

使用NX-OS的Nexus 9000上的組播自動RP故障排除

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[多點傳送概觀](#)

[PIM稀疏模式下的自動RP操作 \(控制平面工作流 \)](#)

[自動RP限制和操作限制](#)

[使用自動RP的優勢](#)

[自動RP對映代理和RP選擇過程](#)

[對映代理\(MA\)選擇](#)

[集結點\(RP\)選擇](#)

[第2層注意事項：多點傳送MAC和窺探](#)

[自動RP配置](#)

[第一跳路由器](#)

[對映代理和RP候選](#)

[IPv4可達性的路由配置](#)

[檢驗運行狀態並排除自動RP故障](#)

[步驟 1: 驗證基本IP連線能力 \(單點傳送底層驗證 \)](#)

[驗證方法](#)

[重要注意事項](#)

[步驟 2: 確定組播角色和端到端拓撲](#)

[拓撲和控制平面感知](#)

[必須清楚理解的內容](#)

[步驟 3: 根據裝置角色驗證自動RP配置](#)

[步驟 4: 驗證所有RP候選程式和所有對映代理的操作](#)

[第4.1步檢驗PIM鄰居鄰接關係](#)

[第4.2步檢驗啟用PIM的介面](#)

[步驟4.3分析show ip pim rp](#)

[第4.4步驗證自動RP選舉流程和RP選擇](#)

[步驟 5: 驗證與RP候選和對映代理的可達性](#)

[檢驗穩定的路由條目](#)

[驗證從PIM介面的Ping可達性](#)

[Ping驗證摘要](#)

[檢驗FHR和LHR上的運行狀態和組播流量轉發](#)

[第1步檢驗FHR和LHR上的RP學習](#)

[第2步在活動組播流量之前檢驗組播路由狀態](#)

[步驟3驗證具有主動多點傳播流量的多點傳播路由狀態](#)

[vPC運行的主要和次要行為](#)

[LHR組播狀態分析](#)

[RP組播狀態分析](#)

簡介

本文檔介紹自動RP如何在帶NX-OS的Cisco Nexus 9000上工作，以及如何驗證組播操作並排除故障。

必要條件

需求

- IP組播基礎知識
- IGMP和PIM稀疏模式的基本知識
- 使用Cisco Nexus 9000和NX-OS CLI的經驗
- 瞭解單播路由和RPF概念

採用元件

- Cisco Nexus 9000 系列交換器
- Cisco NX-OS 軟體
- PIM稀疏模式
- 靜態RP和自動RP

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。

多點傳送概觀

組播是一種一對多通訊模型，其中源向多個感興趣的接收方傳送單個通訊流。網路不是為每個目標建立單獨的副本，而是僅在轉發路徑分支的位置複製流量。這使得組播比廣播或重複的單播傳輸更有效。在IPv4中，多點傳播流量使用來自224.0.0.0/4範圍的目的地位址。

PIM稀疏模式是運行NX-OS的Cisco Nexus交換機支援的組播路由模型。只有在明確瞭解了接收者的興趣後，它才會轉發流量。在任意源組播設計中，接收方最初加入朝向交匯點的共用樹，源向該RP註冊。流量開始流動後，最後一跳路由器可以從共用樹移動到最短路徑樹到源路徑。

定義組播中使用的術語非常重要，因為準確的故障排除取決於瞭解控制平面事件、路由條目和轉發決策是如何表示的。清除術語有助於正確解釋命令輸出，區分共用樹和源樹行為，並確定每個組播元件在端到端轉發過程中的作用。

字詞	定義
組播組地址	224.0.0.0/4範圍內用於標識組播組的IPv4目標地址。
來源位址	將流量傳輸到組播組的傳送方的單播IP地址。
mroute	定義組或源組組合的組播流量處理方式的組播路由條目。
IIF	傳入介面。組播流量預期到達的介面。
OIF	傳出介面。用於向接收方或下游鄰居轉發組播流量的介面。
石油	傳出介面清單。與組播路由條目關聯的傳出介面集。
RPF	反向路徑轉發。檢驗組播流量是否基於通向源或RP的單點傳播路由到達正確的介面。
MDT	組播分發樹。將組播流量從源傳送到所有接收者的邏輯樹。
RPT	RP樹，也稱為共用樹。它將接收器連線到RP並以(*,G)表示。
SPT	最短路徑樹，也稱為源樹。它直接將接收器連線到源，並由(S, G)表示。
FHR	第一跳路由器。組播路由器直接連線到源，負責向RP註冊源。
LHR	最後一跳路由器。組播路由器直接連線到接收者，負責通過IGMP學習接收者興趣後建立組播狀態。
RP	會合點。最初在ASM和PIM稀疏模式中用於連線源和接收器的邏輯會議點。
ASM	任意來源多點傳送。一種組播模型，其中接收機加入一個組時不需要預先指定源。

瞭解眾所周知的保留組播地址非常重要，因為組播故障排除取決於快速確定哪個控制協定正在使用給定的目標組，以及流量在網路中起什麼作用。這些地址有助於區分正常協定操作與異常行為，並且更容易驗證IGMP、PIM和自動RP交換。具體而言，對於自動RP，要識別的最重要組是224.0.1.39(RP-Announce)和224.0.1.40(RP-Discovery)，因為它們包含允許路由器瞭解動態RP對映的資訊。

組播地址	目的
224.0.0.1	本地子網上的所有主機
224.0.0.2	本地子網上的所有路由器
224.0.0.13	所有PIM路由器
224.0.0.22	IGMPv3消息
224.0.1.39	Cisco RP-Announce消息由自動RP使用
224.0.1.40	自動RP使用的Cisco RP-Discovery消息

PIM稀疏模式下的自動RP操作（控制平面工作流）

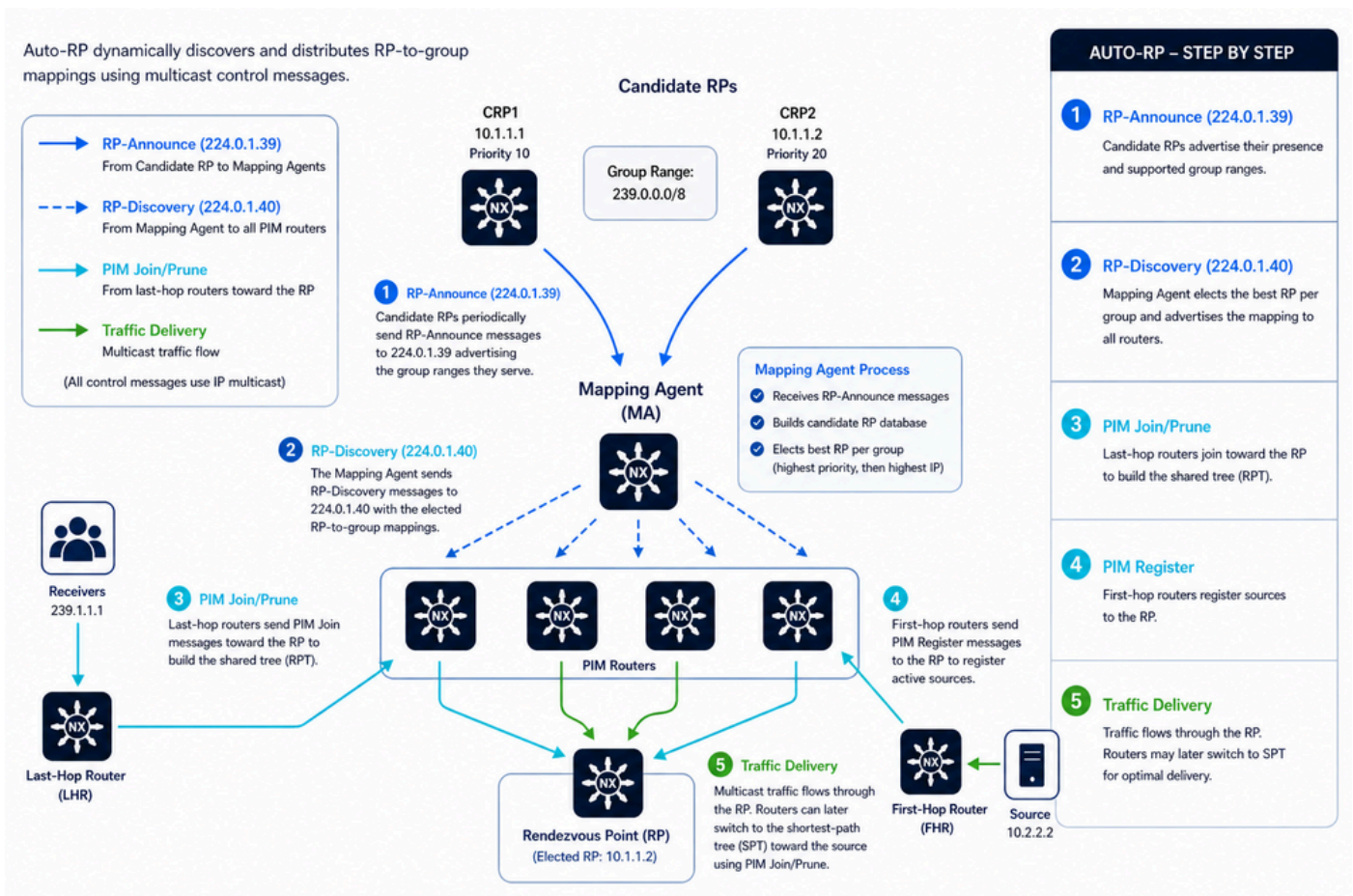
自動RP是一種思科機制，用於協定無關組播稀疏模式，用於在組播域中動態發現和分發集結點 (RP) 資訊。它通過使用基於組播的控制平面消息來通告、選擇和學習RP到組的對映，從而無需進行靜態RP配置。其主要元件是候選RP和對映代理，候選RP為特定組範圍提供RP服務，對映代理收集候選並確定每個組的活動RP。

當每個候選RP定期向224.0.1.39傳送RP-Announce消息 (包括其IP地址、優先順序和支援的組播組範圍) 時，該過程開始。這些消息使用NX-OS中的自動RP偵聽程式在整個網路中泛洪，確保所有對映代理甚至在網路完全以稀疏模式運行之前都接收它們。

對映代理會偵聽這些通告，構建候選RP資料庫，並對每個組執行確定性選擇過程 (通常是最高優先順序，然後是最高IP地址)。選擇最佳的RP後，對映代理生成RP發現消息並將其傳送到224.0.1.40，向域中的所有路由器通告最終的RP到組的對映。

所有PIM路由器都會收到RP-Discovery消息並將對映安裝到其本地RP表中。利用此資訊，最後一跳路由器 (連線到接收器) 向選定RP傳送PIM加入消息以構建共用樹(RPT)，而第一跳路由器 (連線到源) 將組播流量封裝在PIM註冊消息中，以通知RP有關活動源的資訊。

當流量通過RP時，路由器可以選擇使用額外的PIM加入/修整信令直接從共用樹(RPT)切換到最短路徑樹(SPT)。在此過程中，自動RP通過週期性控制消息不斷刷新對映，從而確保了可復原性和自動適應拓撲或RP更改。



自動RP限制和操作限制

- 自動RP僅用於IPv4組播，IPv6(PIM6)不支援該功能，因此IPv6部署需要不同的RP發現機制。
- 自動RP不能與同一PIM域中的BSR共存，因為只能配置一個RP發現機制以避免控制平面衝突。
- 在NX-OS裝置上，預設情況下不會處理或轉發自動RP消息，因此正確操作需要顯式配置。
- 自動RP控制消息的傳播取決於啟用自動RP監聽程式；否則，RP對映資訊不會到達所有路由器。
- RP通告間隔的最小閾值為15秒，這限制了RP更新的通告速度並影響了收斂時間。
- 通過路由對映進行自動RP消息過濾可能會影響操作，因為不正確的策略可能會阻止RP到組的對映。
- 自動RP在VRF上下文中預設禁用，因此必須在多VRF部署中顯式啟用。
- 基於ECMP的多路徑轉發功能預設啟用，允許組播流量使用等價路徑進行負載均衡。
- 支援使用IPsec AH-MD5的PIM鄰居身份驗證。
- PIM監聽不可用。

使用自動RP的優勢

自動RP為PIM稀疏模式組播環境提供動態RP發現和集中RP對映分佈。它無需在每個組播路由器上進行靜態RP配置，降低了操作複雜性，並提高了組播可擴充性。自動RP還支援多個RP候選，從而啟用自動RP故障切換和冗餘。此機制有助於維持一致的組播轉發行為，簡化網路擴展，並允許組播路由器自動獲知域中的RP資訊。

自動RP對映代理和RP選擇過程

自動RP中的選擇過程是確定性的，主要基於IP地址。與其他通訊協定（例如PIMv2 BSR）不同，自動RP不使用可設定的「優先順序」值；相反，它依賴於IP地址層次結構來解決衝突。

對映代理(MA)選擇

在自動RP中，多個對映代理可以在同一個網路中共存以實現冗餘。沒有正式的選舉程式，一個選舉關閉，另一個選舉開啟；所有裝置在技術上均處於活動狀態。但是，網路中的交換機必須決定信任哪些資訊。

- 選擇標準：最高IP地址。
- 流程：
 1. 所有對映代理將其RP發現消息傳送到組播組224.0.1.40。
 2. 客戶端交換機接收這些消息。
 3. 如果交換機從兩個不同的對映代理收到包含衝突資訊的發現消息，則交換機將接受來自源IP地址最高的對映代理的資訊。
 4. Nexus行為：IP地址較低的對映代理通常會在檢測到另一個具有較高IP地址的MA時進入「被動」或「抑制」狀態，以防止網路中的重複控制流量。

集結點(RP)選擇

此過程由對映代理在偵聽來自候選的所有RP-Announce消息（傳送到組224.0.1.39）後執行。

當對映代理收到同一組播組的多個候選時，它按順序應用以下規則：

規則A:最長字首匹配（最具體掩碼）

如果候選通告重疊範圍，則MA會將組分配給通告最小範圍（最長子網掩碼）的RP。

- 範例：
 - 候選者A宣佈整個組播範圍(224.0.0.0/4 (the))。
 - 候選人B宣佈224.10.20.0/24。
 - Result:對於組224.10.20.5，候選B獲勝，因為其範圍更加具體。

規則B:最高IP地址（分路器）

如果兩個或多個候選者通告的組範圍完全相同，則對映代理只能選擇一個。

- 條件：具有最高IP地址的候選者將獲勝。
- 範例：
 - 候選人1:10.2.0.1用於範圍224.10.20.0/24。
 - 候選人2:10.2.0.4用於範圍224.10.20.0/24。
 - Result:對映代理選擇10.2.0.4作為該範圍的官方RP，並在其Discovery消息中通告該範圍。

第2層注意事項：多點傳送MAC和窺探

雖然本文的重點是使用PIM的第3層組播，但第2層行為在故障排除和總體設計中起著至關重要的作用。在第2層，裝置使用MAC地址（分配給網路介面的48位識別符號）進行通訊，組播流量需要特

定的MAC編址方案來將其與單播和廣播流量區分開來。

IPv4組播MAC地址是使用保留字首「01:00:5E」從第3層組播組地址派生的。但是，IP組播地址中只有23位對映到MAC地址，這將產生32:1的重疊，這意味著多達32個不同的組播IP組可以對映到同一個MAC地址。因此，偵聽給定組播MAC地址的主機可以接收多個組播組的流量，即使它們只對其中一個組感興趣。例如，224.1.1.1、225.1.1.1、226.1.1.1、227.1.1.1、228.1.1.1等。

這種重疊對網路效率和故障排除有直接影響。由於完全基於MAC地址的第2層轉發決策無法區分重疊的組播組，因此交換機可以向主機傳輸不必要的流量。然後，這些主機必須依靠高層過濾(IP/IGMP)來丟棄不需要的資料包，從而消耗CPU和緩衝區資源。

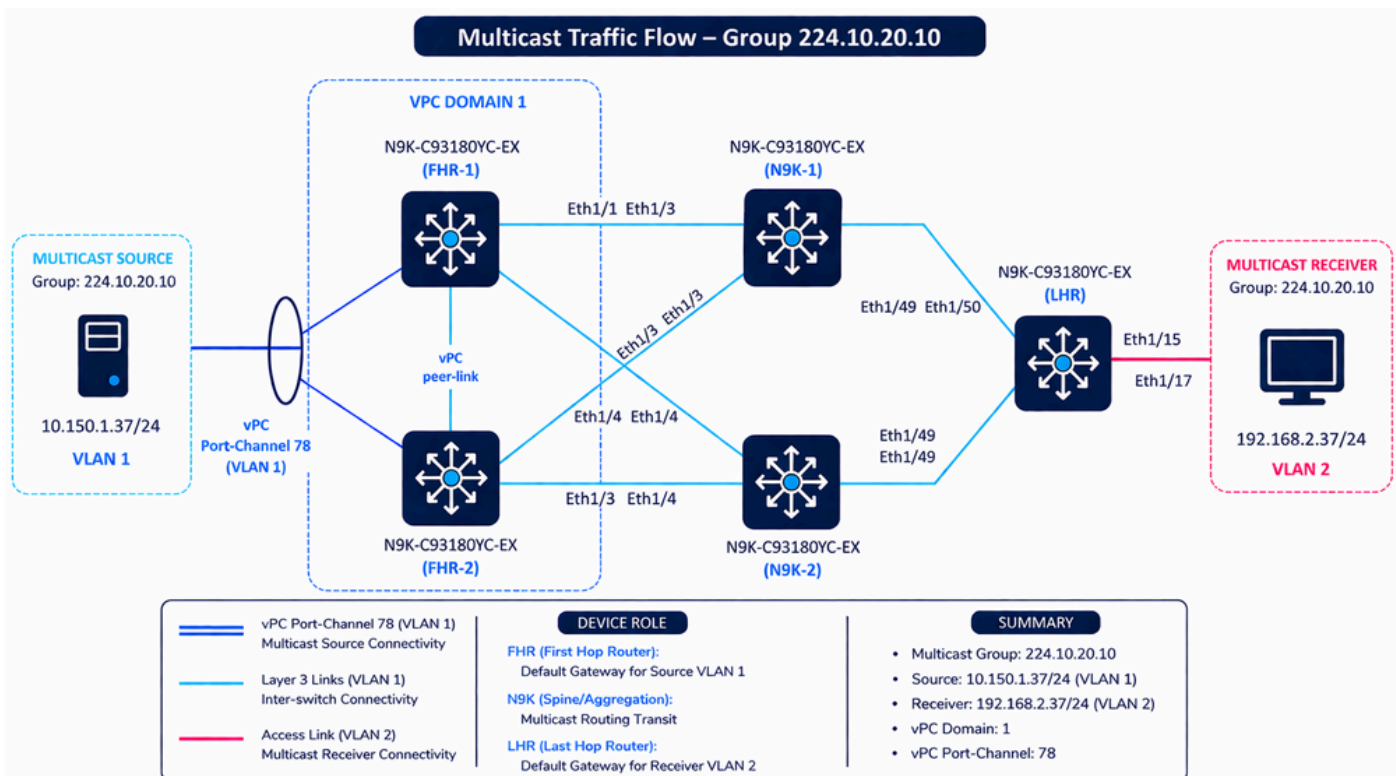
在Cisco Nexus NX-OS中，此限制通過IGMP監聽行為得以緩解。預設情況下，IGMP監聽執行基於IP的查詢，而不是僅支援MAC的轉發，即使多個組播組共用同一個MAC地址，交換機也能做出更精確的轉發決策。這顯著提高了第2層的效率並減少了不必要的流量傳輸。

自動RP配置

本部分提供自動RP配置的詳細說明，並且使用簡單的實現作為參考。在此設定中，組播源通過vPC連線到兩個Nexus交換機，以向接收器傳輸流量。在本設計中，N9K-1和N9K-2同時充當RP候選和對映代理。



注意：vPC埠通道不支援PIM鄰居。



組播流量

第一跳路由器

相同的配置有FHR-2。

```
FHR-1# show running-config pim
feature pim
ip pim auto-rp forward listen
```

```
interface Vlan1
 ip pim sparse-mode
```

```
interface Ethernet1/1
 ip pim sparse-mode
```

```
interface Ethernet1/3
 ip pim sparse-mode
```

指令	目的/說明
功能pim	在交換機上啟用PIM (協定無關組播) 進程。
ip pim auto-rp forward listen	啟用自動RP偵聽程式。這允許交換機接收和轉發自動RP控制消息 (224.0.1.39和224.0.1.40) ，即使交換機尚不知道RP的身份。
ip pim sparse-mode	在特定介面上啟用PIM稀疏模式。在此模式下，組播流量僅轉發到已通過PIM加入消息明確請求組播流量的網段。

對映代理和RP候選

```
N9K-1# show running-config pim
feature pim

ip pim auto-rp rp-candidate loopback0 group-list 224.10.20.0/24 interval 15
ip pim auto-rp mapping-agent loopback1

interface loopback0
 ip pim sparse-mode

interface loopback1
 ip pim sparse-mode

interface Ethernet1/3
 ip pim sparse-mode

interface Ethernet1/4
 ip pim sparse-mode

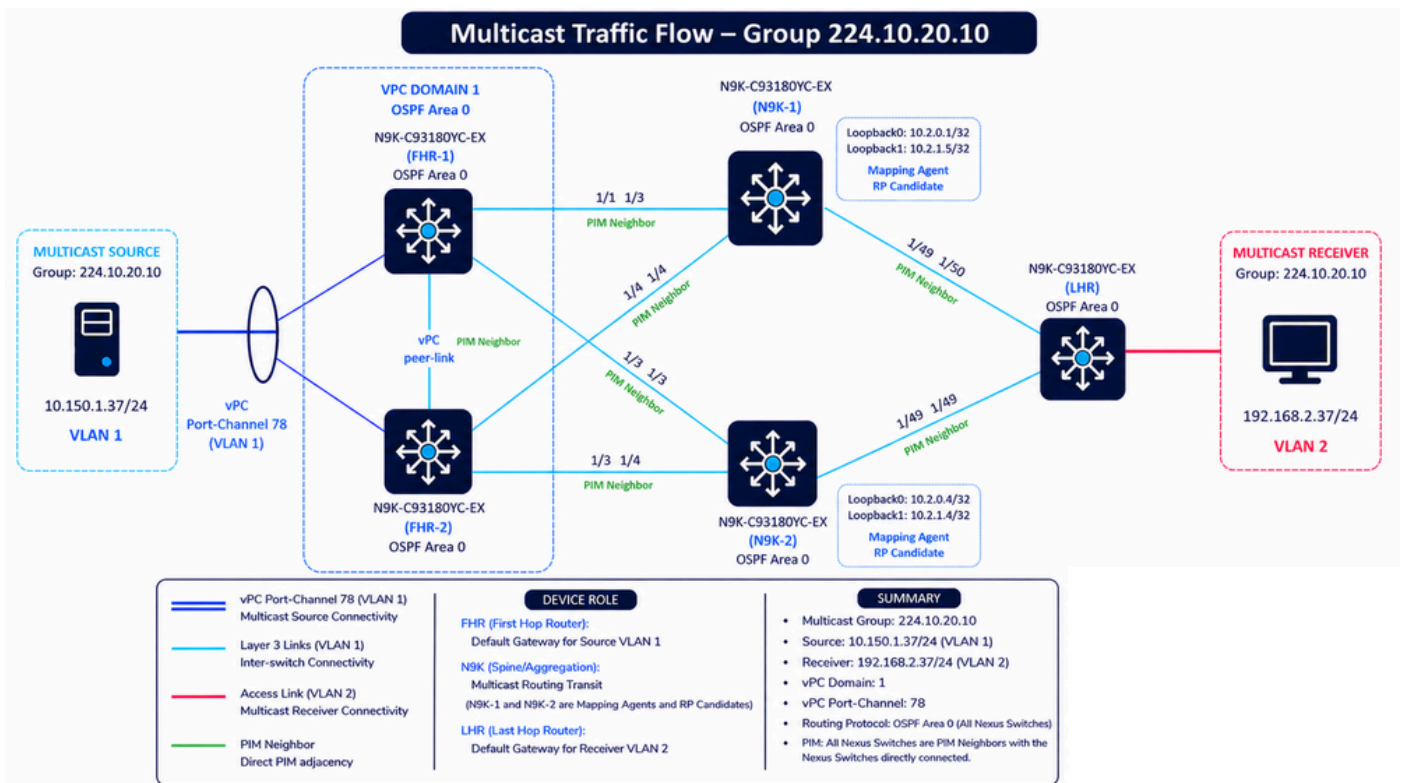
interface Ethernet1/49
 ip pim sparse-mode
```

下表提供了N9K-1的PIM配置的詳細技術細分。此配置在N9K-2上複製。兩台交換機均配置有雙重角色，既充當組播域的RP候選者，又充當對映代理。

指令	詳細技術說明
功能pim	功能啟用：在Nexus交換機上全域性啟用協定無關組播(PIM)引擎。
ip pim auto-rp rp-candidate loopback0 group-list 224.10.20.0/24 interval 15	RP候選角色：將此交換機配置為「志願」作為集結點(RP)。 來源：使用loopback0 IP地址 範圍：它提供處理組播組範圍224.10.20.0/24的功能。 Interval:每15秒向對映代理傳送「通告」消息。保持計時器是此值的三倍。
ip pim auto-rp mapping-agent loopback1	對映代理角色：將交換機配置為自動RP過程的「管理員」。 功能:它偵聽所有RP候選者，解決衝突（使用最高的IP地址作為連線中斷器），並向網路其餘部分廣播「發現」消息，通知他們活動RP是誰。
interface loopback0 / loopback1	邏輯介面：在這些介面上啟用PIM，因為它們作為RP候選和對映代理角色的源IP。它們必須通過單播路由表從所有PIM路由器到達。
interface Ethernet1/3, 1/4, 1/49	物理轉發：在物理埠上啟用PIM稀疏模式。這允許交換機與其他路由器形成PIM鄰居鄰接關係並通過這些特定鏈路轉發組播流量。
ip pim sparse-mode	操作模式：已應用於上面的所有介面。它確保組播流量僅傳送到通過PIM加入消息明確請求組播流量的接收器，從而防止不必要的網路泛洪。

IPv4可達性的路由配置

- 在此拓撲中，所有Nexus交換機都參與一個名為UNDERLAY的OSPF進程，在區域0(0.0.0.0)內運行，以提供整個網路的完整路由可達性。
- 在LHR上，OSPF在用作組播接收器網關的VLAN 2上啟用，在上行鏈路介面Ethernet1/49和Ethernet1/50上啟用，二者均配置為點對點。這樣可確保無需進行DR/BDR選舉，即可有效建立到上游N9K交換機的鄰接關係。
- FHR層 (FHR-1和FHR-2) 共用相同的配置。OSPF在充當組播源網關的VLAN 1上啟用，在通往匯聚層的路由上行鏈路 (Ethernet1/1和Ethernet1/3) 上啟用。這些鏈路還配置為點對點鏈路，從而最佳化收斂和鄰接行為。
- 同樣，N9K-1和N9K-2具有相同的OSPF配置，在OSPF中增加了Loopback0和Loopback1通告
 - 這些環回非常關鍵，因為它們稍後用於PIM RP候選和對映代理角色。



PIM鄰居和OSPF區域0

N9K-1 — OSPF配置

```
N9K-1(config)# show running-config ospf
feature ospf
```

```
router ospf UNDERLAY
router-id 10.2.0.1
```

```
interface loopback0
ip router ospf UNDERLAY area 0.0.0.0
```

```
interface loopback1
  ip router ospf UNDERLAY area 0.0.0.0

interface Ethernet1/3
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0

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

interface Ethernet1/49
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
```

FHR-1 — OSPF配置

```
FHR-1(config)# show running-config ospf
feature ospf
```

```
router ospf UNDERLAY

interface Vlan1
  ip router ospf UNDERLAY area 0.0.0.0

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

interface Ethernet1/3
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
```

LHR - OSPF配置

```
LHR(config)# show running-config ospf
feature ospf
```

```
router ospf UNDERLAY

interface Vlan2
  ip router ospf UNDERLAY area 0.0.0.0

interface Ethernet1/49
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0

interface Ethernet1/50
  ip ospf network point-to-point
  ip router ospf UNDERLAY area 0.0.0.0
```

檢驗運行狀態並排除自動RP故障

步驟 1: 驗證基本IP連線能力 (單點傳送底層驗證)

在分析組播行為之前，必須驗證單播底層 (OSPF區域0) 是否完全可操作。組播控制平面協定 (如 PIM和自動RP) 依賴於單播可達性才能正常工作。

第一個驗證步驟是確認來源和接收者(或其最接近的第3層網關：FHR和LHR)可訪問。

在拓撲中：

- FHR-1/FHR-2最→於組播源(10.150.1.37 - VLAN 1)
- LHR最→於組播接收器(192.168.2.37 - VLAN 2)

驗證方法

1. 執行ICMP可達性測試，測試範圍包括：

- FHR ↔ LHR
- FHR接↔器子網 (VLAN 2網關)
- LHR↔源子網 (VLAN 1網關)

2. 驗證路由表中的源和接收器可達性。對於vPC部署，確保兩個Nexus對等點之間的一致性。請注意，接收方具有ECMP路徑，而來源可透過第2層連線。

FHR-1 — 路由到源

```
FHR-1# show ip route 10.150.1.37
```

```
10.150.1.37/32, ubest/mbest: 1/0, attached  
*via 10.150.1.37, Vlan1, [250/0], 06:57:19, am
```

FHR-1 — 到接收方的路由

```
FHR-1# show ip route 192.168.2.37
```

```
192.168.2.0/24, ubest/mbest: 2/0  
  *via 10.4.0.6, Eth1/3, [110/45], 04:11:08, ospf-UNDERLAY, intra  
  *via 10.4.0.10, Eth1/1, [110/45], 04:11:08, ospf-UNDERLAY, intra
```

FHR-2 — 路由到源

```
FHR-2# show ip route 10.150.1.37
```

```
10.150.1.37/32, ubest/mbest: 1/0, attached  
  *via 10.150.1.37, Vlan1, [250/0], 07:03:45, am
```

FHR-2 — 到接收方的路由

```
FHR-2# show ip route 192.168.2.37
```

```
192.168.2.0/24, ubest/mbest: 2/0  
  *via 10.4.0.13, Eth1/3, [110/45], 04:16:16, ospf-UNDERLAY, intra  
  *via 10.4.0.18, Eth1/4, [110/45], 04:16:16, ospf-UNDERLAY, intra
```

LHR — 路由到源

```
LHR(config)# show ip route 10.150.1.37
```

```
10.150.1.0/24, ubest/mbest: 2/0  
  *via 10.4.0.22, Eth1/49, [110/45], 04:14:52, ospf-UNDERLAY, intra  
  *via 10.4.0.26, Eth1/50, [110/45], 04:14:52, ospf-UNDERLAY, intra
```

LHR — 到接收方的路由

```
LHR(config)# show ip route 192.168.2.37
```

```
192.168.2.37/32, ubest/mbest: 1/0, attached  
  *via 192.168.2.37, Vlan2, [250/0], 06:47:21, am
```

重要注意事項

- 此測試失敗強烈表明存在底層問題。
- 沒有單播可達性，組播無法正常工作，因為：
 - PIM鄰居依賴單播路由
 - RP (環回) 地址必須可訪問
 - RPF (反向路徑轉發) 檢查取決於路由表
- 成功的ping並不能保證組播可以工作，因為：
 - 在阻止組播時可以允許ICMP
 - PIM或自動RP仍可能配置錯誤



提示：在分析自動RP、PIM鄰接關係或RP選擇之前，請始終確保底層路由域是穩定、一致且完全可到達的端到端路由域。

步驟 2: 確定組播角色和端到端拓撲

下一步是明確確定轉發組播流量時涉及的每台裝置的角色。這是強制性的步驟，因為組播故障排除完全取決於瞭解整個拓撲中的流量和預期行為。

至少必須確定以下要素：

- 多點傳送來源：10.150.1.37(VLAN 1)
- 組播組(G):224.10.20.10
- 接收者：192.168.2.37(VLAN 2)
- 第一躍點路由器(FHR):FHR-1/FHR-2 (最靠近源)
- 最後一跳路由器(LHR):LHR (最靠近接收器)

此外，還需要確定控制平面角色：

- RP候選項：N9K-1和N9K-2(Loopback0)
- 對映代理：N9K-1和N9K-2(Loopback1)

拓撲和控制平面感知

要進行組播故障排除，必須提供詳細的網路拓撲。其中包括：

- 物理連線（裝置之間使用的介面）
- 邏輯拓撲（VLAN、路由鏈路、vPC關係）
- 使用的路由協定（本設計中的OSPF區域0）
- 路由域邊界(單IGP與混合協定（如OSPF、EIGRP或BGP）)
- 用於RP和對映代理角色的環回介面
- 啟用PIM的介面和鄰居關係

必須清楚理解的內容

- 從源FHR到RP→LHR→收器→精確→徑
- 哪些裝置負責：
 - 正在傳送PIM暫存器(FHR)
 - 建築(*,G)或(S, G)樹(LHR)
 - 通告RP資訊（對映代理）
- 路由(OSPF)如何確保可達性：
 - 源子網
 - 接收器子網
 - RP環回地址
 - 對映代理環回地址



注意：沒有清晰拓撲的組播故障排除等同於無可見性的調試，這會導致錯誤的假設和誤診。

步驟 3:根據裝置角色驗證自動RP配置

下一步是根據每台裝置在組播拓撲中的角色，驗證是否正確配置了自動RP。這包括確認：

- RP候選(N9K-1 / N9K-2)已正確配置為將其Loopback0通告為組播組範圍的RP。

- 對映代理(N9K-1 / N9K-2)配置為收集RP-Announce消息並使用Loopback1生成RP-Discovery消息。
- FHR和LHR在所有相關介面上啟用了PIM稀疏模式以參與自動RP和接收RP對映。

確保為PIM稀疏模式啟用所有必需介面（包括環回和路由鏈路），並且不存在會阻止RP-Announce(224.0.1.39)和RP-Discovery(224.0.1.40)消息交換的缺少配置，這一點至關重要。



附註：N9K-1和N9K-2配置為組播域內的RP-Candidates和對映代理。



注意：任何缺失或不一致的自動RP配置都會阻止路由器學習RP，從而導致組播流量無法正確轉發。

步驟 4: 驗證所有RP候選程式和所有對映代理的操作

第4.1步檢驗PIM鄰居鄰接關係

第一個驗證步驟是確認所有預期的PIM鄰居都正確建立了組播拓撲。

N9K-1 — 驗證PIM鄰居

```
N9K-1# show ip pim neighbor
```

PIM Neighbor Status for VRF "default"								
Neighbor	Interface	Uptime	Expires	DR	Bidir- Priority	BFD Capable	BFD State	ECMP Redirect Capable
10.4.0.5	Ethernet1/3	23:19:45	00:01:20	1	yes	n/a	no	no
10.4.0.14	Ethernet1/4	23:19:45	00:01:38	1	yes	n/a	no	no
10.4.0.21	Ethernet1/49	23:19:45	00:01:38	1	yes	n/a	no	no

N9K-2 — 驗證PIM鄰居

```
N9K-2# show ip pim neighbor
```

PIM Neighbor Status for VRF "default"								
Neighbor	Interface	Uptime	Expires	DR	Bidir- Priority	BFD Capable	BFD State	ECMP Redirect Capable
10.4.0.9	Ethernet1/3	23:21:18	00:01:29	1	yes	n/a	no	no

10.4.0.17	Ethernet1/4	23:21:18	00:01:23	1	yes	n/a	no
10.4.0.25	Ethernet1/49	23:21:18	00:01:44	1	yes	n/a	no

驗證點

- 檢驗所有期望的鄰居是否出現在每個路由介面上。
- 驗證鄰居正常運行時間穩定且持續增加。
- 驗證到期計時器是否定期刷新。
- 驗證是否未發生意外的鄰接擺動。
- 驗證正確的指定路由器(DR)選擇。
- 驗證組播介面是否成功形成PIM鄰接關係。

在以下拓撲中：

- N9K-1建立三個PIM鄰居。
- N9K-2也建立三個PIM鄰居。
- 所有鄰接關係保持穩定超過23小時，這表示組播控制平面運行穩定。

第4.2步檢驗啟用PIM的介面

下一步是確認所有參與自動RP的介面均已啟用PIM。

此驗證對以下各項特別重要：

- RP-Candidate環回
- 對映代理環回

N9K-1 — 驗證PIM介面

```
N9K-1# show ip pim interface brief
```

```
PIM Interface Status for VRF "default"
```

Interface	IP Address	PIM DR Address	Neighbor Count	Border Interface
Ethernet1/3	10.4.0.6	10.4.0.6	1	no
Ethernet1/4	10.4.0.13	10.4.0.14	1	no
Ethernet1/49	10.4.0.22	10.4.0.22	1	no
loopback0	10.2.0.1	10.2.0.1	0	no
loopback1	10.2.1.5	10.2.1.5	0	no

N9K-2 — 驗證PIM介面

```
N9K-2# show ip pim interface brief
```

```
PIM Interface Status for VRF "default"
```

Interface	IP Address	PIM DR Address	Neighbor Count	Border Interface
Ethernet1/3	10.4.0.10	10.4.0.10	1	no
Ethernet1/4	10.4.0.18	10.4.0.18	1	no
Ethernet1/49	10.4.0.26	10.4.0.26	1	no
loopback0	10.2.0.4	10.2.0.4	0	no
loopback1	10.2.1.4	10.2.1.4	0	no

驗證點：

- 檢驗輸出中是否顯示所有路由的組播介面。
- 驗證RP-Candidate環回是否啟用了PIM。
- 驗證對映代理環回是否已啟用PIM。
- 驗證傳輸介面上的鄰居計數是否正確。
- 檢驗環回介面是否正確顯示Neighbor Count = 0。

環回角色分配：

裝置	功能	回送
N9K-1	RP候選	10.2.0.1
N9K-1	對映代理	10.2.1.5
N9K-2	RP候選	10.2.0.4
N9K-2	對映代理	10.2.1.4

環回正確地顯示為活動PIM介面，即使它們不構成PIM鄰居。之所以會出現此行為，是因為環回介面不直接建立組播鄰接關係。

這些環回的存在證實：

- PIM可以正確生成自動RP控制消息。
- 可以成功生成RP通告。
- 對映代理功能可以正常運行。

步驟4.3分析show ip pim rp

此命令可驗證：

- RP發現
- 自動RP通告
- 對映代理操作
- 組到RP的對映

N9K-1 — RP資訊

```
N9K-1# show ip pim rp
```

```
PIM RP Status Information for VRF "default"
```

```
BSR disabled
```

```
Auto-RP RPA: 10.2.1.5*, next Discovery message in: 00:00:39
```

```
BSR RP Candidate policy: None
```

```
BSR RP policy: None
```

```
Auto-RP Announce policy: None
```

```
Auto-RP Discovery policy: None
```

```
RP: 10.2.0.1*, (0),  
  uptime: 22:18:44  priority: 255,  
  RP-source: 10.2.0.1 (A),  
  group ranges:  
  224.10.20.0/24   , expires: 00:00:37 (A)
```

```
RP: 10.2.0.4, (0),  
  uptime: 23:00:32  priority: 255,  
  RP-source: 10.2.0.4 (A),  
  group ranges:  
  224.10.20.0/24   , expires: 00:00:44 (A)
```

逐行說明

BSR已禁用

```
BSR disabled
```

這確認：

- 未使用啟動路由器(BSR)。
- 組播域以獨佔方式依賴自動RP。

此拓撲中應出現此行為。

自動RP RPA:10.2.1.5*

Auto-RP RPA: 10.2.1.5*

- 10.2.1.5對應於N9K-1上的loopback1。
- *表示本地交換機本身是活動對映代理。

這意味著：

- N9K-1生成RP-Discovery消息。
- N9K-1從RP-Candidates收集RP通告。
- N9K-1將RP對映資訊分發到組播域。

中的下一條發現消息

next Discovery message in: 00:00:39

- 驗證計時器是否持續刷新。
- 驗證發現消息是否定期傳送。

如果此計時器凍結或意外過期，則自動RP通告無法正確傳播。

策略欄位

BSR RP Candidate policy: None
BSR RP policy: None
Auto-RP Announce policy: None
Auto-RP Discovery policy: None

- 未應用任何過濾策略。
- 接受所有RP公告和發現通告。

第一個RP條目

RP: 10.2.0.1*, (0),

- RP地址= 10.2.0.1
- *表示此RP是N9K-1的本地路由器。
- 10.2.0.1對應於N9K-1上的loopback0。

這確認N9K-1已配置為RP-Candidate。

正常運行時間和優先順序

uptime: 22:18:44 priority: 255,

- 穩定的正常運行時間表示穩定的RP通告操作。
- 優先順序255是預設的最高優先順序。

RP源

RP-source: 10.2.0.1 (A),

- RP通告直接源自RP本身。
- (A)表示自動RP獲取的資訊。

組範圍

224.10.20.0/24

- 驗證是否通告了正確的組播範圍。
- 驗證組範圍是否與配置匹配。

第二個RP條目

RP: 10.2.0.4, (0),

- 拓撲中存在另一個RP。
- 10.2.0.4對應於N9K-2上的loopback0。
- 否*顯示，因為此RP遠離N9K-1。

N9K-2 — RP資訊

```
N9K-2# show ip pim rp
```

```
PIM RP Status Information for VRF "default"  
BSR disabled  
Auto-RP RPA: 10.2.1.5, uptime: 22:19:10, expires: 00:02:28
```

```
RP: 10.2.0.4*, (0),  
  uptime: 23:14:14  priority: 255,  
  RP-source: 10.2.1.5 (A),  
  group ranges:  
  224.10.20.0/24   , expires: 00:02:28 (A)
```

N9K-2的主要區別

```
Auto-RP RPA: 10.2.1.5
```

- 否*顯示，因為N9K-2不是對映代理。
- N9K-2從N9K-1遠端學習對映代理資訊。

重要RP差異

```
RP-source: 10.2.1.5 (A),
```

- N9K-2從對映代理獲取RP資訊。
- 對映代理是10.2.1.5。

這確認：

- 自動RP發現消息正常運行。
- 對映代理通告通過組播域成功傳播。

第4.4步驗證自動RP選舉流程和RP選擇

Auto-RP使用兩種不同的組播控制平面功能：

- RP候選
- 對映代理

在PIM稀疏模式環境中驗證組播操作時，瞭解這些函式如何互動至關重要。

在以下拓撲中：

- N9K-1和N9K-2作為RP候選者運行。
- N9K-1作為活動對映代理運行。

RP候選操作

RP-Candidate將自身通告為一個或多個組播組範圍的有效Rendezvous Point。

每個RP候選定期將自動RP通告消息傳送到：

- 224.0.1.39 — 自動RP通告組

這些公告包括：

- RP地址
- 組範圍
- 優先順序機制
- 保持時間

在以下拓撲中：

- N9K-1上的10.2.0.1將自身通告為RP。
- N9K-2上的10.2.0.4將自身通告為RP。

N9K-1 — RP候選資訊

```
N9K-1# show ip pim rp
<snip>
RP: 10.2.0.1*, (0),
  uptime: 23:11:22 priority: 255,
  RP-source: 10.2.0.1 (A),
  group ranges:
  224.10.20.0/24 , expires: 00:00:38 (A)

RP: 10.2.0.4, (0),
  uptime: 23:53:09 priority: 255,
  RP-source: 10.2.0.4 (A),
  group ranges:
  224.10.20.0/24 , expires: 00:00:43 (A)
```

N9K-2 — RP候選資訊

```
N9K-2# show ip pim rp
<snip>
RP: 10.2.0.4*, (0),
  uptime: 1d00h priority: 255,
  RP-source: 10.2.1.5 (A),
  group ranges:
  224.10.20.0/24 , expires: 00:02:52 (A)
```

兩台裝置通告相同的組播組範圍：

- 224.10.20.0/24

兩個RP候選還使用：

- 優先順序 255

這一點非常重要，因為自動RP在RP選擇期間使用優先順序和RP地址。

主動對映代理標識

對映代理從以下邏輯開始為組播組選擇活動RP:

1. 最高RP優先順序獲勝。
2. 如果優先順序相等，則最高RP IP地址將入選。

在以下拓撲中：

- N9K-1 RP地址= 10.2.0.1
- N9K-2 RP地址= 10.2.0.4
- 兩個RP候選都使用優先順序255

因為兩個RP候選具有相同的優先順序：

- 最高RP IP地址成為所選RP。

因此：

- 10.2.0.4成為224.10.20.0/24的活動RP。

選定的RP驗證

N9K-2 — 選定的RP

```
N9K-2# show ip pim rp
<snip>
RP: 10.2.0.4*, (0),
  uptime: 23:14:14  priority: 255,
  RP-source: 10.2.1.5 (A),
  group ranges:
  224.10.20.0/24
```

為什麼N9K-1仍顯示兩個RP條目

在N9K-1上：

```
N9K-1# show ip pim rp
<snip>
RP: 10.2.0.1*, (0),
RP: 10.2.0.4, (0),
```

出現此行為是意料之中的，因為：

- N9K-1作為對映代理運行。
- 對映代理維護所有RP候選的可見性。
- 在分發最終的RP選擇之前，輸出會顯示所有學習的RP通告。



注意：在對映代理上，必須顯示同一組播域內的所有RP候選項。如果缺少任何RP-Candidate，請通過向來自對映代理IP地址（通常是環回介面）的RP-Candidate IP地址傳送ping來驗證可達性。

步驟 5: 驗證與RP候選和對映代理的可達性

參與PIM稀疏模式域的所有組播路由器必須保持穩定的IP可達性，以便：

- 所有RP候選項
- 所有對映代理

此驗證至關重要，因為PIM稀疏模式依賴於單播路由，以便：

- 到達集結點(RP)
- 建立共用多點傳送樹(RPT)
- 從第一跳路由器(FHR)傳送PIM註冊消息
- 接收自動RP發現通告
- 執行反向路徑轉發(RPF)驗證

如果到RP或對映代理的可達性失敗：

- 組播源可能無法成功註冊。
 - 接收者可能無法加入組播組。
 - 自動RP通告無法正確傳播。
 - 可能會發生RPF故障。
 - 組播流量轉發可能變得不穩定或間歇性。
-

檢驗穩定的路由條目

路由表必須包含通向以下目標的穩定單播路由：

- 所有RP環回
- 所有對映代理環回
- 所有組播傳輸介面

路由必須保持連續安裝狀態，並且不會出現路由擺動或過度重新收斂事件。

驗證：

- 正確選擇下一跳
 - 正確的傳出介面
 - 穩定的路由正常運行時間
 - 預期的路由協定源
 - 跨組播域的一致單播可達性
-

FHR-1 — 到RP候選10.2.0.1的路由

```
FHR-1# show ip route 10.2.0.1
```

```
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>
```

```
10.2.0.1/32, ubest/mbest: 1/0  
  *via 10.4.0.6, Eth1/3, [110/5], 1d02h, ospf-UNDERLAY, intra
```

FHR-1 — 到RP候選10.2.0.4的路由

```
FHR-1# show ip route 10.2.0.4
```

```
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>
```

```
10.2.0.4/32, ubest/mbest: 1/0
```

```
  *via 10.4.0.10, Eth1/1, [110/5], 1d02h, ospf-UNDERLAY, intra
```

FHR-1 — 到對映代理10.2.1.5的路由

```
FHR-1# show ip route 10.2.1.5
```

```
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>
```

```
10.2.1.5/32, ubest/mbest: 1/0
```

```
  *via 10.4.0.6, Eth1/3, [110/5], 1d02h, ospf-UNDERLAY, intra
```

FHR-1 — 到對映代理10.2.1.4的路由

```
FHR-1# show ip route 10.2.1.4
```

```
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>
```

```
10.2.1.4/32, ubest/mbest: 1/0
```

```
  *via 10.4.0.10, Eth1/1, [110/5], 1d02h, ospf-UNDERLAY, intra
```

FHR-2 — 到RP候選10.2.0.1的路由

```
FHR-2# show ip route 10.2.0.1
```

```
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>

```
10.2.0.1/32, ubest/mbest: 1/0
  *via 10.4.0.13, Eth1/3, [110/5], 1d02h, ospf-UNDERLAY, intra
```

FHR-2 — 到RP候選10.2.0.4的路由

```
FHR-2# show ip route 10.2.0.4
```

```
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>
```

```
10.2.0.4/32, ubest/mbest: 1/0
  *via 10.4.0.18, Eth1/4, [110/5], 1d02h, ospf-UNDERLAY, intra
```

FHR-2 — 到對映代理10.2.1.5的路由

```
FHR-2# show ip route 10.2.1.5
```

```
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>
```

```
10.2.1.5/32, ubest/mbest: 1/0
  *via 10.4.0.13, Eth1/3, [110/5], 1d02h, ospf-UNDERLAY, intra
```

FHR-2 — 到對映代理10.2.1.4的路由

```
FHR-2# show ip route 10.2.1.4
```

```
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>
```

```
10.2.1.4/32, ubest/mbest: 1/0
  *via 10.4.0.18, Eth1/4, [110/5], 1d02h, ospf-UNDERLAY, intra
```

LHR — 到RP候選10.2.0.1的路由

```
LHR# show ip route 10.2.0.1
```

```
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>
```

```
10.2.0.1/32, ubest/mbest: 1/0
```

```
  *via 10.4.0.22, Eth1/49, [110/2], 1d02h, ospf-UNDERLAY, intra
```

LHR — 到RP候選10.2.0.4的路由

```
LHR# show ip route 10.2.0.4
```

```
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>
```

```
10.2.0.4/32, ubest/mbest: 1/0
```

```
  *via 10.4.0.26, Eth1/50, [110/2], 1d02h, ospf-UNDERLAY, intra
```

LHR — 到對映代理10.2.1.5的路由

```
LHR# show ip route 10.2.1.5
```

```
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>
```

```
10.2.1.5/32, ubest/mbest: 1/0
```

```
  *via 10.4.0.22, Eth1/49, [110/2], 1d02h, ospf-UNDERLAY, intra
```

LHR — 到對映代理10.2.1.4的路由

```
LHR# show ip route 10.2.1.4
```

```
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>

10.2.1.4/32, ubest/mbest: 1/0
*via 10.4.0.26, Eth1/50, [110/2], 1d02h, ospf-UNDERLAY, intra

驗證從PIM介面的Ping可達性

驗證路由表後，驗證端到端IP可達性，以便：

- 所有RP候選項
- 所有對映代理

ping的來源必須是：

- 啟用PIM的介面
- 參與組播路由的環回介面

此驗證非常重要，因為組播路由器會在以下過程中使用這些源地址：

- PIM鄰居建立
- RPF計算
- PIM暫存器封裝
- 自動RP通訊



提示：如果使用未編號的介面（其中多個第3層介面共用來自環回介面的同一IP地址），可達性驗證會變得更簡單，因為單個源IP地址可以一致使用。

Ping驗證摘要

裝置	來源 IP	目的地	功能	結果
FHR-1	10.4.0.5	10.2.0.1	RP候選	成功
FHR-1	10.4.0.5	10.2.0.4	RP候選	成功
FHR-1	10.4.0.5	10.2.1.5	對映代理	成功
FHR-1	10.4.0.5	10.2.1.4	對映代理	成功
FHR-2	10.4.0.9	10.2.0.1	RP候選	成功
FHR-2	10.4.0.9	10.2.0.4	RP候選	成功
FHR-2	10.4.0.9	10.2.1.5	對映代理	成功
FHR-2	10.4.0.9	10.2.1.4	對映代理	成功

LHR	10.4.0.5	10.2.0.1	RP候選	成功
LHR	10.4.0.5	10.2.0.4	RP候選	成功
LHR	10.4.0.5	10.2.1.5	對映代理	成功
LHR	10.4.0.5	10.2.1.4	對映代理	成功

檢驗FHR和LHR上的運行狀態和組播流量轉發

下一個驗證步驟是驗證：

- FHR和LHR成功獲取了預期的RP。
- 組播路由表包含預期組播狀態。
- 組播控制平面在組播流量傳輸開始之前運行正常。
- PIM稀疏模式構建預期的共用樹狀態。

第1步檢驗FHR和LHR上的RP學習

在驗證組播轉發狀態之前，檢驗所有組播路由器是否都獲取了正在驗證的組播組的預期RP。

此步驟至關重要，因為：

- FHR必須知道傳送PIM註冊消息的位置。
- LHR必須知道構建共用樹時要使用的RP。
- 組播域必須維護一致的RP對映。

在以下拓撲中：

- 選定的RP是10.2.0.4。
- 對映代理是10.2.1.5。
- 組播組範圍是224.10.20.0/24。

FHR-1 — 驗證獲知的RP

```
FHR-1# show ip pim rp
```

```
PIM RP Status Information for VRF "default"  
BSR disabled  
Auto-RP RPA: 10.2.1.5, uptime: 1d02h, expires: 00:02:30  
BSR RP Candidate policy: None  
BSR RP policy: None  
Auto-RP Announce policy: None
```

Auto-RP Discovery policy: None

RP: 10.2.0.4, (0),
uptime: 1d03h priority: 255,
RP-source: 10.2.1.5 (A),
group ranges:
224.10.20.0/24 , expires: 00:02:30 (A)

FHR-2 — 驗證獲知的RP

FHR-2# show ip pim rp

PIM RP Status Information for VRF "default"
BSR disabled
Auto-RP RPA: 10.2.1.5, uptime: 1d02h, expires: 00:02:15
BSR RP Candidate policy: None
BSR RP policy: None
Auto-RP Announce policy: None
Auto-RP Discovery policy: None

RP: 10.2.0.4, (0),
uptime: 1d03h priority: 255,
RP-source: 10.2.1.5 (A),
group ranges:
224.10.20.0/24 , expires: 00:02:15 (A)

LHR — 驗證獲知的RP

LHR# show ip pim rp

PIM RP Status Information for VRF "default"
BSR disabled
Auto-RP RPA: 10.2.1.5, uptime: 1d02h, expires: 00:02:07
BSR RP Candidate policy: None
BSR RP policy: None
Auto-RP Announce policy: None
Auto-RP Discovery policy: None

RP: 10.2.0.4, (0),
uptime: 1d03h priority: 255,
RP-source: 10.2.1.5 (A),
group ranges:
224.10.20.0/24 , expires: 00:02:07 (A)

輸出確認：

- 所有組播路由器都一致學習相同的RP。
- 選定的RP是10.2.0.4。
- RP對映通過自動RP獲取。
- 分發RP資訊的對映代理是10.2.1.5。
- RP對映在一天以上的時間內保持穩定。
- 發現過期計時器已正確刷新。

出現此行為是意料之中的，因為：

- 在自動RP選擇過程中，10.2.0.4被選為活動RP。
- 10.2.1.5作為對映代理運行。
- 所有組播路由器都成功接收自動RP發現通告。

在這個階段，組播控制平面運行正常，所有路由器都具有一致的224.10.20.0/24 RP對映

第2步在活動組播流量之前檢驗組播路由狀態

下一步是在組播流量傳輸開始之前驗證組播路由表。

在此情況中：

- 組播接收器已加入組播組。
- 活動組播源流量尚未流動。

此狀態很重要，因為它驗證：

- IGMP成員身份操作
- PIM共用樹建立
- 初始(*,G)組播狀態
- 接收者興趣傳播

FHR-1 — 組播路由表

```
FHR-1# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(* , 232.0.0.0/8), uptime: 23:07:34, pim ip
```

Incoming interface: Null, RPF nbr: 0.0.0.0
Outgoing interface list: (count: 0)

FHR-2 — 組播路由表

```
FHR-2# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(* , 232.0.0.0/8), uptime: 23:07:37, pim ip  
  Incoming interface: Null, RPF nbr: 0.0.0.0  
  Outgoing interface list: (count: 0)
```

FHR組播狀態分析

FHR尚未包含：

- 224.10.20.10的(*,G)狀態
- 組播源的(S, G)狀態

出現此行為是意料之中的，因為：

- 尚無活動的組播源流量。
- 未生成PIM註冊消息。
- FHR尚未啟動組播轉發。

唯一的組播條目是：

- 232.0.0.0/8

這對應於預設的SSM範圍，並由系統自動安裝。

需要以下值：

- 傳入介面：空
- RPF鄰居：0.0.0.0
- 傳出介面計數：0

此條目不表示活動的組播轉發。

LHR — 組播路由表

```
LHR# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(* , 224.10.20.10/32), uptime: 23:07:39, igmp ip pim  
Incoming interface: Ethernet1/50, RPF nbr: 10.4.0.26  
Outgoing interface list: (count: 1)  
Vlan2, uptime: 23:07:39, igmp
```

```
(* , 232.0.0.0/8), uptime: 23:07:39, pim ip  
Incoming interface: Null, RPF nbr: 0.0.0.0  
Outgoing interface list: (count: 0)
```

LHR組播狀態分析

與FHR不同，LHR包含一個活動的(*,G)條目，用於：

- 224.10.20.10

出現此行為是意料之中的，因為：

- 組播接收器已加入組。
- 已在Vlan2上獲知IGMP成員資訊。
- LHR啟動了指向RP的共用樹連線。

組播路由表確認：

- 傳入介面(IIF)是Ethernet1/50。
- RPF鄰居是10.4.0.26。
- 傳出介面清單(OIL)包含Vlan2。

這表示：

- LHR正確建立了指向RP的共用樹。
- 多播接收者興趣已成功向上游傳播。
- 已正確傳輸PIM加入消息。

在此階段：

- 尚無活動的組播源流量。
 - 尚不存在(S, G)狀態。
 - 僅存在共用樹狀態。
-

N9K-1 — 對映代理組播路由表

```
N9K-1# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(* , 232.0.0.0/8), uptime: 1d03h, pim ip  
  Incoming interface: Null, RPF nbr: 0.0.0.0  
  Outgoing interface list: (count: 0)
```

對映代理組播狀態分析

N9K-1僅作為對映代理運行，不參與224.10.20.10的組播轉發

因此，預期會缺少(*, G)entrie和(S, G)條目。

對映代理僅分發RP對映資訊，並不一定參與組播資料轉發，除非組播流量直接流經裝置。

N9K-2 — RP組播路由表

```
N9K-2# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(* , 224.10.20.10/32), uptime: 1d01h, pim ip  
  Incoming interface: loopback0, RPF nbr: 10.2.0.4  
  Outgoing interface list: (count: 1)  
    Ethernet1/49, uptime: 1d01h, pim
```

```
(* , 232.0.0.0/8), uptime: 1d03h, pim ip  
  Incoming interface: Null, RPF nbr: 0.0.0.0  
  Outgoing interface list: (count: 0)
```

RP組播狀態分析

N9K-2用作以下對象的活動RP:

- 224.10.20.0/24

因此，RP包含224.10.20.10的共用樹(*,G)狀態

需要以下值：

- 傳入介面= loopback0
- RPF鄰居= 10.2.0.4
- 傳出介面= Ethernet1/49

這表示：

- RP已成功安裝共用樹狀態。
- RP從下游路由器收到PIM加入消息。
- 正確建立了面向接收者的共用組播樹。

在此階段：

- RP僅包含共用樹狀態。
- 尚無活動的組播源流量。
- 尚不存在(S，G)條目。

步驟3驗證具有主動多點傳播流量的多點傳播路由狀態

組播流量傳輸開始後，組播路由表將從共用樹狀態轉換到活動源特定轉發狀態。

在此情況中：

- 組播源是10.150.1.37。
- 組播組為224.10.20.10。
- N9K-2上的活動RP為10.2.0.4。
- 組播源通過vPC埠通道78連線。

vPC組播注意事項

組播源通過FHR-1和FHR-2形成的vPC域進行連線。

由於源通過vPC成員port-channel進行連線：

- 流量雜湊可以將組播資料包轉發到任一Nexus交換機。
- 兩個FHR需要相同的組播配置。
- 兩個FHR都需要一致的單播路由資訊。
- 兩個FHR必須維護相同的RP對映。
- 兩個FHR必須保持穩定的PIM鄰接關係。

在此特定情況中：

- 組播流向FHR-2雜湊。
- FHR-2成為組播源的活動轉發FHR。

FHR-1 — 組播路由表

```
FHR-1# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(10.150.1.37/32, 224.10.20.10/32), uptime: 00:03:58, ip pim  
  Incoming interface: Vlan1, RPF nbr: 10.150.1.37  
  Outgoing interface list: (count: 0)
```

```
(*, 232.0.0.0/8), uptime: 1d00h, pim ip  
  Incoming interface: Null, RPF nbr: 0.0.0.0  
  Outgoing interface list: (count: 0)
```

FHR-1 — vPC角色

```
FHR-1# show vpc role
```

```
vPC Role status
```

```
-----  
vPC role : primary <<<  
Dual Active Detection Status : 0  
vPC system-mac : 00:23:04:ee:be:01  
vPC system-priority : 32667  
vPC local system-mac : 00:6b:f1:84:02:97  
vPC local role-priority : 32667  
vPC local config role-priority : 32667  
vPC peer system-mac : 6c:b2:ae:ee:5a:97  
vPC peer role-priority : 32667  
vPC peer config role-priority : 32667
```

FHR-1組播狀態分析

FHR-1包含一個活動的(S, G)條目, 用於:

- 來源= 10.150.1.37
- 組= 224.10.20.10

組播路由條目確認:

- 來源是在Vlan1上得知的。
- RPF鄰居是直接連線的組播源。
- OIL中不存在任何傳出介面。

之所以會出現此行為, 是因為組播流未針對出站轉發向FHR-1雜湊。

因此:

- FHR-1僅安裝本地源狀態。
- FHR-1不會通過PIM向上游轉發組播流量。

FHR-2 — 組播路由表

```
FHR-2# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(10.150.1.37/32, 224.10.20.10/32), uptime: 00:16:35, ip pim
  Incoming interface: Vlan1, RPF nbr: 10.150.1.37
  Outgoing interface list: (count: 1)
    Ethernet1/3, uptime: 00:16:35, pim
```

```
(*, 232.0.0.0/8), uptime: 1d00h, pim ip
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)
```

FHR-2 — vPC角色

```
FHR-2# show 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    : 6c:b2:ae:ee:5a:97
```

```
vPC local role-priority      : 32667
vPC local config role-priority : 32667
vPC peer system-mac         : 00:6b:f1:84:02:97
vPC peer role-priority      : 32667
vPC peer config role-priority : 32667
```

FHR-2組播狀態分析

與FHR-1不同，FHR-2包含：

- 活動的傳出介面清單
- 活動PIM轉發路徑

這表示：

- FHR-2成為組播流的操作轉發FHR。
- 組播資料包通過vPC成員port-channel雜湊到FHR-2。
- FHR-2將組播流量封裝到PIM註冊消息中。
- FHR-2將組播流量向上游轉發到RP。

ECMP和組播轉發行為

傳出介面Ethernet1/3匹配通向接收方192.168.2.37的單播路由表

FHR-2 — 到組播接收器的路由

```
FHR-2# show ip route 192.168.2.37
```

```
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>
```

```
192.168.2.0/24, ubest/mbest: 2/0
```

```
*via 10.4.0.13, Eth1/3, [110/45], 1d02h, ospf-UNDERLAY, intra
```

```
*via 10.4.0.18, Eth1/4, [110/45], 1d02h, ospf-UNDERLAY, intra
```

FHR-2包含通向組播接收器子網的兩條等價路由：

- 10.4.0.13至Ethernet1/3
- 10.4.0.18至Ethernet1/4

這確認：

- ECMP路由在組播傳輸拓撲中處於活動狀態。
- 存在通向接收方網路的多個有效單播轉發路徑。
- PIM可以使用MRIB選擇的RPF路徑來做出組播轉發決策。
- 不會在並行ECMP路徑上發生資料包重複。

雖然存在兩條等價路由，但組播轉發對每個組播流使用單個RPF路徑。

在此拓撲中，組播流使用：

- Ethernet1/3到10.4.0.13

此行為與之前觀察到的組播路由表相匹配：

```
(10.150.1.37/32, 224.10.20.10/32)
  Outgoing interface list:
    Ethernet1/3
```

vPC運行的主要和次要行為

vPC操作角色對組播轉發行為的影響不同：

- PIM流量
- IGMP處理

在以下拓撲中：

- FHR-1是運行vPC主路由器。
- FHR-2是運行vPC輔助裝置。

兩台Nexus交換機都可以：

- 處理PIM控制平面流量
- 建立組播路由狀態
- 通過PIM轉發組播流量

但是：

- 只有可操作的vPC主裝置會處理流向接入層的IGMP接收器流量。

這一區別很重要，因為：

- 向路由網路的PIM轉發在兩個對等體上保持活動狀態。
- 向第2層接收器網段的IGMP轉發仍然集中在運行主節點上。

因此：

- 從源到路由組播域的組播流量可通過任一vPC對等體退出。
- 向IGMP接收器的組播流量取決於運行的主要行為。

LHR — 組播路由表

```
LHR# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(* , 224.10.20.10/32), uptime: 1d00h, igmp ip pim
  Incoming interface: Ethernet1/50, RPF nbr: 10.4.0.26
  Outgoing interface list: (count: 1)
    Vlan2, uptime: 1d00h, igmp

(10.150.1.37/32, 224.10.20.10/32), uptime: 00:06:31, ip mrib pim
  Incoming interface: Ethernet1/49, RPF nbr: 10.4.0.22
  Outgoing interface list: (count: 1)
    Vlan2, uptime: 00:06:31, mrib

(* , 232.0.0.0/8), uptime: 1d00h, pim ip
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)
```

LHR組播狀態分析

LHR現在包含兩者：

- (*,G)共用樹狀態
- (S, G)源樹狀態

這確認：

- 組播接收器已成功加入。
- 組播源變為活動狀態。
- LHR從共用樹轉發轉換到源樹轉發。

(S , G)條目確認：

- 組播源路徑已成功獲知。
- RPF鄰居是10.4.0.22。
- 多點傳播流量通過Ethernet1/49到達。
- 組播流量轉發至Vlan2接收器。

此行為確認成功：

- PIM轉發
- RPF驗證
- 源樹構造
- 組播流量傳輸

N9K-1 — 中轉組播路由表

```
N9K-1# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(10.150.1.37/32, 224.10.20.10/32), uptime: 00:06:42, pim ip  
  Incoming interface: Ethernet1/4, RPF nbr: 10.4.0.14  
  Outgoing interface list: (count: 1)  
    Ethernet1/49, uptime: 00:06:42, pim
```

```
(* , 232.0.0.0/8), uptime: 1d04h, pim ip  
  Incoming interface: Null, RPF nbr: 0.0.0.0  
  Outgoing interface list: (count: 0)
```

N9K-1傳輸狀態分析

N9K-1充當活動組播流的傳輸組播路由器。

組播路由條目確認：

- 多點傳播流量從Ethernet1/4到達。
- 多點傳播流量轉送到Ethernet1/49。
- 通向LHR的組播轉發路徑可正常運行。

這確認成功：

- PIM鄰居操作
 - RPF驗證
 - 通過傳輸網路進行組播轉發
-

N9K-2 — RP組播路由表

```
N9K-2# show ip mroute
```

```
IP Multicast Routing Table for VRF "default"
```

```
(* , 224.10.20.10/32), uptime: 1d02h, pim ip  
  Incoming interface: loopback0, RPF nbr: 10.2.0.4  
  Outgoing interface list: (count: 1)  
    Ethernet1/49, uptime: 1d02h, pim
```

```
(10.150.1.37/32, 224.10.20.10/32), uptime: 00:06:50, ip pim mrib  
  Incoming interface: Ethernet1/4, RPF nbr: 10.4.0.17, internal  
  Outgoing interface list: (count: 0)
```

```
(* , 232.0.0.0/8), uptime: 1d04h, pim ip  
  Incoming interface: Null, RPF nbr: 0.0.0.0  
  Outgoing interface list: (count: 0)
```

RP組播狀態分析

N9K-2作為組播組的活動RP運行。

RP包含兩者：

- (*,G)共用樹狀態
- (S, G)源樹狀態

預計(S, G)條目中沒有傳出介面，原因是：

- RP已經將接收器切換到最短路徑樹。
- 活動組播轉發不再需要共用樹轉發路徑。

RCA和運行狀況診斷所需的最低組播資訊

命令清單提供在運行NX-OS的Cisco Nexus 9000系列交換機上執行適當的根本原因分析(RCA)或組播運行狀況診斷所需的最低建議組播資料收集。這些輸出捕獲組播控制平面狀態、MRIB程式設計、轉發資訊、vPC運行狀態和硬體轉發詳細資訊。但是，根據故障情況仍可能需要其他資訊。例如，組播資料包丟失、間歇性流量丟棄、資料包複製問題、硬體轉發不一致或亂序組播轉發通常需要使用Ethanalyzer、SPAN或硬體級捕獲在Nexus交換機上直接捕獲資料包。類似地，瞬時RPF不一致、ECMP轉發更改、ASIC程式設計失敗或IGMP抑制事件也可能發生，而不會生成永久日誌。

因此，將show tech輸出與資料包捕獲和轉發驗證結合使用可顯著改善診斷準確性和RCA品質。雖然此資訊為組播故障排除提供了強大的操作基線，但它不能保證始終僅從這些輸出中識別RCA。某些組播故障需要額外的故障排除、即時流量分析、硬體級驗證、拓撲關聯或擴展資料包捕獲來查明確切的根本原因。



提示：在工作和非工作期間收集此資訊可清晰地從Nexus的角度快速瞭解問題的出現方式，並顯著提高了確定根本原因的能力。

最小組播資料收集命令

```
<#root>
```

```
N9K-1#
```

```
show tech-support multicast >> bootflash:$(SWITCHNAME)-sh-tech-multicast.txt
```

```
N9K-1#
```

```
show tech-support details >> bootflash:$(SWITCHNAME)-sh-tech-det.txt
```

```
N9K-1#
```

```
show tech-support vpc >> bootflash:$(SWITCHNAME)-sh-tech-vpc.txt
```

```
N9K-1#
```

```
show tech-support forwarding multicast >> bootflash:$(SWITCHNAME)-sh-tech-fwd-multicast.txt
```

```
N9K-1#
```

```
show tech-support forwarding l3 multicast detail vdc-all >> bootflash:$(SWITCHNAME)-sh-tech-fwd-l3-multi
```

```
N9K-1#
```

```
show tech-support forwarding l3 unicast detail vdc-all >> bootflash:${SWITCHNAME}-sh-tech-fwd-l3-unicast
```

建立用於匯出的TAR存檔

生成show tech檔案後，將它們整合到單個TAR歸檔中，用於匯出和分析。命令是單行。

```
<#root>
```

```
N9K-1#
```

```
tar create bootflash:${SWITCHNAME}-multicast-logs  
bootflash:${SWITCHNAME}-sh-tech-multicast.txt  
bootflash:${SWITCHNAME}-sh-tech-det.txt  
bootflash:${SWITCHNAME}-sh-tech-vpc.txt  
bootflash:${SWITCHNAME}-sh-tech-fwd-multicast.txt  
bootflash:${SWITCHNAME}-sh-tech-fwd-l3-multicast.txt  
bootflash:${SWITCHNAME}-sh-tech-fwd-l3-unicast-det.txt
```

匯出單個TAR歸檔可以簡化：

- TAC案件上傳
- RCA工作流程
- 集中日誌分析
- 多點傳送轉送關聯
- 歷史故障保留

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）。