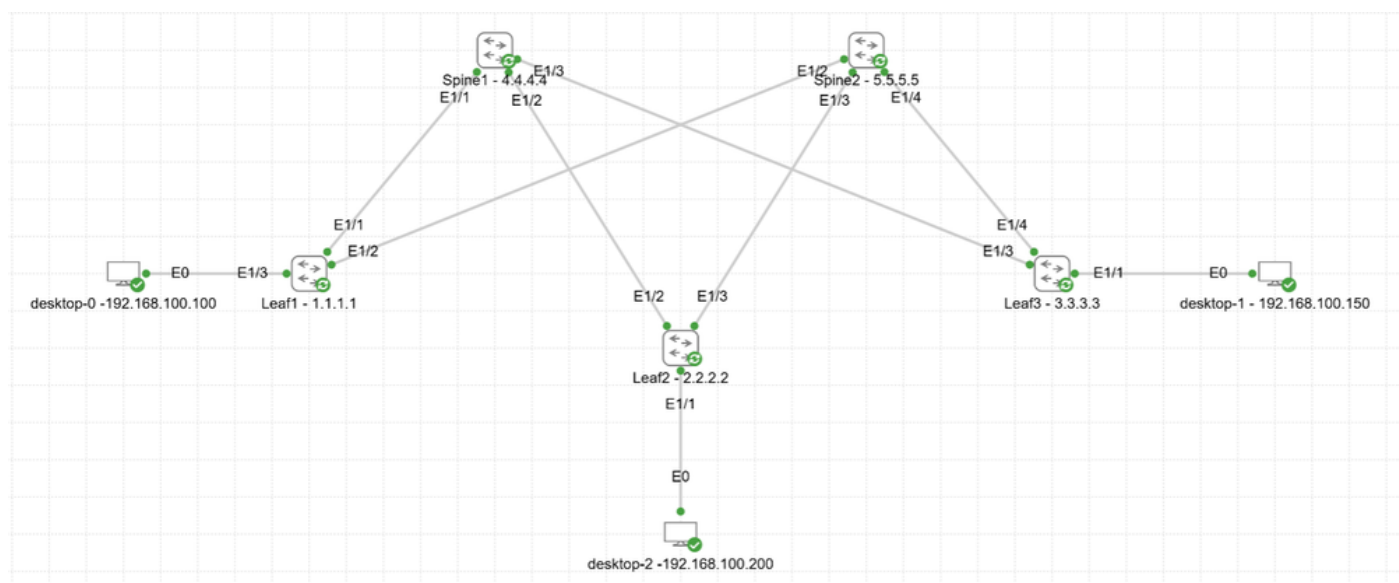
VXLAN網路識別碼(VNID)- VXLAN標頭中的ID，用於識別網路且可對映到VLAN。從轉發角度來看，VNID是一個廣播域。

網路虛擬介面(NVE) — 進行封裝和解除封裝的邏輯介面。

廣播、未知單點傳送和多點傳送(BUM)

## 設定

### 網路圖表



網路連線圖

## 組態

### 步驟 1.

- 啟用開放最短路徑優先(OSPF)功能。
- 向所有裝置新增環回。
- 在乙太網介面和環回上啟用OSPF。

```
leaf1# show running-config interface ethernet 1/1

!Command: show running-config interface Ethernet1/1
!Running configuration last done at: Tue Dec 24 13:12:55 2024
!Time: Wed Dec 25 05:24:23 2024

version 9.3(8) Bios:version

interface Ethernet1/1
  no switchport
  ip address 70.0.0.2/30
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

leaf1# show running-config interface loopback 0

!Command: show running-config interface loopback0
!Running configuration last done at: Tue Dec 24 13:12:55 2024
!Time: Wed Dec 25 05:24:24 2024

version 9.3(8) Bios:version

interface loopback0
  ip address 1.1.1.1/32
  ip router ospf 1 area 0.0.0.0
```

在枝葉介面上啟用OSPF

```

spinel# show running-config interface ethernet 1/1

!Command: show running-config interface Ethernet1/1
!Running configuration last done at: Tue Dec 24 13:16:16 2024
!Time: Wed Dec 25 05:25:46 2024

version 9.3(8) Bios:version

interface Ethernet1/1
  no switchport
  ip address 70.0.0.1/30
  ip router ospf 5 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

spinel# show running-config interface loopback 0

!Command: show running-config interface loopback0
!Running configuration last done at: Tue Dec 24 13:16:16 2024
!Time: Wed Dec 25 05:25:51 2024

version 9.3(8) Bios:version

interface loopback0
  ip address 4.4.4.4/32
  ip router ospf 5 area 0.0.0.0

```

在主幹介面上啟用OSPF

在枝葉和主幹交換機之間建立OSPF鄰居關係。

```

spinel# show ip ospf neighbors
OSPF Process ID 5 VRF default
Total number of neighbors: 3
Neighbor ID      Pri State           Up Time  Address      Interface
1.1.1.1          1 FULL/DR         16:22:51 70.0.0.2     Eth1/1
2.2.2.2          1 FULL/DR         16:22:52 50.0.0.2     Eth1/2
3.3.3.3          1 FULL/DR         16:22:52 30.0.0.2     Eth1/3

```

使用枝葉交換機建立OSPF鄰居

```

leaf1# show ip ospf neighbors
OSPF Process ID 1 VRF default
Total number of neighbors: 2
Neighbor ID      Pri State           Up Time  Address      Interface
4.4.4.4          1 FULL/BDR        16:15:40 70.0.0.1     Eth1/1
5.5.5.5          1 FULL/BDR        16:15:10 80.0.0.1     Eth1/2

```

使用主幹交換機建立OSPF鄰居

```
leaf1# ping 3.3.3.3 source-interface loopback 0
PING 3.3.3.3 (3.3.3.3): 56 data bytes
64 bytes from 3.3.3.3: icmp_seq=0 ttl=253 time=6.616 ms
64 bytes from 3.3.3.3: icmp_seq=1 ttl=253 time=6.695 ms
64 bytes from 3.3.3.3: icmp_seq=2 ttl=253 time=6.018 ms
64 bytes from 3.3.3.3: icmp_seq=3 ttl=253 time=6.52 ms
64 bytes from 3.3.3.3: icmp_seq=4 ttl=253 time=7.768 ms
```

從Leaf1到Leaf3的Ping可達性

## 步驟 2.

在枝葉交換機上新增將用於VXLAN的附加環回介面。此外，檢驗所有枝葉交換機到交換矩陣的可達性。

```
leaf2# show running-config interface loopback 1

!Command: show running-config interface loopback1
!Running configuration last done at: Wed Dec 25 05:41:46 2024
!Time: Wed Dec 25 05:42:33 2024

version 9.3(8) Bios:version

interface loopback1
  description Vxlan loopback
  ip address 20.20.20.20/32
  ip router ospf 2 area 0.0.0.0
```

VXLAN的回送

```
leaf2# ping 10.10.10.10 source 20.20.20.20
PING 10.10.10.10 (10.10.10.10) from 20.20.20.20: 56 data bytes
64 bytes from 10.10.10.10: icmp_seq=0 ttl=253 time=7.187 ms
64 bytes from 10.10.10.10: icmp_seq=1 ttl=253 time=6.248 ms
64 bytes from 10.10.10.10: icmp_seq=2 ttl=253 time=5.472 ms
64 bytes from 10.10.10.10: icmp_seq=3 ttl=253 time=4.741 ms
64 bytes from 10.10.10.10: icmp_seq=4 ttl=253 time=4.887 ms

--- 10.10.10.10 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 4.741/5.707/7.187 ms
```

從枝葉2到枝葉1的覆蓋可達性

## 步驟 3.

在主幹上配置PIM任意源組播(ASM)和任播RP:

- 啟用PIM功能。
- 在所有底層鏈路上啟用PIM。
- 在用於任播RP的脊柱上建立新的環回。
- 將此環回通告到OSPF。
- 配置主幹上的任播RP ( Nexus功能 ) 。
- 在所有裝置上配置RP。

#### 任播RP:

任播RP是一種提供快速RP故障切換和RP負載共用的機制。任播RP涉及在兩個或多個路由器上使用與RP相同的IP地址(rp-address)。必須在內部網關協定(IGP)中通告此IP地址，以便其他路由器可以選擇到rp-address的最佳路徑。出現故障時，收斂時間將與IGP相同。

如果有多個RP具有相同的IP地址，則可以確保源路由器和接收器始終根據單播路由表路由到最近的RP。來自接收器的PIM加入消息可以傳送到一個RP，而PIM指定的路由器將其本地源註冊到另一個RP。

必須在不同的RP之間同步資訊，因為某些傳送方和接收方可以加入路由器1作為RP，而其它傳送方和接收方可以加入路由器2作為RP。如果路由器沒有關於所有源的完整資訊，組播通訊可能會中斷。為了解決此問題，需要一種機制來同步所有作為RP的路由器之間的源資訊。

有兩種協定可以達到此目的：多點傳送來源探索通訊協定(MSDP)和PIM。

```
spinel# show running-config interface loopback 1

!Command: show running-config interface loopback1
!Running configuration last done at: Tue Dec 24 13:16:16 2024
!Time: Wed Dec 25 05:50:44 2024

version 9.3(8) Bios:version

interface loopback1
  description Anycast RP loopback
  ip address 60.60.60.60/32
  ip router ospf 5 area 0.0.0.0
  ip pim sparse-mode

spinel# show running-config | section rp
ip pim rp-address 60.60.60.60 group-list 224.0.0.0/4
ip pim anycast-rp 60.60.60.60 4.4.4.4
ip pim anycast-rp 60.60.60.60 5.5.5.5
```

60.60.60.60是任播RP IP，4.4.4.4/5.5.5.5是骨幹1和骨幹2的環回IP

```

leaf2# show running-config interface ethernet 1/2

!Command: show running-config interface Ethernet1/2
!Running configuration last done at: Wed Dec 25 05:41:46 2024
!Time: Wed Dec 25 05:51:18 2024

version 9.3(8) Bios:version

interface Ethernet1/2
  no switchport
  ip address 50.0.0.2/30
  ip router ospf 2 area 0.0.0.0
  ip pim sparse-mode
  no shutdown

leaf2# show running-config | section rp-address
ip pim rp-address 60.60.60.60 group-list 224.0.0.0/4

```

枝葉交換機上的任播RP配置

```

spinel# show ip pim neighbor
PIM Neighbor Status for VRF "default"
Neighbor      Interface      Uptime      Expires      DR      Bidir-  BFD
ECMP Redirect                                     Priority Capable State
Capable
70.0.0.2      Ethernet1/1    16:44:28    00:01:39    1       yes     n/a
no
50.0.0.2      Ethernet1/2    16:44:28    00:01:44    1       yes     n/a
no
30.0.0.2      Ethernet1/3    16:44:28    00:01:25    1       yes     n/a
no
spinel# 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 60.60.60.60 members:
  4.4.4.4*  5.5.5.5

RP: 60.60.60.60*, (0),
  uptime: 16:57:06  priority: 255,
  RP-source: (local),
  group ranges:
  224.0.0.0/4

```

PIM鄰居關係形成



附註：不要忘記將PIM放在要用於枝葉交換機上VXLAN的環回上。

---

#### 步驟 4.

- 啟用VXLAN功能。
- 啟用功能以將VLAN對應到虛擬網路識別碼(VNI)。
- 建立NVE。
- 配置通往案頭的接入埠。

```
feature vn-segment-vlan-based
feature nv overlay
```

啟用VXLAN功能

```
!Command: show running-config interface nve1
!Running configuration last done at: Wed Dec 25 06:08:01 2024
!Time: Wed Dec 25 06:08:04 2024

version 9.3(8) Bios:version

interface nve1
  no shutdown
  source-interface loopback1
  member vni 10000
  mcast-group 239.0.0.1
```

建立NVE介面

```
vlan 10
  vn-segment 10000
```

VLAN到VN區段對應

```
leaf1# show interface nve 1
nve1 is up
admin state is up, Hardware: NVE
  MTU 9216 bytes
  Encapsulation VXLAN
  Auto-mdix is turned off
  RX
    ucast: 39 pkts, 3346 bytes - mcast: 0 pkts, 0 bytes
  TX
    ucast: 18 pkts, 2216 bytes - mcast: 0 pkts, 0 bytes
```

NVE介面的狀態

從Desktop0 ping Desktop1和Desktop2，以驗證其連通性。

從Desktop0向Desktop1發起地址解析協定(ARP)請求時，該ARP資料包將傳送到Leaf1。然後，Leaf1將使用用於VNI1000的組播地址239.0.0.1將資料包轉發到主幹裝置。主幹裝置會將資料包組播到屬於同一VNI 10000的所有主幹裝置。

從Desktop0(192.168.100.100)對Desktop1(192.168.100.150)和Desktop2(192.168.100.200)執行ping。

```

inserthostname-here:~$ ping 192.168.100.150
PING 192.168.100.150 (192.168.100.150): 56 data bytes
64 bytes from 192.168.100.150: seq=0 ttl=42 time=8.477 ms
64 bytes from 192.168.100.150: seq=1 ttl=42 time=12.791 ms
64 bytes from 192.168.100.150: seq=2 ttl=42 time=8.352 ms
^C
--- 192.168.100.150 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 8.352/9.873/12.791 ms
inserthostname-here:~$
inserthostname-here:~$
inserthostname-here:~$ ping 192.168.100.200
PING 192.168.100.200 (192.168.100.200): 56 data bytes
64 bytes from 192.168.100.200: seq=0 ttl=42 time=15.432 ms
64 bytes from 192.168.100.200: seq=1 ttl=42 time=9.228 ms
64 bytes from 192.168.100.200: seq=2 ttl=42 time=7.133 ms
^C
--- 192.168.100.200 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 7.133/10.597/15.432 ms
inserthostname-here:~$ arp -a
? (192.168.100.150) at 52:54:00:05:84:a2 [ether] on eth0
? (192.168.100.1) at 00:01:00:01:00:01 [ether] on eth0
? (192.168.100.200) at 52:54:00:10:70:ae [ether] on eth0

```

從Desktop0 ping Desktop1和Desktop2

LEAF1正在與LEAF3形成NVE對等體。

<#root>

leaf1# show nve peers

| Interface | Peer-IP | State | LearnType | Uptime | Route |
|-----------|---------|-------|-----------|--------|-------|
|-----------|---------|-------|-----------|--------|-------|

r-Mac

|       |  |  |  |  |  |
|-------|--|--|--|--|--|
| ----- |  |  |  |  |  |
| ----- |  |  |  |  |  |

nve1

30.30.30.30

|    |    |          |     |
|----|----|----------|-----|
| Up | DP | 00:10:23 | n/a |
|----|----|----------|-----|

leaf1# show nve vni 10000

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      | 10000 | 239.0.0.1       | Up    | DP   | L2   | [10]     |       |

LEAF3正在與LEAF1形成NVE對等體。

<#root>

leaf3# show nve peers

| Interface | Peer-IP     | State | LearnType | Uptime   | Route |
|-----------|-------------|-------|-----------|----------|-------|
| r-Mac     |             |       |           |          |       |
| nve1      | 10.10.10.10 | Up    | DP        | 00:10:56 | n/a   |

leaf3# show nve vni 10000

- 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      | 10000 | 239.0.0.1       | Up    | DP   | L2   | [10]     |       |

leaf1# show mac address-table

Legend:

\* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC

age - seconds since last seen,+ - primary entry using vPC Peer-Link,

(T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan

| VLAN                                      | MAC Address    |        | Type | age               | Secure | NTFY        | Ports                         |
|-------------------------------------------|----------------|--------|------|-------------------|--------|-------------|-------------------------------|
| -----+-----+-----+-----+-----+-----+----- |                |        |      |                   |        |             |                               |
| * 10                                      |                |        |      |                   |        |             |                               |
| 5254.0004.8b92                            |                |        |      |                   |        |             |                               |
| dynamic                                   | 0              | F      | F    | Eth1/3            |        |             | ----- MAC Address of De       |
| * 10                                      |                |        |      |                   |        |             |                               |
| 5254.0005.84a2                            |                |        |      |                   |        |             |                               |
| dynamic                                   | 0              | F      | F    | nve1(30.30.30.30) |        |             | ----- MAC Address of Desktop1 |
| G -                                       | 5206.ab8a.1b08 | static | -    | F                 | F      | sup-eth1(R) |                               |

leaf3# show mac address-table

Legend:

\* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC

age - seconds since last seen,+ - primary entry using vPC Peer-Link,

(T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan

| VLAN                                      | MAC Address    |        | Type | age               | Secure | NTFY        | Ports                             |
|-------------------------------------------|----------------|--------|------|-------------------|--------|-------------|-----------------------------------|
| -----+-----+-----+-----+-----+-----+----- |                |        |      |                   |        |             |                                   |
| * 10                                      |                |        |      |                   |        |             |                                   |
| 5254.0004.8b92                            |                |        |      |                   |        |             |                                   |
| dynamic                                   | 0              | F      | F    | nve1(10.10.10.10) |        |             | ----- MAC Address of Desktop0 con |
| * 10                                      |                |        |      |                   |        |             |                                   |
| 5254.0005.84a2                            |                |        |      |                   |        |             |                                   |
| dynamic                                   | 0              | F      | F    | Eth1/1            |        |             | ----- MAC Address of Desk         |
| G -                                       | 5206.0619.1b08 | static | -    | F                 | F      | sup-eth1(R) |                                   |

以下是從Leaf1向組播發起ARP資料包時Wireshark的快照。

| No. | Time     | Source            | Destination       | Protocol | Length | Info                                          |
|-----|----------|-------------------|-------------------|----------|--------|-----------------------------------------------|
| 7   | 5.105615 | 52:54:00:05:84:a2 | 52:54:00:04:8b:92 | ARP      | 110    | 192.168.100.200 is at 52:54:00:05:84:a2       |
| 8   | 7.019252 | 52:54:00:04:8b:92 | ff:ff:ff:ff:ff:ff | ARP      | 110    | Who has 192.168.100.200? Tell 192.168.100.100 |

- ▶ Frame 8: 110 bytes on wire (880 bits), 110 bytes captured (880 bits)
- ▶ Ethernet II, Src: 52:06:ab:8a:1b:08 (52:06:ab:8a:1b:08), Dst: IPv4mcast\_01 (01:00:5e:00:00:01)
- ▶ Internet Protocol Version 4, Src: 10.10.10.10, Dst: 239.0.0.1
- ▶ User Datagram Protocol, Src Port: 50384, Dst Port: 4789
- ▶ Virtual eXtensible Local Area Network
- ▶ Ethernet II, Src: RealtekU\_04:8b:92 (52:54:00:04:8b:92), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
- ▶ Address Resolution Protocol (request)

顯示ARP請求資料包進入組播組的Wireshark捕獲

## 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）。