

排除使用NX-OS的Nexus 9000上的组播自动RP故障

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[组播概述](#)

[PIM稀疏模式下的自动RP操作 \(控制平面工作流程 \)](#)

[自动RP限制和操作限制](#)

[使用自动RP的优势](#)

[自动RP映射代理和RP选择过程](#)

[映射代理\(MA\)选择](#)

[交汇点\(RP\)选择](#)

[第2层注意事项：组播MAC和监听](#)

[自动RP配置](#)

[第一跳路由器](#)

[映射代理和RP候选](#)

[IPv4可达性的路由配置](#)

[检验运行状态并排除自动RP故障](#)

[步骤 1：检验基本IP连通性 \(单播底层验证 \)](#)

[验证方法](#)

[重要注意事项](#)

[步骤 2：确定组播角色和端到端拓扑](#)

[拓扑和控制平面感知](#)

[必须清楚理解的内容](#)

[步骤 3：根据设备角色验证自动RP配置](#)

[步骤 4：验证所有RP候选者和所有映射代理的操作](#)

[第4.1步检验PIM邻居邻接关系](#)

[第4.2步检验启用PIM的接口](#)

[第4.3步分析show ip pim rp](#)

[第4.4步验证自动RP选举过程和RP选择](#)

[步骤 5：验证与RP候选和映射代理的可达性](#)

[检验稳定的路由条目](#)

[检验从PIM接口的Ping可达性](#)

[Ping验证摘要](#)

[检验FHR和LHR上的运行状态和组播流量转发](#)

[第1步检验FHR和LHR上的RP学习](#)

[第2步检验活动组播流量之前的组播路由状态](#)

[第3步使用活动组播流量检验组播路由状态](#)

[vPC运行的主要和次要行为](#)

[LHR组播状态分析](#)

[RP组播状态分析](#)

简介

本文档介绍自动RP如何在带NX-OS的Cisco Nexus 9000上运行，以及如何验证组播操作并排除故障。

先决条件

要求

- IP组播基础知识
- IGMP和PIM稀疏模式的基础知识
- 使用Cisco Nexus 9000和NX-OS CLI的经验
- 了解单播路由和RPF概念

使用的组件

- 思科 Nexus 9000 系列交换机
- Cisco NX-OS 软件
- PIM 稀疏模式
- 静态RP和自动RP

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

组播概述

组播是一种一对多通信模型，其中源向多个感兴趣的接收方发送单个流量流。网络只复制转发路径分支处的流量，而不是为每个目标创建单独的副本。这使得组播比广播或重复的单播传输更有效。在IPv4中，组播流量使用224.0.0.0/4范围内的目标地址。

PIM稀疏模式是运行NX-OS的Cisco Nexus交换机支持的组播路由模型。仅当明确获知了接收方兴趣后，才会转发流量。在任意源组播设计中，接收方最初加入指向交汇点的共享树，并且源注册到该RP。在流量开始流动后，最后一跳路由器可以从共享树移动到通往源的最短路径树。

定义组播中使用的术语非常重要，因为准确的故障排除取决于了解控制平面事件、路由条目和转发决策是如何表示的。清晰的术语有助于正确解释命令输出，区分共享树和源树行为，并确定每个组播组件在端到端转发过程中的作用。

期限	定义
组播组地址	224.0.0.0/4范围内用于标识组播组的IPv4目标地址。
源地址	将流量传输到组播组的发送方的单播IP地址。
mroute	组播路由条目，定义如何处理组或源 — 组组合的组播流量。
IIF	传入接口.组播流量预期到达的接口。
OIF	传出接口.用于向接收方或下游邻居转发组播流量的接口。
石油	传出接口列表。与组播路由条目关联的传出接口集。
RPF	反向路径转发。检验组播流量是否根据指向源或RP的单播路由到达正确的接口。
MDT	组播分布树。将组播流量从源传输到所有接收器的逻辑树。
RPT	RP树，也称为共享树。它将接收器连接到RP，并以(*,G)表示。
SPT	最短路径树，也称为源树。它将接收器直接连接到源，并以(S, G)表示。
FHR	第一跳路由器。组播路由器直接连接到源，负责向RP注册源。
LHR	最后一跳路由器。组播路由器直接连接到接收方，负责通过IGMP了解接收方兴趣后创建组播状态。
RP	集合点。在ASM和PIM稀疏模式中使用的逻辑会议点，用于最初连接源和接收器。
ASM	任意源组播。一种组播模型，接收方无需事先指定源即可加入组。

了解众所周知的保留组播地址非常重要，因为组播故障排除取决于快速确定哪个控制协议正在使用给定的目标组，以及流量在网络中起到什么作用。这些地址有助于区分正常协议操作和异常行为，并使IGMP、PIM和自动RP交换的验证更容易。具体而言，对于自动RP，要识别的最重要组是224.0.1.39 (RP通告) 和224.0.1.40 (RP发现)，因为它们承载的信息允许路由器了解动态RP映射。

组播地址	目的
224.0.0.1	本地子网上的所有主机
224.0.0.2	本地子网上的所有路由器
224.0.0.13	所有PIM路由器
224.0.0.22	IGMPv3消息
224.0.1.39	自动RP使用的思科RP-Announce消息
224.0.1.40	自动RP使用的思科RP发现消息

PIM稀疏模式下的自动RP操作 (控制平面工作流程)

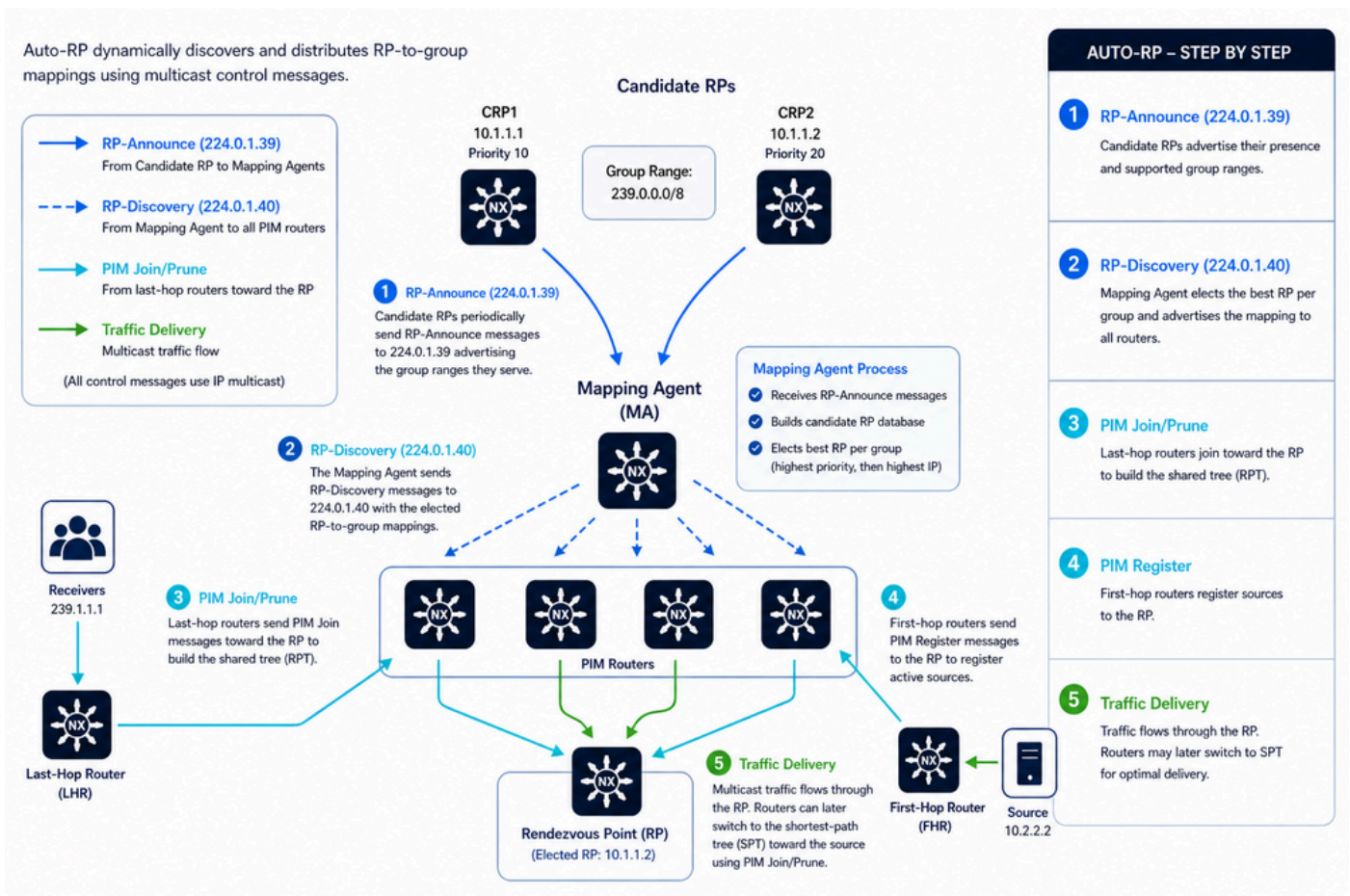
自动RP是协议无关组播稀疏模式中使用的Cisco机制，用于动态发现和分布组播域中的交汇点 (RP)信息。它通过使用基于组播的控制平面消息来通告、选择和学习RP到组的映射，从而无需进行静态RP配置。其主要组件是候选RP和映射代理，候选RP为特定组范围提供RP服务，映射代理收集候选并确定每个组的活动RP。

当每个候选RP定期向224.0.1.39发送RP-Announce消息 (包括其IP地址、优先级和支持的组播组范围) 时，该过程开始。在NX-OS中使用自动RP侦听程序将这些消息泛洪到整个网络，从而确保所有映射代理甚至在网络完全以稀疏模式运行之前收到这些消息。

映射代理监听这些通告，构建候选RP数据库，并对每个组执行确定性选择过程 (通常是最高优先级，然后是最高IP地址)。选择最佳RP后，映射代理生成RP发现消息并将其发送到224.0.1.40，向域中的所有路由器通告最终RP到组的映射。

所有PIM路由器都会收到RP发现消息并将映射安装到其本地RP表中。有了此信息，最后一跳路由器 (连接到接收器) 向选定RP发送PIM加入消息以构建共享树(RPT)，而第一跳路由器 (连接到源) 将组播流量封装在PIM注册消息中，以通知RP有关活动源。

当流量流经RP时，路由器可以选择使用额外的PIM加入/修剪信令直接从共享树(RPT)切换到最短路径树(SPT)，直接流向源。在此过程中，自动RP通过定期控制消息持续刷新映射，从而确保恢复能力和自动适应拓扑或RP更改。



自动RP限制和操作限制

- 自动RP仅用于IPv4组播，不支持IPv6(PIM6)，因此IPv6部署需要不同的RP发现机制。
- 自动RP不能与同一PIM域中的BSR共存，因为只能配置一个RP发现机制以避免控制平面冲突。
- 默认情况下，NX-OS设备上不处理或转发自动RP消息，因此正确操作需要显式配置。
- 自动RP控制消息的传播取决于启用自动RP侦听程序；否则，RP映射信息不会到达所有路由器。
- RP通告间隔的最小阈值为15秒，这限制了RP更新的通告速度并影响了收敛时间。
- 通过路由映射进行自动RP消息过滤可能会影响操作，因为不正确的策略可能会阻止RP到组的映射。
- 默认情况下，在VRF情景中禁用自动RP，因此必须在多VRF部署中显式启用它。
- 默认情况下启用基于ECMP的多路径转发，以允许组播流量使用等价路径进行负载均衡。
- 支持使用IPsec AH-MD5进行PIM邻居身份验证。
- PIM监听不可用。

使用自动RP的优势

自动RP为PIM稀疏模式组播环境提供动态RP发现和集中式RP映射分布。它无需在每台组播路由器上进行静态RP配置，降低了操作复杂性，并提高了组播可扩展性。自动RP还支持多个RP候选，从而启用自动RP故障切换和冗余。此机制有助于保持一致的组播转发行为，简化网络扩展，并允许组播路由器自动获知域中的RP信息。

自动RP映射代理和RP选择过程

自动RP中的选择过程是确定性的，主要基于IP地址。与其他协议（例如PIMv2 BSR）不同，自动RP不使用可配置的“优先级”值；相反，它依赖于IP地址层级来解决冲突。

映射代理(MA)选择

在自动RP中，多个映射代理可以在同一网络中共存以实现冗余。正式选举过程不会出现一个过程关闭，另一个过程开启；所有设备都处于技术活跃状态。但是，网络中的交换机必须决定信任哪些信息。

- 选择标准：最大IP地址。
- 过程：
 1. 所有映射代理将其RP发现消息发送到组播组224.0.1.40。
 2. 客户端交换机接收这些消息。
 3. 如果交换机从两个不同的映射代理收到包含冲突信息的发现消息，则交换机将接受来自源IP地址最高的映射代理的信息。
 4. Nexus行为：具有较低IP地址的映射代理通常会在检测到具有较高IP地址的另一个MA时进入“被动”或“抑制”状态，以防止网络中的重复控制流量。

交汇点(RP)选择

此过程由映射代理在侦听来自候选的所有RP-Announce消息（发送到组224.0.1.39）后执行。

当映射代理收到同一组播组的多个候选时，它按顺序应用以下规则：

规则A:最长前缀匹配（最具体的掩码）

如果候选者通告重叠范围，则MA会将组分配给通告最小范围（最长子网掩码）的RP。

- 示例：
 - 候选A通告224.0.0.0/4 (the整个组播范围)。
 - 候选人B通告224.10.20.0/24。
 - 结果：对于组224.10.20.5，候选B胜出，因为其范围更加具体。

规则B:最高IP地址(Tier-breaker)

如果两个或多个候选者通告完全相同的组范围，则映射代理只能选择一个。

- 条件：具有最高IP地址的候选者优先。
- 示例：
 - 考生1:范围224.10.20.0/24为10.2.0.1。
 - 考生2:范围224.10.20.0/24为10.2.0.4。
 - 结果：映射代理选择10.2.0.4作为该范围的官方RP，并且会在其发现消息中通告该范围。

第2层注意事项：组播MAC和监听

虽然本文的重点是使用PIM的第3层组播，但第2层行为在故障排除和总体设计方面起着关键作用。在第2层，设备使用MAC地址进行通信，MAC地址是分配给网络接口的48位标识符，组播流量需要

特定的MAC编址方案来将其与单播和广播流量区分开来。

IPv4组播MAC地址源自使用保留前缀“01:00:5E”的第3层组播组地址。但是，IP组播地址只有23位映射到MAC地址，这会产生32:1的重叠，这意味着多达32个不同的组播IP组可以映射到同一个MAC地址。因此，侦听给定组播MAC地址的主机可以接收多个组播组的流量，即使它们只对其中一个组感兴趣。例如，224.1.1.1、225.1.1.1、226.1.1.1、227.1.1.1、228.1.1.1等。

这种重叠对网络效率和故障排除有直接影响。由于完全基于MAC地址的第2层转发决策无法区分重叠的组播组，因此交换机可以向主机传输不必要的流量。然后，这些主机必须依靠高层过滤(IP/IGMP)来丢弃不必要的数据包，从而消耗CPU和缓冲区资源。

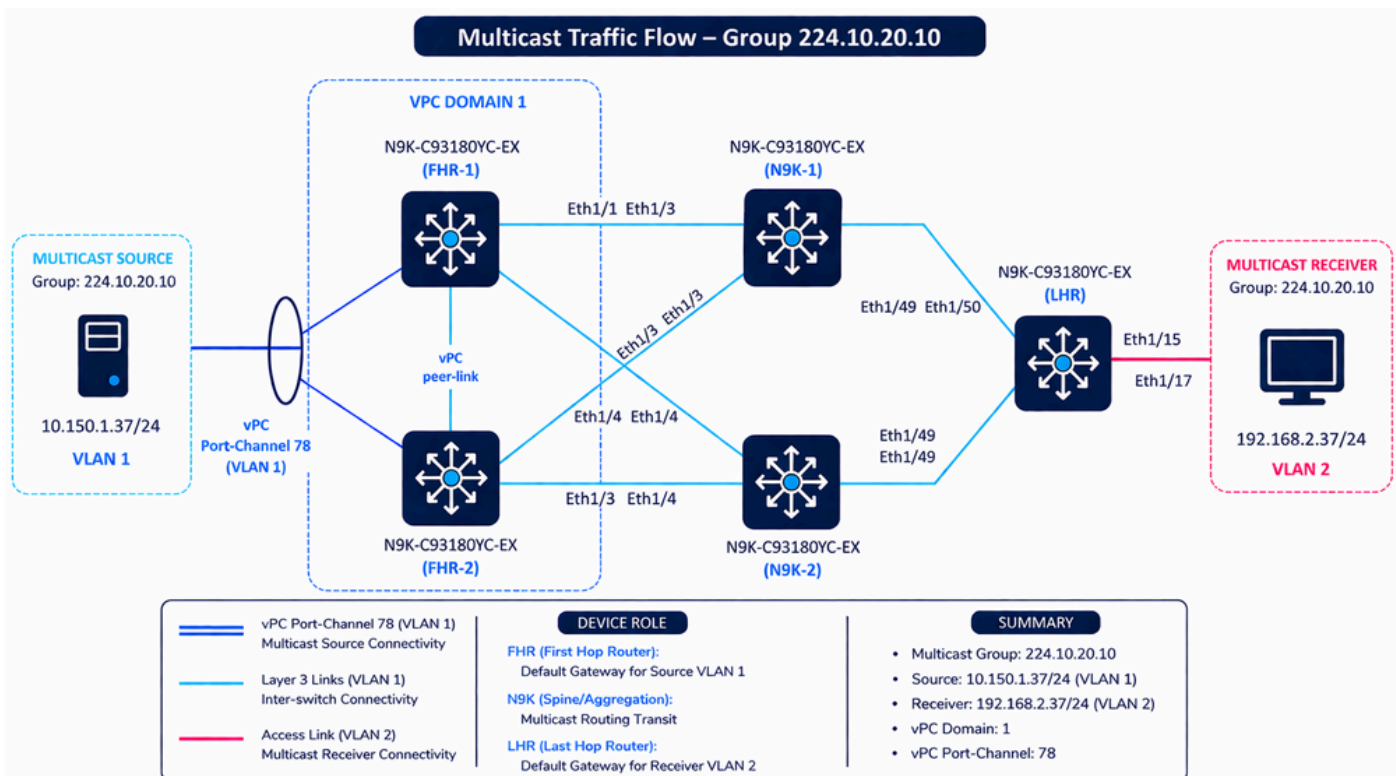
在Cisco Nexus NX-OS中，此限制通过IGMP监听行为得以缓解。默认情况下，IGMP监听执行基于IP的查找，而不是仅MAC转发，即使多个组播组共享同一MAC地址，交换机也能做出更精确的转发决策。这显着提高了第2层的效率并减少不必要的流量传输。

自动RP配置

本部分使用简单的实施作为参考，详细说明了自动RP配置。在此设置中，组播源通过vPC连接到两个Nexus交换机，以向接收方传输流量。在此设计中，N9K-1和N9K-2同时充当RP候选和映射代理。



警告： vPC端口通道不支持PIM邻居。



组播流量

第一跳路由器

相同配置具有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
```

命令	目的/说明
功能PIM	在交换机上启用PIM (协议无关组播) 进程。
ip pim auto-rp forward listen	启用自动RP侦听程序。这样，即使交换机尚不知道RP的身份，交换机也可以接收和转发自动RP控制消息 (224.0.1.39和224.0.1.40) 。
ip pim sparse-mode	在特定接口上启用PIM稀疏模式。在此模式下，组播流量仅转发到已通过PIM加入消息明确请求组播流量的网段。

映射代理和RP候选

```
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

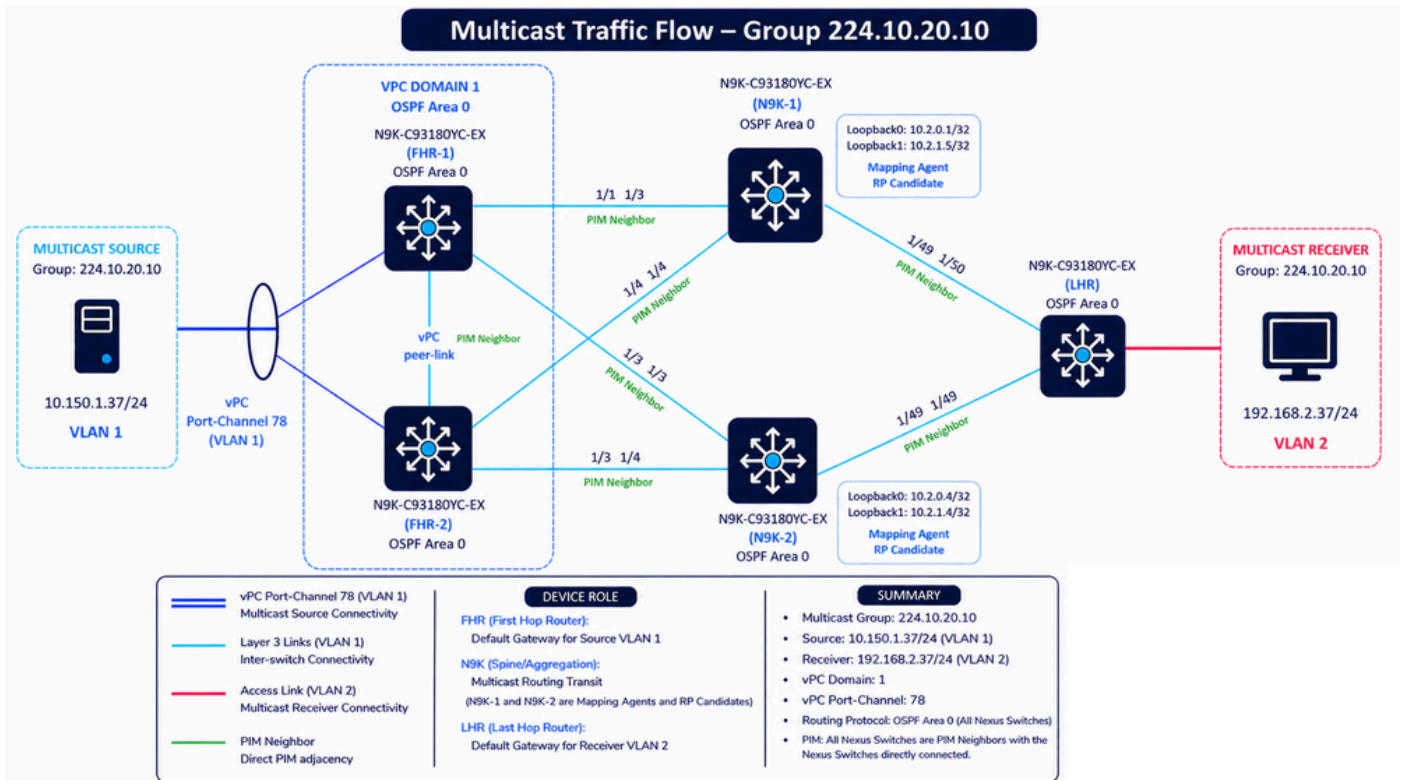
interface Ethernet1/49
 ip pim sparse-mode
```

下表提供了N9K-1的PIM配置的详细技术细分。此配置在N9K-2上复制。两台交换机均配置有双重角色，既充当组播域的RP候选者，又充当映射代理。

命令	详细的技术说明
功能PIM	功能激活：在Nexus交换机上全局启用协议无关组播(PIM)引擎。
ip pim auto-rp rp-candidate loopback0 group-list 224.10.20.0/24 interval 15	RP候选角色：将此交换机配置为“主动”作为交汇点(RP)。来源：使用loopback0的IP地址范围：它提供处理组播组范围224.10.20.0/24。间隔:每15秒向映射代理发送“通告”消息。保持计时器是此值的三倍。
ip pim auto-rp mapping-agent loopback1	映射代理角色：将交换机配置为自动RP进程的“管理员”。功能:它会侦听所有RP候选，解决冲突（使用最高的IP地址作为连接点），并向网络其余部分广播“发现”消息，以通知它们活动RP是谁。
interface loopback0 / loopback1	逻辑接口：在这些接口上启用PIM，因为它们充当RP候选和映射代理角色的源IP。它们必须通过单播路由表从所有PIM路由器到达。
interface Ethernet1/3、1/4、1/49	物理转发：在物理端口上启用PIM稀疏模式。这允许交换机与其他路由器形成PIM邻居邻接关系并通过这些特定链路转发组播流量。
ip pim sparse-mode	Operational Mode:应用于上面的所有接口。它确保组播流量仅发送到通过PIM加入消息明确请求组播流量的接收器，从而防止不必要的网络泛洪。

IPv4可达性的路由配置

- 在此拓扑中，所有Nexus交换机都参与一个名为UNDERLAY的OSPF进程，并在区域0(0.0.0.0)内运行，以便在整个网络中提供完全的路由可达性。
- 在LHR上，OSPF在充当组播接收器网关的VLAN 2上启用，而在配置为点对点的上行链路接口Ethernet1/49和Ethernet1/50上启用。这样可确保不进行DR/BDR选举，即可有效建立到上游N9K交换机的邻接关系。
- FHR层 (FHR-1和FHR-2) 共享相同的配置。OSPF在充当组播源网关的VLAN 1上启用，在通往汇聚层的路由上行链路 (Ethernet1/1和Ethernet1/3) 上启用。这些链路也配置为点对点链路，从而优化融合和邻接行为。
- 同样，N9K-1和N9K-2具有相同的OSPF配置，在OSPF中增加了Loopback0和Loopback1通告。
 - 这些环回非常关键，因为它们稍后用于PIM RP候选和映射代理角色。



PIM邻居和OSPF区域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
```

检验运行状态并排除自动RP故障

步骤 1：检验基本IP连通性（单播底层验证）

在分析组播行为之前，必须验证单播基础（OSPF区域0）是否完全运行。组播控制平面协议（如PIM和自动RP）取决于单播可达性才能正常工作。

第一个验证步骤是确认源和接收器(或其最近的第3层网关：FHR和LHR)可访问。

从拓扑中：

- FHR-1/FHR-2 →最接近组播源(10.150.1.37 - VLAN 1)
- LHR →最靠近组播接收器(192.168.2.37 - VLAN 2)

验证方法

1.执行ICMP可达性测试，测试范围包括：

- FHR ↔ LHR
- FHR ↔接收器子网（VLAN 2网关）
- LHR ↔源子网（VLAN 1网关）

2.验证路由表中的源和接收器可达性。对于vPC部署，确保两个Nexus对等设备的一致性。请注意，接收器具有ECMP路径，而源可通过第2层到达。

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 — 路由到接收器

```
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 — 路由到接收器

```
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 — 路由到接收器

```
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
```

重要注意事项

- 此测试失败强烈表明存在底层问题。
- 没有单播可达性，组播无法正常工作，原因是：
 - PIM邻居依赖单播路由
 - RP (环回) 地址必须可达
 - RPF (反向路径转发) 检查取决于路由表
- ping成功并不保证组播可以正常工作，因为：
 - 在阻止组播时可以允许ICMP
 - PIM或自动RP仍可能配置错误



提示：在分析自动RP、PIM邻接或RP选择之前，始终确保底层路由域是稳定、一致且完全可访问的端到端路由域。

步骤 2：确定组播角色和端到端拓扑

下一步是明确确定转发组播流量时涉及的每台设备的角色。这是强制性的步骤，因为组播故障排除完全取决于了解整个拓扑中的流量和预期行为。

至少必须确定以下要素：

- 组播源：10.150.1.37(VLAN 1)
- 组播组(G):224.10.20.10
- 接收方：192.168.2.37(VLAN 2)
- 第一跳路由器(FHR):FHR-1/FHR-2 (最靠近源)
- 最后一跳路由器(LHR):LHR (最靠近接收方)

此外，还需要确定控制平面角色：

- RP候选：N9K-1和N9K-2 (环回接口0)
- 映射代理：N9K-1和N9K-2(Loopback1)

拓扑和控制平面感知

要进行组播故障排除，必须提供详细的网络拓扑。包括：

- 物理连接（设备之间使用的接口）
- 逻辑拓扑（VLAN、路由链路、vPC关系）
- 使用的路由协议（本设计中的OSPF区域0）
- 路由域边界（单IGP与混合协议，例如OSPF、EIGRP或BGP）
- 用于RP和映射代理角色的环回接口
- 启用PIM的接口和邻居关系

必须清楚理解的内容

- 从源FHR到RP→LHR→收器→精确→径
- 哪些设备负责：
 - 发送PIM注册(FHR)
 - 建筑(*,G)或(S, G)树(LHR)
 - 通告RP信息（映射代理）
- 路由(OSPF)如何确保可达性：
 - 来源子网
 - 接收方子网
 - RP环回地址
 - 映射代理环回地址



警告：没有清晰拓扑的组播故障排除等同于没有可视性的调试，这会导致错误的假设和误诊。

步骤 3：根据设备角色验证自动RP配置

下一步是根据每台设备在组播拓扑中的角色验证自动RP是否正确配置。这包括确认：

- RP候选(N9K-1 / N9K-2)已正确配置为将其Loopback0通告为组播组范围的RP。

- 映射代理(N9K-1 / N9K-2)配置为使用Loopback1收集RP通告消息和生成RP发现消息。
- FHR和LHR在所有相关接口上启用了PIM稀疏模式，以参与自动RP和接收RP映射。

必须确保为PIM稀疏模式启用所有必需接口（包括环回和路由链路），并且不存在会阻止RP-Announce(224.0.1.39)和RP-Discovery(224.0.1.40)消息交换的缺少配置。



注意：N9K-1和N9K-2配置为组播域内的RP候选和映射代理。



警告：任何缺失或不一致的自动RP配置都会阻止路由器学习RP，从而导致组播流量无法正确转发。

步骤 4：验证所有RP候选者和所有映射代理的操作

第4.1步检验PIM邻居邻接关系

第一个验证步骤是确认所有预期的PIM邻居已正确建立跨组播拓扑。

N9K-1 — 检验PIM邻居

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

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

N9K-2 — 检验PIM邻居

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

```
PIM Neighbor Status for VRF "default"
Neighbor      Interface      Uptime    Expires    DR      Bidir-  BFD      ECMP Redirect
              Interface      Uptime    Expires    Priority Capable  State    Capable
10.4.0.9      Ethernet1/3    23:21:18  00:01:29  1       yes     n/a      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
```

验证点

- 检验所有期望的邻居是否出现在每个路由接口上。
- 检验邻居的正常运行时间是否稳定并持续增加。
- 验证Expires计时器定期刷新。
- 确认未发生意外的邻接摆动。
- 检验正确的指定路由器(DR)选择。
- 验证组播接口是否成功形成PIM邻接关系。

在此拓扑中：

- N9K-1建立三个PIM邻居。
- N9K-2还建立了三个PIM邻居。
- 所有邻接关系保持稳定超过23小时，这表示组播控制平面运行稳定。

第4.2步检验启用PIM的接口

下一步是确认所有参与自动RP的接口都启用了PIM。

此验证对于以下方面尤其重要：

- RP候选环回
- 映射代理环回

N9K-1 — 检验PIM接口

```
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 — 检验PIM接口

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

验证点：

- 检验输出中是否显示所有路由组播接口。
- 验证RP-Candidate环回是否启用了PIM。
- 验证映射代理环回是否已启用PIM。
- 检验传输接口上的邻居计数是否正确。
- 检验环回接口是否正确显示Neighbor Count = 0。

环回角色分配：

设备	功能	环回
N9K-1	RP候选	10.2.0.1
N9K-1	映射代理	10.2.1.5
N9K-2	RP候选	10.2.0.4
N9K-2	映射代理	10.2.1.4

环回正确地显示为活动PIM接口，即使它们不形成PIM邻居。由于环回接口不直接建立组播邻接关系，所以此行为是预期行为。

这些环回的存在证实：

- PIM可以正确发送自动RP控制消息。
- 可以成功生成RP通告。
- 映射代理功能可以正常运行。

第4.3步分析show ip pim rp

此命令验证：

- RP发现
- 自动RP通告

- 映射代理操作
- 组到RP的映射

N9K-1 - RP信息

```
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)
```

逐行说明

BSR已禁用

```
BSR disabled
```

这证实：

- 未使用引导路由器(BSR)。
- 组播域完全依赖自动RP。

此行为在此拓扑中是预期行为。

自动RP RPA:10.2.1.5*

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

- 10.2.1.5对应于N9K-1上的loopback1。
- *表示本地交换机本身是活动的映射代理。

这意味着：

- N9K-1生成RP发现消息。
- N9K-1从RP-Candidates收集RP通告。
- N9K-1将RP映射信息分发到组播域。

中的下一个发现消息

next Discovery message in: 00:00:39

- 验证计时器是否持续刷新。
- 检验发现消息是否定期发送。

如果此计时器冻结或意外过期，自动RP通告将无法正确传播。

策略字段

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

- 未应用过滤策略。
- 接受所有RP通告和发现通告。

第一个RP条目

RP: 10.2.0.1*, (0),

- RP地址= 10.2.0.1
- *表示此RP位于N9K-1本地。
- 10.2.0.1对应于N9K-1上的环回接口0。

这确认N9K-1已配置为RP候选。

正常运行时间和优先级

uptime: 22:18:44 priority: 255,

- 稳定的正常运行时间表示稳定的RP通告操作。
- 优先级255是默认的最高优先级。

RP源

RP-source: 10.2.0.1 (A),

- RP通告直接源自RP本身。
- (A)表示自动RP获取的信息。

组范围

224.10.20.0/24

- 检验是否通告了正确的组播范围。
- 检验组范围是否与配置匹配。

第二个RP条目

RP: 10.2.0.4, (0),

- 拓扑中存在另一个RP。
- 10.2.0.4对应于N9K-2上的环回接口0。
- 否*显示，因为此RP从N9K-1远程。

N9K-2 - RP信息

```
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的主要区别

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

- 否*，因为N9K-2不是映射代理。
- N9K-2从N9K-1远程学习映射代理信息。

重要的RP差异

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

- N9K-2从映射代理获取RP信息。
- 映射代理是10.2.1.5。

这证实：

- 自动RP发现消息运行正常。
- 映射代理通告成功传播到组播域。

第4.4步验证自动RP选举过程和RP选择

自动RP使用两种不同的组播控制平面功能：

- RP候选
- 映射代理

在PIM稀疏模式环境中验证组播操作时，了解这些功能如何交互至关重要。

在此拓扑中：

- N9K-1和N9K-2作为RP候选者运行。
- N9K-1作为活动映射代理运行。

RP候选操作

RP候选通告自身作为一个或多个组播组范围的有效交汇点。

每个RP候选定期将自动RP通告消息发送到：

- 224.0.1.39 — 自动RP通告组

这些公告包括：

- RP地址
- 组范围
- 优先级
- 保持时间

在此拓扑中：

- N9K-1上的10.2.0.1将自身通告为RP。
- N9K-2上的10.2.0.4将自身通告为RP。

N9K-1 - RP候选信息

```
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候选信息

```
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)
```

两台设备通告相同的组播组范围：

- 224.10.20.0/24

两个RP候选还使用：

- 第 255 优先级

这一点很重要，因为自动RP在RP选择期间使用优先级和RP地址。

主动映射代理标识

映射代理从以下逻辑开始为组播组选择活动RP:

1. 最高RP优先级优先。
2. 如果优先级相等，则使用最高的RP IP地址。

在此拓扑中：

- N9K-1 RP地址= 10.2.0.1
- N9K-2 RP地址= 10.2.0.4
- 两个RP候选都使用优先级255

因为两个RP候选具有相同的优先级：

- 最大RP IP地址成为所选RP。

因此：

- 10.2.0.4成为224.10.20.0/24的活动RP。

所选RP验证

N9K-2 — 选定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
```

为什么N9K-1仍然显示两个RP条目

在N9K-1上：

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

出现此行为的原因如下：

- N9K-1作为映射代理运行。
- 映射代理维护所有RP候选的可见性。
- 输出在分发最终RP选择之前显示所有已学习的RP通告。



警告：在映射代理上，必须显示同一组播域内的所有RP候选。如果缺少任何RP-Candidate，则通过向源自映射代理IP地址（通常为环回接口）的RP-Candidate IP地址发送ping来验证可达性。

步骤 5：验证与RP候选和映射代理的可达性

参与PIM稀疏模式域的所有组播路由器必须保持稳定的IP可达性，以便：

- 所有RP候选
- 所有映射代理

此验证至关重要，因为PIM稀疏模式依赖于单播路由：

- 到达交汇点(RP)
- 构建共享组播树(RPT)
- 从第一跳路由器(FHR)传送PIM注册消息
- 接收自动RP发现通告
- 执行反向路径转发(RPF)验证

如果到RP或映射代理的连通性失败：

- 组播源可能无法成功注册。
 - 接收方可能无法加入组播组。
 - 自动RP通告无法正确传播。
 - 可能会发生RPF故障。
 - 组播流量转发可能变得不稳定或间歇性。
-

检验稳定的路由条目

路由表必须包含指向以下目标的稳定单播路由：

- 所有RP环回
- 所有映射代理环回
- 所有组播传输接口

路由必须保持连续安装，不会出现路由摆动或过度重新收敛事件。

验证：

- 正确的下一跳选择
 - 正确的传出接口
 - 稳定的路由正常运行时间
 - 预期的路由协议源
 - 跨组播域的一致单播可达性
-

FHR-1 — 到RP候选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候选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 — 到映射代理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 — 到映射代理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候选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候选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 — 到映射代理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 — 到映射代理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候选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候选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 — 到映射代理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 — 到映射代理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

检验从PIM接口的Ping可达性

验证路由表后，验证通向以下目标的端到端IP可达性：

- 所有RP候选
- 所有映射代理

ping的来源必须是：

- 启用PIM的接口
- 参与组播路由的环回接口

此验证非常重要，因为组播路由器在以下过程中使用这些源地址：

- PIM邻居建立
- RPF计算
- PIM寄存器封装
- 自动RP通信



提示：如果使用未编号接口，其中多个第3层接口共享来自环回接口的相同IP地址，可达性验证会变得更简单，因为单个源IP地址可以一致使用。

Ping验证摘要

设备	源 IP	目的地	功能	结果
FHR-1	10.4.0.5	10.2.0.1	RP候选	成功
FHR-1	10.4.0.5	10.2.0.4	RP候选	成功
FHR-1	10.4.0.5	10.2.1.5	映射代理	成功
FHR-1	10.4.0.5	10.2.1.4	映射代理	成功
FHR-2	10.4.0.9	10.2.0.1	RP候选	成功
FHR-2	10.4.0.9	10.2.0.4	RP候选	成功
FHR-2	10.4.0.9	10.2.1.5	映射代理	成功
FHR-2	10.4.0.9	10.2.1.4	映射代理	成功

LHR	10.4.0.5	10.2.0.1	RP候选	成功
LHR	10.4.0.5	10.2.0.4	RP候选	成功
LHR	10.4.0.5	10.2.1.5	映射代理	成功
LHR	10.4.0.5	10.2.1.4	映射代理	成功

检验FHR和LHR上的运行状态和组播流量转发

下一个验证步骤是验证：

- FHR和LHR成功获取了预期的RP。
- 组播路由表包含预期组播状态。
- 组播控制平面在组播流量传输开始之前运行正常。
- PIM稀疏模式构建预期的共享树状态。

第1步检验FHR和LHR上的RP学习

在验证组播转发状态之前，验证所有组播路由器是否都已获知所验证组播组的预期RP。

此步骤至关重要，因为：

- FHR必须知道将PIM注册消息发送到何处。
- LHR必须知道构建共享树时要使用的RP。
- 组播域必须维护一致的RP映射。

在此拓扑中：

- 所选RP是10.2.0.4。
- 映射代理是10.2.1.5。
- 组播组范围是224.10.20.0/24。

FHR-1 — 验证学习的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 — 验证学习的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 — 验证获取的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)

RP学习分析

输出确认：

- 所有组播路由器都一致地获取了相同的RP。
- 所选RP是10.2.0.4。
- RP映射通过自动RP获取。
- 分发RP信息的映射代理是10.2.1.5。
- RP映射在一天以上的时间内保持稳定。
- 发现到期计时器会正确刷新。

出现此行为的原因如下：

- 在自动RP选举过程中，10.2.0.4被选为活动RP。
- 10.2.1.5作为映射代理运行。
- 所有组播路由器都成功接收自动RP发现通告。

在此阶段，组播控制平面运行正常，并且所有路由器都具有一致的224.10.20.0/24 RP映射

第2步检验活动组播流量之前的组播路由状态

下一步是在组播流量传输开始之前验证组播路由表。

在这种情况下：

- 组播接收器已加入组播组。
- 活动组播源流量尚未流动。

此状态非常重要，因为它验证以下情况：

- IGMP成员身份操作
- PIM共享树创建
- 初始(*,G)组播状态
- 接收器兴趣传播

FHR-1 — 组播路由表

```
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 — 组播路由表

```
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组播状态分析

FHR尚未包含：

- 224.10.20.10的(*,G)状态
- 组播源的(S , G)状态

出现此行为的原因如下：

- 组播源流量尚未处于活动状态。
- 未生成PIM注册消息。
- FHR尚未启动组播转发。

唯一存在的组播条目是：

- 232.0.0.0/8

这相当于默认的SSM范围，由系统自动安装。

这些值应为预期值：

- 传入接口:Null
- RPF邻居：0.0.0.0
- 传出接口计数：0

此条目不表示活动的组播转发。

LHR — 组播路由表

```
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组播状态分析

与FHR不同，LHR包含活动(*,G)条目，用于：

- 224.10.20.10

出现此行为的原因如下：

- 组播接收器已加入组。
- IGMP成员信息是在Vlan2上获取的。
- LHR启动了指向RP的共享树连接。

组播路由表确认：

- 传入接口(IIF)是Ethernet1/50。
- RPF邻居为10.4.0.26。
- 传出接口列表(OIL)包含Vlan2。

这表示：

- LHR正确构建了指向RP的共享树。
- 组播接收器兴趣已成功向上游传播。
- 已正确传输PIM加入消息。

在此阶段：

- 组播源流量尚未处于活动状态。
- 尚不存在(S, G)状态。
- 仅存在共享树状态。

N9K-1 — 映射代理组播路由表

```
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)
```

映射代理组播状态分析

N9K-1仅作为映射代理运行，不参与224.10.20.10的组播转发

因此，预计会缺少(*,G)条目和(S, G)条目。

映射代理仅分发RP映射信息，并不一定参与组播数据转发，除非组播流量直接流经设备。

N9K-2 — RP组播路由表

```
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组播状态分析

N9K-2用作以下设备的活动RP:

- 224.10.20.0/24

因此，RP包含224.10.20.10的共享树(*,G)状态

这些值应为预期值：

- 传入接口= loopback0
- RPF邻居= 10.2.0.4
- 传出接口= Ethernet1/49

这表示：

- RP已成功安装共享树状态。
- RP收到来自下游路由器的PIM加入消息。
- 指向接收器的共享组播树已正确构建。

在此阶段：

- RP仅包含共享树状态。
- 组播源流量尚未处于活动状态。
- 尚不存在(S，G)条目。

第3步使用活动组播流量检验组播路由状态

一旦组播流量传输开始，组播路由表就会从共享树状态转换到活动源特定转发状态。

在这种情况下：

- 组播源为10.150.1.37。
- 组播组为224.10.20.10。
- N9K-2上的活动RP是10.2.0.4。
- 组播源通过vPC端口通道78连接。

vPC组播的重要注意事项

组播源通过FHR-1和FHR-2形成的vPC域连接。

由于源通过vPC成员port-channel连接：

- 流量散列可以将组播数据包转发到任一Nexus交换机。
- 两个FHR需要相同的组播配置。
- 两个FHR都需要一致的单播路由信息。
- 两个FHR必须维护相同的RP映射。
- 两个FHR必须保持稳定的PIM邻接关系。

在此特定场景中：

- 组播流向FHR-2散列。
- FHR-2成为组播源的活动转发FHR。

FHR-1 — 组播路由表

```
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组播状态分析

FHR-1包含活动(S, G)条目, 用于:

- 源= 10.150.1.37
- 组= 224.10.20.10

组播路由条目确认:

- 源地址是在Vlan1上获取的。
- RPF邻居是直接连接的组播源。
- OIL中不存在传出接口。

这是预期行为, 因为组播流未针对出站转发向FHR-1散列。

因此:

- FHR-1仅安装本地源状态。
- FHR-1不通过PIM转发上游组播流量。

FHR-2 — 组播路由表

```
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组播状态分析

与FHR-1不同，FHR-2包含：

- 活动的传出接口列表
- 活动PIM转发路径

这表示：

- FHR-2成为组播流的操作转发FHR。
- 组播数据包通过vPC成员port-channel散列到FHR-2。
- FHR-2将组播流量封装到PIM注册消息中。
- FHR-2将组播流量向上游转发到RP。

ECMP和组播转发行为

传出接口Ethernet1/3匹配指向接收方192.168.2.37的单播路由表

FHR-2 — 到组播接收器的路由

```
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包含两个通往组播接收器子网的等价路由：

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

这证实：

- ECMP路由在组播传输拓扑中处于活动状态。
- 存在多条通向接收方网络的有效单播转发路径。
- PIM可以使用MRIB选择的RPF路径做出组播转发决策。
- 不会发生跨并行ECMP路径的数据包重复。

虽然存在两个等价路由，但组播转发对每个组播流使用单个RPF路径。

在此拓扑中，组播流使用：

- Ethernet1/3到10.4.0.13

此行为与之前观察到的组播路由表匹配：

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

vPC运行的主要和次要行为

vPC操作角色对组播转发行为的影响不同：

- PIM流量
- IGMP处理

在此拓扑中：

- FHR-1是运行中的vPC主设备。
- FHR-2是运行中的vPC辅助设备。

两台Nexus交换机都可以：

- 处理PIM控制平面流量
- 构建组播路由状态
- 通过PIM转发组播流量

但是：

- 只有运行中的vPC主设备会处理流向接入层的IGMP接收器流量。

这一区别很重要，因为：

- 向路由网络转发的PIM在两个对等体上均保持活动状态。
- 向第2层接收器网段转发的IGMP仍然集中在运行主网段上。

因此：

- 从源到路由组播域的组播流量可以通过任一vPC对等体退出。
 - 指向IGMP接收器的组播流量依赖于运行的主要行为。
-

LHR — 组播路由表

```
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组播状态分析

LHR现在包含两者：

- (* ,G)共享树状态
- (S , G)源树状态

这将确认：

- 组播接收器已成功加入。
- 组播源变为活动状态。
- LHR从共享树转发转换为源树转发。

(S , G)条目确认 :

- 组播源路径已成功获取。
- RPF邻居是10.4.0.22。
- 组播流量通过Ethernet1/49到达。
- 组播流量转发到Vlan2接收器。

此行为确认成功 :

- PIM转发
 - RPF验证
 - 源树构造
 - 组播流量传输
-

N9K-1 — 传输组播路由表

```
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传输状态分析

N9K-1充当活动组播流的传输组播路由器。

组播路由条目确认 :

- 组播流量从Ethernet1/4到达。
- 组播流量转发到Ethernet1/49。
- 指向LHR的组播转发路径运行正常。

这确认成功：

- PIM邻居操作
 - RPF验证
 - 通过传输网络进行组播转发
-

N9K-2 — RP组播路由表

```
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组播状态分析

N9K-2用作组播组的活动RP。

RP同时包含两者：

- (* ,G)共享树状态
- (S , G)源树状态

(S , G)条目中缺少传出接口，原因如下：

- RP已将接收器切换到最短路径树。
- 活动组播转发不再需要共享树转发路径。

RCA和运行状况诊断所需的最小组播信息

命令列表提供了在运行NX-OS的Cisco Nexus 9000系列交换机上执行适当的根本原因分析(RCA)或组播运行状况诊断所需的最低建议组播数据收集。这些输出捕获组播控制平面状态、MRIB编程、转发信息、vPC运行状态和硬件转发详细信息。但是，根据故障场景，仍可能需要其他信息。例如，组播数据包丢失、间歇性流量丢弃、数据包复制问题、硬件转发不一致或无序组播转发通常需要使用Ethanalyzer、SPAN或硬件级捕获在Nexus交换机上直接捕获数据包。同样，瞬时RPF不一致、ECMP转发更改、ASIC编程故障或IGMP抑制事件也可能发生，而不会生成永久日志。

因此，将show tech输出与数据包捕获和转发验证相结合，可显著提高诊断准确性和RCA质量。虽然此信息为组播故障排除提供了强大的操作基线，但不能保证始终可以仅从这些输出中识别RCA。某些组播故障需要额外的故障排除、实时流量分析、硬件级验证、拓扑关联或扩展数据包捕获来查明确切的根本原因。



提示：在工作和非工作期间收集此信息，可清晰地从Nexus的角度快速了解问题的出现方式，并显著增强确定根本原因的能力。

最小组播数据收集命令

```
<#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
```

创建用于导出的TAR存档

生成show tech文件后，将其整合到单个TAR存档中，以供导出和分析。该命令是单行。

```
<#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
```

导出单个TAR归档可简化：

- TAC案例上传
- RCA工作流程
- 集中日志分析
- 组播转发关联
- 历史故障保留

关于此翻译

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

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

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