H3C S5500-HI 系列以太网交换机

IP 组播配置指导

杭州华三通信技术有限公司
http://www.h3c.com.cn

资料版本：6W102-20131220
产品版本：Release 52xx 系列

前 言

H3C S5500-HI 系列以太网交换机配置指导共分为十一本手册，介绍了 S5500-HI 系列以太网交换机 Release52xx 系列软件版本各软件特性的原理及其配置方法，包含原理简介、配置任务描述和配置举例。《IP 组播配置指导》主要介绍组播业务的相关协议原理和具体配置，包括 IPv4 组播业务配置及 IPv6 组播业务的配置。利用组播技术可以实现网络中点到多点的高效数据传送。

前言部分包含如下内容：

读者对象

本手册主要适用于如下工程师：

- 网络规划人员
- 现场技术支持与维护人员
- 负责网络配置和维护的网络管理员

新增及修改特性说明

本手册对应 S5500-HI 系列以太网交换机的 Release52xx 系列软件版本，各版本间特性差异如下：

- Release5206 与Release5203 版本相比，修改了部分特性，具体请参见 表 1。
- Release5203 与Release5101 版本相比，新增、修改了部分特性，具体请参见 表 2。

表1 Release5206 与 Release5203 版本间特性差异

配置指导	新增及修改特性
组播概述	无
IGMP Snooping	变更特性：支持配置 0100-5Exx-xxxx 范围内的静态组播 MAC 地址表项，其中 x 代表 0～F 的任意一个十六进制数
PIM Snooping	无
组播 VLAN	无
组播路由与转发	无
IGMP	无
PIM	无

配置指导	新增及修改特性
MSDP	无
MBGP	无
组播 VPN	无
MLD Snooping	变更特性：支持配置 3333-xxxx-xxxx 范围内的静态组播 MAC 地址表项，其中 x 代表 0～F 的任意一个十六进制数
IPv6 PIM Snooping	无
IPv6 组播 VLAN	无
IPv6 组播路由与转发	无
MLD	无
IPv6 PIM	无
IPv6 MBGP	无

表2 Release5203 与 Release5101 版本间特性差异

配置指导	新增及修改特性
组播概述	无
IGMP Snooping	新增特性：配置 IGMP Snooping 协议中发送 IGMP 报文的 DSCP 优先级
PIM Snooping	无
组播 VLAN	无
组播路由与转发	无
IGMP	新增特性：配置 IGMP 报文的 DSCP 优先级
PIM	新增特性：配置 PIM 报文的 DSCP 优先级
MSDP	无
MBGP	无
组播 VPN	新增特性：配置 BGP MDT
MLD Snooping	新增特性：配置 MLD Snooping 协议中发送的 MLD 报文的 DSCP 优先级
IPv6 PIM Snooping	无
IPv6 组播 VLAN	无
IPv6 组播路由与转发	无
MLD	新增特性：配置 MLD 报文的 DSCP 优先级
IPv6 PIM	新增特性：配置 IPv6 PIM 报文的 DSCP 优先级
IPv6 MBGP	无

本书约定

1. 命令行格式约定

格　式	意　义
粗体	命令行关键字（命令中保持不变、必须照输的部分）采用**加粗**字体表示。
斜体	命令行参数（命令中必须由实际值进行替代的部分）采用*斜体*表示。
[ ]	表示用“[ ]”括起来的部分在命令配置时是可选的。
{ x | y | ... }	表示从多个选项中仅选取一个。
[x | y | ...]	表示从多个选项中选取一个或者不选。
{ x | y | ... }*	表示从多个选项中至少选取一个。
[x | y | ...]*	表示从多个选项中选取一个、多个或者不选。
&<1-n>	表示符号&前面的参数可以重复输入1～n次。
#	由“#”号开始的行表示为注释行。

2. 图形界面格式约定

格　式	意　义
< >	带尖括号“< >”表示按钮名，如“单击<确定>按钮”。
[ ]	带方括号“[ ]”表示窗口名、菜单名和数据表，如“弹出[新建用户]窗口”。
/	多级菜单用“/”隔开。如[文件/新建/文件夹]多级菜单表示[文件]菜单下的[新建]子菜单下的[文件夹]菜单项。

3. 各类标志

本书还采用各种醒目标志来表示在操作过程中应该特别注意的地方，这些标志的意义如下：

 警告	该标志后的注释需给予格外关注，不当的操作可能会对人身造成伤害。
 注意	提醒操作中应注意的事项，不当的操作可能会导致数据丢失或者设备损坏。
 提示	为确保设备配置成功或者正常工作而需要特别关注的操作或信息。
 说明	对操作内容的描述进行必要的补充和说明。
 窍门	配置、操作、或使用设备的技巧、小窍门。

4. 图标约定

本书使用的图标及其含义如下：

	该图标及其相关描述文字代表一般网络设备，如路由器、交换机、防火墙等。
	该图标及其相关描述文字代表一般意义下的路由器，以及其他运行了路由协议的设备。
	该图标及其相关描述文字代表二、三层以太网交换机，以及运行了二层协议的设备。

5. 端口编号示例约定

本手册中出现的端口编号仅作示例，并不代表设备上实际具有此编号的端口，实际使用中请以设备上存在的端口编号为准。

产品配套资料

H3C S5500-HI 系列以太网交换机的配套资料包括如下部分：

大类	资料名称	内容介绍
产品知识介绍	产品彩页	帮助您了解产品的主要规格参数及亮点
	技术白皮书	帮助您了解产品和特性功能，对于特色及复杂技术从细节上进行介绍
硬件介绍及安装	安全兼容性手册	列出产品的兼容性声明，并对兼容性和安全的细节进行说明
	H3C 设备防雷安装指导手册	帮助您了解防雷接地设计和工程安装方法，以保证交换机具有良好的抗雷击性能
	快速安装指南	指导您对设备进行初始安装，通常针对最常用的情况，减少您的检索时间
	安装指导	帮助您详细了解设备硬件规格和安装方法，指导您对设备进行安装
	风扇安装手册	帮助您了解产品支持的可插拔风扇模块的外观、功能、规格、安装及拆卸方法
	电源手册	帮助您了解产品支持的可插拔电源模块的外观、功能、规格、安装及拆卸方法
	RPS电源用户手册	帮助您了解产品支持的RPS电源的外观、功能、规格
	H3C低端系列以太网交换机RPS电源选购指南	帮助您了解各种RPS电源适用的交换机产品型号及RPS电源配套电缆的相关规格
	接口模块扩展卡用户手册	帮助您了解该接口模块扩展卡的外观、规格、安装及拆卸方法
	H3C低端系列以太网交换机可插拔模块手册	帮助您了解产品支持的可插拔模块类型、外观和规格
	H3C可插拔SFP[SFP+][XFP]模块安装指南	帮助您掌握SFP/SFP+/XFP模块的正确安装方法，避免因操作不当而造成器件损坏

大类	资料名称	内容介绍
业务配置	配置指导	帮助您掌握设备软件功能的配置方法及配置步骤
	命令参考	详细介绍设备的命令，相当于命令字典，方便您查阅各个命令的功能
运行维护	故障处理手册	指导您快速定位并处理软件故障
	版本说明书	帮助您了解产品版本的相关信息（包括：版本配套说明、兼容性说明、特性变更说明、技术支持信息）及软件升级方法

资料获取方式

您可以通过H3C网站（www.h3c.com.cn）获取最新的产品资料：

H3C 网站与产品资料相关的主要栏目介绍如下：

- [服务支持/文档中心]：可以获取硬件安装类、软件升级类、配置类或维护类等产品资料。
- [产品技术]：可以获取产品介绍和技术介绍的文档，包括产品相关介绍、技术介绍、技术白皮书等。
- [解决方案]：可以获取解决方案类资料。
- [服务支持/软件下载]：可以获取与软件版本配套的资料。

技术支持

用户支持邮箱：service@h3c.com

技术支持热线电话：400-810-0504（手机、固话均可拨打）

010-62982107

网址：http://www.h3c.com.cn

资料意见反馈

如果您在使用过程中发现产品资料的任何问题，可以通过以下方式反馈：

E-mail：info@h3c.com

感谢您的反馈，让我们做得更好！

目 录

1 组播概述

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- 如不特别指明，本文中提到的组播均指 IP 组播。

1.1 组播简介

作为一种与单播（Unicast）和广播（Broadcast）并列的通信方式，组播（Multicast）技术能够有效地解决单点发送、多点接收的问题，从而实现了网络中点到多点的高效数据传送，能够节约大量网络带宽、降低网络负载。

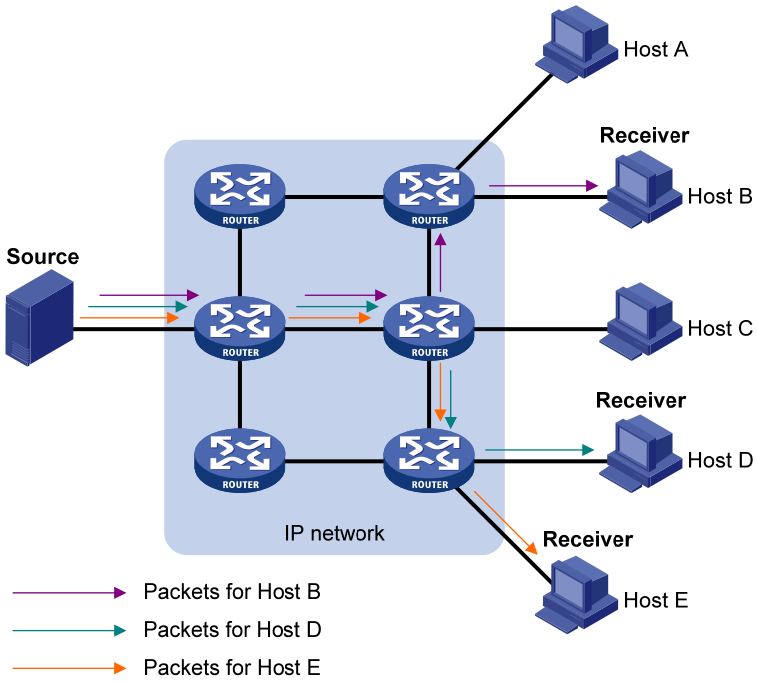
利用组播技术可以方便地提供一些新的增值业务，包括在线直播、网络电视、远程教育、远程医疗、网络电台、实时视频会议等对带宽和数据交互的实时性要求较高的信息服务。

1.1.1 三种信息传输方式的比较

1. 单播方式的信息传输

如 图 1-1 所示，在IP网络中若采用单播的方式，信息源（即Source）要为每个需要信息的主机（即Receiver）都发送一份独立的信息拷贝。

图1-1 单播方式的信息传输



假设 Host B、Host D 和 Host E 需要信息，则 Source 要与 Host B、Host D 和 Host E 分别建立一条独立的信息传输通道。

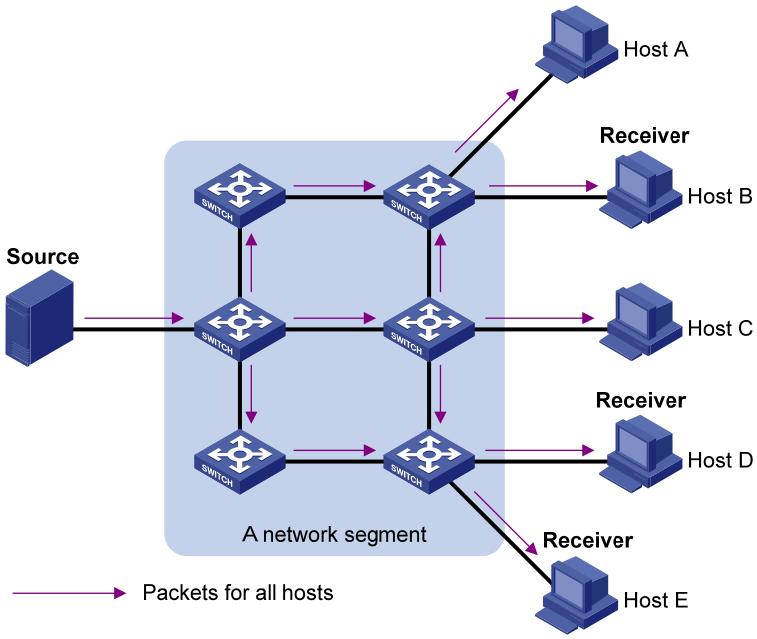
采用单播方式时，网络中传输的信息量与需要该信息的用户量成正比，因此当需要该信息的用户数量较大时，信息源需要将多份内容相同的信息发送给不同的用户，这对信息源以及网络带宽都将造成巨大的压力。

从单播方式的信息传播过程可以看出，该传输方式不利于信息的批量发送。

2. 广播方式的信息传输

如 图 1-2 所示，在一个网段中若采用广播的方式，信息源（即Source）将把信息传送给该网段中的所有主机，而不管其是否需要该信息。

图1-2 广播方式的信息传输



假设只有 Host B、Host D 和 Host E 需要信息，若将该信息在网段中进行广播，则原本不需要信息的 Host A 和 Host C 也将收到该信息，这样不仅信息的安全性得不到保障，而且会造成同一网段中信息的泛滥。

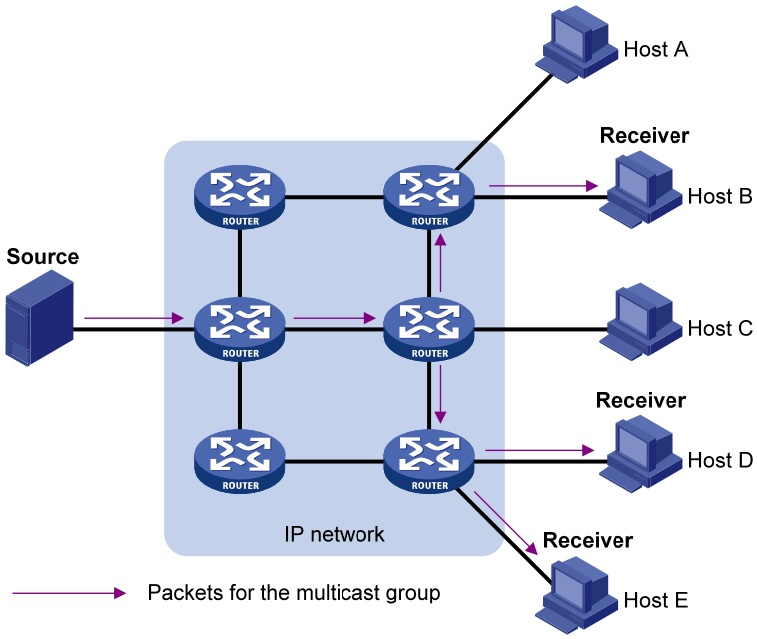
因此，广播方式不利于与特定对象进行数据交互，并且还浪费了大量的带宽。

3. 组播方式的信息传输

综上所述，传统的单播和广播的通信方式均不能以最小的网络开销实现单点发送、多点接收的问题，IP 组播技术的出现及时解决了这个问题。

如 图 1-3 所示，当IP网络中的某些主机（即Receiver）需要信息时，若采用组播的方式，组播源（即Source）仅需发送一份信息，借助组播路由协议建立组播分发树，被传递的信息在距离组播源尽可能远的网络节点才开始复制和分发。

图1-3 组播方式的信息传输



假设只有 Host B、Host D 和 Host E 需要信息，采用组播方式时，可以让这些主机加入同一个组播组（Multicast group），组播源向该组播组只需发送一份信息，并由网络中各路由器根据该组播组中各成员的分布情况对该信息进行复制和转发，最后该信息会准确地发送给 Host B、Host D 和 Host E。

综上所述，组播的优势归纳如下：

- 相比单播来说，组播的优势在于：由于被传递的信息在距信息源尽可能远的网络节点才开始被复制和分发，所以用户的增加不会导致信息源负载的加重以及网络资源消耗的显著增加。
- 相比广播来说，组播的优势在于：由于被传递的信息只会发送给需要该信息的接收者，所以不会造成网络资源的浪费，并能提高信息传输的安全性；另外，广播只能在同一网段中进行，而组播可以实现跨网段的传输。

1.1.2 组播传输的特点

组播传输的特点归纳如下：

- “组播组”是一个用 IP 组播地址进行标识的接收者集合，主机通过加入某组播组成为该组播组的成员，从而可以接收发往该组播组的组播数据。组播源通常不需要加入组播组。
- 信息的发送者称为“组播源”，如 图 1-3 中的Source。一个组播源可以同时向多个组播组发送信息，多个组播源也可以同时向一个组播组发送信息。
- 所有加入某组播组的主机便成为该组播组的成员，如 图 1-3 中的Receiver。组播组中的成员是动态的，主机可以在任何时刻加入或离开组播组。组播组成员可以广泛地分布在网络中的任何地方。
- 支持三层组播功能的路由器或三层交换机统称为“组播路由器”或“三层组播设备”。组播路由器不仅能够提供组播路由功能，也能够在与用户连接的末梢网段上提供组播组成员的管理功能。组播路由器本身也可能是组播组的成员。

为了更好地理解，可以将组播方式的信息传输过程类比于电视节目的传送过程，如 表 1-1 所示。

表1-1 组播信息传输与电视节目传输的类比

步骤	电视节目的传送过程	组播方式的信息传输过程
1	电视台S通过频道G传送电视节目	组播源S向组播组G发送组播数据
2	用户U将电视机的频道调至频道G	接收者U加入组播组G
3	用户U能够收看到由电视台S通过频道G传送的电视节目了	接收者U能够收到由组播源S发往组播组G的组播数据了
4	用户U关闭电视机或切换到其它频道	接收者U离开组播组G

1.1.3 组播中常用的表示法

在组播中，经常出现以下两种表示方式：

- （*，G）：通常用来表示共享树，或者由任意组播源发往组播组 G 的组播报文。其中的“*”代表任意组播源，“G”代表特定组播组 G。
- （S，G）：也称为“组播源组”，通常用来表示最短路径树，或者由组播源 S 发往组播组 G 的组播报文。其中的“S”代表特定组播源 S，“G”代表特定组播组 G。

说明

有关共享树和最短路径树的详细介绍，请参见“IP 组播配置指导”中的“PIM”或“IPv6 PIM”。

1.1.4 组播的优点和应用

1. 组播的优点

组播技术的优点主要在于：

- 提高效率：减轻信息源服务器和网络设备 CPU 的负荷；
- 优化性能：减少冗余流量；
- 分布式应用：使用最少的网络资源实现点到多点应用。

2. 组播的应用

组播技术主要应用于以下几个方面：

- 多媒体、流媒体的应用，如：网络电视、网络电台、实时视/音频会议；
- 培训、联合作业场合的通信，如：远程教育、远程医疗；
- 数据仓库、金融应用（股票）；
- 其它任何“点到多点”的数据发布应用。

1.2 组播模型分类

根据接收者对组播源处理方式的不同，组播模型分为以下三类：

1. ASM模型

简单地说，ASM（Any-Source Multicast，任意信源组播）模型就是任意源组播模型。

在 ASM 模型中，任意一个发送者都可以作为组播源向某个组播组地址发送信息，接收者通过加入由该地址标识的组播组，来接收发往该组播组的组播信息。

在 ASM 模型中，接收者无法预先知道组播源的位置，但可以在任意时间加入或离开该组播组。

2. SFM模型

SFM（Source-Filtered Multicast，信源过滤组播）模型继承了 ASM 模型，从发送者角度来看，两者的组播组成员关系完全相同。

SFM 模型在功能上对 ASM 模型进行了扩展。在 SFM 模型中，上层软件对收到的组播报文的源地址进行检查，允许或禁止来自某些组播源的报文通过。因此，接收者只能收到来自部分组播源的组播数据。从接收者的角度来看，只有部分组播源是有效的，组播源被经过了筛选。

3. SSM模型

在现实生活中，用户可能只对某些组播源发送的组播信息感兴趣，而不愿接收其它源发送的信息。SSM（Source-Specific Multicast，指定信源组播）模型为用户提供了一种能够在客户端指定组播源的传输服务。

SSM 模型与 ASM 模型的根本区别在于：SSM 模型中的接收者已经通过其它手段预先知道了组播源的具体位置。SSM 模型使用与 ASM/SFM 模型不同的组播地址范围，直接在接收者与其指定的组播源之间建立专用的组播转发路径。

1.3 组播框架结构

对于 IP 组播，需要关注下列问题：

- 组播源将组播信息传输到哪里？即组播寻址机制；
- 网络中有哪些接收者？即主机注册；
- 这些接收者需要从哪个组播源接收信息？即组播源发现；
- 组播信息如何传输？即组播路由。

IP 组播属于端到端的服务，组播机制包括以下四个部分：

(1) 寻址机制：借助组播地址，实现信息从组播源发送到一组接收者；

(2) 主机注册：允许接收者主机动态加入和离开某组播组，实现对组播成员的管理；

(3) 组播路由：构建组播报文分发树（即组播数据在网络中的树型转发路径），并通过该分发树将报文从组播源传输到接收者；

(4) 组播应用：组播源与接收者必须安装支持视频会议等组播应用的软件，TCP/IP 协议栈必须支持组播信息的发送和接收。

1.3.1 组播地址

为了让组播源和组播组成员进行通信，需要提供网络层组播地址，即 IP 组播地址。同时必须存在一种技术将 IP 组播地址映射为链路层的组播 MAC 地址。

1. IP组播地址

(1) IPv4 组播地址

IANA（Internet Assigned Numbers Authority，互联网编号分配委员会）将D类地址空间分配给IPv4组播使用，范围从 224.0.0.0 到 239.255.255.255，具体分类及其含义如 表 1-2 所示。

表1-2 IPv4 组播地址的范围及含义

地址范围	含义
224.0.0.0～224.0.0.255	永久组地址。除224.0.0.0保留不做分配外，其它地址供路由协议、拓扑查找和协议维护等使用，常用的永久组地址及其含义如表1-3所示。对于以该范围内组播地址为目的地址的数据包来说，不论其TTL（Time to Live，生存时间）值为多少，都不会被转发出本地网段
224.0.1.0～238.255.255.255	用户组地址，全网范围内有效。包含两种特定的组地址： • 232.0.0.0/8：SSM 组地址 • 233.0.0.0/8：GLOP 组地址
239.0.0.0～239.255.255.255	本地管理组地址，仅在本地管理域内有效。使用本地管理组地址可以灵活定义组播域的范围，以实现不同组播域之间的地址隔离，从而有助于在不同组播域内重复使用相同组播地址而不会引起冲突。详情请参见RFC 2365

说明

- 组播组中的成员是动态的，主机可以在任何时刻加入或离开组播组。
- GLOP 是一种 AS（Autonomous System，自治系统）之间的组播地址分配机制，将 AS 号填入该范围内组播地址的中间两个字节中，每个 AS 都可以得到 255 个组播地址。有关 GLOP 的详细介绍请参见 RFC 2770。

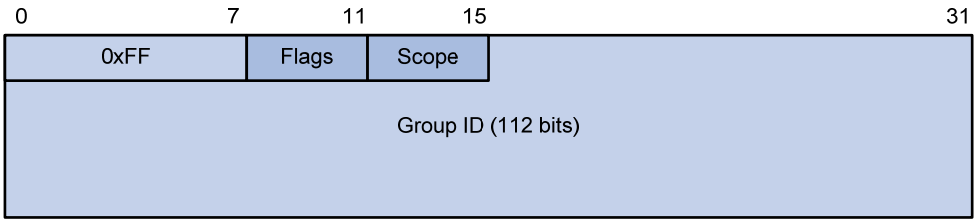
表1-3 常用永久组地址及其含义

永久组地址	含义
224.0.0.1	所有系统，包括主机与路由器
224.0.0.2	所有组播路由器
224.0.0.3	未分配
224.0.0.4	DVMRP（Distance Vector Multicast Routing Protocol，距离矢量组播路由协议）路由器
224.0.0.5	OSPF（Open Shortest Path First，开放最短路径优先）路由器
224.0.0.6	OSPF指定路由器/备用指定路由器
224.0.0.7	ST（Shared Tree，共享树）路由器
224.0.0.8	ST主机
224.0.0.9	RIP-2（Routing Information Protocol version 2，路由信息协议版本2）路由器
224.0.0.11	移动代理
224.0.0.12	DHCP（Dynamic Host Configuration Protocol，动态主机配置协议）服务器/中继代理
224.0.0.13	所有PIM（Protocol Independent Multicast，协议无关组播）路由器
224.0.0.14	RSVP（Resource Reservation Protocol，资源预留协议）封装

永久组地址	含义
224.0.0.15	所有CBT（Core-Based Tree，有核树）路由器
224.0.0.16	指定SBM（Subnetwork Bandwidth Management，子网带宽管理）
224.0.0.17	所有SBM
224.0.0.18	VRRP（Virtual Router Redundancy Protocol，虚拟路由器冗余协议）

(2) IPv6 组播地址

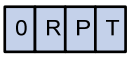
图1-4 IPv6 组播地址格式



如 图 1-4 所示，IPv6 组播地址中各字段的含义如下：

- 0xFF：最高 8 比特为 11111111，标识此地址为 IPv6 组播地址。

图1-5 Flags 字段格式



- Flags：4 比特，如 图 1-5 所示，该字段中各位的取值及含义如 表 1-4 所示。

表1-4 Flags 字段各位的取值及含义

位	取值及含义
0位	保留位，必须取0
R位	• 取 0 表示非内嵌 RP 的 IPv6 组播地址 • 取 1 则表示内嵌 RP 的 IPv6 组播地址（此时 P、T 位也必须置 1）
P位	• 取 0 表示非基于单播前缀的 IPv6 组播地址 • 取 1 则表示基于单播前缀的 IPv6 组播地址（此时 T 位也必须置 1）
T位	• 取 0 表示由 IANA 永久分配的 IPv6 组播地址 • 取 1 则表示非永久分配的 IPv6 组播地址

- Scope：4 比特，标识该IPv6 组播组的应用范围，其可能的取值及其含义如 表 1-5 所示。

表1-5 Scope 字段的取值及其含义

取值	含义
0、F	保留（Reserved）

取值	含义
1	接口本地范围（Interface-Local Scope）
2	链路本地范围（Link-Local Scope）
3	子网本地范围（Subnet-Local Scope）
4	管理本地范围（Admin-Local Scope）
5	站点本地范围（Site-Local Scope）
6、7、9～D	未分配（Unassigned）
8	机构本地范围（Organization-Local Scope）
E	全球范围（Global Scope）

- Group ID：112 比特，IPv6 组播组的标识号，用来在由 Scope 字段所指定的范围内唯一标识 IPv6 组播组。

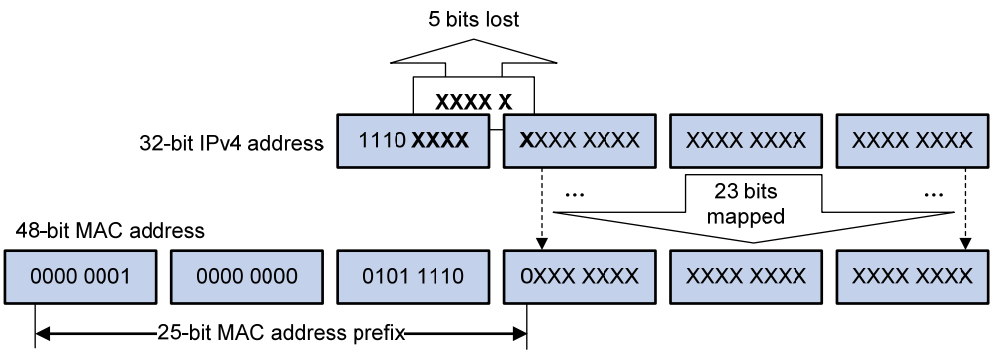
2. 以太网组播MAC地址

以太网组播 MAC 地址用于在链路层上标识属于同一组播组的接收者。

(1) IPv4 组播 MAC 地址

IANA规定，IPv4 组播MAC地址的高 24 位为 0x01005E，第 25 位为 0，低 23 位为IPv4 组播地址的低 23 位。IPv4 组播地址与MAC地址的映射关系如 图 1-6 所示。

图1-6 IPv4 组播地址与 MAC 地址的映射关系

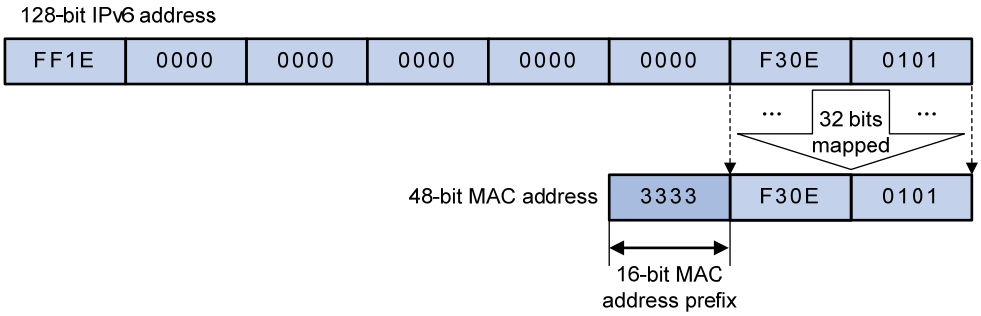


由于 IPv4 组播地址的高 4 位是 1110，代表组播标识，而低 28 位中只有 23 位被映射到 IPv4 组播 MAC 地址，这样 IPv4 组播地址中就有 5 位信息丢失。于是，就有 32 个 IPv4 组播地址映射到了同一个 IPv4 组播 MAC 地址上，因此在二层处理过程中，设备可能要接收一些本 IPv4 组播组以外的组播数据，而这些多余的组播数据就需要设备的上层进行过滤了。

(2) IPv6 组播 MAC 地址

IPv6 组播MAC地址的高 16 位为 0x3333，低 32 位为IPv6 组播地址的低 32 位。如 图 1-7 所示，是 IPv6 组播地址FF1E::F30E:101 的MAC地址映射举例。

图1-7 IPv6 组播地址的 MAC 地址映射举例



1.3.2 组播协议

说明

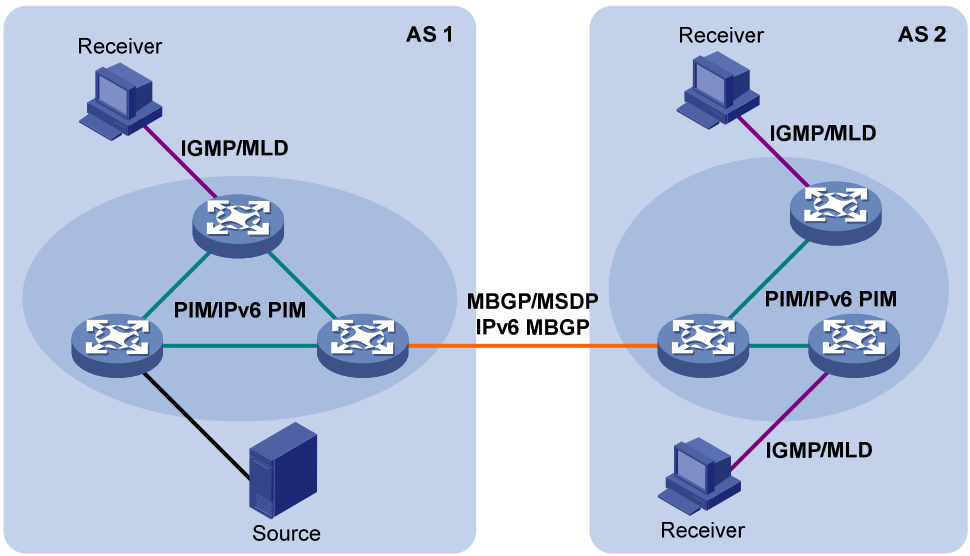
- 通常，我们把工作在网络层的 IP 组播称为“三层组播”，相应的组播协议称为“三层组播协议”，包括 IGMP/MLD、PIM/IPv6 PIM、MSDP、MBGP/IPv6 MBGP 等；把工作在数据链路层的 IP 组播称为“二层组播”，相应的组播协议称为“二层组播协议”，包括 IGMP Snooping/MLD Snooping、PIM Snooping/IPv6 PIM Snooping、组播 VLAN/IPv6 组播 VLAN 等。
- IGMP Snooping、PIM Snooping、组播 VLAN、IGMP、PIM、MSDP 和 MBGP 应用于 IPv4；MLD Snooping、IPv6 PIM Snooping、IPv6 组播 VLAN、MLD、IPv6 PIM 和 IPv6 MBGP 应用于 IPv6。

本节主要针对二、三层组播协议在网络中的应用位置和功能进行总体介绍，有关各协议的详细介绍请分别参见“IP 组播配置指导”中的相应内容。

1. 三层组播协议

三层组播协议包括组播组管理协议和组播路由协议两种类型，它们在网络中的应用位置如 图 1-8 所示。

图1-8 三层组播协议的应用位置



(1) 组播组管理协议

在主机和与其直接相连的三层组播设备之间通常采用组播组的管理协议 IGMP（Internet Group Management Protocol，互联网组管理协议）或 MLD（Multicast Listener Discovery Protocol，组播侦听者发现协议），该协议规定了主机与三层组播设备之间建立和维护组播组成员关系的机制。

(2) 组播路由协议

组播路由协议运行在三层组播设备之间，用于建立和维护组播路由，并正确、高效地转发组播数据包。组播路由建立了从一个数据源端到多个接收端的无环（loop-free）数据传输路径，即组播分发树。

对于 ASM 模型，可以将组播路由分为域内和域间两大类：

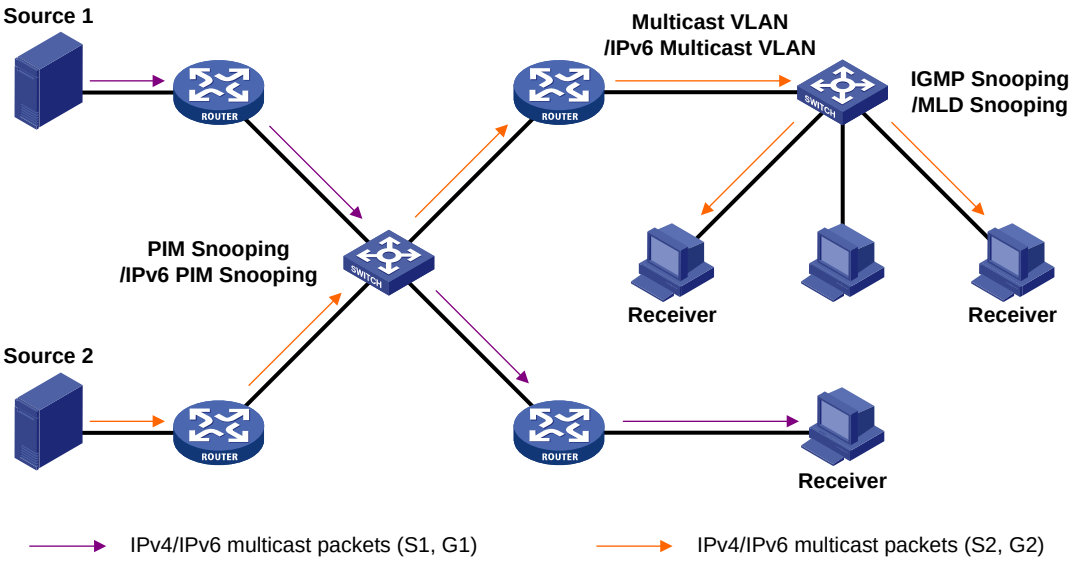
- 域内组播路由用来在 AS 内部发现组播源并构建组播分发树，从而将组播信息传递到接收者。在众多域内组播路由协议中，PIM（Protocol Independent Multicast，协议无关组播）是目前较为典型的一个。按照转发机制的不同，PIM 可以分为 DM（Dense Mode，密集模式）和 SM（Sparse Mode，稀疏模式）两种模式。
- 域间组播路由用来实现组播信息在 AS 之间的传递，目前比较成型的解决方案有：MSDP（Multicast Source Discovery Protocol，组播源发现协议）能够跨越 AS 传播组播源的信息；而 MP-BGP（Multiprotocol Border Gateway Protocol，多协议边界网关协议）的组播扩展 MBGP（Multicast BGP）则能够跨越 AS 传播组播路由。

对于 SSM 模型，没有域内和域间的划分。由于接收者预先知道组播源的具体位置，因此只需要借助 PIM-SM 构建的通道即可实现组播信息的传输。

2. 二层组播协议

二层组播协议包括IGMP Snooping/MLD Snooping、PIM Snooping/IPv6 PIM Snooping、组播VLAN/IPv6 组播VLAN等，它们在网络中的应用位置如 图 1-9 所示。

图1-9 二层组播协议的应用位置



(1) IGMP Snooping/MLD Snooping

IGMP Snooping（Internet Group Management Protocol Snooping，互联网组管理协议窥探）和MLD Snooping（Multicast Listener Discovery Snooping，组播侦听者发现协议窥探）是运行在二层设备上的组播约束机制，通过窥探和分析主机与三层组播设备之间交互的 IGMP 或 MLD 报文来管理和控制组播组，从而可以有效抑制组播数据在二层网络中的扩散。

(2) PIM Snooping/IPv6 PIM Snooping

运行 PIM Snooping（Protocol Independent Multicast Snooping，协议无关组播窥探）或 IPv6 PIM Snooping 的二层设备通过对收到的 PIM 协议报文进行分析，将有接收需求的端口添加到相应的组播转发表项中，以实现组播报文的精确转发。

(3) 组播 VLAN/IPv6 组播 VLAN

在传统的组播点播方式下，当连接在二层设备上、属于不同 VLAN 的用户分别进行组播点播时，三层组播设备需要向该二层设备的每个 VLAN 分别发送一份组播数据；而当二层设备运行了组播 VLAN 或 IPv6 组播 VLAN 之后，三层组播设备只需向该二层设备的组播 VLAN 或 IPv6 组播 VLAN 发送一份组播数据即可，从而既避免了带宽的浪费，也减轻了三层组播设备的负担。

1.4 组播报文的转发机制

在组播模型中，IP 报文的目的地址字段为组播组地址，组播源向以此目的地址所标识的主机群组传送信息。因此，转发路径上的组播路由器为了将组播报文传送到各个方位的接收站点，往往需要将从一个入接口收到的组播报文转发到多个出接口。与单播模型相比，组播模型的复杂性就在于此：

- 为了保证组播报文在网络中的传输，必须依靠单播路由表或者单独提供给组播使用的路由表（如 MBGP 路由表）来指导转发；
- 为了处理同一设备在不同接口上收到来自不同对端的相同组播信息，需要对组播报文的入接口进行 RPF（Reverse Path Forwarding，逆向路径转发）检查，以决定转发还是丢弃该报文。RPF 检查机制是大部分组播路由协议进行组播转发的基础。

说明

有关 RPF 机制的详细介绍，请参见“IP 组播配置指导”中的“组播路由与转发”或“IPv6 组播路由与转发”。

1.5 多实例组播

说明

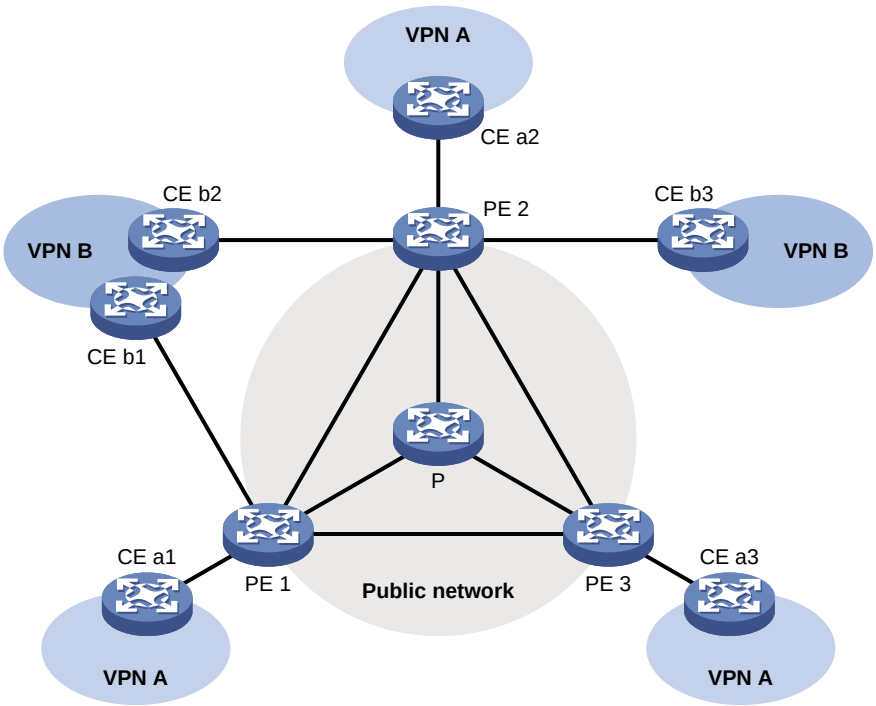
目前，IPv6 不支持多实例组播。

简单地说，多实例组播就是指在 VPN（Virtual Private Network，虚拟专用网络）中应用组播。

1.5.1 多实例简介

各VPN网络之间、VPN网络与公共网络之间要求信息隔离。如 图 1-10 所示，VPN A和VPN B通过PE设备接入公共网络。

图1-10 典型的 VPN 组网



- P 设备专属于公网，各 CE 设备则专属于某一 VPN。每台设备只为其专属的网络服务，仅维护一套转发机制。
- PE 设备同时接入公网和 VPN 网络，同时为多个网络服务。在设备上必须严格区分各个网络的信息并为各个网络独立维护一套转发机制。这时，PE 设备上为同一网络服务的一套软硬件

设施统称为一个实例（Instance）。PE 设备上同时存在多个实例，而同一个实例也可以分布在多台 PE 设备上。在 PE 设备上，我们将服务于公网的实例称为公网实例，服务于 VPN 的实例则称为 VPN 实例。

1.5.2 多实例在组播中的应用

在 PE 设备上应用多实例组播之后，将具备以下功能：

- 每个实例都独立维护一套组播转发机制：支持各种组播协议，拥有各自独立的 PIM 邻居列表、组播路由表等信息。每个实例转发组播数据时只查找本实例的转发表或路由表；
- 保证各实例之间相互隔离；
- 实现公网实例和 VPN 实例之间的信息交流和数据转换。

组播VPN是基于MPLS L3VPN网络来实现组播传输的技术。以 图 1-10 中的VPN实例A为例，当VPN A中的组播源向某组播组发送组播数据时，在网络中所有可能的接收者中，只有属于VPN A的组播组成员才能收到该组播源发来的组播数据。组播数据在VPN A中以组播方式进行传输，在公网中也以组播方式进行传输。

目 录

1 IGMP Snooping配置

1.1 IGMP Snooping简介

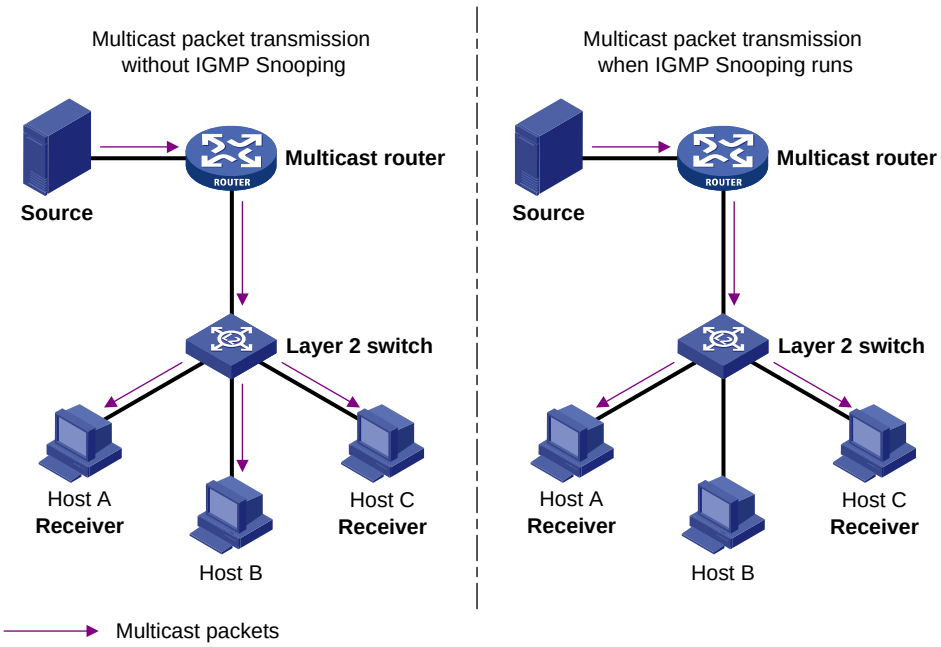
IGMP Snooping 是 Internet Group Management Protocol Snooping（互联网组管理协议窥探）的简称，它是运行在二层设备上的组播约束机制，用于管理和控制组播组。

1.1.1 IGMP Snooping原理

运行 IGMP Snooping 的二层设备通过对收到的 IGMP 报文进行分析，为端口和 MAC 组播地址建立起映射关系，并根据这样的映射关系转发组播数据。

如 图 1-1 所示，当二层设备没有运行IGMP Snooping时，组播数据在二层被广播；当二层设备运行了IGMP Snooping后，已知组播组的组播数据不会在二层被广播，而在二层被组播给指定的接收者。

图1-1 二层设备运行 IGMP Snooping 前后的对比



IGMP Snooping 通过二层组播将信息只转发给有需要的接收者，可以带来以下好处：

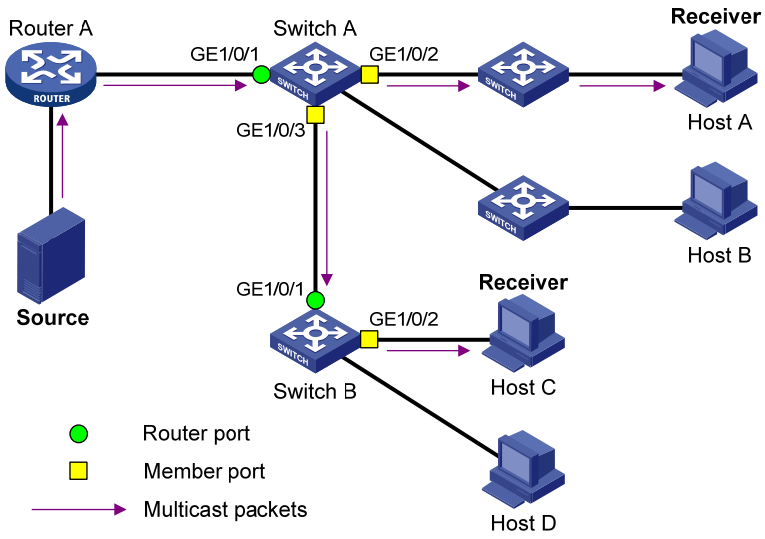
- 减少了二层网络中的广播报文，节约了网络带宽；
- 增强了组播信息的安全性；
- 为实现对每台主机的单独计费带来了方便。

1.1.2 IGMP Snooping基本概念

1. IGMP Snooping相关端口

如 图 1-2 所示，Router A连接组播源，在Switch A和Switch B上分别运行IGMP Snooping，Host A和Host C为接收者主机（即组播组成员）。

图1-2 IGMP Snooping 相关端口



结合 图 1-2，介绍一下IGMP Snooping相关的端口概念：

- 路由器端口（Router Port）：交换机上朝向三层组播设备（DR 或 IGMP 查询器）一侧的端口，如 Switch A 和 Switch B 各自的 GigabitEthernet1/0/1 端口。交换机将本设备上的所有路由器端口都记录在路由器端口列表中。
- 成员端口（Member Port）：又称组播组成员端口，表示交换机上朝向组播组成员一侧的端口，如 Switch A 的 GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 端口，以及 Switch B 的 GigabitEthernet1/0/2 端口。交换机将本设备上的所有成员端口都记录在 IGMP Snooping 转发表中。

说明

- 本文中提到的路由器端口都是指交换机上朝向组播路由器的端口，而不是指路由器上的端口。
- 如不特别指明，本文中提到的路由器/成员端口均包括动态和静态端口。
- 在运行了 IGMP Snooping 的交换机上，所有收到源地址不为 0.0.0.0 的 IGMP 普遍组查询报文或 PIM Hello 报文的端口都将被视为动态路由器端口。有关 PIM Hello 报文的详细介绍，请参见“IP 组播配置指导”中的“PIM”。

---

2. IGMP Snooping动态端口老化定时器

表1-1 IGMP Snooping 动态端口老化定时器

定时器	说明	超时前应收到的报文	超时后交换机的动作
动态路由器端口老化定时器	交换机为其每个动态路由器端口都启动一个定时器，其超时时间就是动态路由器端口老化时间	源地址不为0.0.0.0的IGMP普遍组查询报文或PIM Hello报文	将该端口从路由器端口列表中删除
动态成员端口老化定时器	当一个端口动态加入某组播组时，交换机为该端口启动一个定时器，其超时时间就是动态成员端口老化时间	IGMP成员关系报告报文	将该端口从IGMP Snooping转发表中删除

说明

IGMP Snooping 端口老化机制只针对动态端口，静态端口永不老化。

1.1.3 IGMP Snooping工作机制

运行了 IGMP Snooping 的交换机对不同 IGMP 动作的具体处理方式如下：

注意

本节中所描述的增删端口动作均只针对动态端口，静态端口只能通过相应的配置进行增删，具体步骤请参见“1.4.3 配置静态端口”。

1. 普遍组查询

IGMP 查询器定期向本地网段内的所有主机与路由器（224.0.0.1）发送 IGMP 普遍组查询报文，以查询该网段有哪些组播组的成员。

在收到 IGMP 普遍组查询报文时，交换机将其通过 VLAN 内除接收端口以外的其它所有端口转发出去，并对该报文的接收端口做如下处理：

- 如果在路由器端口列表中已包含该动态路由器端口，则重置其老化定时器。
- 如果在路由器端口列表中尚未包含该动态路由器端口，则将其添加到路由器端口列表中，并启动其老化定时器。

2. 报告成员关系

以下情况，主机会向 IGMP 查询器发送 IGMP 成员关系报告报文：

- 当组播组的成员主机收到 IGMP 查询报文后，会回复 IGMP 成员关系报告报文。
- 如果主机要加入某个组播组，它会主动向 IGMP 查询器发送 IGMP 成员关系报告报文以声明加入该组播组。

在收到 IGMP 成员关系报告报文时，交换机将其通过 VLAN 内的所有路由器端口转发出去，从该报文中解析出主机要加入的组播组地址，并对该报文的接收端口做如下处理：

- 如果不存在该组播组所对应的转发表项，则创建转发表项，将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；

- 如果已存在该组播组所对应的转发表项，但其出端口列表中不包含该端口，则将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；
- 如果已存在该组播组所对应的转发表项，且其出端口列表中已包含该动态成员端口，则重置其老化定时器。

说明

交换机不会将 IGMP 成员关系报告报文通过非路由器端口转发出去，因为根据主机上的 IGMP 成员关系报告抑制机制，如果非路由器端口下还有该组播组的成员主机，则这些主机在收到该报告报文后便抑制了自身的报告，从而使交换机无法获知这些端口下还有该组播组的成员主机。有关主机上的 IGMP 成员关系报告抑制机制的详细介绍，请参见“IP 组播配置指导”中的“IGMP”。

---

3. 离开组播组

运行 IGMPv1 的主机离开组播组时不会发送 IGMP 离开组报文，因此交换机无法立即获知主机离开的信息。但是，由于主机离开组播组后不会再发送 IGMP 成员关系报告报文，因此当其对应的动态成员端口的老化定时器超时后，交换机就会将该端口对应的转发表项从转发表中删除。

运行 IGMPv2 或 IGMPv3 的主机离开组播组时，会通过发送 IGMP 离开组报文，以通知组播路由器自己离开了某个组播组。当交换机从某动态成员端口上收到 IGMP 离开组报文时，首先判断要离开的组播组所对应的转发表项是否存在，以及该组播组所对应转发表项的出端口列表中是否包含该接收端口：

- 如果不存在该组播组对应的转发表项，或者该组播组对应转发表项的出端口列表中不包含该端口，交换机不会向任何端口转发该报文，而将其直接丢弃；
- 如果存在该组播组对应的转发表项，且该组播组对应转发表项的出端口列表中包含该端口，交换机会将该报文通过 VLAN 内的所有路由器端口转发出去。同时，由于并不知道该接收端口下是否还有该组播组的其它成员，所以交换机不会立刻把该端口从该组播组所对应转发表项的出端口列表中删除，而是重置其老化定时器。

当 IGMP 查询器收到 IGMP 离开组报文后，从中解析出主机要离开的组播组的地址，并通过接收端口向该组播组发送 IGMP 特定组查询报文。交换机在收到 IGMP 特定组查询报文后，将其通过 VLAN 内的所有路由器端口和该组播组的所有成员端口转发出去。对于 IGMP 离开组报文的接收端口（假定为动态成员端口），交换机在其老化时间内：

- 如果从该端口收到了主机响应该特定组查询的 IGMP 成员关系报告报文，则表示该端口下还有该组播组的成员，于是重置其老化定时器；
- 如果没有从该端口收到主机响应特定组查询的 IGMP 成员关系报告报文，则表示该端口下已没有该组播组的成员，则在其老化时间超时后，将其从该组播组所对应转发表项的出端口列表中删除。

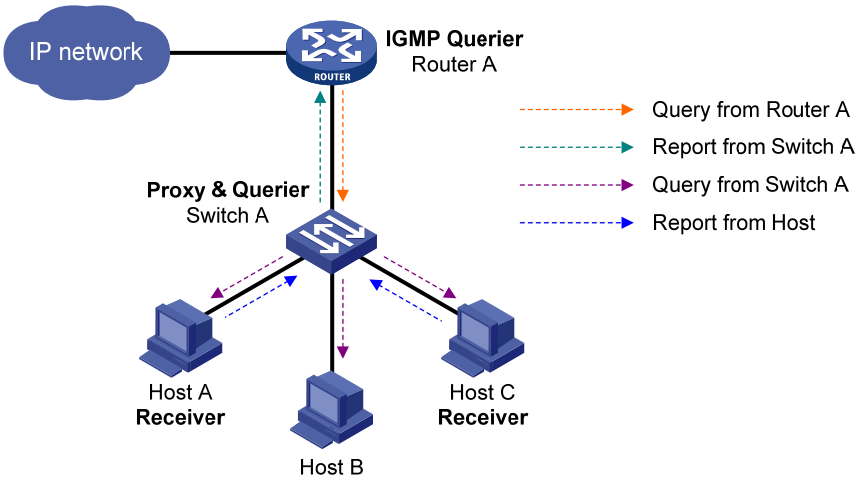
1.1.4 IGMP Snooping Proxying

为了减少上游设备收到的 IGMP 报告报文和离开报文的数量，可以通过在边缘设备上配置 IGMP Snooping Proxying（IGMP Snooping 代理）功能，使其能够代理下游主机来向上游设备发送报告报文和离开报文。配置了 IGMP Snooping Proxying 功能的设备称为 IGMP Snooping 代理设备，在其上游设备看来，它就相当于一台主机。

说明

尽管在其上游设备看来，IGMP Snooping 代理设备相当于一台主机，但主机上的 IGMP 成员关系报告抑制机制在 IGMP Snooping 代理设备上并不会生效。有关主机上的 IGMP 成员关系报告抑制机制的详细介绍，请参见“IP 组播配置指导”中的“IGMP”。

图1-3 IGMP Snooping Proxying 组网图



如 图 1-3 所示，作为IGMP Snooping代理设备的Switch A，对其上游的IGMP查询器Router A来说相当于一台主机，代理下游主机向Router A发送报告报文和离开报文。

IGMP Snooping代理设备对IGMP报文的处理方式如 表 1-2 所示。

表1-2 IGMP Snooping 代理设备对 IGMP 报文的处理方式

IGMP 报文类型	处理方式
普遍组查询报文	收到普遍组查询报文后，向本VLAN内除接收端口以外的所有端口转发；同时根据本地维护的组成员关系生成报告报文，并向所有路由器端口发送
特定组查询报文	收到针对某组播组的特定组查询报文时，若该组对应的转发表项中还有成员端口，则向所有路由器端口回复该组的报告报文
报告报文	从某端口收到某组播组的报告报文时，若已存在该组对应的转发表项，且其出端口列表中已包含该动态成员端口，则重置其老化定时器；若已存在该组对应的转发表项，但其出端口列表中不包含该端口，则将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；若尚不存在该组对应的转发表项，则创建转发表项，将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器，然后向所有路由器端口发送该组的报告报文
离开报文	从某端口收到某组播组的离开报文后，向该端口发送针对该组的特定组查询报文。只有当删除某组播组对应转发表项中的最后一个成员端口时，才会向所有路由器端口发送该组的离开报文

1.1.5 协议规范

与 IGMP Snooping 相关的协议规范有：

- RFC 4541：Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches

1.2 IGMP Snooping配置任务简介

表1-3 IGMP Snooping 配置任务简介

配置任务		说明	详细配置
配置IGMP Snooping基本功能	使能IGMP Snooping	必选	1.3.2
	配置IGMP Snooping版本	可选	1.3.3
	配置IGMP Snooping转发表项的最大数量	可选	1.3.4
	配置静态组播MAC地址表项	可选	1.3.5
配置IGMP Snooping端口功能	配置动态端口老化定时器	可选	1.4.2
	配置静态端口	可选	1.4.3
	配置模拟主机加入	可选	1.4.4
	配置端口快速离开	可选	1.4.5
	禁止端口成为动态路由器端口	可选	1.4.6
配置IGMP Snooping查询器	使能IGMP Snooping查询器	可选	1.5.2
	配置IGMP查询和响应	可选	1.5.3
	配置IGMP查询报文源IP地址	可选	1.5.4
配置IGMP Snooping Proxying	使能IGMP Snooping Proxying	可选	1.6.2
	配置代理发送IGMP报文的源IP地址	可选	1.6.3
配置IGMP Snooping策略	配置组播组过滤器	可选	1.7.2
	配置组播数据报文源端口过滤	可选	1.7.3
	配置丢弃未知组播数据报文	可选	1.7.4
	配置IGMP成员关系报告报文抑制	可选	1.7.5
	配置端口加入的组播组最大数量	可选	1.7.6
	配置组播组替换	可选	1.7.7
	配置IGMP报文的802.1p优先级	可选	1.7.8
	配置组播用户控制策略	可选	1.7.9
	配置IGMP Snooping主机跟踪功能	可选	1.7.10
	配置IGMP Snooping协议中发送IGMP报文的DSCP优先级	可选	1.7.11

说明

对于 IGMP Snooping 的相关配置来说:

- IGMP-Snooping 视图下的配置对所有 VLAN 都有效,VLAN 视图下的配置只对当前 VLAN 有效。对于某 VLAN 来说，优先采用该 VLAN 视图下的配置，只有当在该 VLAN 视图下没有进行配置时，才采用 IGMP-Snooping 视图下的相应配置。
- IGMP-Snooping 视图下的配置对所有端口都有效；二层以太网端口视图下的配置只对当前端口有效；二层聚合接口视图下的配置只对当前接口有效；端口组视图下的配置对当前端口组中的所有端口有效。对于某端口来说，优先采用二层以太网端口视图、二层聚合接口视图或端口组视图下的配置，只有当在上述视图下没有进行配置时，才采用 IGMP-Snooping 视图下的相应配置。
- 二层聚合接口与其各成员端口上的配置互不影响,且成员端口上的配置只有当该端口退出聚合组后才会生效，二层聚合接口上的配置也不会参与聚合计算。

1.3 配置IGMP Snooping基本功能

1.3.1 配置准备

在配置 IGMP Snooping 基本功能之前，需完成以下任务：

- 配置相应 VLAN

在配置 IGMP Snooping 基本功能之前，需准备以下数据：

- IGMP Snooping 的版本

1.3.2 使能IGMP Snooping

表1-4 使能 IGMP Snooping

操作	命令	说明
进入系统视图	**system-view**	-
全局使能IGMP Snooping，并进入IGMP-Snooping视图	**igmp-snooping**	必选 缺省情况下，IGMP Snooping处于关闭状态
退回系统视图	**quit**	-
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能IGMP Snooping	**igmp-snooping enable**	必选 缺省情况下，VLAN内的IGMP Snooping处于关闭状态

说明

- 在 VLAN 内使能 IGMP Snooping 之前，必须先在系统视图下全局使能 IGMP Snooping，否则将无法在 VLAN 内使能 IGMP Snooping。
- 在 VLAN 内使能 IGMP Snooping 之后，不允许在该 VLAN 所对应的 VLAN 接口上再使能 IGMP 和 PIM，反之亦然。
- 在指定 VLAN 内使能了 IGMP Snooping 之后，IGMP Snooping 功能只在属于该 VLAN 的端口上生效。

1.3.3 配置IGMP Snooping版本

配置 IGMP Snooping 的版本，实际上就是配置 IGMP Snooping 可以处理的 IGMP 报文的版本：

- 当 IGMP Snooping 的版本为 2 时，IGMP Snooping 能够对 IGMPv1 和 IGMPv2 的报文进行处理，对 IGMPv3 的报文则不进行处理，而是在 VLAN 内将其广播；
- 当 IGMP Snooping 的版本为 3 时，IGMP Snooping 能够对 IGMPv1、IGMPv2 和 IGMPv3 的报文进行处理。

表1-5 配置 IGMP Snooping 版本

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置IGMP Snooping的版本	**igmp-snooping version** *version-number*	必选 缺省情况下，IGMP Snooping的版本为2

注意

当 IGMP Snooping 的版本由版本 3 切换到版本 2 时，系统将清除所有通过动态加入的 IGMP Snooping 转发表项；对于在版本 3 下通过手工配置而静态加入的 IGMP Snooping 转发表项，则分为以下两种情况进行不同的处理：

- 如果配置的仅仅是静态加入组播组，而没有指定组播源，则这些转发表项将不会被清除；
- 如果配置的是指定了组播源的静态加入组播源组，则这些转发表项将会被清除，并且当再次切换回版本 3 时，这些转发表项将被重新恢复。

有关静态加入的详细配置，请参见“1.4.3 配置静态端口”。

1.3.4 配置IGMP Snooping转发表项的最大数量

用户可以调整 IGMP Snooping 转发表项的最大数量，当设备上维护的表项数量达到最大数量后，将不再创建新的表项，直至有表项被老化或被手工删除。

表1-6 配置全局 IGMP Snooping 转发表项的最大数量

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
配置IGMP Snooping转发表项的最大数量	**entry-limit** *limit*	必选 缺省情况下， IGMP Snooping转发表项的最大数量为2000

说明

- 在对 IGMP Snooping 转发表项的最大数量进行配置时，如果设备上维护的表项数量已超过了配置值，系统将提示用户删除多余表项，但并不会自动删除任何已存在的表项，同时也不再继续创建新的表项。
- 组播 VLAN 建立的转发表项不受 **entry-limit** 命令的限制，可通过 **multicast-vlan entry-limit** 命令进行限制，详情请参见“IP 组播配置指导”中的“组播 VLAN 配置”。

1.3.5 配置静态组播MAC地址表项

在二层组播中，除了可通过二层组播协议（如 IGMP Snooping）动态建立组播 MAC 地址表项外，还可以通过手工方式配置组播 MAC 地址表项，将端口与组播 MAC 地址进行静态绑定，以便灵活控制组播信息送达的目的端口。

1. 系统视图下配置静态组播MAC地址表项

表1-7 系统视图下配置静态组播 MAC 地址表项

操作	命令	说明
进入系统视图	**system-view**	-
配置静态组播MAC地址表项	**mac-address multicast** *mac-address* **interface** *interface-list* **vlan** *vlan-id*	必选 缺省情况下，没有配置静态组播MAC地址表项

2. 接口视图下配置静态组播MAC地址表项

表1-8 接口视图下配置静态组播 MAC 地址表项

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一 进入以太网端口或二层聚合接口

操作		命令	说明
	进入端口组视图	**port-group manual** *port-group-name*	视图后，下面进行的配置只在当前接口生效；进入端口组视图后，下面进行的配置将在端口组的所有接口生效
配置静态组播MAC地址表项		**mac-address multicast** *mac-address* **vlan** *vlan-id*	必选 缺省情况下，没有配置静态组播MAC地址表项

说明

- 系统视图下的配置对指定接口有效，而接口视图下的配置只对当前接口（或当前端口组内的所有接口）有效。
- 对于 R5206 之前的版本，可手工配置的组播 MAC 地址表项为除 0100-5Exx-xxxx 以外的任意合法的组播 MAC 地址（组播 MAC 地址就是最高字节的最低比特位为 1 的 MAC 地址），其中 x 代表 0～F 的任意一个十六进制数。对于 R5206 及以上版本，可手工配置的组播 MAC 地址表项为任意合法的组播 MAC 地址。

1.4 配置IGMP Snooping端口功能

1.4.1 配置准备

在配置 IGMP Snooping 端口功能之前，需完成以下任务：

- 在 VLAN 内使能 IGMP Snooping
- 配置相应端口组

在配置 IGMP Snooping 端口功能之前，需准备以下数据：

- 动态路由器端口老化时间
- 动态成员端口老化时间
- 组播组和组播源的地址

1.4.2 配置动态端口老化定时器

对于动态路由器端口，如果在其老化时间超时前没有收到 IGMP 普遍组查询报文或者 PIM Hello 报文，交换机将把该端口从路由器端口列表中删除。

对于动态成员端口，如果在其老化时间超时前没有收到该组播组的 IGMP 成员关系报告报文，交换机将把该端口从该组播组所对应转发表项的出端口列表中删除。

如果组播组成员的变动比较频繁，可以把动态成员端口老化时间设置小一些，反之亦然。

1. 全局配置动态端口老化定时器

表1-9 全局配置动态端口老化定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
配置动态路由器端口老化时间	**router-aging-time** *interval*	必选 缺省情况下，动态路由器端口的老化时间为105秒
配置动态成员端口老化时间	**host-aging-time** *interval*	必选 缺省情况下，动态成员端口的老化时间为260秒

2. 在VLAN内配置动态端口老化定时器

表1-10 在 VLAN 内配置动态端口老化定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置动态路由器端口老化时间	**igmp-snooping router-aging-time** *interval*	必选 缺省情况下，动态路由器端口的老化时间为105秒
配置动态成员端口老化时间	**igmp-snooping host-aging-time** *interval*	必选 缺省情况下，动态成员端口的老化时间为260秒

1.4.3 配置静态端口

如果某端口所连接的主机需要固定接收发往某组播组或组播源组的组播数据，可以配置该端口静态加入该组播组或组播源组，成为静态成员端口。

可以通过将交换机上的端口配置为静态路由器端口，从而使交换机上所有收到的组播数据可以通过该端口被转发出去。

表1-11 配置静态端口

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	

操作	命令	说明
配置静态成员端口	**igmp-snooping static-group** *group-address* [**source-ip** *source-address*] **vlan** *vlan-id*	必选 缺省情况下，端口不是静态成员端口
配置静态路由器端口	**igmp-snooping static-router-port vlan** *vlan-id*	必选 缺省情况下，端口不是静态路由器端口

- 静态成员端口不会对 IGMP 查询器发出的查询报文进行响应；当配置静态成员端口或取消静态成员端口的配置时，端口也不会主动发送 IGMP 成员关系报告报文或 IGMP 离开组报文。
- 静态成员端口和静态路由器端口都不会老化，只能通过相应的 **undo** 命令删除。

1.4.4 配置模拟主机加入

通常情况下，运行 IGMP 的主机会对 IGMP 查询器发出的查询报文进行响应。如果主机由于某种原因无法响应，就可能导致组播路由器认为该网段没有该组播组的成员，从而取消相应的转发路径。为避免这种情况的发生，可以将交换机的端口配置成为组播组成员（即配置模拟主机加入）。当收到 IGMP 查询报文时由模拟主机进行响应，从而保证该交换机能够继续收到组播报文。

模拟主机加入功能的实现原理如下：

- 在某端口上使能模拟主机加入功能时，交换机会通过该端口主动发送一个 IGMP 成员关系报告报文；
- 在某端口上使能了模拟主机加入功能后，当收到 IGMP 普遍组查询报文时，交换机会通过该端口响应一个 IGMP 成员关系报告报文；
- 在某端口上关闭模拟主机加入功能时，交换机会通过该端口发送一个 IGMP 离开组报文。

表1-12 配置模拟主机加入

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置模拟主机加入组播组或组播源组		**igmp-snooping host-join** *group-address* [**source-ip** *source-address*] **vlan** *vlan-id*	必选 缺省情况下，没有配置模拟主机加入组播组或组播源组

说明

- 每配置一次模拟主机加入，即相当于启动了一台独立的主机。例如，当收到 IGMP 查询报文时，每条配置所对应的模拟主机将分别进行响应。
- 与静态成员端口不同，配置了模拟主机加入的端口会作为动态成员端口而参与动态成员端口的老化过程。

1.4.5 配置端口快速离开

端口快速离开是指当交换机从某端口收到主机发送的离开某组播组的 IGMP 离开组报文时，直接把该端口从对应转发表项的出端口列表中删除。此后，当交换机收到对该组播组的 IGMP 特定组查询报文时，交换机将不再向该端口转发。

在交换机上，在只连接有一个接收者的端口上，可以通过使能端口快速离开功能来节约带宽和资源；而在连接有多个接收者的端口上，如果交换机或该端口所在的 VLAN 已使能了丢弃未知组播数据报文功能，则不要再使能端口快速离开功能，否则，一个接收者的离开将导致该端口下属于同一组播组的其它接收者无法收到组播数据。

1. 全局配置端口快速离开

表1-13 全局配置端口快速离开

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
使能端口快速离开功能	**fast-leave** [**vlan** *vlan-list*]	必选 缺省情况下，端口快速离开功能处于关闭状态

2. 在端口上配置端口快速离开

表1-14 在端口上配置端口快速离开

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
使能端口快速离开功能		**igmp-snooping fast-leave** [**vlan** *vlan-list*]	必选 缺省情况下，端口快速离开功能处于关闭状态

1.4.6 禁止端口成为动态路由器端口

目前，在组播用户接入网络中存在以下问题：

- 如果交换机收到了某用户主机发来的 IGMP 普遍组查询报文或 PIM Hello 报文，那么该主机所在的端口就将成为动态路由器端口，从而使 VLAN 内的所有组播报文都会向该端口转发，导致该用户主机收到的组播报文失控。
- 同时，用户主机发送 IGMP 普遍组查询报文或 PIM Hello 报文，也会影响该接入网络中三层设备上的组播路由协议状态（如影响 IGMP 查询器或 DR 的选举），严重时可能导致网络中断。

当禁止某端口成为动态路由器端口后，即使该端口收到了 IGMP 普遍组查询报文或 PIM Hello 报文，该端口也不会成为动态路由器端口，从而能够有效解决上述问题，提高网络的安全性和对组播用户的控制能力。

表1-15 禁止端口成为动态路由器端口

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
禁止端口成为动态路由器端口		**igmp-snooping router-port-deny** [**vlan** *vlan-list*]	必选 缺省情况下，不禁止端口成为动态路由器端口

说明

本配置与静态路由器端口的配置互不影响。

1.5 配置IGMP Snooping查询器

1.5.1 配置准备

在配置 IGMP Snooping 查询器之前，需完成以下任务：

- 在 VLAN 内使能 IGMP Snooping

在配置 IGMP Snooping 查询器之前，需准备以下数据：

- 发送 IGMP 普遍组查询报文的时间间隔
- 发送 IGMP 特定组查询报文的时间间隔
- IGMP 普遍组查询的最大响应时间
- IGMP 普遍组查询报文的源 IP 地址
- IGMP 特定组查询报文的源 IP 地址

1.5.2 使能IGMP Snooping查询器

在运行了 IGMP 的组播网络中，会有一台三层组播设备充当 IGMP 查询器，负责发送 IGMP 查询报文，使三层组播设备能够在网络层建立并维护组播转发表项，从而在网络层正常转发组播数据。

但是，在一个没有三层组播设备的网络中，由于二层设备并不支持 IGMP，因此无法实现 IGMP 查询器的相关功能。为了解决这个问题，可以在二层设备上使能 IGMP Snooping 查询器，使二层设备能够在数据链路层建立并维护组播转发表项，从而在数据链路层正常转发组播数据。

表1-16 使能 IGMP Snooping 查询器

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
使能IGMP Snooping查询器	**igmp-snooping querier**	必选 缺省情况下，IGMP Snooping查询器处于关闭状态

注意

尽管 IGMP Snooping 查询器并不参与 IGMP 查询器的选举，但在运行了 IGMP 的组播网络中，配置 IGMP Snooping 查询器不但没有实际的意义，反而可能会由于其发送的 IGMP 普遍组查询报文的源 IP 地址较小而影响 IGMP 查询器的选举。

有关 IGMP 查询器的详细介绍，请参见“IP 组播配置指导”中的“IGMP”。

1.5.3 配置IGMP查询和响应

可以根据网络的实际情况来修改发送 IGMP 普遍组查询报文的时间间隔。

在收到 IGMP 查询报文（包括普遍组查询和特定组查询）后，主机会为其所加入的每个组播组都启动一个定时器，定时器的值在 0 到最大响应时间（该时间值由主机从所收到的 IGMP 查询报文的最大响应时间字段获得）中随机选定，当定时器的值减为 0 时，主机就会向该定时器对应的组播组发送 IGMP 成员关系报告报文。

合理配置 IGMP 查询的最大响应时间，既可以使主机对 IGMP 查询报文做出快速响应，又可以减少由于定时器同时超时，造成大量主机同时发送报告报文而引起的网络拥塞：

- 对于 IGMP 普遍组查询报文来说，通过配置 IGMP 普遍组查询的最大响应时间来填充其最大响应时间字段；
- 对于 IGMP 特定组查询报文来说，所配置的发送 IGMP 特定组查询报文的时间间隔将被填充到其最大响应时间字段。也就是说，IGMP 特定组查询的最大响应时间从数值上与发送 IGMP 特定组查询报文的时间间隔相同。

1. 全局配置IGMP查询和响应

表1-17 全局配置 IGMP 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
配置IGMP普遍组查询的最大响应时间	**max-response-time** *interval*	必选 缺省情况下，IGMP普遍组查询的最大响应时间为10秒
配置发送IGMP特定组查询报文的时间间隔	**last-member-query-interval** *interval*	必选 缺省情况下，发送IGMP特定组查询报文的时间间隔为1秒

2. 在VLAN内配置IGMP查询和响应

表1-18 在 VLAN 内配置 IGMP 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置发送IGMP普遍组查询报文的时间间隔	**igmp-snooping query-interval** *interval*	必选 缺省情况下，发送IGMP普遍组查询报文的时间间隔为60秒
配置IGMP普遍组查询的最大响应时间	**igmp-snooping max-response-time** *interval*	必选 缺省情况下，IGMP普遍组查询的最大响应时间为10秒
配置发送IGMP特定组查询报文的时间间隔	**igmp-snooping last-member-query-interval** *interval*	必选 缺省情况下，发送IGMP特定组查询报文的时间间隔为1秒

注意

应确保发送 IGMP 普遍组查询报文的时间间隔大于 IGMP 普遍组查询的最大响应时间，否则有可能造成对组播组成员的误删。

1.5.4 配置IGMP查询报文源IP地址

对于收到源 IP 地址为 0.0.0.0 的查询报文的端口，交换机不会将其设置为动态路由器端口，从而影响数据链路层组播转发表项的建立，最终导致组播数据无法正常转发。

当由二层设备充当 IGMP Snooping 查询器时，可以把 IGMP 查询报文的源 IP 地址配置为一个有效的 IP 地址以避免上述问题的出现。

表1-19 配置 IGMP 查询报文源 IP 地址

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置IGMP普遍组查询报文源IP地址	**igmp-snooping general-query source-ip** { *ip-address* | **current-interface** }	必选 缺省情况下，IGMP普遍组查询报文的源IP地址为0.0.0.0
配置IGMP特定组查询报文源IP地址	**igmp-snooping special-query source-ip** { *ip-address* | **current-interface** }	必选 缺省情况下，IGMP特定组查询报文的源IP地址为0.0.0.0

注意

IGMP 查询报文源 IP 地址的改变可能会影响网段内 IGMP 查询器的选举。

1.6 配置IGMP Snooping Proxying

1.6.1 配置准备

在配置 IGMP Snooping Proxying 之前，需完成以下任务：

- 在 VLAN 内使能 IGMP Snooping

在配置 IGMP Snooping Proxying 之前，需准备以下数据：

- 代理发送 IGMP 报告报文的源 IP 地址
- 代理发送 IGMP 离开报文的源 IP 地址

1.6.2 使能IGMP Snooping Proxying

当在有组播需求的 VLAN 内使能了 IGMP Snooping 代理功能后，该设备就成为该 VLAN 内的 IGMP Snooping 代理设备。

表1-20 使能 IGMP Snooping Proxying

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能IGMP Snooping代理功能	**igmp-snooping proxying enable**	必选 缺省情况下，VLAN内的IGMP Snooping代理功能处于关闭状态

1.6.3 配置代理发送IGMP报文的源IP地址

通过本配置可以改变代理发送的 IGMP 报告报文和离开报文的源 IP 地址。

表1-21 配置代理发送 IGMP 报文的源 IP 地址

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置代理发送IGMP报告报文的源IP地址	**igmp-snooping report source-ip** { *ip-address* | **current-interface** }	必选 缺省情况下，代理发送IGMP报告报文的源IP地址为0.0.0.0
配置代理发送IGMP离开报文的源IP地址	**igmp-snooping leave source-ip** { *ip-address* | **current-interface** }	必选 缺省情况下，代理发送IGMP离开报文的源IP地址为0.0.0.0

1.7 配置IGMP Snooping策略

1.7.1 配置准备

在配置 IGMP Snooping 策略之前，需完成以下任务：

- 在 VLAN 内使能 IGMP Snooping

在配置 IGMP Snooping 策略之前，需准备以下数据：

- 组播组过滤的 ACL 规则
- 端口加入的组播组最大数量
- IGMP 报文的 802.1p 优先级

1.7.2 配置组播组过滤器

在使能了 IGMP Snooping 的交换机上，通过配置组播组过滤器，可以限制用户对组播节目的点播。在实际应用中，当用户点播某个组播节目时，主机会发起一个 IGMP 成员关系报告报文，该报文到达交换机后，进行 ACL 检查：如果该接收端口可以加入这个组播组，则将其列入到 IGMP Snooping 转发表中；否则交换机就丢弃该报文。这样，未通过 ACL 检查的组播数据就不会送到该端口，从而达到控制用户点播组播节目的目的。

1. 全局配置组播组过滤器

表1-22 全局配置组播组过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-

操作	命令	说明
配置组播组过滤器	**group-policy** *acl-number* [**vlan** *vlan-list*]	必选 缺省情况下，没有配置全局组播组过滤器，即各VLAN内主机可以加入任意合法的组播组

2. 在端口上配置组播组过滤器

表1-23 在端口上配置组播组过滤器

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置组播组过滤器		**igmp-snooping group-policy** *acl-number* [**vlan** *vlan-list*]	必选 缺省情况下，端口上没有配置组播组过滤器，即该端口下的主机可以加入任意合法的组播组

说明

- 当在组播 VLAN 中配置组播组过滤器时，需在 sub-vlan 上进行配置，在主 VLAN 上配置不会生效。
- 当使用 IGMP v3 的用户点播多个组播组时，组播组过滤器功能可能不能正确对用户点播的组播组进行过滤，因为 IGMP v3 会将多个组播组的 report 信息在一个报文中发送，从而导致组播组过滤的失效。

1.7.3 配置组播数据报文源端口过滤

通过配置组播数据报文源端口过滤功能，可以允许或禁止端口作为组播源端口：

- 使能了该功能后，端口下不能连接组播源，因为该端口将过滤掉所有的组播数据报文（但允许组播协议报文通过），只能连接组播数据接收者；
- 关闭了该功能后，端口下既可以连接组播源，也可以连接组播数据接收者。

1. 全局配置组播数据报文源端口过滤

表1-24 全局配置组播数据报文源端口过滤

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-

操作	命令	说明
使能组播数据报文源端口过滤功能	**source-deny port** *interface-list*	必选 缺省情况下，组播数据报文源端口过滤功能处于关闭状态

2. 在端口上配置组播数据报文源端口过滤

表1-25 在端口上配置组播数据报文源端口过滤

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网端口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
使能组播数据报文源端口过滤功能		**igmp-snooping source-deny**	必选 缺省情况下，组播数据报文源端口过滤功能处于关闭状态

1.7.4 配置丢弃未知组播数据报文

未知组播数据报文是指在 IGMP Snooping 转发表中不存在对应转发表项的那些组播数据报文，当交换机收到发往未知组播组的报文时，数据报文会在未知组播数据报文所属的 VLAN 内广播，这样会占用大量的网络带宽，影响转发效率。

此时用户可以在交换机上启动丢弃未知组播数据报文功能，当交换机收到未知组播数据报文时，只向其路由器端口转发，不在 VLAN 内广播。如果交换机没有路由器端口，数据报文会被丢弃，不再转发

表1-26 在 VLAN 内配置丢弃未知组播数据报文

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
使能丢弃未知组播数据报文功能	**igmp-snooping drop-unknown**	必选 缺省情况下，丢弃未知组播数据报文的功能处于关闭状态，即对未知组播数据报文进行广播

1.7.5 配置IGMP成员关系报告报文抑制

当二层设备收到来自某组播组成员的 IGMP 成员关系报告报文时，会将该报文转发给与其直连的三层设备。这样，当二层设备上存在属于某组播组的多个成员时，与其直连的三层设备会收到这些成员发送的相同 IGMP 成员关系报告报文。

当使能了 IGMP 成员关系报告报文抑制功能后，在一个查询间隔内二层设备只会把收到的某组播组内的第一个 IGMP 成员关系报告报文转发给三层设备，而不继续向三层设备转发来自同一组播组的其它 IGMP 成员关系报告报文，这样可以减少网络中的报文数量。

表1-27 配置 IGMP 成员关系报告报文抑制

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
使能IGMP成员关系报告报文抑制功能	**report-aggregation**	必选 缺省情况下，IGMP成员关系报告报文抑制功能处于使能状态

注意

在 IGMP Snooping 代理设备上，不论是否使能了 IGMP 成员关系报告报文抑制功能，只要存在某组播组对应的转发表项，就会将从下游收到的针对该组的报告报文都抑制掉。

1.7.6 配置端口加入的组播组最大数量

通过配置端口加入的组播组最大数量，可以限制用户点播组播节目的数量，从而控制了端口上的数据流量。

表1-28 配置端口加入的组播组最大数量

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置端口加入的组播组最大数量		**igmp-snooping group-limit** *limit* [**vlan** *vlan-list*]	必选 缺省情况下，端口加入的组播组最大数量为2000

说明

在配置端口加入的组播组最大数量时，如果当前端口上的组播组数量已超过配置值，系统将把该端口相关的所有转发表项从IGMP Snooping转发表中删除，该端口下的主机都需要重新加入组播组，直至该端口上的组播组数量达到限制值为止。其中，如果该端口已配置为静态成员端口，系统会将静态成员端口的配置重新生效一次；如果在该端口上配置了模拟主机加入，系统在收到模拟主机发来的报告报文之后才会重新建立相应的转发表项。

1.7.7 配置组播组替换

由于某些特殊的原因，当前交换机或端口上通过的组播组数目有可能会超过交换机或该端口的限定；另外，在某些特定的应用中，交换机上新加入的组播组需要自动替换已存在的组播组（一个典型的应用就是“频道切换”，即用户通过加入一个新的组播组就能完成离开原组播组并切换到新组播组的动作）。

针对以上情况，可以在交换机或者某些端口上使能组播组替换功能。当交换机或端口上加入的组播组数量已达到限定值时：

- 若使能了组播组替换功能，则新加入的组播组会自动替代已存在的组播组，替代规则是替代IP地址最小的组播组；
- 若没有使能组播组替换功能，则自动丢弃新的IGMP成员关系报告报文。

1. 全局配置组播组替换

表1-29 全局配置组播组替换

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
使能组播组替换功能	**overflow-replace** [**vlan** *vlan-list*]	必选 缺省情况下，组播组替换功能处于关闭状态

2. 在端口上配置组播组替换

表1-30 在端口上配置组播组替换

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
使能组播组替换功能		**igmp-snooping overflow-replace** [**vlan** *vlan-list*]	必选 缺省情况下，组播组替换功能处于关闭状态

注意

在使能组播组替换功能之前，必须首先配置端口通过的组播组的最大数量（具体配置过程请参见“1.7.6 配置端口加入的组播组最大数量”），否则组播组替换功能将不会生效。

1.7.8 配置IGMP报文的 802.1p优先级

可以通过本配置来改变 IGMP 报文的 802.1p 优先级。当交换机的出端口发生拥塞时，交换机通过识别报文的 802.1p 优先级，优先发送优先级较高的报文。

1. 全局配置IGMP报文的 802.1p优先级

表1-31 全局配置 IGMP 报文的 802.1p 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
配置IGMP报文的802.1p优先级	**dot1p-priority** *priority-number*	必选 缺省情况下，IGMP报文的802.1p优先级为0

2. 在VLAN内配置IGMP报文的 802.1p优先级

表1-32 在 VLAN 内配置 IGMP 报文的 802.1p 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置IGMP报文的802.1p优先级	**igmp-snooping dot1p-priority** *priority-number*	必选 缺省情况下，IGMP报文的802.1p优先级为0

1.7.9 配置组播用户控制策略

组播用户控制策略通常配置在接入交换机上，是基于用户权限控制来实现的，即只有通过授权的用户才能收到相应的组播流，从而达到限制用户对组播节目点播的目的。在实际应用中，用户先要通过接入交换机向 RADIUS 服务器发起认证（如 802.1X 认证），当认证通过后再根据用户的点播行为进行策略检查：

- 当用户点播组播节目时，主机会发送 IGMP 成员关系报告报文，接入交换机收到该报文后对其携带的组播组和组播源地址进行策略检查，若该报文通过检查则允许该用户加入该组播组；否则，接入交换机将丢弃该报文。

- 当用户停止点播组播节目时，主机会发送 IGMP 离开报文，接入交换机收到该报文后对其携带的组播组和组播源地址进行策略检查，若该报文通过检查则允许该用户离开该组播组；否则，接入交换机将丢弃该报文。

表1-33 配置组播用户控制策略

操作	命令	说明
进入系统视图	**system-view**	-
创建User Profile，并进入User-Profile视图	**user-profile** *profile-name*	-
配置组播用户控制策略	**igmp-snooping access-policy** *acl-number*	必选 缺省情况下，没有配置组播用户控制策略，即用户可以加入/离开任意合法的组播组
退回系统视图	**quit**	-
激活该User Profile	**user-profile** *profile-name* **enable**	必选 缺省情况下，User Profile处于未激活状态

说明

- 有关 **user-profile** 和 **user-profile enable** 命令的详细介绍，请参见“安全命令参考”中的“User Profile”。
- 组播用户控制策略与组播组过滤器在功能上类似，二者的区别在于：前者是基于用户权限的组播控制，需要与认证、授权功能结合使用，能够控制组播用户的加入与离开；后者是基于设备端口的组播控制，无需与认证、授权功能结合使用，只能控制组播用户的加入。

1.7.10 配置IGMP Snooping主机跟踪功能

通过使能 IGMP Snooping 主机跟踪功能，可以使交换机能够记录正在接收组播数据的成员主机信息（包括主机的 IP 地址、运行时间和超时时间等），以便于网络管理员对这些主机进行监控和管理。

1. 全局配置IGMP Snooping主机跟踪功能

表1-34 全局配置 IGMP Snooping 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
全局使能IGMP Snooping主机跟踪功能	**host-tracking**	必选 缺省情况下，IGMP Snooping主机跟踪功能处于关闭状态

2. 在VLAN内配置IGMP Snooping主机跟踪功能

表1-35 在 VLAN 内配置 IGMP Snooping 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
在 VLAN 内使能 IGMP Snooping主机跟踪功能	**igmp-snooping host-tracking**	必选 缺省情况下，IGMP Snooping主机跟踪功能处于关闭状态

1.7.11 配置IGMP Snooping协议中发送IGMP报文的DSCP优先级

在 IPv4 报文头中，包含一个 8bit 的 ToS 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 IGMP Snooping 协议中发送的 MLD 报文的 DSCP 优先级进行配置。

表1-36 配置发送的 IGMP 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入IGMP-Snooping视图	**igmp-snooping**	-
配置发送的IGMP报文的DSCP优先级	**dscp** *dscp-value*	必选 缺省情况下，发送的IGMP报文的DSCP优先级为48

说明

此配置仅用于配置交换机自身生成的 IGMP 报文的 DSCP 优先级，不会对交换机转发的 IGMP 报文的 DSCP 优先级进行修改。

1.8 IGMP Snooping显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IGMP Snooping 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除组播组信息。

表1-37 IGMP Snooping 显示和维护

操作	命令
查看IGMP Snooping组的信息	**display igmp-snooping group** [**vlan** *vlan-id*] [**slot** *slot-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
查看IGMP Snooping跟踪的主机信息	**display igmp-snooping host vlan** *vlan-id* **group** *group-address* [**source** *source-address*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看静态组播 MAC 地址表信息	**display mac-address** [*mac-address* [**vlan** *vlan-id*] | [**multicast**] [**vlan** *vlan-id*] [**count**]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IGMP Snooping监听到的IGMP报文的统计信息	**display igmp-snooping statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IGMP Snooping组的动态加入记录	**reset igmp-snooping group** { *group-address* | **all** } [**vlan** *vlan-id*]
清除IGMP Snooping监听到的所有IGMP报文的统计信息	**reset igmp-snooping statistics**

说明

- **reset igmp-snooping group** 命令只对使能了 IGMP Snooping 的 VLAN 有效，而对 VLAN 接口上使能了 IGMP 的 VLAN 无效。
- **reset igmp-snooping group** 命令只能清除动态加入记录，而无法清除静态加入记录。

1.9 IGMP Snooping典型配置举例

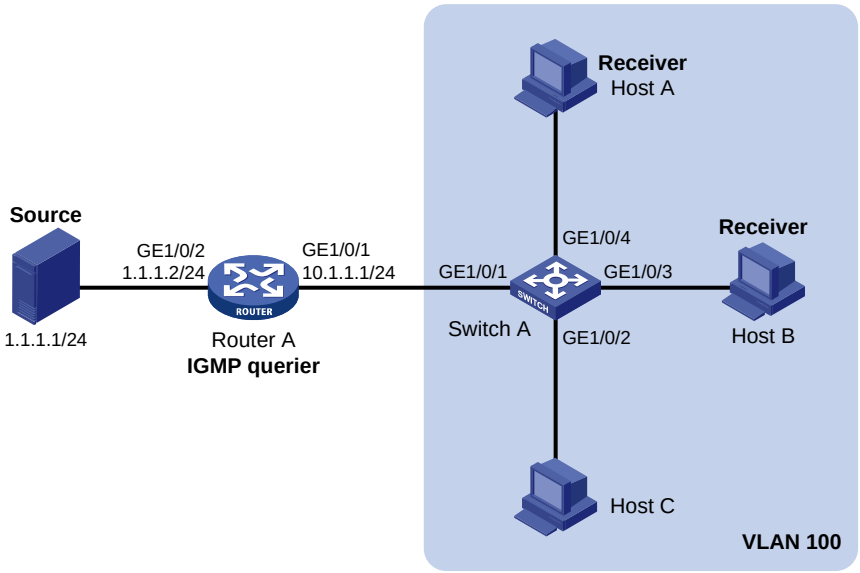
1.9.1 组策略及模拟主机加入配置举例

1. 组网需求

- 如 图 1-4 所示，Router A通过GigabitEthernet1/0/2 端口连接组播源（Source），通过GigabitEthernet1/0/1 端口连接Switch A；Router A上运行IGMPv2，Switch A上运行版本 2 的IGMP Snooping，并由Router A充当IGMP查询器。
- 通过配置，使 Host A 和 Host B 能且只能接收发往组播组 224.1.1.1 的组播数据，并且当 Host A 和 Host B 即使发生意外而临时中断接收组播数据时，发往组播组 224.1.1.1 组播数据也能不间断地通过 Switch A 的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/4 转发出去；同时，使 Switch A 将收到的未知组播数据直接丢弃，避免在其所属的 VLAN 内广播。

2. 组网图

图1-4 组策略及模拟主机加入配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-4 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 配置 Router A

使能 IP 组播路由，在各接口上使能 PIM-DM，并在端口 GigabitEthernet1/0/1 上使能 IGMP。

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] igmp enable
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch A

全局使能 IGMP Snooping。

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中；在该 VLAN 内使能 IGMP Snooping，并使能丢弃未知组播数据报文功能。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] igmp-snooping drop-unknown
[SwitchA-vlan100] quit
```

配置组播组过滤器，以限定 VLAN 100 内的主机只能加入组播组 224.1.1.1。

```
[SwitchA] acl number 2001
[SwitchA-acl-basic-2001] rule permit source 224.1.1.1 0
[SwitchA-acl-basic-2001] quit
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] group-policy 2001 vlan 100
[SwitchA-igmp-snooping] quit
```

在 GigabitEthernet1/0/3 和 GigabitEthernet1/0/4 上分别配置模拟主机加入组播组 224.1.1.1。

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] igmp-snooping host-join 224.1.1.1 vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
[SwitchA] interface gigabitethernet 1/0/4
[SwitchA-GigabitEthernet1/0/4] igmp-snooping host-join 224.1.1.1 vlan 100
[SwitchA-GigabitEthernet1/0/4] quit
```

(4) 检验配置效果

查看 Switch A 上 VLAN 100 内 IGMP Snooping 组的详细信息。

```
[SwitchA] display igmp-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port(s).
            GE1/0/1                (D) ( 00:01:30 )
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:224.1.1.1
        (0.0.0.0, 224.1.1.1):
          Attribute:    Host Port
          Host port(s):total 2 port(s).
            GE1/0/3                (D) ( 00:03:23 )
            GE1/0/4                (D) ( 00:04:10 )
    MAC group(s):
      MAC group address:0100-5e01-0101
          Host port(s):total 2 port(s).
            GE1/0/3
            GE1/0/4
```

由此可见，Switch A 上的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/4 已经加入了组播组 224.1.1.1。

1.9.2 静态端口配置举例

1. 组网需求

- 如 图 1-5 所示，Router A通过GigabitEthernet1/0/2 端口连接组播源（Source），通过GigabitEthernet1/0/1 端口连接Switch A；Router A上运行IGMPv2，Switch A、Switch B和Switch C上都运行版本 2 的IGMP Snooping，并由Router A充当IGMP查询器。
- Host A 和 Host C 均为组播组 224.1.1.1 的固定接收者（Receiver），通过将 Switch C 上的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/5 配置为组播组 224.1.1.1 的静态成员端口，可以增强组播数据在传输过程中的可靠性。
- 假设由于受 STP 等链路层协议的影响，为了避免出现环路，Switch A—Switch C 的转发路径在正常情况下是阻断的，组播数据只能通过 Switch A—Switch B—Switch C 的路径传递给连接在 Switch C 上的接收者；要求通过将 Switch A 的端口 GigabitEthernet1/0/3 配置为静态路由器端口，以保证当 Switch A—Switch B—Switch C 的路径出现阻断时，组播数据可以几乎不间断地通过 Switch A—Switch C 的新路径传递给接收者。

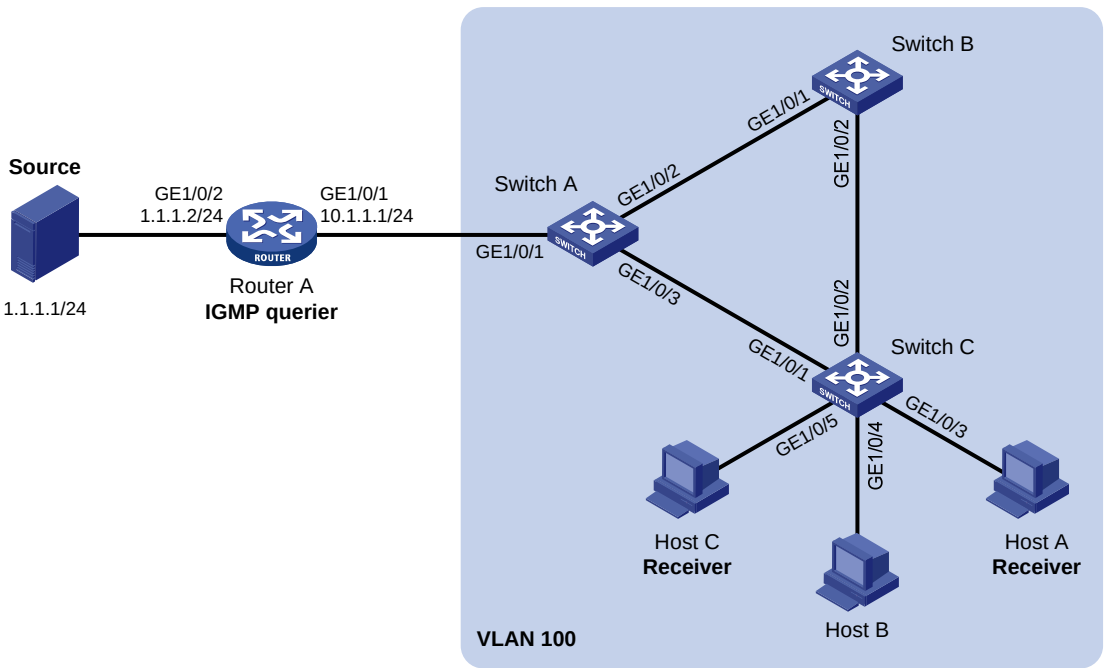
说明

如果没有配置静态路由器端口，那么当 Switch A—Switch B—Switch C 的路径出现阻断时，至少需要等待一个 IGMP 查询和响应周期完成后，组播数据才能通过 Switch A—Switch C 的新路径传递给接收者，组播数据的传输在这个过程中将中断。

有关 STP（Spanning Tree Protocol，生成树协议）的详细介绍，请参见“二层技术-以太网交换配置指导”中的“生成树”。

2. 组网图

图1-5 静态端口配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-5 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 配置 Router A

使能 IP 组播路由，在各接口上使能 PIM-DM，并在端口 GigabitEthernet1/0/1 上使能 IGMP。

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] igmp enable
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch A

全局使能 IGMP Snooping。

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/3 添加到该 VLAN 中，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] quit
```

把 GigabitEthernet1/0/3 配置为静态路由器端口。

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] igmp-snooping static-router-port vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
```

(4) 配置 Switch B

全局使能 IGMP Snooping。

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/2 添加到该 VLAN 中，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchB] vlan 100
[SwitchB-vlan100] port gigabitethernet 1/0/1 gigabitethernet 1/0/2
[SwitchB-vlan100] igmp-snooping enable
[SwitchB-vlan100] quit
```

(5) 配置 Switch C

全局使能 IGMP Snooping。

```
<SwitchC> system-view
[SwitchC] igmp-snooping
```

```
[SwitchC-igmp-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/5 添加到该 VLAN 中，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchC] vlan 100
[SwitchC-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/5
[SwitchC-vlan100] igmp-snooping enable
[SwitchC-vlan100] quit
```

分别在端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/5 上配置静态加入组播组 224.1.1.1。

```
[SwitchC] interface gigabitethernet 1/0/3
[SwitchC-GigabitEthernet1/0/3] igmp-snooping static-group 224.1.1.1 vlan 100
[SwitchC-GigabitEthernet1/0/3] quit
[SwitchC] interface gigabitethernet 1/0/5
[SwitchC-GigabitEthernet1/0/5] igmp-snooping static-group 224.1.1.1 vlan 100
[SwitchC-GigabitEthernet1/0/5] quit
```

(6) 检验配置效果

查看 Switch A 上 VLAN 100 内 IGMP Snooping 组的详细信息。

```
[SwitchA] display igmp-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 2 port(s).
            GE1/0/1                (D) ( 00:01:30 )
            GE1/0/3                (S)
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:224.1.1.1
        (0.0.0.0, 224.1.1.1):
          Attribute:    Host Port
          Host port(s):total 1 port(s).
            GE1/0/2                (D) ( 00:03:23 )
    MAC group(s):
      MAC group address:0100-5e01-0101
          Host port(s):total 1 port(s).
            GE1/0/2
```

由此可见，Switch A 上的端口 GigabitEthernet1/0/3 已经成为了静态路由器端口。

查看 Switch C 上 VLAN 100 内 IGMP Snooping 组的详细信息。

```
[SwitchC] display igmp-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
```

```
Port flags: D-Dynamic port, S-Static port, C-Copy port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
          GE1/0/2                (D) ( 00:01:23 )
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:224.1.1.1
      (0.0.0.0, 224.1.1.1):
        Attribute:    Host Port
        Host port(s):total 2 port(s).
          GE1/0/3                (S)
          GE1/0/5                (S)
  MAC group(s):
    MAC group address:0100-5e01-0101
        Host port(s):total 2 port(s).
          GE1/0/3
          GE1/0/5
```

由此可见，Switch C 上的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/5 已经成为了组播组 224.1.1.1 的静态成员端口。

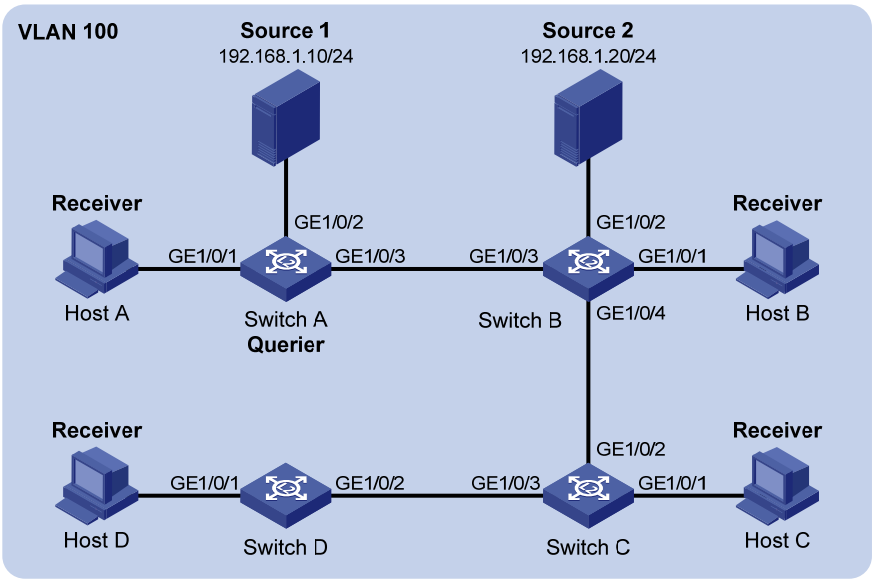
1.9.3 IGMP Snooping查询器配置举例

1. 组网需求

- 如 图 1-6 所示，在一个没有三层设备的纯二层网络环境中，组播源Source 1 和Source 2 分别向组播组 224.1.1.1 和 225.1.1.1 发送组播数据，Host A和Host C是组播组 224.1.1.1 的接收者（Receiver），Host B和Host D则是组播组 225.1.1.1 的接收者；所有接收者均使用IGMPv2，所有交换机上都运行版本 2 的IGMP Snooping，并选择距组播源较近的Switch A来充当IGMP Snooping查询器。
- 为防止交换机在没有二层组播转发表项时将组播数据在 VLAN 内广播，在所有交换机上都使能丢弃未知组播数据报文功能；同时，由于交换机不会将收到源 IP 地址为 0.0.0.0 的 IGMP 查询报文的端口设置为动态路由器端口，从而会影响二层组播转发表项的建立并导致组播数据无法正常转发，因此要求通过改变缺省的 IGMP 查询报文源 IP 地址（即 0.0.0.0）以避免此问题出现。

2. 组网图

图1-6 IGMP Snooping 查询器配置组网图



3. 配置步骤

(1) 配置 Switch A

全局使能 IGMP Snooping。

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

创建 VLAN 100，并把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/3 添加到该 VLAN 中。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
```

在 VLAN 100 内使能 IGMP Snooping，并使能丢弃未知组播数据报文功能。

```
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] igmp-snooping drop-unknown
```

在 VLAN 100 内使能 IGMP Snooping 查询器。

```
[SwitchA-vlan100] igmp-snooping querier
```

在 VLAN 100 内把 IGMP 普遍组查询和特定组查询报文的源 IP 地址均设置为 192.168.1.1。

```
[SwitchA-vlan100] igmp-snooping general-query source-ip 192.168.1.1
[SwitchA-vlan100] igmp-snooping special-query source-ip 192.168.1.1
[SwitchA-vlan100] quit
```

(2) 配置 Switch B

全局使能 IGMP Snooping。

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

创建 VLAN 100，并把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中。

```
[SwitchB] vlan 100
```

```
[SwitchB-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
```

在 VLAN 100 内使能 IGMP Snooping，并使能丢弃未知组播数据报文功能。

```
[SwitchB-vlan100] igmp-snooping enable
[SwitchB-vlan100] igmp-snooping drop-unknown
[SwitchB-vlan100] quit
```

Switch C 和 Switch D 的配置与 Switch B 相似，配置过程略。

(3) 检验配置效果

当 IGMP Snooping 查询器开始工作之后，除查询器以外的所有交换机都能收到 IGMP 普遍组查询报文。通过使用 **display igmp-snooping statistics** 命令可以查看 IGMP 报文的统计信息，例如：

查看 Switch B 上收到的 IGMP 报文的统计信息。

```
[SwitchB] display igmp-snooping statistics
  Received IGMP general queries:3.
  Received IGMPv1 reports:0.
  Received IGMPv2 reports:12.
  Received IGMP leaves:0.
  Received IGMPv2 specific queries:0.
  Sent     IGMPv2 specific queries:0.
  Received IGMPv3 reports:0.
  Received IGMPv3 reports with right and wrong records:0.
  Received IGMPv3 specific queries:0.
  Received IGMPv3 specific sg queries:0.
  Sent     IGMPv3 specific queries:0.
  Sent     IGMPv3 specific sg queries:0.
  Received error IGMP messages:0.
```

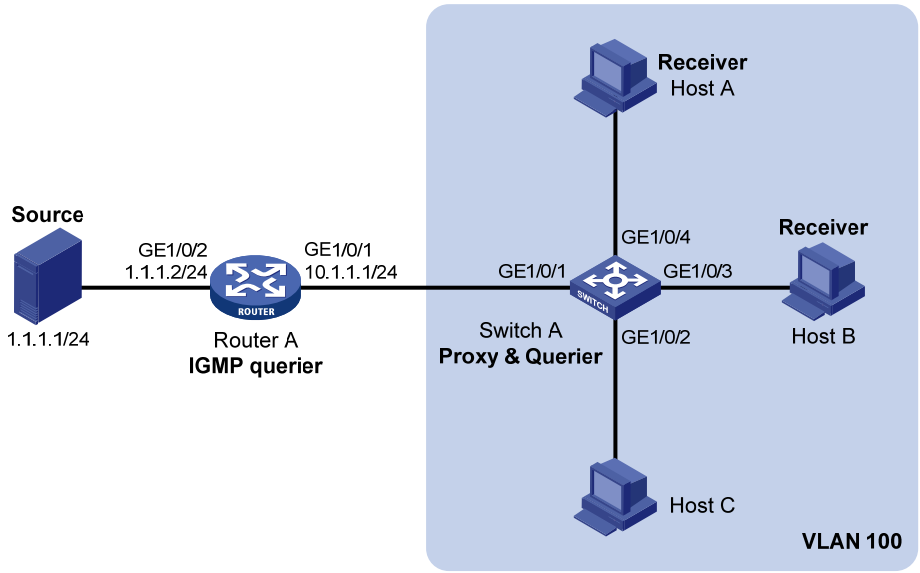
1.9.4 IGMP Snooping Proxying配置举例

1. 组网需求

- 如 图 1-7 所示，Router A通过GigabitEthernet1/0/2 端口连接组播源（Source），通过GigabitEthernet1/0/1 端口连接Switch A；Router A上运行IGMPv2，Switch A上运行版本 2 的IGMP Snooping，并由Router A充当IGMP查询器。
- 通过配置，使 Switch A 能够代理下游主机向 Router A 发送的 IGMP 报告报文和离开报文，以及响应 Router A 发来的 IGMP 查询报文并向下游主机转发。

2. 组网图

图1-7 IGMP Snooping Proxying 配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-7 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 配置 Router A

使能 IP 组播路由，在各接口上使能 PIM-DM，并在端口 GigabitEthernet1/0/1 上使能 IGMP。

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] igmp enable
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch A

全局使能 IGMP Snooping。

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中；在该 VLAN 内使能 IGMP Snooping，并使能 IGMP Snooping Proxying。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] igmp-snooping proxying enable
[SwitchA-vlan100] quit
```

(4) 检验配置效果

当配置完成后，Host A 和 Host B 分别发送组地址为 224.1.1.1 的 IGMP 加入报文，Switch A 收到该报文后通过其路由器端口 GigabitEthernet1/0/1 向 Router A 也发送该组的加入报文。通过使用 **display igmp-snooping group** 和 **display igmp group** 命令可以分别查看 IGMP Snooping 组和 IGMP 组的信息，例如：

查看 Switch A 上 IGMP Snooping 组的信息。

```
[SwitchA] display igmp-snooping group
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Port flags: D-Dynamic port, S-Static port, C-Copy port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port.
            GE1/0/1                (D)
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:224.1.1.1
        (0.0.0.0, 224.1.1.1):
          Host port(s):total 2 port.
            GE1/0/3                (D)
            GE1/0/4                (D)
    MAC group(s):
      MAC group address:0100-5e01-0101
          Host port(s):total 2 port.
            GE1/0/3
            GE1/0/4
```

查看 Router A 上 IGMP 组的信息。

```
[RouterA] display igmp group
Total 1 IGMP Group(s).
Interface group report information of VPN-Instance: public net
 GigabitEthernet1/0/1(10.1.1.1):
   Total 1 IGMP Group reported
    Group Address    Last Reporter    Uptime    Expires
    224.1.1.1        0.0.0.0          00:00:06  00:02:04
```

当 Host A 离开组播组 224.1.1.1 时，向 Switch A 发送该组的 IGMP 离开报文，但由于 Host B 仍未离开该组，因此 Switch A 并不会删除该组，也不会向 Router A 发送该组的离开报文，只是在该组对应转发表项的成员端口列表中将端口 GigabitEthernet1/0/3 删除。通过使用 **display igmp-snooping group** 命令可以查看 IGMP Snooping 组的信息，例如：

查看 Switch A 上 IGMP Snooping 组的信息。

```
[SwitchA] display igmp-snooping group
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
```

```
Port flags: D-Dynamic port, S-Static port, C-Copy port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port.
          GE1/0/1                 (D)
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:224.1.1.1
      (0.0.0.0, 224.1.1.1):
        Host port(s):total 1 port.
          GE1/0/4                 (D)
  MAC group(s):
    MAC group address:0100-5e01-0101
        Host port(s):total 1 port.
          GE1/0/4
```

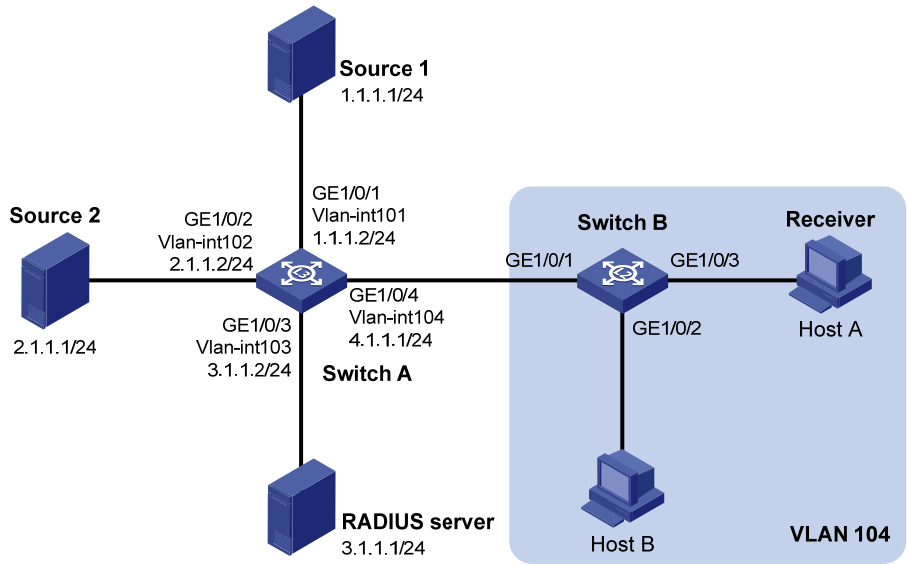
1.9.5 组播源与组播用户控制策略配置举例

1. 组网需求

- 如 图 1-8 所示，三层交换机Switch A分别通过接口Vlan-interface101 和Vlan-interface102 连接组播源Source 1 和Source 2，通过接口Vlan-interface103 连接RADIUS服务器，并通过接口Vlan-interface104 连接二层交换机Switch B；Switch A上运行IGMPv2，Switch B上运行版本 2 的IGMP Snooping，连接在Switch A上的组播源和连接在Switch B上的主机都使用802.1X进行认证。
- 通过在 Switch A 上配置组播源控制策略，禁止由 Source 2 发往组播组 224.1.1.1 的组播数据进入网络；通过在 Switch B 上配置组播用户控制策略，使接收者（Receiver）Host A 只能加入/离开组播组 224.1.1.1。

2. 组网图

图1-8 组播源与组播用户控制策略配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-8 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 配置 Switch A

分别创建 VLAN 101～104，并将端口 GigabitEthernet 1/0/1～GigabitEthernet 1/0/4 依次加入 VLAN 101～104 中。

```
<SwitchA> system-view
[SwitchA] vlan 101
[SwitchA-vlan101] port gigabitethernet 1/0/1
[SwitchA-vlan101] quit
[SwitchA] vlan 102
[SwitchA-vlan102] port gigabitethernet 1/0/2
[SwitchA-vlan102] quit
[SwitchA] vlan 103
[SwitchA-vlan103] port gigabitethernet 1/0/3
[SwitchA-vlan103] quit
[SwitchA] vlan 104
[SwitchA-vlan104] port gigabitethernet 1/0/4
[SwitchA-vlan104] quit
```

使能 IP 组播路由，在接口 Vlan-interface101、Vlan-interface102 和 Vlan-interface104 上使能 PIM-DM，并在接口 Vlan-interface104 上使能 IGMP。

```
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
```

```
[SwitchA-Vlan-interface102] pim dm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim dm
[SwitchA-Vlan-interface104] igmp enable
[SwitchA-Vlan-interface104] quit
```

创建名为 policy1 的 QoS 策略，不允许来自组播源 2.1.1.1 的组播数据向组播组 224.1.1.1 发送。

```
[SwitchA] acl number 3001
[SwitchA-acl-adv-3001] rule permit udp source 2.1.1.1 0 destination 224.1.1.1 0
[SwitchA-acl-adv-3001] quit
[SwitchA] traffic classifier classifier1
[SwitchA-classifier-classifier1] if-match acl 3001
[SwitchA-classifier-classifier1] quit
[SwitchA] traffic behavior behavior1
[SwitchA-behavior-behavior1] filter deny
[SwitchA-behavior-behavior1] quit
[SwitchA] qos policy policy1
[SwitchA-qospolicy-policy1] classifier classifier1 behavior behavior1
[SwitchA-qospolicy-policy1] quit
```

创建名为 profile1 的 User Profile，在该 User Profile 下应用 QoS 策略 policy1，并将该 User Profile 激活。

```
[SwitchA] user-profile profile1
[SwitchA-user-profile-profile1] qos apply policy policy1 inbound
[SwitchA-user-profile-profile1] quit
[SwitchA] user-profile profile1 enable
```

创建名为 scheme1 的 RADIUS 认证方案，配置 RADIUS 服务器的类型为 extended 类型；指定主 RADIUS 认证/授权服务器和计费服务器的 IP 地址均为 3.1.1.1，共享密钥均为 123321；并配置发送给 RADIUS 服务器的用户名不需要携带域名。

```
[SwitchA] radius scheme scheme1
[SwitchA-radius-scheme1] server-type extended
[SwitchA-radius-scheme1] primary authentication 3.1.1.1
[SwitchA-radius-scheme1] key authentication 123321
[SwitchA-radius-scheme1] primary accounting 3.1.1.1
[SwitchA-radius-scheme1] key accounting 123321
[SwitchA-radius-scheme1] user-name-format without-domain
[SwitchA-radius-scheme1] quit
```

创建名为 domain1 的 ISP 域，指定 lan-access 用户的 RADIUS 认证、授权和计费方案均为 scheme1；并将该域指定为系统缺省的 ISP 域。

```
[SwitchA] domain domain1
[SwitchA-isp-domian1] authentication lan-access radius-scheme scheme1
[SwitchA-isp-domian1] authorization lan-access radius-scheme scheme1
[SwitchA-isp-domian1] accounting lan-access radius-scheme scheme1
[SwitchA-isp-domian1] quit
[SwitchA] domain default enable domain1
```

全局使能 802.1X，并在端口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/2 上分别使能 802.1X。

```
[SwitchA] dot1x
[SwitchA] interface gigabitethernet 1/0/1
```

```
[SwitchA-GigabitEthernet1/0/1] dot1x
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] dot1x
[SwitchA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch B

全局使能 IGMP Snooping。

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

创建 VLAN 104，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/3 添加到该 VLAN 中；在该 VLAN 内使能 IGMP Snooping。

```
[SwitchB] vlan 104
[SwitchB-vlan104] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchB-vlan104] igmp-snooping enable
[SwitchB-vlan104] quit
```

创建名为 profile2 的 User Profile，在该 User Profile 下配置只允许用户加入/离开组播组 224.1.1.1，并将该 User Profile 激活。

```
[SwitchB] acl number 2001
[SwitchB-acl-basic-2001] rule permit source 224.1.1.1 0
[SwitchB-acl-basic-2001] quit
[SwitchB] user-profile profile2
[SwitchB-user-profile-profile2] igmp-snooping access-policy 2001
[SwitchB-user-profile-profile2] quit
[SwitchB] user-profile profile2 enable
```

创建名为 scheme2 的 RADIUS 认证方案，配置 RADIUS 服务器的类型为 extended 类型；指定主 RADIUS 认证/授权服务器和计费服务器的 IP 地址均为 3.1.1.1，共享密钥均为 321123；并配置发送给 RADIUS 服务器的用户名不需要携带域名。

```
[SwitchB] radius scheme scheme2
[SwitchB-radius-scheme2] server-type extended
[SwitchB-radius-scheme2] primary authentication 3.1.1.1
[SwitchB-radius-scheme2] key authentication 321123
[SwitchB-radius-scheme2] primary accounting 3.1.1.1
[SwitchB-radius-scheme2] key accounting 321123
[SwitchB-radius-scheme2] user-name-format without-domain
[SwitchB-radius-scheme2] quit
```

创建名为 domain2 的 ISP 域，指定 lan-access 用户的 RADIUS 认证、授权和计费方案均为 scheme2；并将该域指定为系统缺省的 ISP 域。

```
[SwitchB] domain domain2
[SwitchB-isp-domian2] authentication lan-access radius-scheme scheme2
[SwitchB-isp-domian2] authorization lan-access radius-scheme scheme2
[SwitchB-isp-domian2] accounting lan-access radius-scheme scheme2
[SwitchB-isp-domian2] quit
[SwitchB] domain default enable domain2
```

全局使能 802.1X，并在端口 GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 上分别使能 802.1X。

```
[SwitchB] dot1x
```

```
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] dot1x
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] dot1x
[SwitchB-GigabitEthernet1/0/3] quit
```

(4) 配置 RADIUS 服务器

请在 RADIUS 服务器上进行与 Switch A 和 Switch B 相对应的配置，具体配置过程可参考 RADIUS 服务器的配置手册。

(5) 检验配置效果

配置完成后，所有的组播源和用户主机都向 RADIUS 服务器发起 802.1X 认证。当认证通过后，Source 1 向组播组 224.1.1.1、Source 2 向组播组 224.1.1.2 分别发送不同的组播数据，而 Host A 则分别发起了对组播组 224.1.1.1 和 224.1.1.2 的加入。通过使用 **display igmp-snooping group** 命令可以查看 IGMP Snooping 组的信息：

查看 Switch B 上 VLAN 104 内 IGMP Snooping 组的详细信息。

```
[SwitchB] display igmp-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port(s).
            GE1/0/1                (D) ( 00:01:30 )
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:224.1.1.1
        (0.0.0.0, 224.1.1.1):
          Attribute:    Host Port
          Host port(s):total 1 port(s).
            GE1/0/3                (D) ( 00:04:10 )
    MAC group(s):
      MAC group address:0100-5e01-0101
          Host port(s):total 1 port(s).
            GE1/0/3
```

由此可见，Switch B 上的端口 GigabitEthernet1/0/3 只加入了组播组 224.1.1.1，而未加入组播组 224.1.1.2。

此后由于管理员的操作失误，Source 2 也开始向组播组 224.1.1.1 发送组播数据。通过使用 **display multicast forwarding-table** 命令可以查看组播转发表的信息：

查看 Switch A 上有关组播组 224.1.1.1 的组播转发表信息。

```
[SwitchA] display multicast forwarding-table 224.1.1.1
Multicast Forwarding Table of VPN-Instance: public net
```

```
Total 1 entry

Total 1 entry matched
00001. (1.1.1.1, 224.1.1.1)
     MID: 0, Flags: 0x0:0
     Uptime: 00:08:32, Timeout in: 00:03:26
     Incoming interface: Vlan-interface101
     List of 1 outgoing interfaces:
       1: Vlan-interface104
     Matched 19648 packets(20512512 bytes), Wrong If 0 packets
     Forwarded 19648 packets(20512512 bytes)
```

由此可见，Switch A 上只有组播源组（1.1.1.1，224.1.1.1）的转发表项，而没有组播源组（2.1.1.1，224.1.1.1）的转发表项，说明 Source 2 发往组播组 224.1.1.1 的组播数据没能进入网络。

1.10 常见配置错误举例

1.10.1 交换机不能实现二层组播

1. 故障现象

交换机不能实现 IGMP Snooping 二层组播功能。

2. 分析

IGMP Snooping 没有使能。

3. 处理过程

(1) 使用 **display current-configuration** 命令查看 IGMP Snooping 的运行状态。

(2) 如果是没有使能 IGMP Snooping，则需先在系统视图下使用 **igmp-snooping** 命令全局使能 IGMP Snooping，然后在 VLAN 视图下使用 **igmp-snooping enable** 命令使能 VLAN 内的 IGMP Snooping。

(3) 如果只是没有在相应 VLAN 下使能 IGMP Snooping，则只需在 VLAN 视图下使用 **igmp-snooping enable** 命令使能 VLAN 内的 IGMP Snooping。

1.10.2 配置的组播组策略不生效

1. 故障现象

配置了组播组策略，只允许主机加入某些特定的组播组，但主机仍然可以收到发往其它组播组的组播数据。

2. 分析

- ACL 规则配置不正确；
- 组播组策略应用不正确；
- 没有使能丢弃未知组播数据报文的功能，使得属于过滤策略之外的组播数据报文（即未知组播数据报文）被广播。

3. 处理过程

(1) 使用 **display acl** 命令查看所配置的 ACL 规则，检查其是否与所要实现的组播组过滤策略相符合。

(2) 在 IGMP-Snooping 视图或相应的端口视图下使用 **display this** 命令查看是否应用了正确的组播组策略。如果没有，则使用 **group-policy** 或 **igmp-snooping group-policy** 命令应用正确的组播组策略。

(3) 使用 **display current-configuration** 命令查看是否已使能丢弃未知组播数据报文的功能。如果没有使能，则使用 **igmp-snooping drop-unknown** 命令使能丢弃未知组播数据报文功能。

1.11 附录

1.11.1 交换机对组播协议报文的特殊处理规则

支持 IGMP Snooping 功能的交换机使能了三层组播路由后，在不同的情况下对各种组播协议报文的处理有所差异。具体规则如下：

(1) 如果交换机上只使能了 IGMP，或同时使能了 IGMP 和 PIM：

- 对于 IGMP 报文，根据报文类型为其维护相应的动态成员端口或动态路由器端口；
- 对于 PIM Hello 报文，为其维护相应的动态路由器端口。

(2) 如果交换机上只使能了 PIM：

- 对于 IGMP 报文，将其当作未知报文在 VLAN 内进行广播；
- 对于 PIM Hello 报文，为其维护相应的动态路由器端口。

(3) 在交换机上关闭 IGMP 时：

- 如果未使能 PIM，则删除所有动态成员端口和动态路由器端口；
- 如果已使能 PIM，则删除动态成员端口，只保留动态路由器端口。

说明

在使能了三层组播路由的交换机上，请使用命令 **display igmp group port-info** 来查看二层端口信息。有关 **display igmp group port-info** 命令的详细介绍，请参见“IP 组播命令参考”中的“IGMP”。

---

(4) 在交换机上关闭 PIM 时：

- 如果未使能 IGMP，则删除动态路由器端口；
- 如果已使能 IGMP，则保留所有动态成员端口和动态路由器端口。

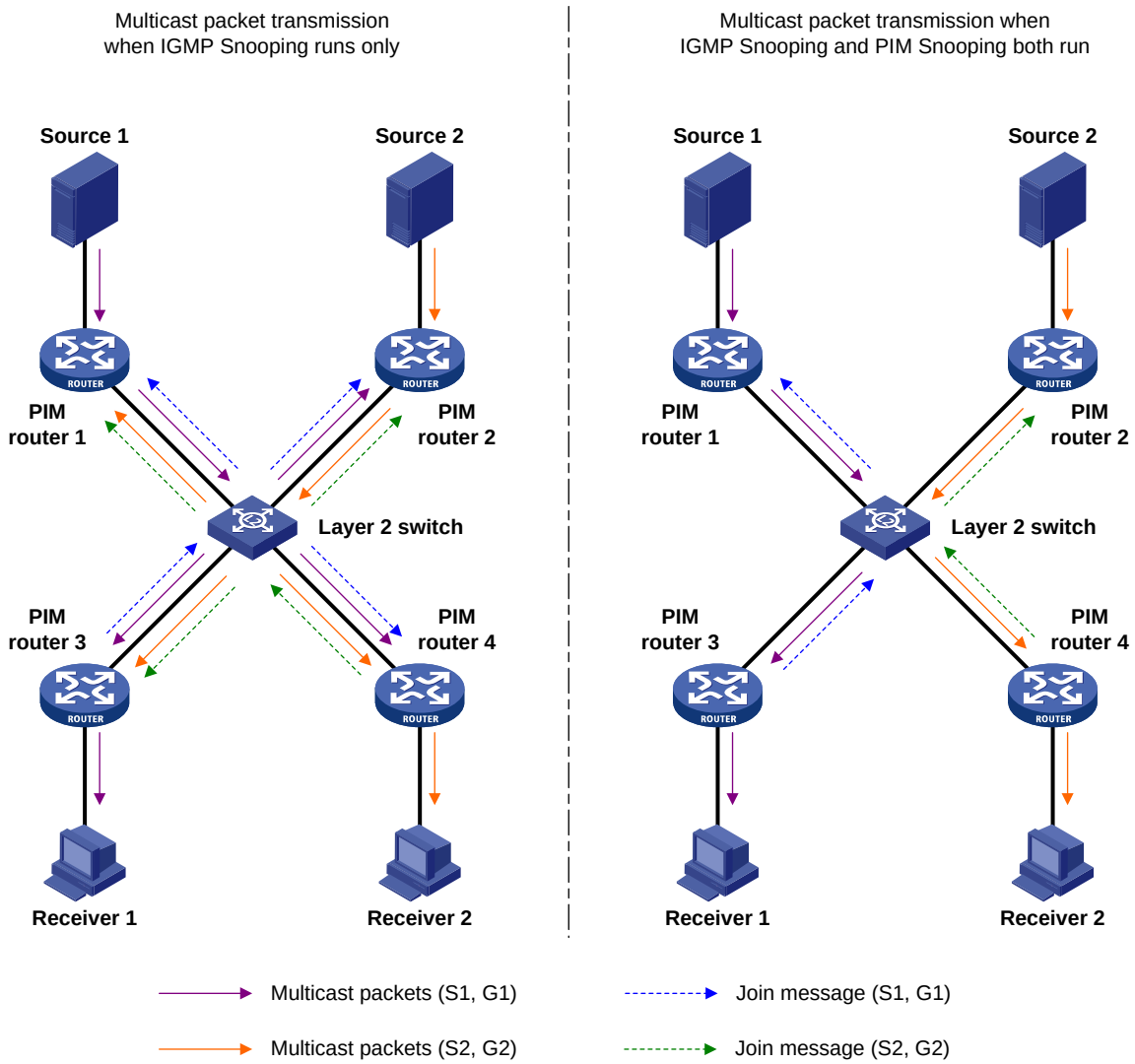
目 录

1 PIM Snooping配置

1.1 PIM Snooping简介

PIM Snooping 是 Protocol Independent Multicast Snooping（协议无关组播窥探）的简称，运行 PIM Snooping 的二层设备通过对收到的 PIM 协议报文进行分析，将有接收需求的端口添加到相应的组播转发表项中，以实现组播报文的精确转发。

图1-1 二层设备运行 PIM Snooping 前后的对比



如 图 1-1 所示，组播源Source 1 和Source 2 分别向组播组G1 和G2 发送组播数据，而Receiver 1 和Receiver 2 则分别是G1 和G2 的接收者，二层设备上连接各PIM路由器的端口都属于同一个VLAN：

- 当二层设备只运行 IGMP Snooping 时，它通过监听 PIM 路由器发出的 PIM Hello 报文来维护路由器端口，并将所有组播数据报文向 VLAN 内的所有路由器端口转发；而对于其它类型的

PIM 协议报文，则在 VLAN 内广播。因此，无论 PIM 路由器是否有接收需求，都会收到所有的 PIM 协议报文和组播数据报文。

- 当二层设备同时运行了 IGMP Snooping 和 PIM Snooping 时，它通过监听 PIM 路由器发出的 PIM 协议报文来了解其接收需求，从而将有接收需求的 PIM 路由器所在的端口加入到相应的组播转发表项中，使 PIM 协议报文和组播数据报文能够被精确转发给有接收需求的 PIM 路由器，从而节约了网络带宽。

说明

- 有关 IGMP Snooping 和路由器端口的详细介绍，请参见“IP 组播配置指导”中的“IGMP Snooping”。
- 有关 PIM 的详细介绍，请参见“IP 组播配置指导”中的“PIM”。

1.2 配置PIM Snooping

在配置某 VLAN 内的 PIM Snooping 时，需要在交换机上先全局使能 IGMP Snooping，然后在该 VLAN 内使能 IGMP Snooping 和 PIM Snooping。

表1-1 配置 PIM Snooping

操作	命令	说明
进入系统视图	**system-view**	-
全局使能IGMP Snooping，并进入IGMP-Snooping视图	**igmp-snooping**	必选 缺省情况下，IGMP Snooping处于关闭状态
退回系统视图	**quit**	-
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能IGMP Snooping	**igmp-snooping enable**	必选 缺省情况下，VLAN内的IGMP Snooping处于关闭状态
在VLAN内使能PIM Snooping	**pim-snooping enable**	必选 缺省情况下，VLAN内的PIM Snooping处于关闭状态

说明

- 有关 **igmp-snooping** 和 **igmp-snooping enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“IGMP Snooping”。
- 在 VLAN 内使能了 PIM Snooping 之后，PIM Snooping 功能只在属于该 VLAN 的端口上生效。
- 在组播 VLAN 的子 VLAN 内使能 PIM Snooping 无效。有关组播 VLAN 的详细介绍，请参见“IP 组播配置指导”中的“组播 VLAN”。
- 在部署了配置 PIM Snooping 功能的交换机的网络中，请在接收者侧的边缘 PIM 设备上配置加入/剪枝报文（**jp-pkt-size**）的最大长度不能大于路径 MTU，有关配置加入/剪枝报文的详细内容，请参见“IP 组播配置指导”中的“PIM”。

1.3 PIM Snooping显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 PIM Snooping 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 PIM Snooping 的统计信息。

表1-2 PIM Snooping 显示和维护

操作	命令
查看PIM Snooping的邻居信息	**display pim-snooping neighbor** [**vlan** *vlan-id*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看PIM Snooping的路由信息	**display pim-snooping routing-table** [**vlan** *vlan-id*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看PIM Snooping监听到的PIM报文的统计信息	**display pim-snooping statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除PIM Snooping监听到的PIM报文的统计信息	**reset pim-snooping statistics**

1.4 PIM Snooping典型配置举例

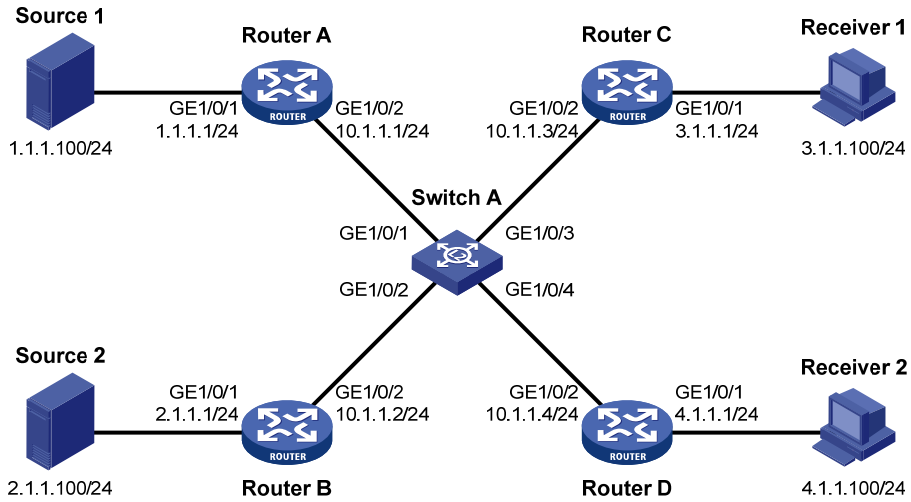
1. 组网需求

- 如 图 1-2 所示，Router A和Router B各自的GigabitEthernet1/0/1 接口分别连接组播源Source 1 和Source 2；Router C和Router D各自的GigabitEthernet1/0/1 接口分别连接接收者Receiver 1 和Receiver 2；Router A、Router B、Router C和Router D各自的GigabitEthernet1/0/2 接口都通过Switch A互连。
- Source 1 和 Source 2 分别通过组播组 224.1.1.1 和 225.1.1.1 发送组播数据，Receiver 1 和 Receiver 2 则分别接收来自组播组 224.1.1.1 和 225.1.1.1 的组播数据；Router C 和 Router D 各自的 GigabitEthernet1/0/1 接口上都运行 IGMP，Router A、Router B、Router C 和 Router D 上都运行 PIM-SM，并由 Router A 的 GigabitEthernet1/0/2 接口充当 C-BSR 和 C-RP。

- 通过在 Switch A 上配置 IGMP Snooping 和 PIM Snooping，使 Switch A 将 PIM 协议报文和组播数据报文只转发给有接收需求的路由器。

2. 组网图

图1-2 PIM Snooping 典型配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-2 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 配置 Router A

使能 IP 组播路由，在各接口上使能 PIM-SM，并将接口 GigabitEthernet1/0/2 配置为 C-BSR 和 C-RP。

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet1/0/1
[RouterA-GigabitEthernet1/0/1] pim sm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet1/0/2
[RouterA-GigabitEthernet1/0/2] pim sm
[RouterA-GigabitEthernet1/0/2] quit
[RouterA] pim
[RouterA-pim] c-bsr gigabitethernet1/0/2
[RouterA-pim] c-rp gigabitethernet1/0/2
```

(3) 配置 Router B

使能 IP 组播路由，在各接口上使能 PIM-SM。

```
<RouterB> system-view
[RouterB] multicast routing-enable
[RouterB] interface gigabitethernet1/0/1
[RouterB-GigabitEthernet1/0/1] pim sm
[RouterB-GigabitEthernet1/0/1] quit
[RouterB] interface gigabitethernet1/0/2
[RouterB-GigabitEthernet1/0/2] pim sm
```

(4) 配置 Router C

使能 IP 组播路由，在各接口上使能 PIM-SM，并在接口 GigabitEthernet1/0/1 上使能 IGMP。

```
<RouterC> system-view
[RouterC] multicast routing-enable
[RouterC] interface gigabitethernet1/0/1
[RouterC-GigabitEthernet1/0/1] pim sm
[RouterC-GigabitEthernet1/0/1] igmp enable
[RouterC-GigabitEthernet1/0/1] quit
[RouterC] interface gigabitethernet1/0/2
[RouterC-GigabitEthernet1/0/2] pim sm
```

(5) 配置 Router D

Router D 的配置与 Router C 相似，配置过程略。

(6) 配置 Switch A

全局使能 IGMP Snooping。

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中；在该 VLAN 内使能 IGMP Snooping 和 PIM Snooping。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet1/0/1 to gigabitethernet1/0/4
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] pim-snooping enable
[SwitchA-vlan100] quit
```

(7) 检验配置效果

查看 Switch A 上 VLAN 100 内 PIM Snooping 的邻居信息。

```
[SwitchA] display pim-snooping neighbor vlan 100
  Total number of neighbors: 4

  VLAN ID: 100
    Total number of neighbors: 4
    Neighbor        Port                    Expires   Option Flags
    10.1.1.1        GE1/0/1                  02:02:23  LAN Prune Delay
    10.1.1.2        GE1/0/2                  03:00:05  LAN Prune Delay
    10.1.1.3        GE1/0/3                  02:22:13  LAN Prune Delay
    10.1.1.4        GE1/0/4                  03:07:22  LAN Prune Delay
```

由此可见，Router A、Router B、Router C 和 Router D 之间都建立起了 PIM Snooping 邻居关系。

查看 Switch A 上 VLAN 100 内 PIM Snooping 的路由信息。

```
[SwitchA] display pim-snooping routing-table vlan 100 slot 1
  Total 2 entry(ies)
  FSM Flag: NI-no info, J-join, PP-prune pending

  VLAN ID: 100
    Total 2 entry(ies)
    (*, 224.1.1.1)
```

```
    Upstream neighbor: 10.1.1.1
      Upstream port: GE1/0/1
      Total number of downstream ports: 1
        1: GE1/0/3
           Expires: 00:03:01, FSM: J
(*, 225.1.1.1)
    Upstream neighbor: 10.1.1.2
      Upstream port: GE1/0/2
      Total number of downstream ports: 1
        1: GE1/0/4
           Expires: 00:01:05, FSM: J
```

由此可见，Switch A 将向 Router C 转发组播组 224.1.1.1 的组播数据，向 Router D 转发组播组 225.1.1.1 的组播数据。

1.5 常见配置错误举例

1.5.1 交换机不能实现PIM Snooping功能

1. 故障现象

交换机不能实现 PIM Snooping 功能。

2. 分析

IGMP Snooping 或 PIM Snooping 没有使能。

3. 处理过程

(1) 使用 **display current-configuration** 命令查看 IGMP Snooping 和 PIM Snooping 的运行状态。

(2) 如果没有使能 IGMP Snooping，请先在系统视图下使用 **igmp-snooping** 命令全局使能 IGMP Snooping，然后在 VLAN 视图下分别使用 **igmp-snooping enable** 和 **pim-snooping enable** 命令使能 VLAN 内的 IGMP Snooping 和 PIM Snooping。

(3) 如果没有使能 PIM Snooping，请在 VLAN 视图下使用 **pim-snooping enable** 命令使能 VLAN 内的 PIM Snooping。

1.5.2 部分下游PIM路由器无法收到组播数据

1. 故障现象

在有分片加入/剪枝报文的网络中，部分下游 PIM 路由器无法收到组播数据。

2. 分析

PIM Snooping 不能对分片报文进行重组，因此无法维护分片加入/剪枝报文中携带的下游状态。为了保证系统功能正常，只能将分片加入/剪枝报文在 VLAN 内广播，因此需要在 VLAN 内连接 PIM Snooping 交换机的所有 PIM 路由器上都禁止加入报文抑制能力（即使能邻居跟踪功能），以保证加入报文不被广播的分片加入/剪枝报文所抑制。假如存在未禁止该能力的 PIM 路由器，被广播的分片加入/剪枝报文就会影响其它 PIM 路由器的加入状态：如果某 PIM 路由器有组播接收需求，但其发送的加入报文被抑制，那么该路由器将无法收到组播数据。

3. 处理过程

(1) 在 PIM 路由器连接 PIM Snooping 交换机的接口上使用 **pim hello-option neighbor-tracking** 命令使能邻居跟踪功能。

(2) 如果存在不能够使能邻居跟踪功能的 PIM 路由器，则需关闭 PIM Snooping 交换机上的 PIM Snooping 功能。

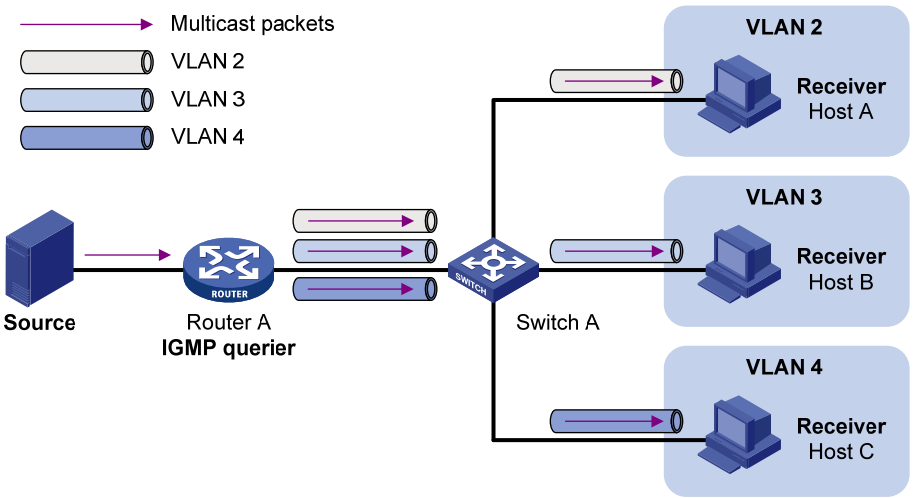
目 录

1 组播VLAN配置

1.1 组播VLAN简介

如 图 1-1 所示，在传统的组播点播方式下，当属于不同VLAN的主机Host A、Host B和Host C同时点播同一组播组时，三层设备（Router A）需要把组播数据在每个用户VLAN（即主机所属的VLAN）内都复制一份发送给二层设备（Switch A）。这样既造成了带宽的浪费，也给三层设备增加了额外的负担。

图1-1 未运行组播 VLAN 时的组播数据传输



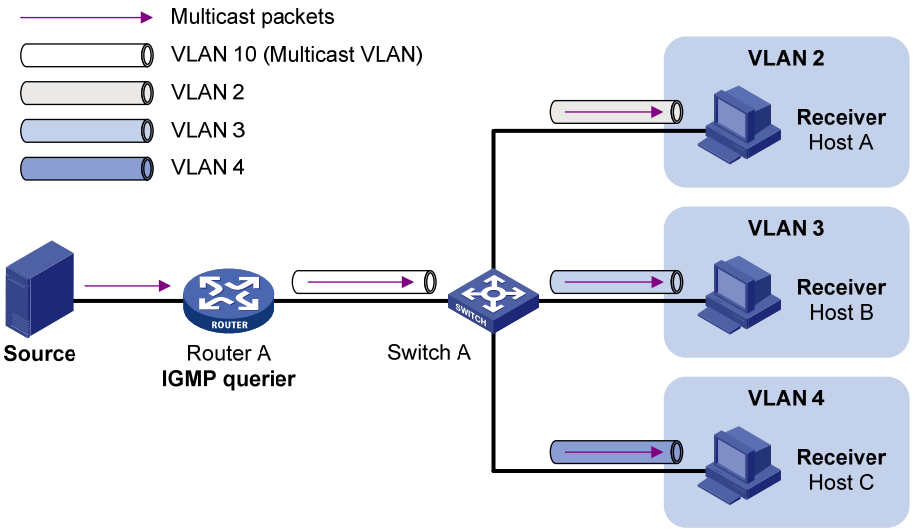
可以使用组播 VLAN 功能解决这个问题。在二层设备上配置了组播 VLAN 后，三层设备只需把组播数据在组播 VLAN 内复制一份发送给二层设备，而不必在每个用户 VLAN 内都复制一份，从而节省了网络带宽，也减轻了三层设备的负担。

组播 VLAN 有以下两种实现和配置方式：

1. 基于子VLAN的组播VLAN

如 图 1-2 所示，接收者主机Host A、Host B和Host C分属不同的用户VLAN。在Switch A上配置VLAN 10 为组播VLAN，将所有的用户VLAN都配置为该组播VLAN的子VLAN，并在组播VLAN内使能IGMP Snooping。

图1-2 基于子 VLAN 的组播 VLAN 示意图

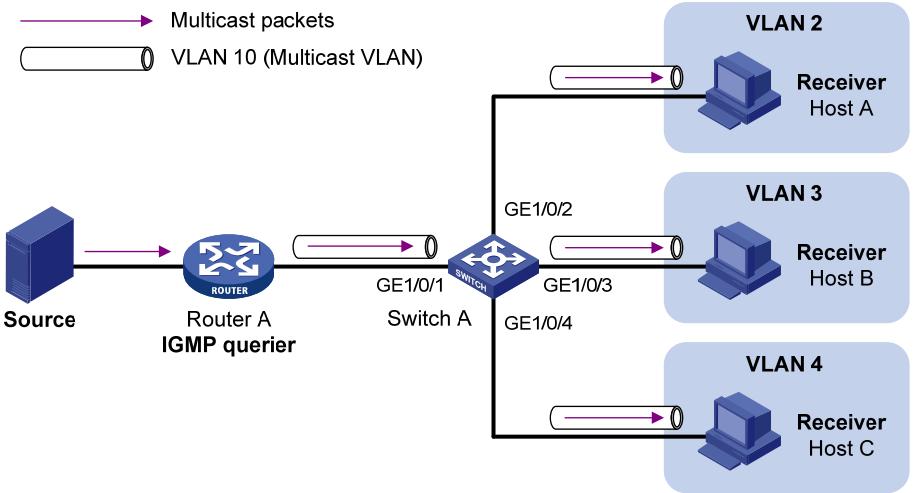


配置完成后，IGMP Snooping 将在组播 VLAN 中对路由器端口进行维护，而在各子 VLAN 中对成员端口进行维护。这样，Router A 只需把组播数据在组播 VLAN 内复制一份发送给 Switch A 即可，Switch A 会将其分发给该组播 VLAN 内那些有接收者的子 VLAN。

2. 基于端口的组播VLAN

如 图 1-3 所示，接收者主机Host A、Host B和Host C分属不同的用户VLAN，Switch A上的所有用户端口（即连接主机的端口）均为Hybrid类型。在Switch A上配置VLAN 10 为组播VLAN，将所有用户端口都添加到该组播VLAN内，并在组播VLAN和所有用户VLAN内都使能IGMP Snooping。

图1-3 基于端口的组播 VLAN 示意图



配置完成后，当 Switch A 上的用户端口收到来自主机的 IGMP 报文时，会为其打上组播 VLAN 的 Tag 并上送给 IGMP 查询器，于是 IGMP Snooping 就可以在组播 VLAN 中对路由器端口和成员端口进行统一的维护。这样，Router A 只需把组播数据在组播 VLAN 内复制一份发送给 Switch A 即可，Switch A 会将其分发给该组播 VLAN 内的所有成员端口。

说明

- 有关 IGMP Snooping 的相关配置，以及路由器端口和成员端口的详细介绍，请参见“IP 组播配置指导”中的“IGMP Snooping”。
- 有关 VLAN Tag 的详细介绍，请参见“二层技术-以太网交换配置指导”中的“VLAN”。

1.2 组播VLAN配置任务简介

表1-1 组播 VLAN 配置任务简介

配置任务		说明	详细配置
配置基于子VLAN的组播VLAN		二者必选其一	1.3
配置基于端口的组播VLAN	配置用户端口属性		1.4.2
	配置组播VLAN端口		1.4.3
配置组播VLAN内IGMP Snooping转发表项的最大数量		可选	1.5

说明

若在设备上同时配置了基于子 VLAN 和基于端口的组播 VLAN，则基于端口的组播 VLAN 将优先生效。

1.3 配置基于子VLAN的组播VLAN

1.3.1 配置准备

在配置基于子 VLAN 的组播 VLAN 之前，需完成以下任务：

- 创建相应的 VLAN
- 在欲配置为组播 VLAN 的 VLAN 内使能 IGMP Snooping

1.3.2 配置基于子VLAN的组播VLAN

首先需要把某个 VLAN 配置为组播 VLAN，再将用户 VLAN 添加到该组播 VLAN 内，使其成为组播 VLAN 的子 VLAN。

表1-2 配置基于子 VLAN 的组播 VLAN

操作	命令	说明
进入系统视图	**system-view**	-
配置指定VLAN为组播VLAN，并进入组播VLAN视图	**multicast-vlan** *vlan-id*	必选 缺省情况下，VLAN不是组播VLAN

操作	命令	说明
向组播VLAN内添加子VLAN	**subvlan** *vlan-list*	必选 缺省情况下，组播VLAN内没有子VLAN

说明

- 在已使能了 IP 组播路由的设备上不允许再配置组播 VLAN。
- 要配置为组播 VLAN 的指定 VLAN 必须存在。
- 要添加到组播 VLAN 内的子 VLAN 必须存在，且不能是组播 VLAN 或其它组播 VLAN 的子 VLAN。
- 组播 VLAN 内子 VLAN 的总数不得超过系统限制。

1.4 配置基于端口的组播VLAN

在配置基于端口的组播 VLAN 时，需要先配置各用户端口的属性，然后再将配置好的用户端口添加到组播 VLAN 内。

说明

- 只允许将以太网端口或二层聚合接口类型的用户端口配置为组播 VLAN 的端口。
- 二层以太网端口视图下的配置只对当前端口有效；二层聚合接口视图下的配置对当前接口有效；端口组视图下的配置对当前端口组中的所有端口有效。

1.4.1 配置准备

在配置基于端口的组播 VLAN 之前，需完成以下任务：

- 创建相应的 VLAN
- 在欲配置为组播 VLAN 的 VLAN 内使能 IGMP Snooping
- 在所有的用户 VLAN 内都内使能 IGMP Snooping

1.4.2 配置用户端口属性

首先，配置用户端口为 Hybrid 类型，允许用户 VLAN 的报文通过，缺省 VLAN 为其所属的用户 VLAN。然后，配置用户端口允许组播 VLAN 的报文不带 Tag 通过。这样，当二层设备通过组播 VLAN 收到来自上游、打有组播 VLAN Tag 的组播数据报文时，会将其 Tag 去掉后再向下游转发。

表1-3 配置用户端口属性

操作	命令	说明
进入系统视图	**system-view**	-

操作		命令	说明
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置用户端口的链路类型为Hybrid类型		**port link-type hybrid**	必选 缺省情况下，端口的链路类型为Access类型
配置用户端口的缺省VLAN为其所属的用户VLAN		**port hybrid pvid vlan** *vlan-id*	必选 缺省情况下，Hybrid端口的缺省VLAN为VLAN 1
允许组播VLAN通过用户端口，且不携带Tag		**port hybrid vlan** *vlan-id-list* **untagged**	必选 缺省情况下，Hybrid端口只允许VLAN 1通过

说明

有关 **port link-type**、**port hybrid pvid vlan** 和 **port hybrid vlan** 命令的详细介绍，请参见“二层技术-以太网交换命令参考”中的“VLAN”。

1.4.3 配置组播VLAN端口

首先需要把某个 VLAN 配置为组播 VLAN，再将用户端口添加到该组播 VLAN 内——既可以在组播 VLAN 内添加端口，也可以在端口上指定其所属的组播 VLAN——这两种配置方式是等效的。

1. 在组播VLAN内配置组播VLAN端口

表1-4 在组播 VLAN 内配置组播 VLAN 端口

操作	命令	说明
进入系统视图	**system-view**	-
配置指定VLAN为组播VLAN，并进入组播VLAN视图	**multicast-vlan** *vlan-id*	必选 缺省情况下，VLAN不是组播VLAN
向组播VLAN内添加端口	**port** *interface-list*	必选 缺省情况下，组播VLAN内没有端口

2. 在端口上配置组播VLAN端口

表1-5 在端口上配置组播 VLAN 端口

操作	命令	说明
进入系统视图	**system-view**	-

操作		命令	说明
配置指定VLAN为组播VLAN，并进入组播VLAN视图		**multicast-vlan** *vlan-id*	必选 缺省情况下，VLAN不是组播VLAN
退回系统视图		**quit**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
指定端口所属的组播VLAN		**port multicast-vlan** *vlan-id*	必选 缺省情况下，端口不属于任何组播VLAN

说明

- 在已使能了 IP 组播路由的设备上不允许再配置组播 VLAN。
- 要配置为组播 VLAN 的指定 VLAN 必须存在。
- 一个端口只能属于一个组播 VLAN。

1.5 配置组播VLAN内IGMP Snooping转发表项的最大数量

用户可以调整组播 VLAN 内 IGMP Snooping 转发表项的最大数量，当所有组播 VLAN 内维护的表项总数达到最大数量后，将不再创建新的表项，直至有表项被老化或被手工删除。

表1-6 配置组播 VLAN 内 IGMP Snooping 转发表项的最大数量

操作	命令	说明
进入系统视图	**system-view**	-
配置组播VLAN内IGMP Snooping转发表项的最大数量	**multicast-vlan entry-limit** *limit*	必选 缺省情况下，组播VLAN内IGMP Snooping转发表项的最大数量为2000

说明

在对组播 VLAN 内 IGMP Snooping 转发表项的最大数量进行配置时，如果所有组播 VLAN 内的表项总数已超过了配置值，系统将提示用户删除多余表项，但并不会自动删除任何已存在的表项，同时也不再继续创建新的表项。

1.6 组播VLAN显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后组播 VLAN 的运行情况，通过查看显示信息验证配置的效果。

表1-7 组播 VLAN 显示和维护

操作	命令
查看组播VLAN的信息	**display multicast-vlan** [*vlan-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.7 组播VLAN典型配置举例

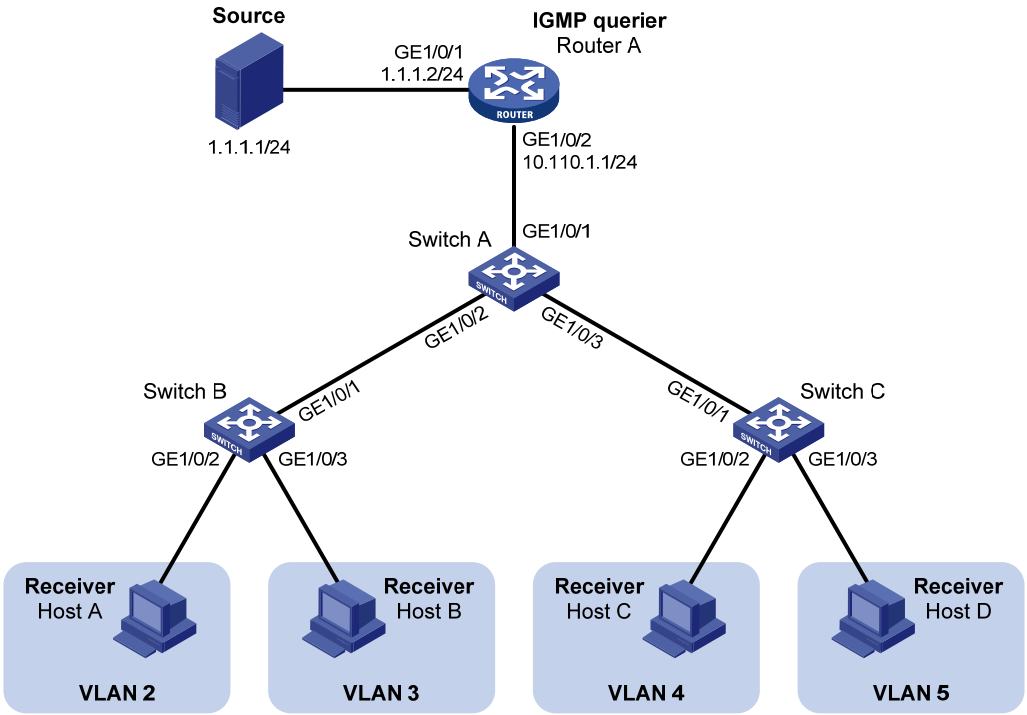
1.7.1 基于子VLAN的组播VLAN配置举例

1. 组网需求

- 如 图 1-4 所示，Router A通过端口GigabitEthernet1/0/1 连接组播源（Source），通过端口GigabitEthernet1/0/2 连接Switch A；Router A上运行IGMPv2，Switch A～Switch C上都运行版本 2 的IGMP Snooping，并由Router A充当IGMP查询器。
- 组播源向组播组 224.1.1.1 发送组播数据，Host A～Host D 都是该组播组的接收者（Receiver），且分别属于 VLAN 2～VLAN 5。
- 通过在 Switch A 上配置基于子 VLAN 的组播 VLAN，使 Router A 通过组播 VLAN 向 Switch A 下分属不同用户 VLAN 的主机分发组播数据。

2. 组网图

图1-4 基于子 VLAN 的组播 VLAN 配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-4 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 配置 Router A

使能IP组播路由，在各接口上使能PIM-DM，并在主机侧端口GigabitEthernet1/0/2上使能IGMP。

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] igmp enable
```

(3) 配置 Switch A

全局使能 IGMP Snooping。

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

创建 VLAN 2～VLAN 5。

```
[SwitchA] vlan 2 to 5
```

配置端口 GigabitEthernet1/0/2 为 Trunk 端口，并允许 VLAN 2 和 VLAN 3 通过。

```
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type trunk
[SwitchA-GigabitEthernet1/0/2] port trunk permit vlan 2 3
[SwitchA-GigabitEthernet1/0/2] quit
```

配置端口 GigabitEthernet1/0/3 为 Trunk 端口，并允许 VLAN 4 和 VLAN 5 通过。

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-GigabitEthernet1/0/3] port trunk permit vlan 4 5
[SwitchA-GigabitEthernet1/0/3] quit
```

创建 VLAN 10，把端口 GigabitEthernet1/0/1 添加到该 VLAN 中，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] igmp-snooping enable
[SwitchA-vlan10] quit
```

配置 VLAN 10 为组播 VLAN，并把 VLAN 2～VLAN 5 都配置为该组播 VLAN 的子 VLAN。

```
[SwitchA] multicast-vlan 10
[SwitchA-mvlan-10] subvlan 2 to 5
[SwitchA-mvlan-10] quit
```

(4) 配置 Switch B

全局使能 IGMP Snooping。

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

创建 VLAN 2，把端口 GigabitEthernet1/0/2 添加到该 VLAN 中，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/2
[SwitchB-vlan2] igmp-snooping enable
[SwitchB-vlan2] quit
```

创建 VLAN 3，把端口 GigabitEthernet1/0/3 添加到该 VLAN 中，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchB] vlan 3
[SwitchB-vlan3] port gigabitethernet 1/0/3
[SwitchB-vlan3] igmp-snooping enable
[SwitchB-vlan3] quit
```

配置端口 GigabitEthernet1/0/1 为 Trunk 端口，并允许 VLAN 2 和 VLAN 3 通过。

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port link-type trunk
[SwitchB-GigabitEthernet1/0/1] port trunk permit vlan 2 3
```

(5) 配置 Switch C

Switch C 的配置与 Switch B 相似，配置过程略。

(6) 检验配置效果

查看 Switch A 上所有组播 VLAN 的信息。

```
[SwitchA] display multicast-vlan
```

```
 Total 1 multicast-vlan(s)

 Multicast vlan 10
   subvlan list:
    vlan 2-5
   port list:
    no port
```

查看 Switch A 上 IGMP Snooping 组播组的信息。

```
[SwitchA] display igmp-snooping group
  Total 5 IP Group(s).
  Total 5 IP Source(s).
  Total 5 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):2.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 0 port(s).
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:224.1.1.1
        (0.0.0.0, 224.1.1.1):
          Host port(s):total 1 port(s).
            GE1/0/2                (D)
    MAC group(s):
      MAC group address:0100-5e01-0101
          Host port(s):total 1 port(s).
            GE1/0/2

  Vlan(id):3.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 0 port(s).
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:224.1.1.1
        (0.0.0.0, 224.1.1.1):
          Host port(s):total 1 port(s).
            GE1/0/2                (D)
    MAC group(s):
      MAC group address:0100-5e01-0101
          Host port(s):total 1 port(s).
            GE1/0/2

  Vlan(id):4.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
```

```
  Total 1 MAC Group(s).
  Router port(s):total 0 port(s).
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:224.1.1.1
      (0.0.0.0, 224.1.1.1):
        Host port(s):total 1 port(s).
          GE1/0/3               (D)
  MAC group(s):
    MAC group address:0100-5e01-0101
        Host port(s):total 1 port(s).
          GE1/0/3

Vlan(id):5.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 0 port(s).
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:224.1.1.1
      (0.0.0.0, 224.1.1.1):
        Host port(s):total 1 port(s).
          GE1/0/3               (D)
  MAC group(s):
    MAC group address:0100-5e01-0101
        Host port(s):total 1 port(s).
          GE1/0/3

Vlan(id):10.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
          GE1/0/1               (D)
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:224.1.1.1
      (0.0.0.0, 224.1.1.1):
        Host port(s):total 0 port(s).
  MAC group(s):
    MAC group address:0100-5e01-0101
        Host port(s):total 0 port(s).
```

由此可见，IGMP Snooping 在组播 VLAN（VLAN 10）中维护路由器端口，而在各子 VLAN（VLAN 2～VLAN 5）中维护各自的成员端口。

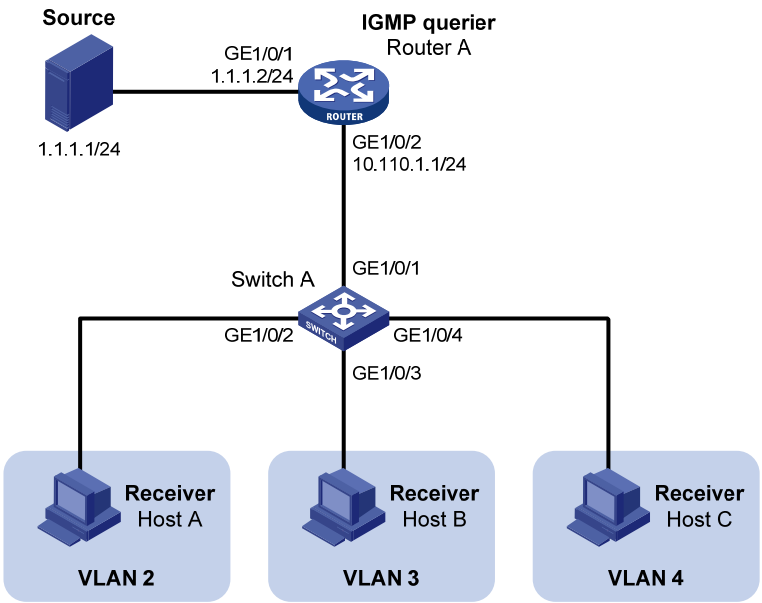
1.7.2 基于端口的组播VLAN配置举例

1. 组网需求

- 如 图 1-5 所示，Router A通过端口GigabitEthernet1/0/1 连接组播源（Source），通过端口GigabitEthernet1/0/2 连接Switch A；Router A上运行IGMPv2，Switch A上运行版本 2 的IGMP Snooping，并由Router A充当IGMP查询器。
- 组播源向组播组 224.1.1.1 发送组播数据，Host A～Host C 都是该组播组的接收者（Receiver），且分别属于 VLAN 2～VLAN 4。
- 通过在 Switch A 上配置基于端口的组播 VLAN，使 Router A 通过组播 VLAN 向 Switch A 下分属不同用户 VLAN 的主机分发组播数据。

2. 组网图

图1-5 基于端口的组播 VLAN 配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-5 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 配置 Router A

使能IP组播路由，在各接口上使能PIM-DM，并在主机侧端口GigabitEthernet1/0/2上使能IGMP。

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] igmp enable
```

(3) 配置 Switch A

全局使能 IGMP Snooping。

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

创建 VLAN 10，把端口 GigabitEthernet1/0/1 添加到该 VLAN 中，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] igmp-snooping enable
[SwitchA-vlan10] quit
```

创建 VLAN 2，并在该 VLAN 内使能 IGMP Snooping。

```
[SwitchA] vlan 2
[SwitchA-vlan2] igmp-snooping enable
[SwitchA-vlan2] quit
```

VLAN 3 与 VLAN 4 的配置与 VLAN 2 相似，配置过程略。

配置端口 GigabitEthernet1/0/2 为 Hybrid 类型，缺省 VLAN 为 VLAN 2；允许 VLAN 2 和 VLAN 10 的报文通过，且均不携带 Tag。

```
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type hybrid
[SwitchA-GigabitEthernet1/0/2] port hybrid pvid vlan 2
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 2 untagged
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 10 untagged
[SwitchA-GigabitEthernet1/0/2] quit
```

GigabitEthernet1/0/3 与 GigabitEthernet1/0/4 的配置与 GigabitEthernet1/0/2 相似，配置过程略。

配置 VLAN 10 为组播 VLAN。

```
[SwitchA] multicast-vlan 10
```

将端口 GigabitEthernet1/0/2 到 GigabitEthernet1/0/3 添加到组播 VLAN 10 内。

```
[SwitchA-mvlan-10] port gigabitethernet 1/0/2 to gigabitethernet 1/0/3
[SwitchA-mvlan-10] quit
```

配置端口 GigabitEthernet1/0/4 也属于组播 VLAN 10。

```
[SwitchA] interface gigabitethernet 1/0/4
[SwitchA-GigabitEthernet1/0/4] port multicast-vlan 10
[SwitchA-GigabitEthernet1/0/4] quit
```

(4) 检验配置效果

查看 Switch A 上所有组播 VLAN 的信息。

```
[SwitchA] display multicast-vlan
 Total 1 multicast-vlan(s)

 Multicast vlan 10
   subvlan list:
    no subvlan
   port list:
    GE1/0/2                 GE1/0/3                 GE1/0/4
```

查看 Switch A 上 IGMP Snooping 组播组的信息。

```
[SwitchA] display igmp-snooping group
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):10.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port(s).
            GE1/0/1                (D)
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:224.1.1.1
        (0.0.0.0, 224.1.1.1):
          Host port(s):total 3 port(s).
            GE1/0/2                (D)
            GE1/0/3                (D)
            GE1/0/4                (D)
    MAC group(s):
      MAC group address:0100-5e01-0101
          Host port(s):total 3 port(s).
            GE1/0/2
            GE1/0/3
            GE1/0/4
```

由此可见，IGMP Snooping 统一在组播 VLAN（VLAN 10）中维护路由器端口和成员端口。

目 录

1 组播路由与转发配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- 组播路由与转发功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

1.1 组播路由与转发简介

在组播实现中，组播路由和转发分为三种表：

- 每个组播路由协议都有一个协议自身的路由表，如 PIM 路由表（PIM Routing-Table）；
- 各组播路由协议的组播路由信息经过综合形成一个总的组播路由表（Multicast Routing-Table）；
- 组播转发表（Multicast Forwarding-Table）直接用于控制组播数据包的转发。

组播路由表由一组（S，G）表项组成，其中（S，G）表示由源 S 向组播组 G 发送组播数据的路由信息。如果路由器支持多种组播路由协议，则其组播路由表中将包括由多种协议生成的组播路由。路由器根据组播路由和转发策略，从组播路由表中选出最优的组播路由，并下发到组播转发表中。

1.1.1 RPF检查机制

组播路由协议依赖于现有的单播路由信息、MBGP 路由或组播静态路由来创建组播路由表项。组播路由协议在创建组播路由表项时，运用了 RPF（Reverse Path Forwarding，逆向路径转发）检查机制，以确保组播数据能够沿正确的路径传输，同时还能避免由于各种原因而造成的环路。

1. RPF检查过程

执行 RPF 检查的依据是单播路由、MBGP 路由或组播静态路由：

- 单播路由表中汇集了到达各个目的网段的最短路径；
- MBGP 路由表直接提供组播路由信息；
- 组播静态路由表中列出了用户通过手工静态配置指定的 RPF 路由信息。

在执行 RPF 检查时，路由器查找单播路由表、MBGP 路由表和组播静态路由表，具体过程如下：

(1) 首先，分别从单播路由表、MBGP 路由表和组播静态路由表中各选出一条最优路由：

- 以“报文源”的 IP 地址为目的地址查找单播路由表，自动选取一条最优单播路由。对应表项中的出接口为 RPF 接口，下一跳为 RPF 邻居。路由器认为来自 RPF 邻居且由该 RPF 接口收到的组播报文所经历的路径是从源 S 到本地的最短路径。
- 以“报文源”的 IP 地址为目的地址查找 MBGP 路由表，自动选取一条最优 MBGP 路由。对应表项中的出接口为 RPF 接口，下一跳为 RPF 邻居。

- 以“报文源”的 IP 地址为指定源地址查找组播静态路由表，自动选取一条最优组播静态路由。对应表项明确指定了 RPF 接口和 RPF 邻居。

(2) 然后，从这三条最优路由中选择一条作为 RPF 路由：

- 如果配置了按照最长匹配选择路由，则从这三条路由中选出最长匹配的那条路由；如果这三条路由的掩码一样，则选择其中优先级最高的那条路由；如果它们的优先级也相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。
- 如果没有配置按照最长匹配选择路由，则从这三条路由中选出优先级最高的那条路由；如果它们的优先级相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。

说明

根据组播报文传输的具体情况不同，“报文源”所代表的具体含义也不同：

- 如果当前报文沿从组播源到接收者或 RP（Rendezvous Point，汇集点）的 SPT（Shortest Path Tree，最短路径树）进行传输，则以组播源为“报文源”进行 RPF 检查；
- 如果当前报文沿从 RP 到接收者的 RPT（Rendezvous Point Tree，共享树）进行传输，或者沿从组播源到 RP 的组播源侧 RPT 进行传输，则都以 RP 为“报文源”进行 RPF 检查；
- 如果当前报文为 BSR（Bootstrap Router，自举路由器）报文，沿从 BSR 到各路由器的路径进行传输，则以 BSR 为“报文源”进行 RPF 检查。

有关 SPT、RPT、组播源侧 RPT、RP 和 BSR 的详细介绍，请参见“IP 组播配置指导”中的“PIM”。

---

2. RPF检查在组播转发中的应用

对每一个收到的组播数据报文都进行 RPF 检查会给路由器带来较大负担，而利用组播转发表可以解决这个问题。在建立组播路由和转发表时，会把组播数据报文（S，G）的 RPF 接口记录为（S，G）表项的入接口。当路由器收到组播数据报文（S，G）后，查找组播转发表：

(1) 如果组播转发表中不存在（S，G）表项，则对该报文执行 RPF 检查，将其 RPF 接口作为入接口，结合相关路由信息创建相应的表项，并下发到组播转发表中：

- 若该报文实际到达的接口正是其 RPF 接口，则 RPF 检查通过，向所有的出接口转发该报文；
- 若该报文实际到达的接口不是其 RPF 接口，则 RPF 检查失败，丢弃该报文。

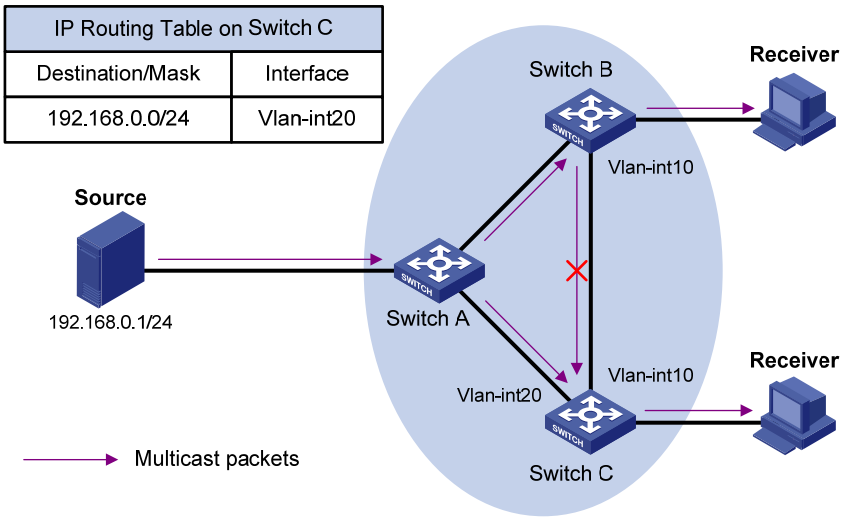
(2) 如果组播转发表中已存在（S，G）表项，且该报文实际到达的接口与入接口相匹配，则向所有的出接口转发该报文。

(3) 如果组播转发表中已存在（S，G）表项，但该报文实际到达的接口与入接口不匹配，则对此报文执行 RPF 检查：

- 若其 RPF 接口与入接口一致，则说明（S，G）表项正确，丢弃这个来自错误路径的报文；
- 若其 RPF 接口与入接口不符，则说明（S，G）表项已过时，于是把入接口更新为 RPF 接口。如果该报文实际到达的接口正是其 RPF 接口，则向所有的出接口转发该报文，否则将其丢弃。

如 图 1-1 所示，假设网络中单播路由畅通，未配置MBGP，Switch C上也未配置组播静态路由。组播报文（S，G）沿从组播源（Source）到接收者（Receiver）的SPT进行传输。假定Switch C上的组播转发表中已存在（S，G）表项，其记录的入接口为Vlan-interface20。

图1-1 RPF 检查过程



- 如果该组播报文从接口 Vlan-interface20 到达 Switch C，与（S，G）表项的入接口相匹配，则向所有的出接口转发该报文。
- 如果该组播报文从接口 Vlan-interface10 到达 Switch C，与（S，G）表项的入接口不匹配，则对其执行 RPF 检查：通过查找单播路由表发现到达 Source 的出接口（即 RPF 接口）是 Vlan-interface20，与（S，G）表项的入接口一致。这说明（S，G）表项是正确的，该报文来自错误的路径，RPF 检查失败，于是丢弃该报文。

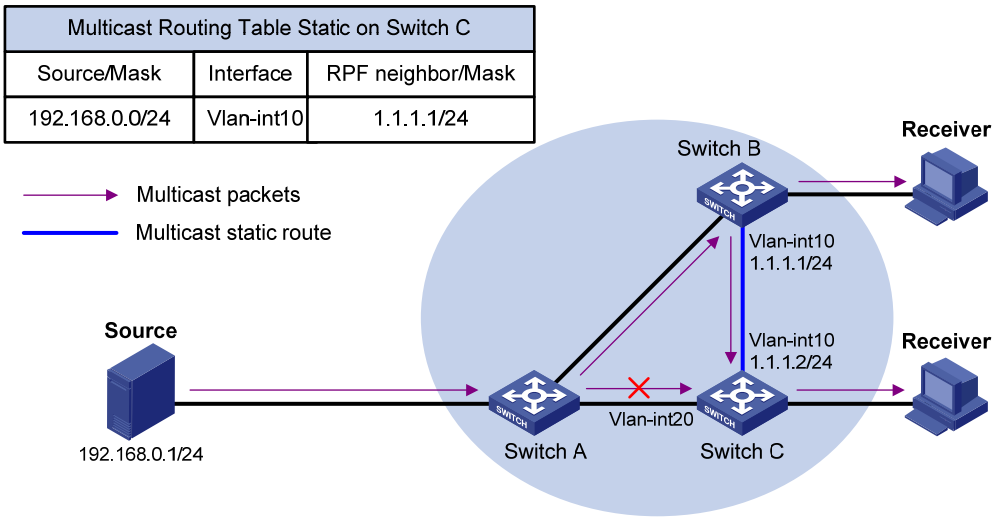
1.1.2 组播静态路由

组播静态路由是 RPF 检查的重要依据之一。根据具体应用环境的不同，组播静态路由有以下两种主要用途：

1. 改变RPF路由

通常，组播的网络拓扑结构与单播相同，组播数据的传输路径也与单播相同。可以通过配置组播静态路由以改变 RPF 路由，从而为组播数据创建一条与单播不同的传输路径。

图1-2 改变 RPF 路由示意图

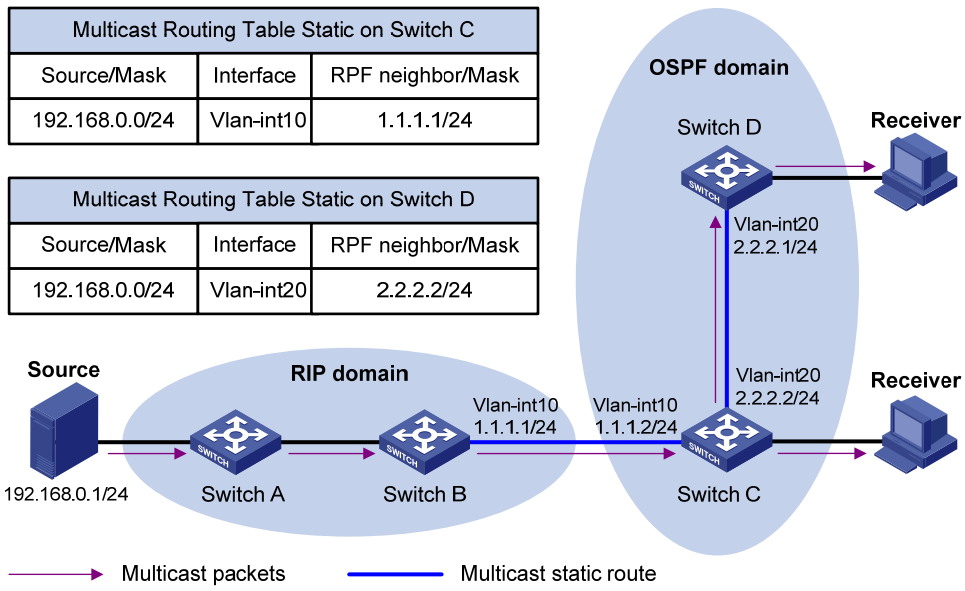


如 图 1-2 所示，当网络中没有配置组播静态路由时，Switch C到组播源（Source）的RPF邻居为Switch A，从Source发出的组播信息沿Switch A—Switch C的路径传输，与单播路径一致；当在Switch C上配置了组播静态路由，指定从Switch C到Source的RPF邻居为Switch B之后，从Source发出的组播信息将改变传输路径，沿Switch A—Switch B—Switch C的新路径传输。

2. 衔接RPF路由

当网络中的单播路由被阻断时，由于没有 RPF 路由而无法进行包括组播数据在内的数据转发。可以通过配置组播静态路由以生成 RPF 路由，从而创建组播路由表项以指导组播数据的转发。

图1-3 衔接 RPF 路由示意图



如 图 1-3 所示，RIP域与OSPF域之间实行单播路由隔离。当网络中没有配置组播静态路由时，OSPF域内的接收者（Receiver）不能收到RIP域内的组播源（Source）所发出的组播信息；当在Switch C

和Switch D上均配置了组播静态路由，分别指定从Switch C到Source的RPF邻居为Switch B、从Switch D到Source的RPF邻居为Switch C之后，Receiver便能收到Source发出的组播信息了。

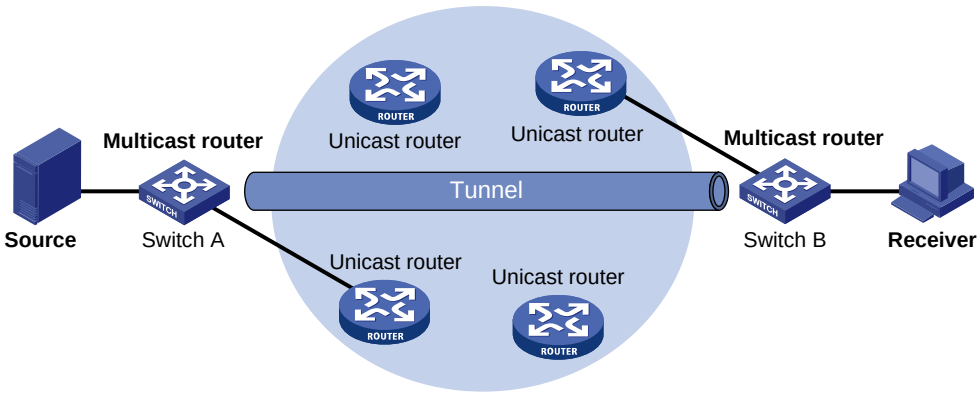
说明

- 组播静态路由的作用只在于影响 RPF 检查，而不能用于指导组播数据转发，故又称为 RPF 静态路由；
- 组播静态路由仅在所配置的组播路由器上生效，不会以任何方式被广播或者引入给其它路由器。

1.1.3 跨越单播网段的组播转发

网络中可能存在不支持组播协议的路由器。从组播源（Source）发出的组播数据沿组播路由器逐跳转发，当下一跳路由器不支持组播协议时，组播转发路径将被阻断。此时，通过在处于单播网段两端的组播路由器之间建立隧道，例如 IPv4 over IPv4 隧道、GRE（Generic Routing Encapsulation，通用路由封装）隧道等，可以实现跨越单播网段的组播数据交换。有关隧道和 GRE 隧道的详细介绍，请参见“三层技术-IP 业务配置指导”中的“隧道”和“GRE”。

图1-4 使用隧道方式传送组播数据



如 图 1-4 所示，在Switch A和Switch B之间建立起GRE隧道。Switch A将组播数据包封装在单播IP报文中，经由单播路由器转发，传送到隧道另一端的Switch B。然后，Switch B将单播IP报文头剥掉，继续进行组播传输。

如果在隧道的两端配置了单播静态路由，则任意单播数据包都可以通过该隧道传输。为了将该隧道专用于组播数据包的传输，可以在隧道两端只配置组播静态路由，从而使单播数据包的传输不能再使用此隧道。

1.1.4 组播路径跟踪

Multicast traceroute（组播路径跟踪）用来跟踪组播数据从第一跳到最后一跳路由器所经过的路径。

1. Multicast traceroute基本概念

(1) 最后一跳路由器（Last-hop Router）：如果某路由器有一个接口的 IP 地址与指定地址在同一个网段内，具备组播功能，且能够向该网段转发特定组播源发来的组播流，则称该路由器为最后一跳路由器。

(2) 第一跳路由器（First-hop Router）：与组播源直连的路由器。

(3) 查询器（Querier）：触发组播路径跟踪的路由器。

2. Multicast traceroute报文简介

Multicast traceroute 报文是一种特殊的 IGMP 报文，与普通 IGMP 报文的区别在于其类型字段为 0x1F/0x1E，且其目的 IP 地址为单播地址。Multicast traceroute 报文分为以下三种类型：

- Query 报文：类型字段为 0x1F
- Request 报文：类型字段为 0x1F
- Response 报文：类型字段为 0x1E

3. Multicast traceroute执行步骤

(1) 查询器向最后一跳路由器发送 Query 报文；

(2) 最后一跳路由器在收到的 Query 报文后加上本地响应数据块转换成 Request 报文，查找到上游邻居后向其单播发送该 Request 报文；

(3) 最后一跳路由器到组播源之间的每一跳都在 Request 报文之后附加一个响应数据块，并向其上游邻居单播转发；

(4) 第一跳路由器在收到 Request 报文后，将其报文类型改为 Response 报文，向查询器单播发送完整的报文。

1.2 组播路由与转发配置任务简介

表1-1 组播路由与转发配置任务简介

配置任务		说明	详细配置
使能IP组播路由		必选	1.3
配置组播路由与转发	配置组播静态路由	可选	1.4.2
	配置组播路由策略	可选	1.4.3
	配置组播转发范围	可选	1.4.4
	配置组播转发表容量	可选	1.4.5
	跟踪组播数据的传输路径	可选	1.4.6

注意

IP 组播不支持从 IP 地址网段应用。也就是说，当为接口配置了从 IP 地址后，组播数据并不能通过从 IP 地址所属的网段进行路由和转发，而只能通过主 IP 地址进行路由和转发。有关主、从 IP 地址的详细介绍，请参见“三层技术-IP 业务配置指导”中的“IP 地址”。

1.3 使能IP组播路由

在配置各项三层组播功能之前，必须首先使能 IP 组播路由。

1. 使能公网实例中的IP组播路由

表1-2 使能公网实例中的 IP 组播路由

操作	命令	说明
进入系统视图	**system-view**	-
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态

2. 使能VPN实例中的IP组播路由

表1-3 使能 VPN 实例中的 IP 组播路由

操作	命令	说明
进入系统视图	**system-view**	-
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态


说明

有关 **ip vpn-instance** 和 **route-distinguisher** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。

1.4 配置组播路由与转发

1.4.1 配置准备

在配置组播路由与转发之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 配置 PIM-DM（或 PIM-SM）

在配置组播路由与转发之前，需准备以下数据：

- 单条组播转发表项的最大下行节点数目
- 组播转发表的最大表项数

1.4.2 配置组播静态路由

通过配置组播静态路由，可以为来自特定组播源的组播报文指定 RPF 接口或 RPF 邻居。在删除已配置好的组播静态路由时，既可通过 **undo ip rpf-route-static** 命令删除指定的组播静态路由，也可通过 **delete ip rpf-route-static** 命令删除所有的组播静态路由。

表1-4 配置组播静态路由

操作	命令	说明
进入系统视图	**system-view**	-
配置组播静态路由	**ip rpf-route-static** [**vpn-instance** *vpn-instance-name*] *source-address* { *mask* | *mask-length* } [*protocol* [*process-id*]] [**route-policy** *policy-name*] { *rpf-nbr-address* | *interface-type interface-number* } [**preference** *preference*] [**order** *order-number*]	必选 缺省情况下，没有配置组播静态路由
删除所有组播静态路由	**delete ip rpf-route-static** [**vpn-instance** *vpn-instance-name*]	可选

注意

在配置组播静态路由时，若 RPF 邻居的接口类型是三层以太网端口、三层聚合接口、Loopback 接口或 VLAN 接口时，则不能用指定接口（*interface-type interface-number*）的方式来指定 RPF 邻居，而只能用指定地址（*rpf-nbr-address*）的方式。

1.4.3 配置组播路由策略

可以配置组播路由器按照最长匹配原则来选择RPF路由，有关RPF路由选择的详细介绍，请参见“1.1.1 1. RPF检查过程”一节。此外，通过配置根据组播源或组播源组进行组播流量的负载分担，还可以优化存在多条组播数据流时的网络流量。

1. 配置公网实例中的组播路由策略

表1-5 配置公网实例中的组播路由策略

操作	命令	说明
进入系统视图	**system-view**	-
配置按照最长匹配选择RPF路由	**multicast longest-match**	可选 缺省情况下，选择优先级最高的路由作为RPF路由
配置对组播流量进行负载分担	**multicast load-splitting { source | source-group }**	可选 缺省情况下，不对组播流量进行负载分担

2. 配置VPN实例中的组播路由策略

表1-6 配置 VPN 实例中的组播路由策略

操作	命令	说明
进入系统视图	**system-view**	-
进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置按照最长匹配选择RPF路由	**multicast longest-match**	可选 缺省情况下，选择优先级最高的路由作为RPF路由
配置对组播流量进行负载分担	**multicast load-splitting** { **source** | **source-group** }	可选 缺省情况下，不对组播流量进行负载分担

注意

multicast load-splitting 命令对双向 PIM 不生效。

1.4.4 配置组播转发范围

组播信息在网络中的转发并不是漫无边际的，每个组播组对应的组播信息都必须在确定的范围内传递。目前可以在所有支持组播转发的接口上配置针对某个组播组的转发边界。组播转发边界为指定范围的组播组划定了边界条件，如果组播报文的目的地址与边界条件匹配，就停止转发。当在一个接口上配置了组播转发边界后，将不能从该接口转发组播报文（包括本机发出的组播报文），也不能从该接口接收组播报文。

表1-7 配置组播转发范围

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置组播转发边界	**multicast boundary** *group-address* { *mask* | *mask-length* }	必选 缺省情况下，没有配置组播转发边界

1.4.5 配置组播转发表容量

路由器为每个收到的组播数据报文都维护相应的转发表项。但是，组播转发表项过多可能会耗尽路由器内存，从而导致路由器性能下降。用户可以根据实际组网情况和业务性能要求对组播转发表中的表项数量进行限制。如果组播转发表最大表项数的配置值小于当前值，则超出数目的表项并不会立刻被删除，而必须由组播路由协议来删除，同时也无法添加新的组播转发表项。

路由器为每个下行节点复制一份组播数据报文并发送出去，每个下行节点就形成组播分发树的一条分支。用户可以根据实际组网情况和业务性能要求对组播转发表中单条表项的下行节点数目（即出

接口数目）进行限制，以缓解路由器的复制压力。如果单条组播转发表项的最大下行节点数目的配置值小于当前值，则超出数目的下行节点并不会被立刻删除，而必须由组播路由协议来删除，同时新增的下行节点将无法添加到该表项中。

1. 配置公网实例中的组播转发表容量

表1-8 配置公网实例中的组播转发表容量

操作	命令	说明
进入系统视图	**system-view**	-
配置组播转发表的最大表项数	**multicast forwarding-table route-limit** *limit*	可选 缺省情况下，组播转发表的最大表项数为2000
配置单条组播转发表项的最大下行节点数目	**multicast forwarding-table downstream-limit** *limit*	可选 缺省情况下，单条组播转发表项的最大下行节点数目为128

2. 配置VPN实例中的组播转发表容量

表1-9 配置 VPN 实例中的组播转发表容量

操作	命令	说明
进入系统视图	**system-view**	-
进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置组播转发表的最大表项数	**multicast forwarding-table route-limit** *limit*	可选 缺省情况下，组播转发表的最大表项数为2000
配置单条组播转发表项的最大下行节点数目	**multicast forwarding-table downstream-limit** *limit*	可选 缺省情况下，单条组播转发表项的最大下行节点数目为128

1.4.6 跟踪组播数据的传输路径

可在任意视图下执行 **mtracert** 命令来跟踪组播数据从第一跳到最后一跳路由器所经过的路径。

表1-10 跟踪组播数据的传输路径

操作	命令	说明
跟踪组播数据的传输路径	**mtracert** *source-address* [[*last-hop-router-address*] *group-address*]	必选 可在任意视图下执行

1.5 组播路由与转发显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后组播路由与转发的信息，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除组播路由与转发的统计信息。

表1-11 组播路由与转发显示和维护

操作	命令
查看组播边界信息	**display multicast** [**all-instance** | **vpn-instance** *vpn-instance-name*] **boundary** [*group-address* [*mask* | *mask-length*]] [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看组播转发表信息	**display multicast** [**all-instance** | **vpn-instance** *vpn-instance-name*] **forwarding-table** [*source-address* [**mask** { *mask* | *mask-length* }] | *group-address* [**mask** { *mask* | *mask-length* }] | **incoming-interface** { *interface-type interface-number* | **register** } | **outgoing-interface** { **exclude** | **include** | **match** } { *interface-type interface-number* | **register** } | **statistics** | **slot** *slot-number*] * [**port-info**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看组播转发表的DF信息	**display multicast** [**all-instance** | **vpn-instance** *vpn-instance-name*] **forwarding-table df-info** [*rp-address*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看组播路由表信息	**display multicast** [**all-instance** | **vpn-instance** *vpn-instance-name*] **routing-table** [*source-address* [**mask** { *mask* | *mask-length* }] | *group-address* [**mask** { *mask* | *mask-length* }] | **incoming-interface** { *interface-type interface-number* | **register** } | **outgoing-interface** { **exclude** | **include** | **match** } { *interface-type interface-number* | **register** }] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看组播静态路由信息	**display multicast routing-table** [**all-instance** | **vpn-instance** *vpn-instance-name*] **static** [*source-address* { *mask-length* | *mask* }] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看组播源的RPF信息	**display multicast** [**all-instance** | **vpn-instance** *vpn-instance-name*] **rpf-info** *source-address* [*group-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除组播转发表中的转发项	**reset multicast** [**all-instance** | **vpn-instance** *vpn-instance-name*] **forwarding-table** { { *source-address* [**mask** { *mask* | *mask-length* }] | *group-address* [**mask** { *mask* | *mask-length* }] | **incoming-interface** { *interface-type interface-number* | **register** } } * | **all** }
清除组播路由表中的路由项	**reset multicast** [**all-instance** | **vpn-instance** *vpn-instance-name*] **routing-table** { { *source-address* [**mask** { *mask* | *mask-length* }] | *group-address* [**mask** { *mask* | *mask-length* }] | **incoming-interface** { *interface-type interface-number* | **register** } } * | **all** }

说明

有关 DF（Designated Forwarder，指定转发者）的详细介绍，请参见“IP 组播配置指导”中的“PIM”。

注意

- 执行 **reset** 命令将清除组播路由表或组播转发表中的信息，可能导致组播信息无法正常传输；
- 清除组播路由表中的路由项后，组播转发表中的相应表项也将随之删除；
- 清除组播转发表中的转发项后，组播路由表中的相应表项也将随之删除。

1.6 组播路由与转发典型配置举例

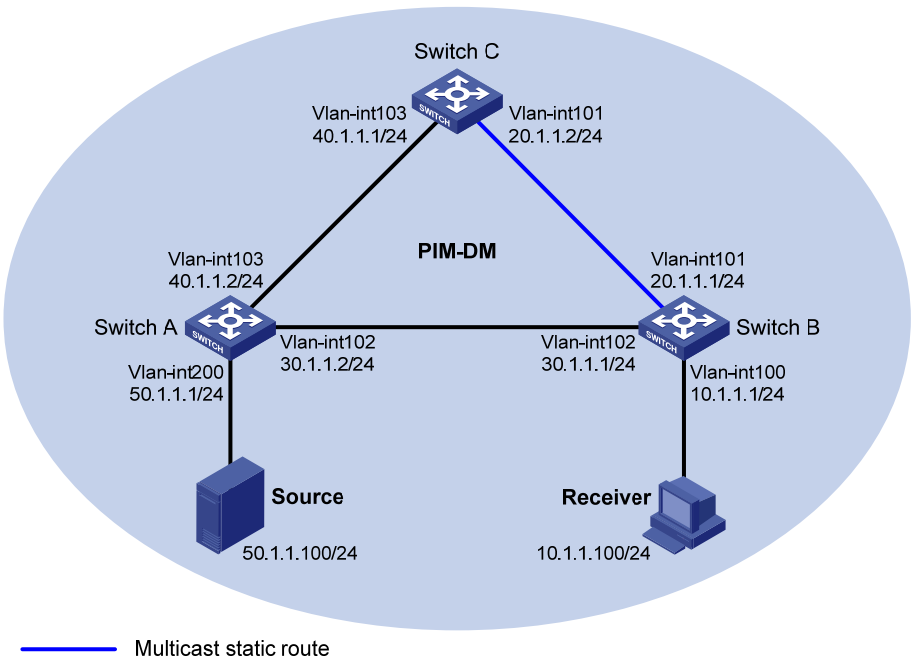
1.6.1 改变RPF路由配置举例

1. 组网需求

- 网络中运行 PIM-DM，所有交换机都支持组播功能；
- Switch A、Switch B 和 Switch C 之间运行 OSPF 协议；
- 通常情况下，Receiver 能通过 Switch A—Switch B 这条与单播路径相同的路径接收来自 Source 的组播信息；
- 要求通过配置，使 Receiver 能通过 Switch A—Switch C—Switch B 这条与单播路径不同的路径接收来自 Source 的组播信息。

2. 组网图

图1-5 改变 RPF 路由配置举例



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-5 配置各接口的IP地址和掩码，具体配置过程略。

配置 PIM-DM 域内的各交换机之间采用 OSPF 协议进行互连，确保 PIM-DM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-DM 和 IGMP

在 Switch B 上使能 IP 组播路由，在各接口上使能 PIM-DM，并在主机侧接口 Vlan-interface100 上使能 IGMP。

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] igmp enable
[SwitchB-Vlan-interface100] pim dm
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim dm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim dm
[SwitchB-Vlan-interface102] quit
```

在 Switch A 上使能 IP 组播路由，并在各接口上使能 PIM-DM。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] pim dm
[SwitchA-Vlan-interface200] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim dm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim dm
[SwitchA-Vlan-interface103] quit
```

Switch C 上的配置与 Switch A 相似，配置过程略。

在 Switch B 上使用 **display multicast rpf-info** 命令查看到 Source 的 RPF 信息。

```
[SwitchB] display multicast rpf-info 50.1.1.100
 RPF information about source 50.1.1.100:
     RPF interface: Vlan-interface102, RPF neighbor: 30.1.1.2
     Referenced route/mask: 50.1.1.0/24
     Referenced route type: igp
     Route selection rule: preference-preferred
     Load splitting rule: disable
```

Switch B 上当前的 RPF 路由来源于单播路由，RPF 邻居是 Switch A。

(3) 配置组播静态路由

在 Switch B 上配置组播静态路由，指定到 Source 的 RPF 邻居为 Switch C。

```
[SwitchB] ip rpf-route-static 50.1.1.100 24 20.1.1.2
```

(4) 检验配置效果

在 Switch B 上使用 **display multicast rpf-info** 命令查看到 Source 的 RPF 信息。

```
[SwitchB] display multicast rpf-info 50.1.1.100
```

```
RPF information about source 50.1.1.100:
    RPF interface: Vlan-interface101, RPF neighbor: 20.1.1.2
    Referenced route/mask: 50.1.1.0/24
    Referenced route type: multicast static
    Route selection rule: preference-preferred
    Load splitting rule: disable
```

与配置组播静态路由前相比，Switch B 上的 RPF 路由已经产生了变化，其来源变为组播静态路由，RPF 邻居变为 Switch C。

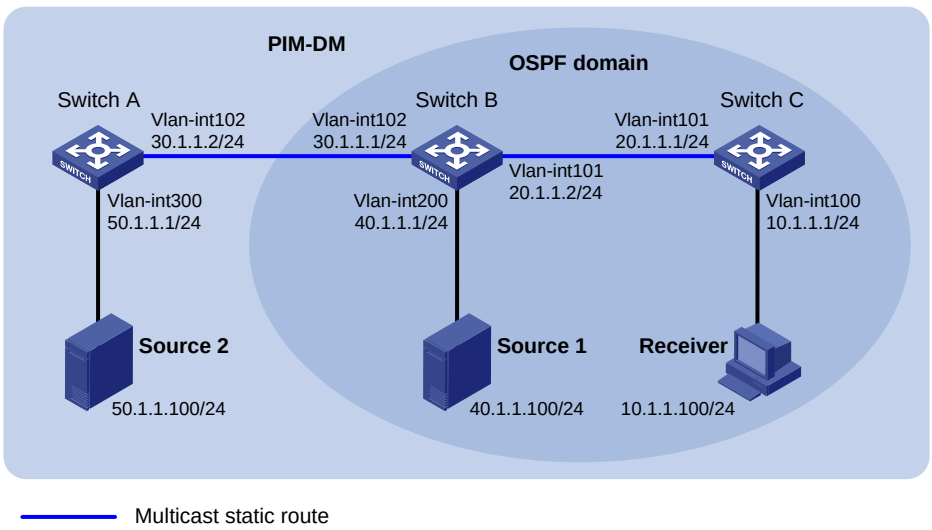
1.6.2 衔接RPF路由配置举例

1. 组网需求

- 网络中运行 PIM-DM，所有交换机都支持组播功能；
- Switch B 和 Switch C 之间运行 OSPF 协议，并与 Switch A 单播路由隔离；
- 通常情况下，Receiver 能接收来自 OSPF 域内 Source 1 的组播信息；
- 要求通过配置，使 Receiver 也可以接收来自 OSPF 域外 Source 2 的组播信息。

2. 组网图

图1-6 衔接 RPF 路由配置组网图



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-6 配置各接口的IP地址和掩码，具体配置过程略。

配置 Switch B 和 Switch C 之间采用 OSPF 协议进行互连，确保 Switch B 和 Switch C 之间在网络层互通，并且能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-DM 和 IGMP

在 Switch C 上使能 IP 组播路由，在各接口上使能 PIM-DM，并在主机侧接口 Vlan-interface100 上使能 IGMP。

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
```

```
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] igmp enable
[SwitchC-Vlan-interface100] pim dm
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 101
[SwitchC-Vlan-interface101] pim dm
[SwitchC-Vlan-interface101] quit
```

在 Switch A 上使能 IP 组播路由，并在各接口上使能 PIM-DM。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 300
[SwitchA-Vlan-interface300] pim dm
[SwitchA-Vlan-interface300] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim dm
[SwitchA-Vlan-interface102] quit
```

Switch B 上的配置与 Switch A 相似，配置过程略。

在 Switch B 和 Switch C 上分别使用 **display multicast rpf-info** 命令查看到 Source 2 的 RPF 信息。

```
[SwitchB] display multicast rpf-info 50.1.1.100
[SwitchC] display multicast rpf-info 50.1.1.100
```

没有显示信息输出，说明在 Switch B 和 Switch C 上都没有到 Source 2 的 RPF 路由。

(3) 配置组播静态路由

在 Switch B 上配置组播静态路由，指定到 Source 2 的 RPF 邻居为 Switch A。

```
[SwitchB] ip rpf-route-static 50.1.1.100 24 30.1.1.2
```

在 Switch C 上配置组播静态路由，指定到 Source 2 的 RPF 邻居为 Switch B。

```
[SwitchC] ip rpf-route-static 50.1.1.100 24 20.1.1.2
```

(4) 检验配置效果

在 Switch B 和 Switch C 上分别使用 **display multicast rpf-info** 命令查看到 Source 2 的 RPF 信息。

```
[SwitchB] display multicast rpf-info 50.1.1.100
 RPF information about source 50.1.1.100:
     RPF interface: Vlan-interface102, RPF neighbor: 30.1.1.2
     Referenced route/mask: 50.1.1.0/24
     Referenced route type: multicast static
     Route selection rule: preference-preferred
     Load splitting rule: disable
[SwitchC] display multicast rpf-info 50.1.1.100
 RPF information about source 50.1.1.100:
     RPF interface: Vlan-interface101, RPF neighbor: 20.1.1.2
     Referenced route/mask: 50.1.1.0/24
     Referenced route type: multicast static
     Route selection rule: preference-preferred
     Load splitting rule: disable
```

与配置组播静态路由前相比，Switch B 和 Switch C 上都有了到 Source 2 的 RPF 路由，且其均来源于组播静态路由。

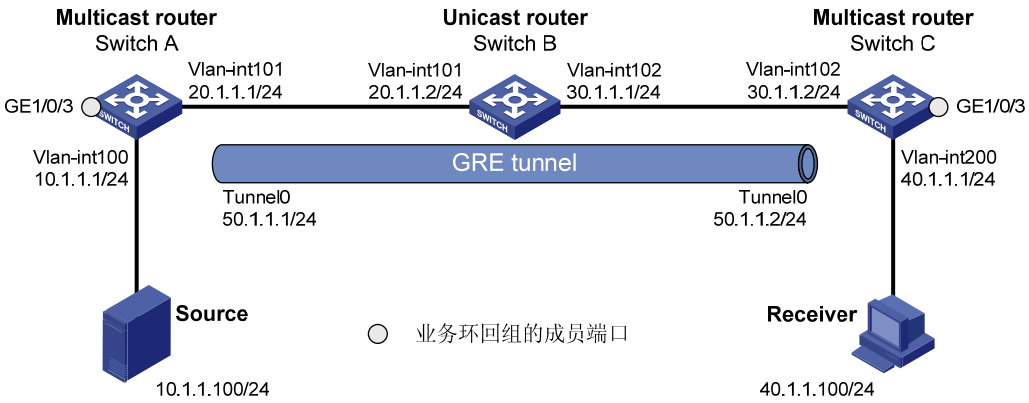
1.6.3 利用GRE隧道实现组播转发配置举例

1. 组网需求

- Switch A 和 Switch C 支持组播功能并运行 PIM-DM，但 Switch B 不支持组播功能；
- Switch A、Switch B 和 Switch C 之间运行 OSPF 协议；
- 要求通过配置，使 Receiver 能够接收来自 Source 的组播信息。

2. 组网图

图1-7 利用 GRE 隧道实现组播转发配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-7 配置各接口的IP地址和掩码，具体配置过程略。

(2) 配置 GRE 隧道

在 Switch A 上创建业务环回组 1，并指定其业务类型为 Tunnel 类型。

```
<SwitchA> system-view
[SwitchA] service-loopback group 1 type tunnel
```

在 Switch A 的端口 GigabitEthernet1/0/3（该端口不属于 VLAN 100 和 101）上关闭生成树协议、LLDP 功能和 NDP 功能，并将其加入业务环回组 1。

```
[SwitchA] interface gigabitethernet1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

在 Switch A 上创建接口 Tunnel0，并为其配置 IP 地址和掩码，并在该接口上引用业务环回组 1。

```
[SwitchA] interface tunnel 0
[SwitchA-Tunnel0] ip address 50.1.1.1 24
[SwitchA-Tunnel0] service-loopback-group 1
```

在 Switch A 上配置 Tunnel0 接口采用 GRE over IPv4 隧道模式，并为该接口指定源地址和目的地址。

```
[SwitchA-Tunnel0] tunnel-protocol gre
[SwitchA-Tunnel0] source 20.1.1.1
[SwitchA-Tunnel0] destination 30.1.1.2
[SwitchA-Tunnel0] quit
```

在 Switch C 上创建业务环回组 1，并指定其业务类型为 Tunnel 类型。

```
<SwitchC> system-view
[SwitchC] service-loopback group 1 type tunnel
```

在 Switch C 的端口 GigabitEthernet1/0/3（该端口不属于 VLAN 200 和 102）上关闭生成树协议、LLDP 功能和 NDP 功能，并将其加入业务环回组 1。

```
[SwitchC] interface gigabitethernet1/0/3
[SwitchC-GigabitEthernet1/0/3] undo stp enable
[SwitchC-GigabitEthernet1/0/3] undo ndp enable
[SwitchC-GigabitEthernet1/0/3] undo lldp enable
[SwitchC-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchC-GigabitEthernet1/0/3] quit
```

在 Switch C 上创建接口 Tunnel0，并为其配置 IP 地址和掩码，并在该接口上引用业务环回组 1。

```
[SwitchC] interface tunnel 0
[SwitchC-Tunnel0] ip address 50.1.1.2 24
[SwitchC-Tunnel0] service-loopback-group 1
```

在 Switch C 上配置 Tunnel0 接口采用 GRE over IPv4 隧道模式，并为该接口指定源地址和目的地址。

```
[SwitchC-Tunnel0] tunnel-protocol gre
[SwitchC-Tunnel0] source 30.1.1.2
[SwitchC-Tunnel0] destination 20.1.1.1
[SwitchC-Tunnel0] quit
```

(3) 配置 OSPF 协议

在 Switch A 上配置 OSPF 协议。

```
[SwitchA] ospf 1
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 20.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 50.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

在 Switch B 上配置 OSPF 协议。

```
<SwitchB> system-view
[SwitchB] ospf 1
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 20.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 30.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

在 Switch C 上配置 OSPF 协议。

```
[SwitchC] ospf 1
```

```
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 30.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 40.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 50.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

(4) 使能 IP 组播路由，并使能 PIM-DM 和 IGMP

在 Switch A 上使能 IP 组播路由，并在各接口上使能 PIM-DM。

```
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface tunnel 0
[SwitchA-Tunnel0] pim dm
[SwitchA-Tunnel0] quit
```

在 Switch C 上使能 IP 组播路由，在各接口上使能 PIM-DM，并在主机侧接口 Vlan-interface200 上使能 IGMP。

```
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] igmp enable
[SwitchC-Vlan-interface200] pim dm
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim dm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface tunnel 0
[SwitchC-Tunnel0] pim dm
[SwitchC-Tunnel0] quit
```

(5) 配置组播静态路由

在 Switch C 上配置组播静态路由，指定到 Source 的 RPF 邻居为 Switch A 的 Tunnel0 接口。

```
[SwitchC] ip rpf-route-static 10.1.1.0 24 50.1.1.1
```

(6) 检验配置效果

组播源向组播组 225.1.1.1 发送组播数据，接收者加入该组播组后能够收到组播源发来的组播数据。通过使用 **display pim routing-table** 命令可以查看交换机的 PIM 路由表信息。例如：

查看 Switch C 上的 PIM 路由表信息。

```
[SwitchC] display pim routing-table
 VPN-Instance: public net
 Total 1 (*, G) entry; 1 (S, G) entry

 (*, 225.1.1.1)
     Protocol: pim-dm, Flag: WC
     UpTime: 00:04:25
     Upstream interface: NULL
```

```
        Upstream neighbor: NULL
        RPF prime neighbor: NULL
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface200
            Protocol: igmp, UpTime: 00:04:25, Expires: never

(10.1.1.100, 225.1.1.1)
    Protocol: pim-dm, Flag: ACT
    UpTime: 00:06:14
    Upstream interface: Tunnel0
        Upstream neighbor: 50.1.1.1
        RPF prime neighbor: 50.1.1.1
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface200
            Protocol: pim-dm, UpTime: 00:04:25, Expires: never
```

Switch C 的 RPF 邻居为 Switch A，组播数据通过 GRE 隧道直接由 Switch A 发往 Switch C。

1.7 常见配置错误举例

1.7.1 组播静态路由失败

1. 故障现象

路由器没有配置动态路由协议，接口的物理状态与链路层协议状态都显示为 up；但是组播静态路由失败。

2. 分析

- 如果没有正确配置或更新与当前网络情况相匹配的组播静态路由，则组播路由表中不存在此路由项以及组播静态路由的配置信息；
- 如果查询到有比组播静态路由更优的路由，也可能导致组播静态路由失败。

3. 处理过程

(1) 使用 **display multicast routing-table static** 命令在组播路由表中查看组播静态路由的信息，以确定是否正确配置了对应的路由并存在于组播路由表中；

(2) 检查组播静态路由下一跳接口的接口类型。若为非点到点接口，则配置组播静态路由时，出接口必须使用下一跳地址的形式配置；

(3) 检查组播静态路由是否匹配指定的路由协议。如果配置组播静态路由时指定了协议，则使用 **display ip routing-table** 命令检查该协议是否添加了相同的路由；

(4) 检查组播静态路由是否匹配指定的路由策略。如果配置组播静态路由时指定了路由策略，则使用 **display route-policy** 命令检查配置的路由策略。

1.7.2 组播数据无法到达接收者

1. 故障现象

组播数据可以到达一些路由器，但无法到达最后一跳路由器。

2. 分析

通过在接口上使用 **multicast boundary** 命令可以设置组播转发边界，组播数据是无法跨越该边界的。

3. 处理过程

(1) 使用 **display pim routing-table** 命令查看各路由器上是否有（S，G）表项：如果有则表示收到了组播数据；否则表示没有收到组播数据。

(2) 使用 **display multicast boundary** 命令来查看接口的组播边界信息。使用 **multicast boundary** 命令来更改组播转发边界。

(3) 若采用了 PIM-SM，使用 **display current-configuration** 命令检查是否配置了 BSR 和 RP。

目 录

1 IGMP配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- IGMP 功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

1.1 IGMP简介

IGMP 是 Internet Group Management Protocol（互联网组管理协议）的简称。它是 TCP/IP 协议族中负责 IP 组播成员管理的协议，用来在 IP 主机和与其直接相邻的组播路由器之间建立、维护组播组成员关系。

1.1.1 IGMP的版本

到目前为止，IGMP 有三个版本：

- IGMPv1（由 RFC 1112 定义）
- IGMPv2（由 RFC 2236 定义）
- IGMPv3（由 RFC 3376 定义）

所有版本的 IGMP 都支持 ASM（Any-Source Multicast，任意信源组播）模型；IGMPv3 可以直接应用于 SSM（Source-Specific Multicast，指定信源组播）模型，而 IGMPv1 和 IGMPv2 则需要在 IGMP SSM Mapping 技术的支持下才能应用于 SSM 模型。

说明

有关 ASM 和 SSM 模型的介绍，请参见“IP 组播配置指导”中的“组播概述”。

1.1.2 IGMPv1 工作机制

IGMPv1 主要基于查询和响应机制来完成对组播组成员的管理。

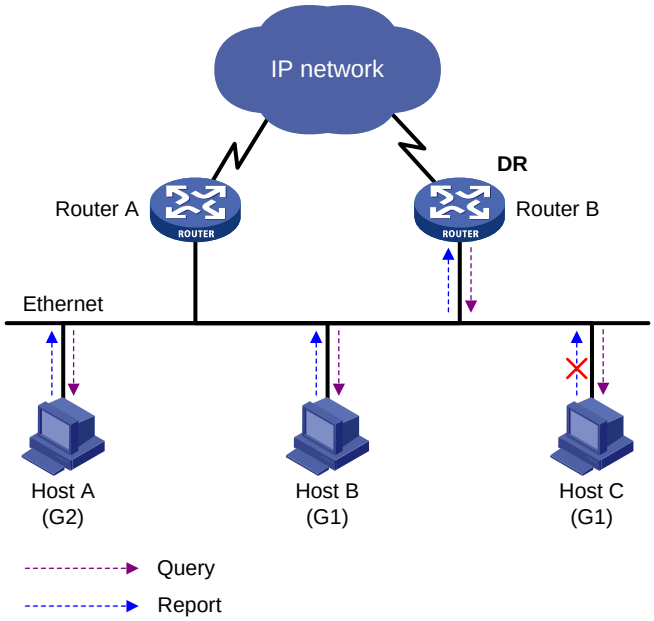
当一个网段内有多台组播路由器时，由于它们都能从主机那里收到 IGMP 成员关系报告报文（Membership Report Message），因此只需其中一台路由器发送 IGMP 查询报文（Query Message）即可，该路由器就称为 IGMP 查询器（Querier）。这就需要有一个查询器的选举机制来确定由哪台路由器作为 IGMP 查询器。

对于 IGMPv1 来说，由组播路由协议（如 PIM）选举出唯一的组播信息转发者 DR（Designated Router，指定路由器）作为 IGMP 查询器。

说明

有关 DR 的介绍，请参见“IP 组播配置指导”中的“PIM”。

图1-1 IGMP 查询响应示意图



如 图 1-1 所示，假设Host B与Host C想要收到发往组播组G1 的组播数据，而Host A想要收到发往组播组G2 的组播数据，那么主机加入组播组以及IGMP查询器（Router B）维护组播组成员关系的基本过程如下：

(1) 主机会主动向其要加入的组播组发送 IGMP 成员关系报告报文以声明加入，而不必等待 IGMP 查询器发来的 IGMP 查询报文；

(2) IGMP 查询器周期性地以组播方式向本地网段内的所有主机与路由器发送 IGMP 查询报文（目的地址为 224.0.0.1）；

(3) 在收到该查询报文后，关注 G1 的 Host B 与 Host C 其中之一（这取决于谁的延迟定时器先超时）——譬如 Host B 会首先以组播方式向 G1 发送 IGMP 成员关系报告报文，以宣告其属于 G1。由于本地网段中的所有主机和路由器都能收到 Host B 发往 G1 的报告报文，因此当 Host C 收到该报告报文后，将不再发送同样针对 G1 的报告报文，因为 IGMP 路由器（Router A 和 Router B）已知道本地网段中有对 G1 感兴趣的主机了。这个机制称为主机上的 IGMP 成员关系报告抑制机制，该机制有助于减少本地网段的信息流量；

(4) 与此同时，由于 Host A 关注的是 G2，所以它仍将以组播方式向 G2 发送报告报文，以宣告其属于 G2；

(5) 经过以上的查询和响应过程，IGMP 路由器了解到本地网段中有 G1 和 G2 的成员，于是由组播路由协议（如 PIM）生成（*，G1）和（*，G2）组播转发项作为组播数据的转发依据，其中的“*”代表任意组播源；

(6) 当由组播源发往 G1 或 G2 的组播数据经过组播路由到达 IGMP 路由器时，由于 IGMP 路由器上存在（*，G1）和（*，G2）组播转发项，于是将该组播数据转发到本地网段，接收者主机便能收到该组播数据了。

IGMPv1 没有专门定义离开组播组的报文。当运行 IGMPv1 的主机离开某组播组时，将不会向其要离开的组播组发送报告报文。当网段中不再存在该组播组的成员后，IGMP 路由器将收不到任何发往该组播组的报告报文，于是 IGMP 路由器在一段时间之后便删除该组播组所对应的组播转发项。

1.1.3 IGMPv2 的改进

与 IGMPv1 相比，IGMPv2 增加了查询器选举机制和离开组机制。

1. 查询器选举机制

在 IGMPv1 中，当某共享网段上存在多个组播路由器时，由组播路由协议（如 PIM）选举的指定路由器充当查询器。

在 IGMPv2 中，增加了独立的查询器选举机制，其选举过程如下：

(1) 所有 IGMPv2 路由器在初始时都认为自己是查询器，并向本地网段内的所有主机和路由器发送 IGMP 普遍组查询（General Query）报文（目的地址为 224.0.0.1）；

(2) 本地网段中的其它 IGMPv2 路由器在收到该报文后，将报文的源 IP 地址与自己的接口地址作比较。通过比较，IP 地址最小的路由器将成为查询器，其它路由器成为非查询器（Non-Querier）；

(3) 所有非查询器上都会启动一个定时器（即其它查询器存在时间定时器 Other Querier Present Timer）。在该定时器超时前，如果收到了来自查询器的 IGMP 查询报文，则重置该定时器；否则，就认为原查询器失效，并发起新的查询器选举过程。

2. 离开组机制

在 IGMPv1 中，主机离开组播组时不会向组播路由器发出任何通知，导致组播路由器只能依靠组播组成员查询的响应超时来获知组播组成员的离开。

而在 IGMPv2 中，当一个主机离开某组播组时：

(1) 该主机向本地网段内的所有组播路由器（目的地址为 224.0.0.2）发送离开组（Leave Group）报文；

(2) 当查询器收到该报文后，向该主机所声明要离开的那个组播组发送特定组查询（Group-Specific Query）报文（目的地址字段和组地址字段均填充为所要查询的组播组地址）；

(3) 如果该网段内还有该组播组的其它成员，则这些成员在收到特定组查询报文后，会在该报文中所设定的最大响应时间（Max Response Time）内发送成员关系报告报文；

(4) 如果在最大响应时间内收到了该组播组其它成员发送的成员关系报告报文，查询器就会继续维护该组播组的成员关系；否则，查询器将认为该网段内已无该组播组的成员，于是不再维护这个组播组的成员关系。

1.1.4 IGMPv3 的改进

IGMPv3 在兼容和继承 IGMPv1 和 IGMPv2 的基础上，进一步增强了主机的控制能力，并增强了查询和报告报文的功能。

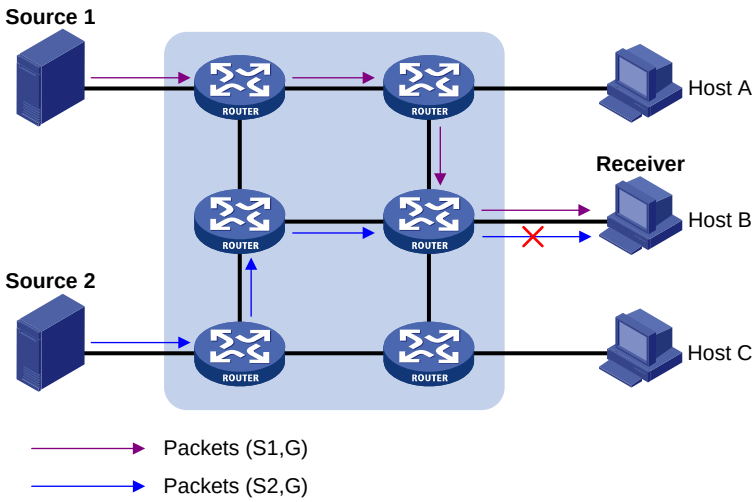
1. 主机控制能力的增强

IGMPv3 增加了针对组播源的过滤模式（INCLUDE/EXCLUDE），使主机在加入某组播组 G 的同时，能够明确要求接收或拒绝来自某特定组播源 S 的组播信息。当主机加入组播组时：

- 若要求只接收来自指定组播源如 S1、S2、……的组播信息，则其报告报文中可以标记为 INCLUDE Sources（S1，S2，……）；
- 若拒绝接收来自指定组播源如 S1、S2、……的组播信息，则其报告报文中可以标记为 EXCLUDE Sources（S1，S2，……）。

如 图 1-2 所示，网络中存在Source 1（S1）和Source 2（S2）两个组播源，均向组播组G发送组播报文。Host B仅对从Source 1 发往G的信息感兴趣，而对来自Source 2 的信息没有兴趣。

图1-2 指定源组的组播流路经



如果主机与路由器之间运行的是 IGMPv1 或 IGMPv2，Host B 加入组播组 G 时无法对组播源进行选择，因此无论 Host B 是否需要，来自 Source 1 和 Source 2 的组播信息都将传递给 Host B。

当主机与路由器之间运行了 IGMPv3 之后，Host B 就可以要求只接收来自 Source 1、发往 G 的组播信息（S1，G），或要求拒绝来自 Source 2、发往 G 的组播信息（S2，G），这样就只有来自 Source 1 的组播信息才能传递给 Host B 了。

2. 查询和报告报文功能的增强

(1) 携带源地址的查询报文

IGMPv3 不仅支持 IGMPv1 的普遍组查询和 IGMPv2 的特定组查询，而且还增加了对特定源组查询的支持：

- 普遍组查询报文中，既不携带组地址，也不携带源地址；
- 特定组查询报文中，携带组地址，但不携带源地址；
- 特定源组查询报文中，既携带组地址，还携带一个或多个源地址。

(2) 包含多组记录的报告报文

IGMPv3 报告报文的目的地址为 224.0.0.22，可以携带一个或多个组记录。在每个组记录中，包含有组播组地址和组播源地址列表。组记录可以分为多种类型，如下：

- IS_IN：表示组播组与组播源列表之间的过滤模式为 INCLUDE，即只接收从指定组播源列表发往该组播组的组播数据。

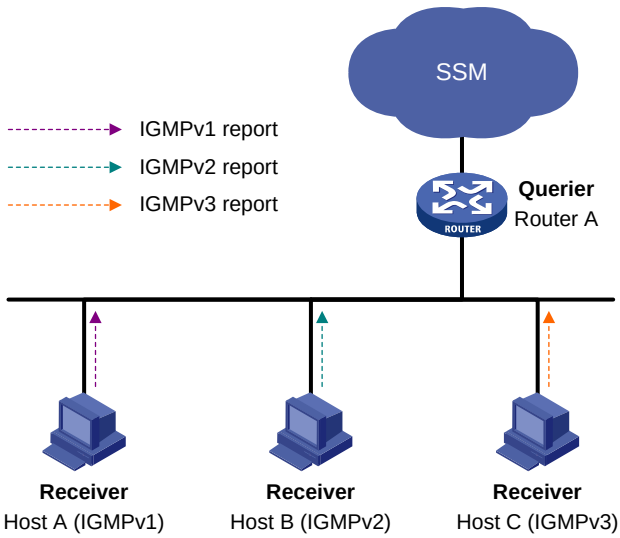
- IS_EX：表示组播组与组播源列表之间的过滤模式为 EXCLUDE，即只接收从指定组播源列表之外的组播源发往该组播组的组播数据。
- TO_IN：表示组播组与组播源列表之间的过滤模式由 EXCLUDE 转变为 INCLUDE。
- TO_EX：表示组播组与组播源列表之间的过滤模式由 INCLUDE 转变为 EXCLUDE。
- ALLOW：表示在现有状态的基础上，还希望从某些组播源接收组播数据。如果当前的对应关系为 INCLUDE，则向现有组播源列表中添加这些组播源；如果当前的对应关系为 EXCLUDE，则从现有组播源列表中删除这些组播源。
- BLOCK：表示在现有状态的基础上，不再希望从某些组播源接收组播数据。如果当前的对应关系为 INCLUDE，则从现有组播源列表中删除这些组播源；如果当前的对应关系为 EXCLUDE，则向现有组播源列表中添加这些组播源。

1.1.5 IGMP SSM Mapping

IGMP SSM Mapping 通过在路由器上配置 SSM 静态映射规则，从而为运行 IGMPv1 或 IGMPv2 的接收者主机提供对 SSM 模型的支持。

SSM 模型要求在接收者主机所在的网段，路由器能够了解主机加入组播组时所指定的组播源。如果接收者主机上运行的是 IGMPv3，则可以在 IGMPv3 的报告报文中直接指定组播源的地址；如果某些接收者主机只能运行 IGMPv1 或 IGMPv2，则在 IGMPv1 或 IGMPv2 的报告报文中无法指定组播源的地址。这种情况下需要通过在路由器上配置 IGMP SSM Mapping 功能，将 IGMPv1 或 IGMPv2 报告报文中所包含的（*，G）信息映射为（G，INCLUDE，（S1，S2...））信息。

图1-3 IGMP SSM Mapping 组网图



在如 图 1-3 所示的SSM网络中，Host A、Host B和Host C上分别运行IGMPv1、IGMPv2 和IGMPv3。在不允许将Host A和Host B升级为IGMPv3 的情况下，若要为Host A和Host B也提供SSM组播服务，则需在Router A上使能IGMP SSM Mapping并配置相应的映射规则。

配置完成后，当 Router A 收到来自主机的 IGMPv1 或 IGMPv2 报告报文时，首先检查该报文中所携带的组播组地址 G，然后根据检查结果的不同分别进行处理：

(1) 如果 G 不在 SSM 组地址范围内，则提供 ASM 组播服务。

(2) 如果 G 在 SSM 组地址范围内：

- 若 Router A 上没有 G 对应的 IGMP SSM Mapping 规则，则无法提供 SSM 组播服务，丢弃该报文；
- 若 Router A 上有 G 对应的 IGMP SSM Mapping 规则，则依据规则将报告报文中所包含的(*，G）信息映射为（G，INCLUDE，（S1，S2...））信息，可以提供 SSM 组播服务。

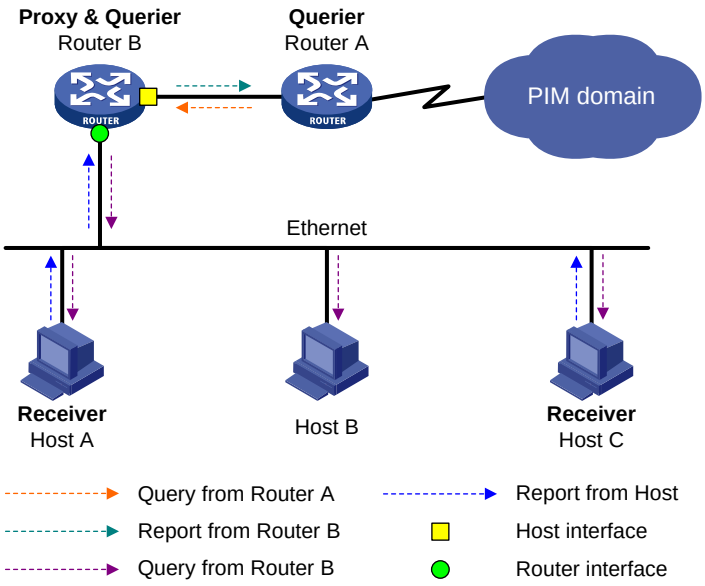
说明

- IGMP SSM Mapping 不对 IGMPv3 的报告报文进行处理。
- 有关 SSM 组地址范围的介绍，请参见“IP 组播配置指导”中的“PIM”。

1.1.6 IGMP Proxying

在一些简单的树型网络拓扑中，边缘设备上并不需要运行复杂的组播路由协议（如 PIM），可以通过在这些设备上配置 IGMP Proxying（IGMP 代理）功能，使其代理下游主机来发送 IGMP 报文及维护组成员关系，并基于该关系进行组播转发。在上游设备看来，配置了 IGMP Proxying 功能的设备（称为 IGMP 代理设备）不再是一个 PIM 邻居，而只是一台主机。

图1-4 IGMP Proxying 组网图



如 图 1-4 所示，IGMP Proxying中定义了以下两种接口类型：

- 上行接口：又称代理接口，指 IGMP 代理设备上运行 IGMP Proxying 功能的接口，即朝向组播分发树树根方向的接口。由于该接口执行 IGMP 协议的主机行为，因此也称为主机接口（Host Interface）。
- 下行接口：指 IGMP 代理设备上除上行接口外其它运行 IGMP 协议的接口，即背向组播分发树树根方向的接口。由于该接口执行 IGMP 协议的路由器行为，因此也称为路由器接口（Router Interface）。

IGMP 代理设备上维护着一个组成员关系数据库（Membership Database），将所有下行接口维护的组成员关系记录都存到这个数据库中。组成员关系记录的结构如下：（Multicast-address，Filter-mode，Source-list），每条记录都是各下行接口上具有相同组地址的成员关系记录的合集。

上行接口正是依据这个数据库来执行主机行为——当收到查询报文时根据当前数据库状态响应报告报文，或者当数据库变化时主动发送报告或离开报文；而下行接口则执行路由器行为——参与查询器的选举、发送查询报文并根据报告报文维护组成员关系等。

1.1.7 多实例的IGMP

IGMP 依据接口来维护组成员关系，各实例的 IGMP 根据接口所属的实例来处理协议报文的收发。当路由器收到 IGMP 报文时，需要区分该报文所属的实例，并在该实例范围内对其进行处理。当某实例内的 IGMP 需要和其它组播协议交互信息时，只会通知本实例内的其它组播协议。

1.1.8 协议规范

与 IGMP 相关的协议规范有：

- RFC 1112：Host Extensions for IP Multicasting
- RFC 2236：Internet Group Management Protocol, Version 2
- RFC 3376：Internet Group Management Protocol, Version 3
- RFC 4605：Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")

1.2 IGMP配置任务简介

表1-1 IGMP 配置任务简介

配置任务		说明	详细配置
配置IGMP基本功能	使能IGMP	必选	1.3.2
	配置IGMP版本	可选	1.3.3
	配置静态加入	可选	1.3.4
	配置组播组过滤器	可选	1.3.5
	配置接口加入的组播组最大数量	可选	1.3.6
调整IGMP性能	配置IGMP报文选项	可选	1.4.2
	配置IGMP查询和响应	可选	1.4.3
	配置IGMP快速离开	可选	1.4.4
	配置IGMP主机跟踪功能	可选	1.4.5
	配置IGMP报文的DSCP优先级	可选	1.4.6
配置IGMP SSM Mapping	使能IGMP SSM Mapping	可选	1.5.2
	配置IGMP SSM Mapping规则	可选	1.5.3
配置IGMP Proxying	使能IGMP Proxying	可选	1.6.2

配置任务		说明	详细配置
	配置下行接口的组播转发能力	可选	1.6.3

说明

- IGMP 视图下的配置具有全局性，接口视图下的配置仅对所在接口有效。
- 若没有在接口视图下进行配置，则该接口继承 IGMP 视图下的全局配置；若两个视图都进行了配置，则优先采用该接口视图下所作的配置。

1.3 配置IGMP基本功能

1.3.1 配置准备

在配置 IGMP 基本功能之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 配置 PIM-DM（或 PIM-SM）

在配置 IGMP 基本功能之前，需准备以下数据：

- IGMP 的版本
- 以静态方式加入的组播组和组播源的地址
- 组播组过滤的 ACL 规则
- 接口加入的组播组最大数量

1.3.2 使能IGMP

在需要建立和维护组播组成员关系的接口上使能 IGMP。

1. 使能公网实例中的IGMP

表1-2 使能公网实例中的 IGMP

操作	命令	说明
进入系统视图	**system-view**	-
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能IGMP	**igmp enable**	必选 缺省情况下，IGMP处于关闭状态

2. 使能VPN实例中的IGMP

表1-3 使能 VPN 实例中的 IGMP

操作	命令	说明
进入系统视图	**system-view**	-
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
将接口与VPN实例进行关联	**ip binding vpn-instance** *vpn-instance-name*	必选 缺省情况下，接口只属于公网，未与任何VPN实例关联
使能IGMP	**igmp enable**	必选 缺省情况下，IGMP处于关闭状态

说明

- 有关 **ip vpn-instance**、**route-distinguisher** 和 **ip binding vpn-instance** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。
- 有关 **multicast routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

1.3.3 配置IGMP版本

由于不同版本 IGMP 协议的报文结构与种类不同，因此需要为同一网段上的所有路由器配置相同版本的 IGMP，否则 IGMP 将不能正常运行。

1. 全局配置IGMP版本

表1-4 全局配置 IGMP 版本

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例IGMP视图	**igmp** [**vpn-instance** *vpn-instance-name*]	-
配置IGMP的版本	**version** *version-number*	必选 缺省情况下，IGMP的版本为IGMPv2

2. 在接口上配置IGMP版本

表1-5 在接口上配置 IGMP 版本

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置IGMP的版本	**igmp version** *version-number*	必选 缺省情况下，IGMP的版本为IGMPv2

1.3.4 配置静态加入

在配置了静态加入组播组或组播源组后，接口将作为该组播组的虚拟组成员存在，从而可以接收发往该组的组播数据，以测试组播数据的转发。

表1-6 配置静态加入

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置静态加入组播组或组播源组	**igmp static-group** *group-address* [**source** *source-address*]	必选 缺省情况下，接口没有以静态方式加入任何组播组或组播源组

说明

- 在运行 PIM-SM 的设备上配置静态加入时，如果待配接口上使能了 PIM-SM，则该接口必须为 PIM-SM 的 DR，否则该接口将不能加入组播组或组播源组；如果待配接口上未使能 PIM-SM 但使能了 IGMP，则该接口必须为 IGMP 查询器，否则该接口也不能加入组播组或组播源组。有关 PIM-SM 和 DR 的介绍，请参见“IP 组播配置指导”中的“PIM”。
- 在配置了静态加入后，接口并不会对 IGMP 查询器发出的查询报文进行响应；当配置静态加入或取消静态加入的配置时，接口也不会主动发送 IGMP 成员关系报告报文或 IGMP 离开组报文。也就是说，该接口并没有真正成为该组播组的成员。

1.3.5 配置组播组过滤器

如果不希望接口所在网段上的主机加入某些组播组，可在该接口上配置 ACL 规则作为过滤器，接口将按照该规则对收到的 IGMP 成员关系报告报文进行过滤，只为该规则所允许的组播组维护组成员关系。

表1-7 配置组播组过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置组播组过滤器	**igmp group-policy** *acl-number* [*version-number*]	必选 缺省情况下，接口上没有配置组播组过滤器，即该接口下的主机可以加入任意合法的组播组

1.3.6 配置接口加入的组播组最大数量

通过配置接口加入的组播组最大数量，可以方便用户灵活控制接口所能加入的组播组数量。

表1-8 配置接口加入的组播组最大数量

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置接口加入的组播组最大数量	**igmp group-limit** *limit*	必选 缺省情况下，接口加入的组播组最大数量为2000

说明

本配置只限制接口动态加入的组播组数量，而不限制其静态加入的组播组数量。

1.4 调整IGMP性能

说明

对于本节中的配置任务来说：

- 在 IGMP 视图下所作的配置对所有接口生效，在接口视图下所作的配置仅对当前接口生效；
- 如果在这两个视图下进行了相同功能或参数的配置，则不论配置先后，接口视图下的配置将被优先采用。

1.4.1 配置准备

在调整 IGMP 性能之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通

- 配置 IGMP 基本功能

在调整 IGMP 性能之前，需准备以下数据：

- IGMP 查询器的启动查询间隔
- IGMP 查询器的启动查询次数
- 发送 IGMP 普遍组查询报文的时间间隔
- IGMP 查询器的健壮系数
- IGMP 普遍组查询的最大响应时间
- 最后组成员查询间隔
- IGMP 其它查询器的存在时间
- IGMP 报文的 DSCP 优先级

1.4.2 配置IGMP报文选项

由于 IGMP 特定组查询报文和 IGMP 特定源组查询报文的存在，而组播组是千变万化的，设备不可能加入所有的组，因此 IGMP 需要借助 Router-Alert（路由器报警）选项来将送达本地、但没有加入的组播报文送往上层协议进行处理。有关 Router-Alert 选项的详细介绍，请参考 RFC 2113。

对于 IP 头中是否携带 Router-Alert 选项的 IGMP 报文，设备会做出不同的处理：

- 出于兼容性考虑，缺省情况下设备不对 Router-Alert 选项进行检查，即处理所有收到的 IGMP 报文。此时，IGMP 报文中无论是否携带有 Router-Alert 选项，设备都会将其送给上层协议进行处理。
- 为了提高设备性能、减少不必要的开支，同时出于协议安全性的考虑，可以配置设备丢弃未携带 Router-Alert 选项的 IGMP 报文，此时，当设备收到 IGMP 报文时，会检查该报文的 Router-Alert 选项，如果没有携带该选项，就丢弃该报文。

1. 全局配置IGMP报文选项

表1-9 全局配置 IGMP 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例IGMP视图	**igmp** [**vpn-instance** *vpn-instance-name*]	-
配置丢弃未携带Router-Alert选项的IGMP报文	**require-router-alert**	必选 缺省情况下，设备不对Router-Alert选项进行检查
配置在发送的IGMP报文中携带Router-Alert选项	**send-router-alert**	必选 缺省情况下，在发送的IGMP报文中携带Router-Alert选项

2. 在接口上配置IGMP报文选项

表1-10 在接口上配置 IGMP 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置丢弃未携带Router-Alert选项的IGMP报文	**igmp require-router-alert**	必选 缺省情况下，设备不对Router-Alert选项进行检查
配置在发送的IGMP报文中携带Router-Alert选项	**igmp send-router-alert**	必选 缺省情况下，在发送的IGMP报文中携带Router-Alert选项

1.4.3 配置IGMP查询和响应

IGMP 查询器的健壮系数是为了弥补可能发生的网络丢包而设置的报文重传次数，健壮系数越大，IGMP 查询器就越“健壮”，但是组播组超时所需的时间也就越长。

当 IGMPv1/v2/v3 查询器启动时，会以“IGMP 查询器启动查询间隔”为时间间隔发送“IGMP 查询器启动查询次数”次 IGMP 普遍组查询报文。

IGMPv1/v2/v3 查询器会周期性地发送 IGMP 普遍组查询报文，以判断网络上是否有组播组成员，发送间隔即为“发送 IGMP 普遍组查询报文的时间间隔”。可以根据网络的实际情况来修改周期性发送 IGMP 普遍组查询报文的时间间隔。

当 IGMPv2 查询器收到 IGMP 离开组报文后，会以“最后组成员查询间隔”为时间间隔发送“最后组成员查询次数”次 IGMP 特定组查询报文；当 IGMPv3 查询器收到改变组播组与组播源列表关系的 IGMP 报告报文后，也会以“最后组成员查询间隔”为时间间隔发送“最后组成员查询次数”次 IGMP 特定源组查询报文。最后组成员查询次数在数值上等于 IGMP 查询器的健壮系数。

在收到 IGMP 查询报文（包括普遍组查询、特定组查询和特定源组查询）后，主机会为其所加入的每个组播组都启动一个延迟定时器，其值在 0 到最大响应时间（该时间值从 IGMP 查询报文的最大响应时间字段获得）中随机选定，当定时器的值减为 0 时，主机就会向该定时器对应的组播组发送 IGMP 成员关系报告报文。

合理配置 IGMP 查询的最大响应时间，既可以使主机对 IGMP 查询报文做出快速响应，又可以减少由于定时器同时超时，造成大量主机同时发送报告报文而引起的网络拥塞：

- 对于 IGMP 普遍组查询报文来说，通过配置 IGMP 普遍组查询的最大响应时间来填充其最大响应时间字段；
- 对于 IGMP 特定组查询报文和 IGMP 特定源组查询报文来说，所配置的最后组成员查询间隔将被填充到其最大响应时间字段。也就是说，IGMP 特定组查询和 IGMP 特定源组查询的最大响应时间在数值上等于最后组成员查询间隔。

当同一网段上有多台组播路由器时，由 IGMP 查询器负责发送 IGMP 查询报文。如果非查询器在“其它查询器存在时间”超时前没有收到来自查询器的 IGMP 查询报文，就会认为原有查询器失效，从而触发新的查询器选举过程；否则，非查询器将重置“其它查询器存在时间定时器”。

1. 全局配置IGMP查询和响应

表1-11 全局配置 IGMP 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例IGMP视图	**igmp** [**vpn-instance** *vpn-instance-name*]	-
配置IGMP查询器的健壮系数	**robust-count** *robust-value*	必选 缺省情况下，IGMP查询器的健壮系数为2
配置IGMP查询器的启动查询间隔	**startup-query-interval** *interval*	必选 缺省情况下，IGMP查询器的启动查询间隔为发送IGMP普遍组查询报文时间间隔的1/4
配置IGMP查询器的启动查询次数	**startup-query-count** *value*	必选 缺省情况下，IGMP查询器的启动查询次数等于IGMP查询器的健壮系数
配置发送IGMP普遍组查询报文的时间间隔	**timer query** *interval*	必选 缺省情况下，发送IGMP普遍组查询报文的时间间隔为60秒
配置IGMP普遍组查询的最大响应时间	**max-response-time** *interval*	必选 缺省情况下，IGMP普遍组查询的最大响应时间为10秒
配置最后组成员查询间隔	**last-member-query-interval** *interval*	必选 缺省情况下，最后组成员查询间隔为1秒
配置IGMP其它查询器的存在时间	**timer other-querier-present** *interval*	必选 缺省情况下，IGMP其它查询器的存在时间＝发送IGMP普遍组查询报文的时间间隔×IGMP查询器的健壮系数＋IGMP普遍组查询的最大响应时间÷2

2. 在接口上配置IGMP查询和响应

表1-12 在接口上配置 IGMP 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置IGMP查询器的健壮系数	**igmp robust-count** *robust-value*	必选 缺省情况下，IGMP查询器的健壮系数为2
配置IGMP查询器的启动查询间隔	**igmp startup-query-interval** *interval*	必选 缺省情况下，IGMP查询器的启动查询间隔为发送IGMP普遍组查询报文时间间隔的1/4

操作	命令	说明
配置IGMP查询器的启动查询次数	**igmp startup-query-count** *value*	必选 缺省情况下，IGMP查询器的启动查询次数等于IGMP查询器的健壮系数
配置发送IGMP普遍组查询报文的时间间隔	**igmp timer query** *interval*	必选 缺省情况下，发送IGMP普遍组查询报文的时间间隔为60秒
配置IGMP普遍组查询的最大响应时间	**igmp max-response-time** *interval*	必选 缺省情况下，IGMP普遍组查询的最大响应时间为10秒
配置最后组成员查询间隔	**igmp last-member-query-interval** *interval*	必选 缺省情况下，最后组成员查询间隔为1秒
配置IGMP其它查询器的存在时间	**igmp timer other-querier-present** *interval*	必选 缺省情况下，IGMP其它查询器的存在时间＝发送IGMP普遍组查询报文的时间间隔×IGMP查询器的健壮系数＋IGMP普遍组查询的最大响应时间÷2

注意

- 应确保 IGMP 其它查询器的存在时间大于发送 IGMP 普遍组查询报文的时间间隔，否则有可能导致网络内的 IGMP 查询器反复变化。
- 应确保发送 IGMP 普遍组查询报文的时间间隔大于 IGMP 普遍组查询的最大响应时间，否则有可能造成对组播组成员的误删。
- 对 IGMP 普遍组查询的最大响应时间、最后组成员查询间隔以及 IGMP 其它查询器的存在时间所做的配置，只有当设备运行在 IGMPv2 或 IGMPv3 时才有效。

1.4.4 配置IGMP快速离开

在某些应用（如 ADSL 拨号上网）中，IGMP 查询器的一个端口唯一对应着一台接收者主机，当主机在多个组播组间频繁切换（如进行电视选台）时，为了快速响应主机的离开组报文，可以在 IGMP 查询器上开启 IGMP 快速离开功能。

在开启了 IGMP 快速离开功能之后，当 IGMP 查询器收到来自主机的离开组报文时，不再发送 IGMP 特定组查询报文或 IGMP 特定源组查询报文，而是直接向上游发送离开通告，这样一方面减小了响应延迟，另一方面也节省了网络带宽。

1. 全局配置IGMP快速离开

表1-13 全局配置 IGMP 快速离开

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入公网实例或VPN实例IGMP视图	**igmp** [**vpn-instance** *vpn-instance-name*]	-
配置组播组成员快速离开	**fast-leave** [**group-policy** *acl-number*]	必选 缺省情况下，组播组成员快速离开功能是关闭的

2. 在接口上配置IGMP快速离开

表1-14 在接口上配置 IGMP 快速离开

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置组播组成员快速离开	**igmp fast-leave** [**group-policy** *acl-number*]	必选 缺省情况下，组播组成员快速离开功能是关闭的

注意

- 只有当设备运行在 IGMPv2 或 IGMPv3 时，配置 IGMP 快速离开功能才有效。
- IGMP 快速离开配置只对除 VLAN 接口以外的其它三层接口生效，如三层以太网端口、三层聚合接口及 Tunnel 接口。

1.4.5 配置IGMP主机跟踪功能

通过使能 IGMP 主机跟踪功能，可以使设备能够记录正在接收组播数据的成员主机信息（包括主机的 IP 地址、运行时间和超时时间等），以便于网络管理员对这些主机进行监控和管理。

1. 全局配置IGMP主机跟踪功能

表1-15 全局配置 IGMP 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例IGMP视图	**igmp** [**vpn-instance** *vpn-instance-name*]	-
全局使能IGMP主机跟踪功能	**host-tracking**	必选 缺省情况下，IGMP主机跟踪功能处于关闭状态

2. 在接口上配置IGMP主机跟踪功能

表1-16 在接口上配置 IGMP 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
在接口上使能IGMP主机跟踪功能	**igmp host-tracking**	必选 缺省情况下，IGMP主机跟踪功能处于关闭状态

1.4.6 配置IGMP报文的DSCP优先级

在 IPv4 报文头中，包含一个 8bit 的 ToS 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 IGMP 报文的 DSCP 优先级进行配置。

表1-17 配置 IGMP 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例IGMP视图	**igmp** [**vpn-instance** *vpn-instance-name*]	-
配置IGMP报文的DSCP优先级	**dscp** *dscp-value*	可选 缺省情况下，IGMP报文的DSCP优先级为48

1.5 配置IGMP SSM Mapping

在 SSM 网络中，由于各种可能的限制，某些接收者主机只能运行 IGMPv1 或 IGMPv2。为了向这些仅支持 IGMPv1 或 IGMPv2 的接收者主机提供 SSM 服务，可以在路由器上配置 IGMP SSM Mapping 功能。

1.5.1 配置准备

在配置 IGMP SSM Mapping 之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 配置 IGMP 基本功能

1.5.2 使能IGMP SSM Mapping

表1-18 使能 IGMP SSM Mapping

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能IGMP SSM Mapping功能	**igmp ssm-mapping enable**	必选 缺省情况下，接口上的IGMP SSM Mapping功能处于关闭状态

说明

为保证本网段内运行任意版本 IGMP 的接收者主机都能得到 SSM 服务，建议在处于该网段的接口上运行 IGMPv3。

1.5.3 配置IGMP SSM Mapping规则

通过多次配置可以实现同一组播组到多个组播源的映射。

表1-19 配置 IGMP SSM Mapping 规则

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例的IGMP视图	**igmp** [**vpn-instance** *vpn-instance-name*]	-
配置IGMP SSM Mapping规则	**ssm-mapping** *group-address* { *mask* | *mask-length* } *source-address*	必选 缺省情况下，未配置IGMP SSM Mapping规则

注意

若 VLAN 接口上运行的是 IGMPv3，则在使用 **igmp-snooping host-join** 命令配置模拟主机加入时即便不指定组播源，模拟主机也仍将发送 IGMPv3 的报告报文，这种情况下不会依据 IGMP SSM Mapping 规则创建相应的组播组。有关 **igmp-snooping host-join** 命令的详细介绍，请参见“IP 组播命令参考”中的“IGMP Snooping”。

1.6 配置IGMP Proxying

1.6.1 配置准备

在配置 IGMP Proxying 之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 使能 IP 组播路由

1.6.2 使能IGMP Proxying

在设备朝向组播分发树树根方向的接口上使能了 IGMP 代理功能之后，该设备就成为了 IGMP 代理设备。

表1-20 使能 IGMP Proxying

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能IGMP代理功能	**igmp proxying enable**	必选 缺省情况下，接口上的IGMP代理功能处于关闭状态

说明

- 一台设备上只允许有一个接口使能 IGMP 代理功能（多实例情况下，每个实例中只允许有一个接口使能 IGMP 代理功能）。
- 在进行 IGMP 的相关配置时，在已使能 IGMP 代理功能的接口上不允许再使能 IGMP，且只有 **igmp require-router-alert**、**igmp send-router-alert** 和 **igmp version** 命令可在该接口上生效。
- 在已使能 IGMP 代理功能的设备上不允许再使能其它组播路由协议（如不能在接口上使能 PIM-DM 或 PIM-SM，但在 PIM 视图下配置的 **source-lifetime**、**source-policy** 和 **ssm-policy** 命令仍会生效），反之亦然；同时，由于 IGMPv1 的查询器要由 PIM 协议选举出的 DR 来充当，因此若 IGMP 代理设备的下行接口运行的是 IGMPv1，那么该接口将无法成为 DR，从而也就无法充当 IGMP 查询器。
- 若某 VLAN 内已使能了 IGMP Snooping，则该 VLAN 对应的 VLAN 接口上不能再使能 IGMP 代理功能，反之亦然。

1.6.3 配置下行接口的组播转发能力

通常，组播数据只能被查询器转发，非查询器不具备组播转发能力，这样可以避免组播数据被重复转发。对于 IGMP 代理设备也同样如此，只有当其下行接口是查询器时，组播数据才能通过该接口转发给下游主机。

但在某些情况下，IGMP 代理设备的下行接口不能在查询器竞选中获胜，此时应使能该下行接口在非查询器状态下的组播转发能力。

表1-21 配置下行接口的组播转发能力

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能下行接口在非查询器状态下的组播转发能力	**igmp proxying forwarding**	必选 缺省情况下，当IGMP代理设备的下行接口处于非查询器状态时，不转发组播数据

注意

在共享网段内存在多台 IGMP 代理设备的情况下，当其中一台 IGMP 代理设备的下行接口已竞选为查询器时，不应再在其它 IGMP 代理设备的下行接口上进行此配置，否则该网段内将收到多份重复的组播数据流。

1.7 IGMP显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IGMP 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 IGMP 的统计信息。

表1-22 IGMP 显示和维护

操作	命令
查看IGMP组的信息	**display igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **group** [*group-address* | **interface** *interface-type interface-number*] [**static** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IGMP组的二层端口信息	**display igmp group port-info** [**vlan** *vlan-id*] [**slot** *slot-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上IGMP跟踪的主机信息	**display igmp host interface** *interface-type interface-number* **group** *group-address* [**source** *source-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看二层端口上IGMP跟踪的主机信息	**display igmp host port-info vlan** *vlan-id* **group** *group-address* [**source** *source-address*] [**vlan** *vlan-id*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上IGMP配置和运行信息	**display igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **interface** [*interface-type interface-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IGMP代理组的信息	**display igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **proxying group** [*group-address*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IGMP路由表的信息	**display igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **routing-table** [*source-address* [**mask** { *mask* | *mask-length* }] | *group-address* [**mask** { *mask* | *mask-length* }] | **flags** { **act** | **suc** }] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IGMP SSM Mapping规则	**display igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **ssm-mapping** *group-address* [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看依据IGMP SSM Mapping规则创建的组播组信息	**display igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **ssm-mapping group** [*group-address* | **interface** *interface-type interface-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上依据IGMP SSM Mapping规则加入的主机信息	**display igmp ssm-mapping host interface** *interface-type interface-number* **group** *group-address* **source** *source-address* [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IGMP组的动态加入记录	**reset igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **group** { **all** | **interface** *interface-type interface-number* { **all** | *group-address* [**mask** { *mask* | *mask-length* }] [*source-address* [**mask** { *mask* | *mask-length* }]] } }
清除IGMP组二层端口的动态加入记录	**reset igmp group port-info** { **all** | *group-address* } [**vlan** *vlan-id*]
清除依据IGMP SSM Mapping规则创建的组播组信息	**reset igmp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **ssm-mapping group** { **all** | **interface** *interface-type interface-number* { **all** | *group-address* [**mask** { *mask* | *mask-length* }] [*source-address* [**mask** { *mask* | *mask-length* }]] } }

说明

- **reset igmp group** 和 **reset igmp group port-info** 命令只能清除动态加入记录，而无法清除静态加入记录。
- **display igmp host interface** 只能查看除 Vlan 接口外的三层接口上 IGMP 跟踪的主机信息。
- **display igmp ssm-mapping host interface** 只能查看除 Vlan 接口外的三层接口上依据 IGMP SSM Mapping 规则加入的主机信息。

注意

执行 **reset igmp group** 命令可能导致接收者中断组播信息的接收。

1.8 IGMP典型配置举例

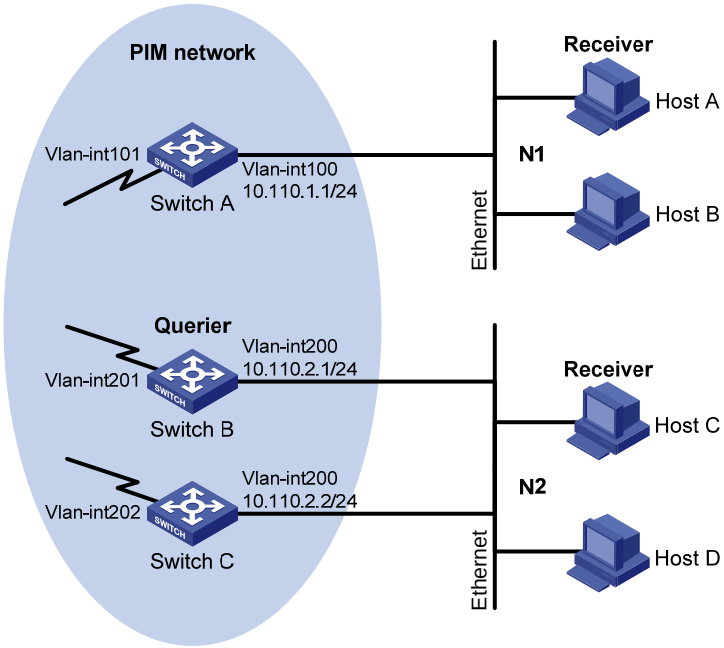
1.8.1 IGMP基本功能配置举例

1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者组成末梢网络 N1 和 N2，Host A 与 Host C 分别为 N1 和 N2 中的组播信息接收者。
- PIM 网络中的 Switch A 连接 N1，Switch B 与 Switch C 共同连接 N2。
- Switch A 通过 Vlan-interface100 连接 N1，通过 Vlan-interface101 连接 PIM 网络中的其它设备。
- Switch B 与 Switch C 分别通过各自的 Vlan-interface200 连接 N2，并分别通过 Vlan-interface201 和 Vlan-interface202 连接 PIM 网络中的其它设备。
- Switch A 与 N1 之间运行 IGMPv2；Switch B、Switch C 与 N2 之间也分别运行 IGMPv2，且由于 Switch B 的接口 IP 地址较小，因此在 N2 中通常由其来充当 IGMP 查询器。
- 通过配置，使 N1 中的主机只能加入组播组 224.1.1.1，而对 N2 中的主机则无任何限制。

2. 组网图

图1-5 IGMP 基本功能配置组网图



3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-5 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 PIM 网络内各交换机之间采用 OSPF 协议进行互连，确保 PIM 网络内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-DM 和 IGMP

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-DM，并在主机侧接口 Vlan-interface100 上使能 IGMP。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim dm
[SwitchA-Vlan-interface101] quit
```

在 Switch B 上使能 IP 组播路由，在各接口上使能 PIM-DM，并在主机侧接口 Vlan-interface200 上使能 IGMP。

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] igmp enable
```

```
[SwitchB-Vlan-interface200] pim dm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 201
[SwitchB-Vlan-interface201] pim dm
[SwitchB-Vlan-interface201] quit
```

在 Switch C 上使能 IP 组播路由，在各接口上使能 PIM-DM，并在主机侧接口 Vlan-interface200 上使能 IGMP。

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] igmp enable
[SwitchC-Vlan-interface200] pim dm
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 202
[SwitchC-Vlan-interface202] pim dm
[SwitchC-Vlan-interface202] quit
```

(3) 配置组播组过滤器

在 Switch A 上限定接口 Vlan-interface100 下的主机只能加入组播组 224.1.1.1。

```
[SwitchA] acl number 2001
[SwitchA-acl-basic-2001] rule permit source 224.1.1.1 0
[SwitchA-acl-basic-2001] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp group-policy 2001
[SwitchA-Vlan-interface100] quit
```

(4) 检验配置效果

通过使用 **display igmp interface** 命令可以查看各交换机接口上 IGMP 的配置和运行情况。例如：

查看 Switch B 在 Vlan-interface200 上的 IGMP 信息。

```
[SwitchB] display igmp interface vlan-interface 200
 Vlan-interface200(10.110.2.1):
   IGMP is enabled
   Current IGMP version is 2
   Value of query interval for IGMP(in seconds): 60
   Value of other querier present interval for IGMP(in seconds): 125
   Value of maximum query response time for IGMP(in seconds): 10
   Querier for IGMP: 10.110.2.1 (this router)
  Total 1 IGMP Group reported
```

1.8.2 IGMP SSM Mapping功能配置举例

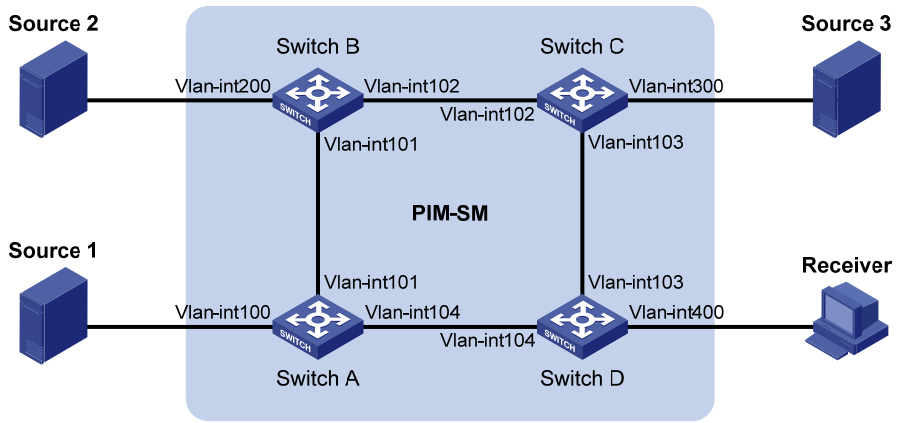
1. 组网需求

- PIM-SM 网络中同时采用 ASM 和 SSM 方式提供组播服务，将 Switch D 的接口 Vlan-interface104 配置为 C-BSR 和 C-RP，SSM 组播组的范围为 232.1.1.0/24。
- Switch D 的接口 Vlan-interface400 上运行 IGMPv3，接收者主机上运行 IGMPv2，且不能升级至 IGMPv3，因此该主机在加入组播组时无法指定组播源。

- Source 1、Source 2 和 Source 3 都向 SSM 组范围内的组播组发送组播数据，要求通过在 Switch D 上配置 IGMP SSM Mapping 功能，使接收者主机只能接收来自 Source 1 和 Source 3 的组播数据。

2. 组网图

图1-6 IGMP SSM Mapping 功能配置组网图



设备	接口	IP地址	设备	接口	IP地址
Source 1	-	133.133.1.1/24	Source 3	-	133.133.3.1/24
Source 2	-	133.133.2.1/24	Receiver	-	133.133.4.1/24
Switch A	Vlan-int100	133.133.1.2/24	Switch C	Vlan-int300	133.133.3.2/24
	Vlan-int101	192.168.1.1/24		Vlan-int103	192.168.3.1/24
	Vlan-int104	192.168.4.2/24		Vlan-int102	192.168.2.2/24
Switch B	Vlan-int200	133.133.2.2/24	Switch D	Vlan-int400	133.133.4.2/24
	Vlan-int101	192.168.1.2/24		Vlan-int103	192.168.3.2/24
	Vlan-int102	192.168.2.1/24		Vlan-int104	192.168.4.1/24

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-6 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 PIM-SM 域内的各交换机之间采用 OSPF 协议进行互连，确保 PIM-SM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-SM、IGMP 和 IGMP SSM Mapping 功能

在 Switch D 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在主机侧接口 Vlan-interface400 上使能 IGMP 和 IGMP SSM Mapping 功能，配置 IGMP 版本为 3。

```
<SwitchD> system-view
[SwitchD] multicast routing-enable
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] igmp enable
[SwitchD-Vlan-interface400] igmp version 3
[SwitchD-Vlan-interface400] igmp ssm-mapping enable
[SwitchD-Vlan-interface400] pim sm
[SwitchD-Vlan-interface400] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim sm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 104
```

```
[SwitchD-Vlan-interface104] pim sm
[SwitchD-Vlan-interface104] quit
```

在 Switch A 上使能 IP 组播路由，并在各接口上使能 PIM-SM。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim sm
[SwitchA-Vlan-interface104] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP

在 Switch D 上配置 C-BSR 和 C-RP 的位置。

```
[SwitchD] pim
[SwitchD-pim] c-bsr vlan-interface 104
[SwitchD-pim] c-rp vlan-interface 104
[SwitchD-pim] quit
```

(4) 配置 SSM 组播组的地址范围

在 Switch D 上配置 SSM 组播组地址范围为 232.1.1.0/24。

```
[SwitchD] acl number 2000
[SwitchD-acl-basic-2000] rule permit source 232.1.1.0 0.0.0.255
[SwitchD-acl-basic-2000] quit
[SwitchD] pim
[SwitchD-pim] ssm-policy 2000
[SwitchD-pim] quit
```

Switch A、Switch B 和 Switch C 的配置与 Switch D 相似，配置过程略。

(5) 配置 IGMP SSM Mapping 规则

在 Switch D 上配置 IGMP SSM Mapping 规则。

```
[SwitchD] igmp
[SwitchD-igmp] ssm-mapping 232.1.1.0 24 133.133.1.1
[SwitchD-igmp] ssm-mapping 232.1.1.0 24 133.133.3.1
[SwitchD-igmp] quit
```

(6) 检验配置效果

通过使用 **display igmp ssm-mapping** 命令可以查看交换机上的 IGMP SSM Mapping 规则。例如：

查看 Switch D 上组播组 232.1.1.1 的 IGMP SSM Mapping 规则。

```
[SwitchD] display igmp ssm-mapping 232.1.1.1
Vpn-Instance: public net
Group: 232.1.1.1
Source list:
       133.133.1.1
       133.133.3.1
```

通过使用 **display igmp ssm-mapping group** 命令可以查看交换机上依据 IGMP SSM Mapping 规则创建的组播组信息。例如：

查看 Switch D 上依据 IGMP SSM Mapping 规则创建的组播组信息。

```
[SwitchD] display igmp ssm-mapping group
Total 1 IGMP SSM-mapping Group(s).
Interface group report information of VPN-Instance: public net
 Vlan-interface400(133.133.4.2):
  Total 1 IGMP SSM-mapping Group reported
   Group Address       Last Reporter   Uptime      Expires
   232.1.1.1           133.133.4.1     00:02:04    off
```

通过使用 **display pim routing-table** 命令可以查看交换机的 PIM 路由表信息。例如：

查看 Switch D 上的 PIM 路由表信息。

```
[SwitchD] display pim routing-table
Vpn-instance: public net
Total 0 (*, G) entry; 2 (S, G) entry

 (133.133.1.1, 232.1.1.1)
     Protocol: pim-ssm, Flag:
     UpTime: 00:13:25
     Upstream interface: Vlan-interface104
         Upstream neighbor: 192.168.4.2
         RPF prime neighbor: 192.168.4.2
     Downstream interface(s) information:
       Total number of downstreams: 1
         1: Vlan-interface400
                 Protocol: igmp, UpTime: 00:13:25, Expires: -

 (133.133.3.1, 232.1.1.1)
     Protocol: pim-ssm, Flag:
     UpTime: 00:13:25
     Upstream interface: Vlan-interface103
         Upstream neighbor: 192.168.3.1
         RPF prime neighbor: 192.168.3.1
     Downstream interface(s) information:
       Total number of downstreams: 1
         1: Vlan-interface400
                 Protocol: igmp, UpTime: 00:13:25, Expires: -
```

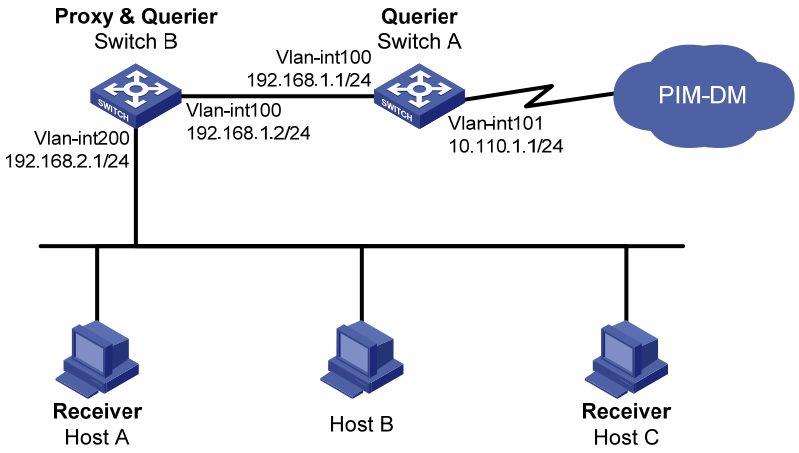
1.8.3 IGMP Proxying功能配置举例

1. 组网需求

- 核心网络中运行 PIM-DM，末梢网络中的接收者 Host A 和 Host C 通过组播组 224.1.1.1 点播视频节目。
- 要求通过在 Switch B 上配置 IGMP Proxying 功能，使其在不运行 PIM-DM 的情况下实现组成员关系的维护和组播数据的正常转发。

2. 组网图

图1-7 IGMP Proxying 功能配置组网图



3. 配置步骤

(1) 配置 IP 地址

请按照 图 1-7 配置各接口的IP地址和子网掩码，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-DM、IGMP 和 IGMP Proxying

在 Switch A 上使能 IP 组播路由，在接口 Vlan-interface101 上使能 PIM-DM，并在接口 Vlan-interface100 上使能 IGMP。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim dm
[SwitchA-Vlan-interface100] quit
```

在 Switch B 上使能 IP 组播路由，在接口 Vlan-interface100 上使能 IGMP Proxying，并在接口 Vlan-interface200 上使能 IGMP。

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] igmp proxying enable
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] igmp enable
[SwitchB-Vlan-interface200] quit
```

(3) 检验配置效果

通过使用 **display igmp interface** 命令可以查看各交换机接口上 IGMP 的配置和运行情况。例如：

查看 Switch B 在 Vlan-interface100 上 IGMP 配置和运行的详细信息。

```
[SwitchB] display igmp interface vlan-interface 100 verbose
```

```
 Vlan-interface100(192.168.1.2):
   IGMP proxy is enabled
   Current IGMP version is 2
   Multicast routing on this interface: enabled
   Require-router-alert: disabled
   Version1-querier-present-timer-expiry: 00:00:20
```

通过使用 **display igmp group** 命令可以查看 IGMP 组的信息。例如：

查看 Switch A 上 IGMP 组的信息。

```
[SwitchA] display igmp group
Total 1 IGMP Group(s).
Interface group report information of VPN-Instance: public net
 Vlan-interface100(192.168.1.1):
  Total 1 IGMP Groups reported
   Group Address   Last Reporter   Uptime      Expires
   224.1.1.1       192.168.1.2     00:02:04    00:01:15
```

由此可见，主机的 IGMP 报告通过 Switch B 的代理接口 Vlan-interface100 发给 Switch A。

1.9 常见配置错误举例

1.9.1 接收者侧路由器上无组成员信息

1. 故障现象

当某主机发送了加入组播组 G 的报文后，离该主机最近的路由器上却没有组播组 G 的组成员信息。

2. 故障分析

- 组网、接口连线的正确与否以及接口的协议层是否 up 将直接影响组播组成员信息的生成；
- 在路由器上必须使能组播路由，在连接主机的接口上必须使能 IGMP；
- 如果路由器接口上运行的 IGMP 版本比主机的低，那么路由器将无法识别主机发来的较高版本的 IGMP 报告报文；
- 如果在接口上使用命令 **igmp group-policy** 对加入组播组 G 进行了限制后，该接口将不再接收未通过过滤的那些要求加入组播组 G 的报文。

3. 处理过程

(1) 检查组网是否正确，接口间的连线是否正确，以及接口状态是否正常，是否配置了正确的 IP 地址。通过命令 **display igmp interface** 查看接口信息。若无接口信息输出，说明接口状态异常，原因通常是接口上配置了 **shutdown** 命令，或者接口连线不正确，或者接口上没有配置正确的 IP 地址。

(2) 检查是否使能了组播路由。通过命令 **display current-configuration** 查看是否配置了命令 **multicast routing-enable**。若缺少该配置，则需要在系统视图下执行命令 **multicast routing-enable** 使能 IP 组播路由，同时也需要在相应接口上使能 IGMP。

(3) 检查接口上运行的 IGMP 版本。通过命令 **display igmp interface** 来检查接口上运行的 IGMP 版本是否低于主机所使用的版本。

(4) 检查接口上是否配置了 ACL 规则来限制主机加入组播组 G。通过命令 **display current-configuration interface** 观察是否配置了 **igmp group-policy** 命令。如果配置的 ACL 规则对加入组播组 G 进行了限制，则需要修改该 ACL 规则，允许接受组播组 G 的报告报文。

1.9.2 同一网段各路由器上组成员关系不一致

1. 故障现象

在同一网段的不同 IGMP 路由器上，各自维护的组成员关系不一致。

2. 故障分析

- 运行 IGMP 的路由器为每个接口维护多个参数，各参数之间相互影响，非常复杂。如果同一网段路由器的 IGMP 接口参数配置不一致，必然导致组成员关系的混乱；
- 另外，IGMP 目前有 3 个版本，版本不同的 IGMP 路由器与主机之间虽然可以兼容，但是连接在同一网段的所有路由器必须运行相同版本的 IGMP。如果同一网段路由器的 IGMP 版本不一致，也将导致 IGMP 组成员关系的混乱。

3. 处理过程

(1) 检查 IGMP 配置。通过命令 **display current-configuration** 观察接口上 IGMP 的配置信息。

(2) 在同一网段的所有路由器上执行命令 **display igmp interface** 来检查 IGMP 相关定时器的参数，确保配置一致。

(3) 通过命令 **display igmp interface** 来检查各路由器上运行的 IGMP 版本是否一致。

目 录

1 PIM配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- PIM 功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

1.1 PIM简介

PIM 是 Protocol Independent Multicast（协议无关组播）的简称，表示可以利用单播静态路由或者任意单播路由协议（包括 RIP、OSPF、IS-IS、BGP 等）所生成的单播路由表为 IP 组播提供路由。组播路由与所采用的单播路由协议无关，只要能够通过单播路由协议产生相应的组播路由表项即可。PIM 借助 RPF（Reverse Path Forwarding，逆向路径转发）机制实现对组播报文的转发。当组播报文到达本地设备时，首先对其进行 RPF 检查：若 RPF 检查通过，则创建相应的组播路由表项，从而进行组播报文的转发；若 RPF 检查失败，则丢弃该报文。有关 RPF 的详细介绍，请参见“IP 组播配置指导”中的“组播路由与转发”。

根据实现机制的不同，PIM 分为以下几种类型：

- PIM-DM（Protocol Independent Multicast-Dense Mode，协议无关组播—密集模式）
- PIM-SM（Protocol Independent Multicast-Sparse Mode，协议无关组播—稀疏模式）
- BIDIR-PIM（Bidirectional Protocol Independent Multicast，双向协议无关组播，简称双向 PIM）
- PIM-SSM（Protocol Independent Multicast Source-Specific Multicast，协议无关组播－指定源组播）

说明

为了描述的方便，本文中把由支持 PIM 协议的组播路由器所组成的网络简称为“PIM 域”。

1.1.1 PIM-DM简介

PIM-DM 属于密集模式的组播路由协议，使用“推（Push）模式”传送组播数据，通常适用于组播组成员相对比较密集的小型网络，其基本原理如下：

- PIM-DM 假设网络中的每个子网都存在至少一个组播组成员，因此组播数据将被扩散（Flooding）到网络中的所有节点。然后，PIM-DM 对没有组播数据转发的分支进行剪枝（Prune），只保留包含接收者的分支。这种“扩散—剪枝”现象周期性地发生，被剪枝的分支也可以周期性地恢复成转发状态。

- 当被剪枝分支的节点上出现了组播组的成员时，为了减少该节点恢复成转发状态所需的时间，PIM-DM 使用嫁接（Graft）机制主动恢复其对组播数据的转发。

一般说来，密集模式下数据包的转发路径是有源树（Source Tree，即以组播源为“根”、组播组成员为“枝叶”的一棵转发树）。由于有源树使用的是从组播源到接收者的最短路径，因此也称为最短路径树（Shortest Path Tree，SPT）。

PIM-DM 的工作机制可以概括如下：

- 邻居发现
- 构建 SPT
- 嫁接
- 断言

1. 邻居发现

在 PIM 域中，路由器通过周期性地向本网段的所有 PIM 路由器（224.0.0.13）以组播方式发送 PIM Hello 报文（以下简称 Hello 报文），以发现 PIM 邻居，维护各路由器之间的 PIM 邻居关系，从而构建和维护 SPT。

说明

路由器每个运行了 PIM 协议的接口都会周期性地发送 Hello 报文，从而了解与该接口相关的 PIM 邻居信息。

---

2. 构建SPT

构建 SPT 的过程也就是“扩散—剪枝”的过程：

(1) 在 PIM-DM 域中，组播源 S 向组播组 G 发送组播报文时，首先对组播报文进行扩散：路由器对该报文的 RPF 检查通过后，便创建一个（S，G）表项，并将该报文向网络中的所有下游节点转发。经过扩散，PIM-DM 域内的每个路由器上都会创建（S，G）表项。

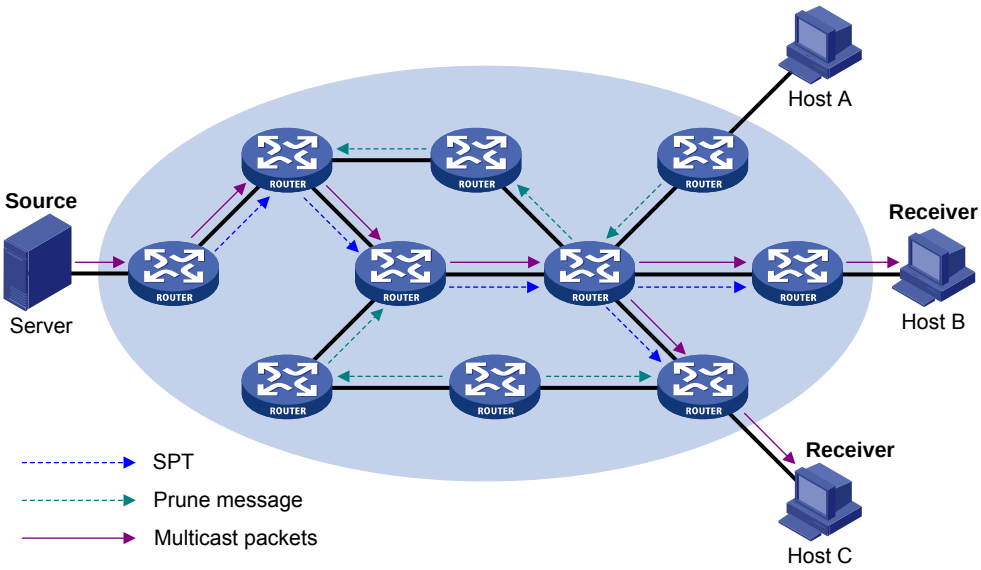
(2) 然后对那些下游没有接收者的节点进行剪枝：由没有接收者的下游节点向上游节点发剪枝报文（Prune Message），以通知上游节点将相应的接口从其组播转发表项（S，G）所对应的出接口列表中删除，并不再转发该组播组的报文至该节点。

说明

- （S，G）表项包括组播源的地址 S、组播组的地址 G、出接口列表和入接口等。
- 路由器上收到组播数据的接口称为“上游”，转发组播数据的接口称为“下游”。

剪枝过程最先由叶子路由器发起，如 图 1-1 所示，没有接收者（Receiver）的路由器（如与Host A直连的路由器）主动发起剪枝，并一直持续到PIM-DM域中只剩下必要的分支，这些分支共同构成了SPT。

图1-1 PIM-DM 中构建 SPT 示意图



“扩散—剪枝”的过程是周期性发生的。各个被剪枝的节点提供超时机制，当剪枝超时后便重新开始这一过程。

说明

剪枝在 PIM-SM 中有着相似的应用。

3. 嫁接

当被剪枝的节点上出现了组播组的成员时，为了减少该节点恢复成转发状态所需的时间，PIM-DM 使用嫁接机制主动恢复其对组播数据的转发，过程如下：

(1) 需要恢复接收组播数据的节点向其上游节点发送嫁接报文（Graft Message）以申请重新加入到 SPT 中；

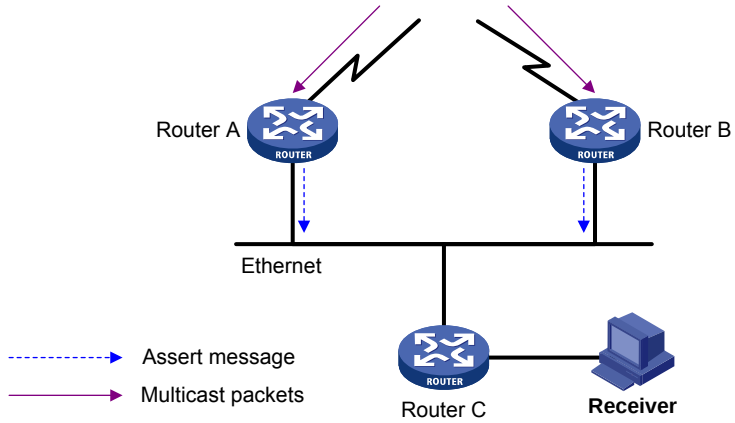
(2) 当上游节点收到该报文后恢复该下游节点的转发状态，并向其回应一个嫁接应答报文（Graft-Ack Message）以进行确认；

(3) 如果发送嫁接报文的下游节点没有收到来自其上游节点的嫁接应答报文，将重新发送嫁接报文直到被确认为止。

4. 断言

在一个网段内如果存在多台组播路由器，则相同的组播报文可能会被重复发送到该网段。为了避免出现这种情况，就需要通过断言（Assert）机制来选定唯一的组播数据转发者。

图1-2 Assert 机制示意图



如 图 1-2 所示，当Router A和Router B从上游节点收到（S，G）组播报文后，都会向本地网段转发该报文，于是处于下游的节点Router C就会收到两份相同的组播报文，Router A和Router B也会从各自的下游接口收到对方转发来的该组播报文。此时，Router A和Router B会通过其下游接口向本网段的所有PIM路由器（224.0.0.13）以组播方式发送断言报文（Assert Message），该报文中携带有以下信息：组播源地址S、组播组地址G、到组播源的单播路由/MBGP路由/组播静态路由的优先级和度量值。通过一定的规则对这些参数进行比较后，Router A和Router B中的获胜者将成为（S，G）组播报文在本网段的转发者，比较规则如下：

(1) 到组播源的优先级较高者获胜；

(2) 如果到组播源的优先级相等，那么到组播源的度量值较小者获胜；

(3) 如果到组播源的度量值也相等，则下游接口 IP 地址较大者获胜。

1.1.2 PIM-SM简介

PIM-DM 使用以“扩散—剪枝”方式构建的 SPT 来传送组播数据。尽管 SPT 的路径最短，但是其建立的过程效率较低，并不适合大中型网络。

PIM-SM 属于稀疏模式的组播路由协议，使用“拉（Pull）模式”传送组播数据，通常适用于组播组成员分布相对分散、范围较广的大中型网络，其基本原理如下：

- PIM-SM 假设所有主机都不需要接收组播数据，只向明确提出需要组播数据的主机转发。PIM-SM 实现组播转发的核心任务就是构造并维护 RPT（Rendezvous Point Tree，共享树），RPT 选择 PIM 域中某台路由器作为公用的根节点 RP（Rendezvous Point，汇集点），组播数据通过 RP 沿着 RPT 转发给接收者；
- 连接接收者的路由器向某组播组对应的 RP 发送加入报文（Join Message），该报文被逐跳送达 RP，所经过的路径就形成了 RPT 的分支；
- 组播源如果要向某组播组发送组播数据，首先由组播源侧 DR（Designated Router，指定路由器）负责向 RP 进行注册，把注册报文（Register Message）通过单播方式发送给 RP，该报文到达 RP 后触发建立 SPT。之后组播源把组播数据沿着 SPT 发向 RP，当组播数据到达 RP 后，被复制并沿着 RPT 发送给接收者。

说明

复制仅发生在分发树的分支处，这个过程能够自动重复直到数据包最终到达接收者。

PIM-SM 的工作机制可以概括如下：

- 邻居发现
- DR 选举
- RP 发现
- 构建 RPT
- 组播源注册
- SPT 切换
- 断言

1. 邻居发现

PIM-SM使用与PIM-DM类似的邻居发现机制，具体请参见“1.1.1 1. 邻居发现”一节。

2. DR选举

借助 Hello 报文还可以为共享网络（如 Ethernet）选举 DR，DR 将作为该共享网络中组播数据的唯一转发者。

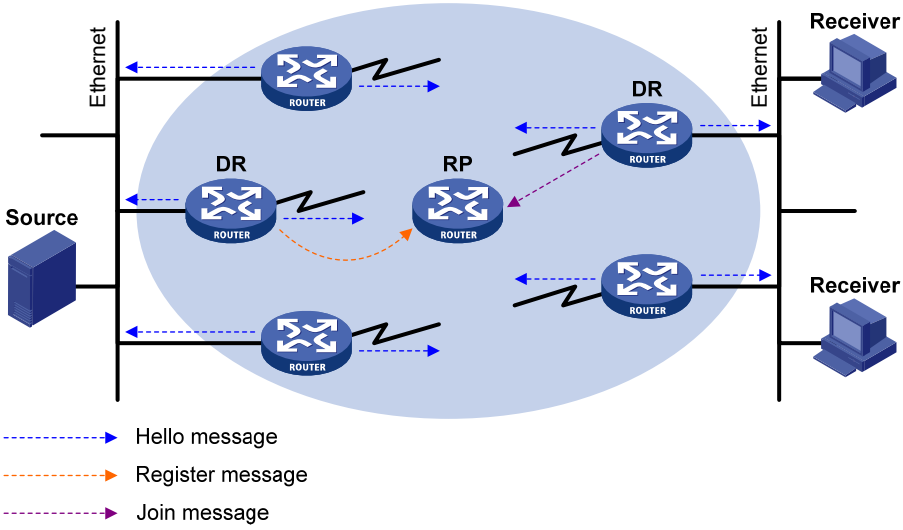
无论是与组播源相连的网络，还是与接收者相连的网络，都需要选举 DR。接收者侧的 DR 负责向 RP 发送加入报文；组播源侧的 DR 负责向 RP 发送注册报文。

说明

- 各路由器之间通过比较 Hello 报文中所携带的优先级和 IP 地址，可以为多路由器网段选举 DR。选举出的 DR 对于 PIM-SM 有实际的意义；而对于 PIM-DM 来说，其本身其实并不需要 DR，但如果 PIM-DM 域中的共享网络上运行了 IGMPv1，则需要选举出 DR 来充当共享网络上的 IGMPv1 查询器。
- 在充当接收者侧 DR 的设备上必须使能 IGMP，否则连接在该 DR 上的接收者将不能通过该 DR 加入组播组。

有关 IGMP 的介绍，请参见“IP 组播配置指导”中的“IGMP”。

图1-3 DR 选举示意图



如 图 1-3 所示，DR的选举过程如下：

(1) 共享网络上的各路由器相互之间发送 Hello 报文（携带有竞选 DR 优先级的参数），拥有最高优先级的路由器将成为 DR；

(2) 如果优先级相同，或者网络中至少有一台路由器不支持在 Hello 报文中携带竞选 DR 优先级的参数，则根据各路由器的 IP 地址大小来竞选 DR，IP 地址最大的路由器将成为 DR。

当 DR 出现故障时，其余路由器在超时后仍没有收到来自 DR 的 Hello 报文，则会触发新的 DR 选举过程。

3. RP发现

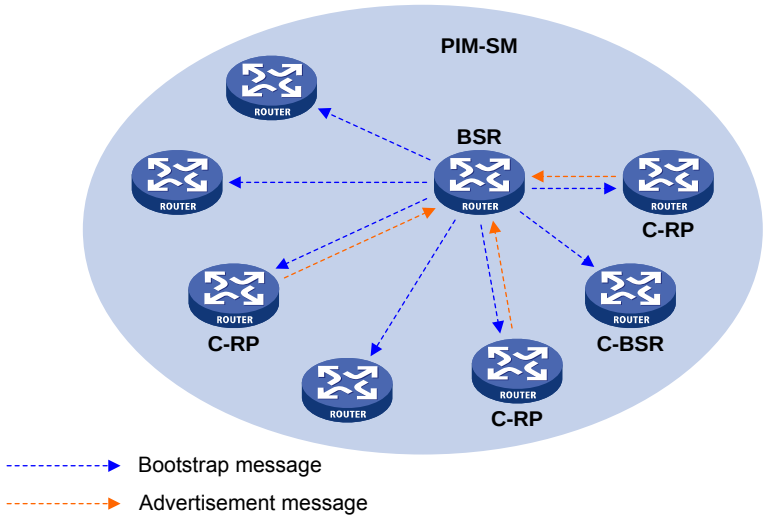
RP 是 PIM-SM 域中的核心设备。在结构简单的小型网络中，组播信息量少，整个网络仅依靠一个 RP 进行组播信息的转发即可，此时可以在 PIM-SM 域中的各路由器上静态指定 RP 的位置；但是在更多的情况下，PIM-SM 域的规模都很大，通过 RP 转发的组播信息量巨大。为了缓解 RP 的负担并优化 RPT 的拓扑结构，可以在 PIM-SM 域中配置多个 C-RP（Candidate-RP，候选 RP），通过自举机制来动态选举 RP，使不同的 RP 服务于不同的组播组，此时需要配置 BSR（Bootstrap Router，自举路由器）。BSR 是 PIM-SM 域的管理核心，一个 PIM-SM 域内只能有一个 BSR，但可以配置多个 C-BSR（Candidate-BSR，候选 BSR）。这样，一旦 BSR 发生故障，其余 C-BSR 能够通过自动选举产生新的 BSR，从而确保业务免受中断。

说明

- 一个 RP 可以同时服务于多个组播组，但一个组播组只能唯一对应一个 RP。
- 一台设备可以同时充当 C-RP 和 C-BSR。

如 图 1-4 所示，BSR负责收集网络中由C-RP发来的宣告报文（Advertisement Message），该报文中携带有C-RP的地址和优先级以及其服务的组范围，BSR将这些信息汇总为RP-Set（RP集，即组播组与RP的映射关系数据库），封装在自举报文（Bootstrap Message，BSM）中并发布到整个PIM-SM域。

图1-4 RP 与 BSR 信息交互示意图



网络中的各路由器将依据 RP-Set 提供的信息，使用相同的规则从众多 C-RP 中为特定组播组选择其对应的 RP，具体规则如下：

(1) 首先比较 C-RP 所服务的组范围，所服务的组范围较小者获胜。

(2) 若服务的组范围相同，再比较 C-RP 的优先级，优先级较高者获胜。

(3) 若优先级也相同，再使用哈希（Hash）函数计算哈希值，哈希值较大者获胜。

(4) 若哈希值也相同，则 C-RP 的 IP 地址较大者获胜。

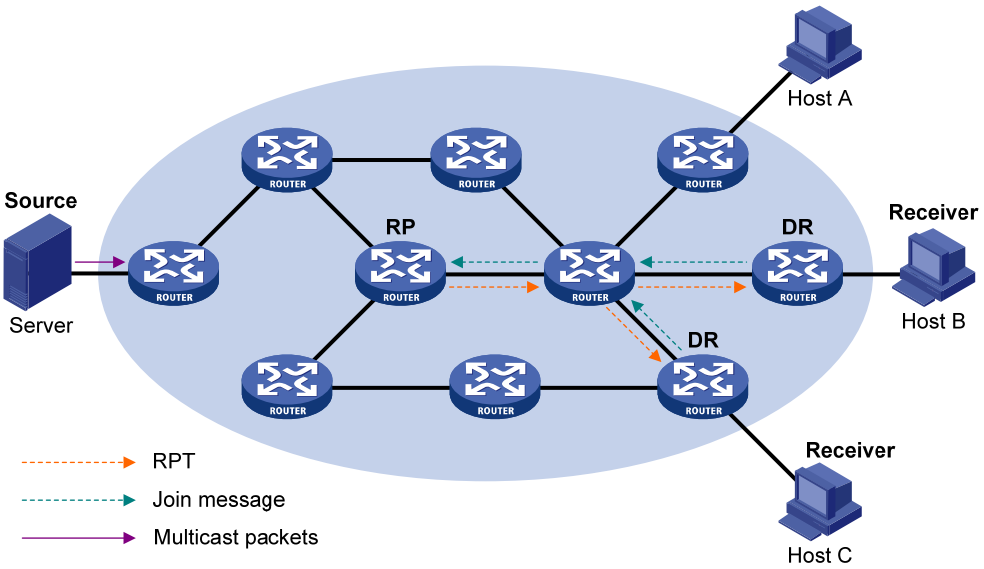
哈希函数的表达式为：Value (G, M, C_i) = (1103515245 * ((1103515245 * (G & M) + 12345) XOR C_i) + 12345) mod 2^{31}，其中各符号的含义如 表 1-1 所示。

表1-1 哈希函数中各符号含义

符号	含义
Value	哈希值
G	IP组播组的地址
M	哈希掩码长度（Hash Mask Length）
C_i	C-RP的IP地址
&	逻辑运算符，表示与运算
XOR	逻辑运算符，表示异或运算
mod	算术运算符，表示整除取余

4. 构建RPT

图1-5 PIM-SM 中构建 RPT 示意图



如 图 1-5 所示，RPT的构建过程如下：

(1) 当接收者加入一个组播组 G 时，先通过 IGMP 报文通知与其直连的 DR；

(2) DR 掌握了组播组 G 的接收者的信息后，向该组所对应的 RP 方向逐跳发送加入报文；

(3) 从 DR 到 RP 所经过的路由器就形成了 RPT 的分支，这些路由器都在其转发表中生成了（*，G）表项，这里的“*”表示来自任意组播源。RPT 以 RP 为根，以 DR 为叶子。

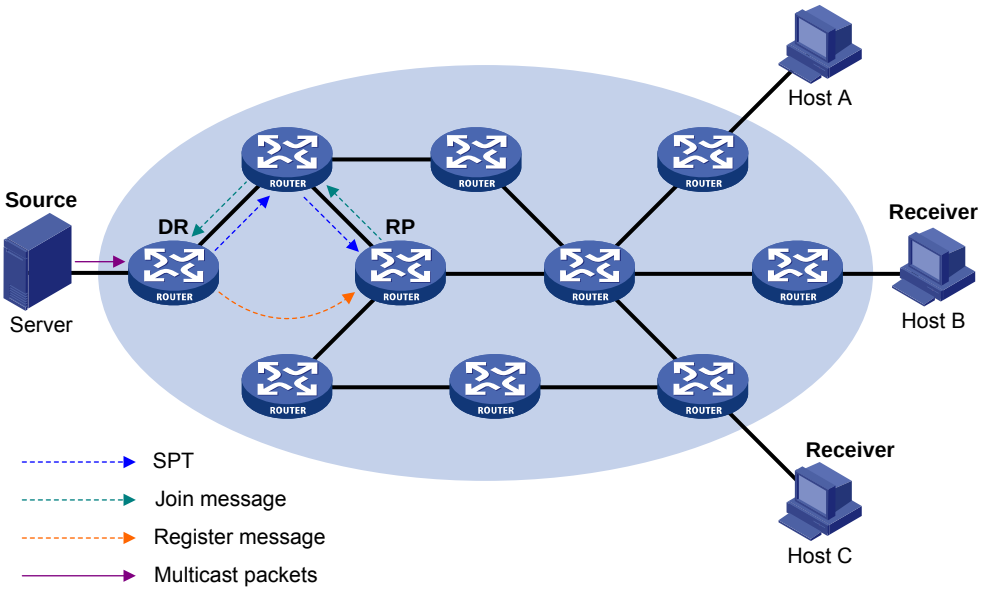
当发往组播组 G 的组播数据流经 RP 时，数据就会沿着已建立好的 RPT 到达 DR，进而到达接收者。

当某接收者对组播组 G 的信息不再感兴趣时，与其直连的 DR 会逆着 RPT 向该组的 RP 方向逐跳发送剪枝报文；上游节点收到该报文后在其出接口列表中删除与下游节点相连的接口，并检查自己是否拥有该组播组的接收者，如果没有则继续向其上游转发该剪枝报文。

5. 组播源注册

组播源注册的目的是向 RP 通知组播源的存在。

图1-6 组播源注册示意图



如 图 1-6 所示，组播源向RP注册的过程如下：

(1) 当组播源 S 向组播组 G 发送了一个组播报文时，与组播源直连的 DR 在收到该报文后，就将其封装成注册报文，并通过单播方式发送给相应的 RP；

(2) 当 RP 收到该报文后，一方面解封装注册报文并将封装在其中的组播报文沿着 RPT 转发给接收者，另一方面向组播源方向逐跳发送（S，G）加入报文。这样，从 RP 到组播源所经过的路由器就形成了 SPT 的分支，这些路由器都在其转发表中生成了（S，G）表项。SPT 以组播源侧的 DR 为根，以 RP 为叶子。

(3) 组播源发出的组播数据沿着已建立好的 SPT 到达 RP，然后由 RP 把组播数据沿着 RPT 向接收者进行转发。当 RP 收到沿着 SPT 转发来的组播数据后，通过单播方式向与组播源直连的 DR 发送注册停止报文（Register-Stop Message），组播源注册过程结束。

说明

上述描述中假定允许 RP 发起 SPT 切换，否则组播源侧 DR 将一直用注册报文封装组播报文，注册过程不会结束，除非 RP 上（S，G）表项的出接口变为空。

6. SPT切换

在 PIM-SM 域中，一个组播组唯一对应一个 RP 和一棵 RPT。在 SPT 切换前，所有发往该组的组播报文都必须先由组播源侧 DR 封装在注册报文中发往 RP，由 RP 解封装后再沿 RPT 分发给接收者侧的 DR，RP 是所有组播数据必经的中转站。这个过程存在以下三个问题：

- 组播源侧的 DR 和 RP 必须对组播数据进行繁琐的封装/解封装处理。
- 组播数据的转发路径不一定是从组播源到接收者的最短路径。
- 当组播流量变大时，RP 负担增大，容易引发故障。

为了解决上述问题，PIM-SM 允许由 RP 或接收者侧的 DR 发起 SPT 切换：

(1) RP 发起的 SPT 切换

RP 收到第一个组播数据包后，立即向组播源方向发送（S，G）加入报文，在组播源侧 DR 与 RP 之间建立起 SPT 分支，后续的组播报文都直接沿该分支到达 RP。

说明

由RP发起的SPT切换的详细过程，请参见“1.1.2 5. 组播源注册”一节。

(2) 接收者侧 DR 发起的 SPT 切换

当接收者侧 DR 收到第一个组播数据包后，立即发起 SPT 切换，过程如下：

- 首先，接收者侧 DR 向组播源方向发送（S，G）加入报文，并最终送达组播源侧 DR，沿途经过的所有路由器在其转发表中都生成了（S，G）表项，从而建立了 SPT 分支；
- 随后，当组播数据沿 SPT 到达 RPT 与 SPT 分叉的路由器时，该路由器开始丢弃沿 RPT 到达的组播数据，同时向 RP 逐跳发送含 RP 位的剪枝报文，RP 收到该报文后继续向组播源方向发送剪枝报文（假设此时只有这一个接收者），从而完成了 SPT 切换；
- 最终，组播数据将沿 SPT 从组播源到达到接收者。

通过 SPT 切换，PIM-SM 能够以比 PIM-DM 更经济的方式建立 SPT。

7. 断言

PIM-SM使用与PIM-DM类似的断言机制，具体请参见“1.1.1 4. 断言”一节。

1.1.3 双向PIM简介

在某些组网应用（譬如多方电视电话会议）中，同时存在多个接收者和多个组播源，在这种情况下，如果使用传统的 PIM-DM 或 PIM-SM 按 SPT 转发组播数据，需在每台路由器上针对每个组播源都创建（S，G）表项，这将占用大量的系统资源。为了解决这个问题，提出了双向 PIM 的概念。双向 PIM 由 PIM-SM 发展而来，它通过建立以 RP 为中心、分别连接组播源和接收者的双向 RPT，使组播数据沿着双向 RPT 从组播源经由 RP 转发到接收者。这样，在每台路由器上只需维护（*，G）表项即可，从而节约了系统资源。

双向 PIM 主要适用于组播源和接收者都比较密集的网络，其工作机制可以概括如下：

- 邻居发现
- RP 发现
- DF 选举
- 构建双向 RPT

1. 邻居发现

双向PIM使用与PIM-SM完全相同的邻居发现机制，具体请参见“1.1.2 1. 邻居发现”一节。

2. RP发现

双向PIM使用与PIM-SM完全相同的RP发现机制，具体请参见“1.1.2 3. RP发现”一节。

PIM-SM 的 RP 必须指定为一个实际存在的 IP 地址，而双向 PIM 的 RP 则可以指定为一个实际不存在的 IP 地址，简称 RPA（Rendezvous Point Address，汇集点地址）。RPA 所属网段对应的链路就称为 RPL（Rendezvous Point Link，汇集点链路），连接到 RPL 上的所有接口都可以充当 RP，且互为备份。

说明

双向 PIM 中的 RPF 接口是指向 RP 的接口，RPF 邻居自然就是到达 RP 的下一跳地址。

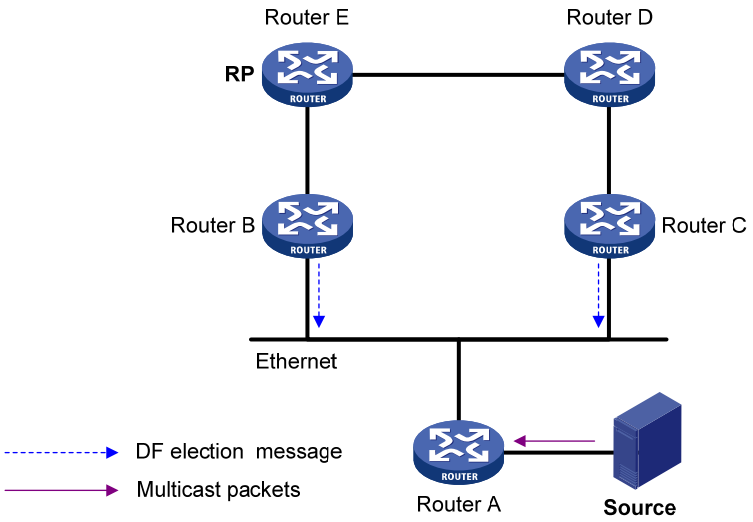
3. DF选举

DF（Designated Forwarder，指定转发者）是双向 PIM 中的重要角色，组播数据由组播源向 RP 转发的动力来自于 DF，也就是说只有 DF 才有能力将组播数据向 RP 方向转发。因此，每个 RP 在每个网段都需要有其对应的 DF，以负责将该网段的组播数据向该 RP 转发；此外，在有多台组播路由器的网段，DF 的唯一性也可以避免相同的组播报文被重复发往 RP。

说明

在 RPL 上不需要选举 DF。

图1-7 DF 选举示意图



如 图 1-7 所示，Router B和Router C都可以从Router A收到由组播源向组播组G发送的组播报文，如果它们都向下游节点转发该报文，RP最终将收到两份相同的组播报文。因此，Router B和Router C一旦获得RP的信息，就会为该RP发起DF的选举：Router B和Router C将分别向本网段的所有PIM路由器（224.0.0.13）以组播方式发送DF选举报文（DF Election Message），该报文携带有以下信息：RP的地址、到RP的单播路由/MBGP路由/组播静态路由的优先级和度量值。通过一定规则对这些参数进行比较后，Router B和Router C中的获胜者将成为DF，具体的比较规则如下：

(1) 到 RP 的优先级较高者获胜；

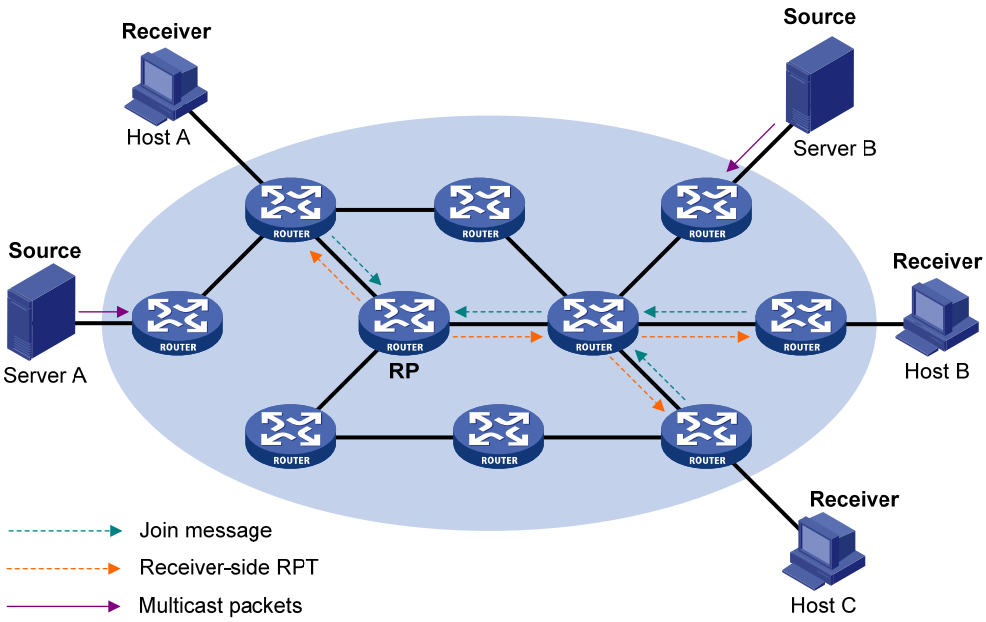
(2) 如果到 RP 的优先级相等，那么到 RP 的度量值较小者获胜；

(3) 如果到 RP 的度量值也相等，则接口的 IP 地址较大者获胜。

4. 构建双向RPT

双向 RPT 由两部分构成：一部分是以 RP 为根、以直连接收者的路由器为叶子的 RPT，简称接收者侧 RPT；而另一部分则是以 RP 为根、以直连组播源的路由器为叶子的 RPT，简称组播源侧 RPT。这两部分 RPT 的构建过程不同，下面分别加以介绍。

图1-8 接收者侧 RPT 构建示意图



接收者侧RPT的构建过程与PIM-SM中RPT的构建过程类似，如 图 1-8 所示，其构建过程如下：

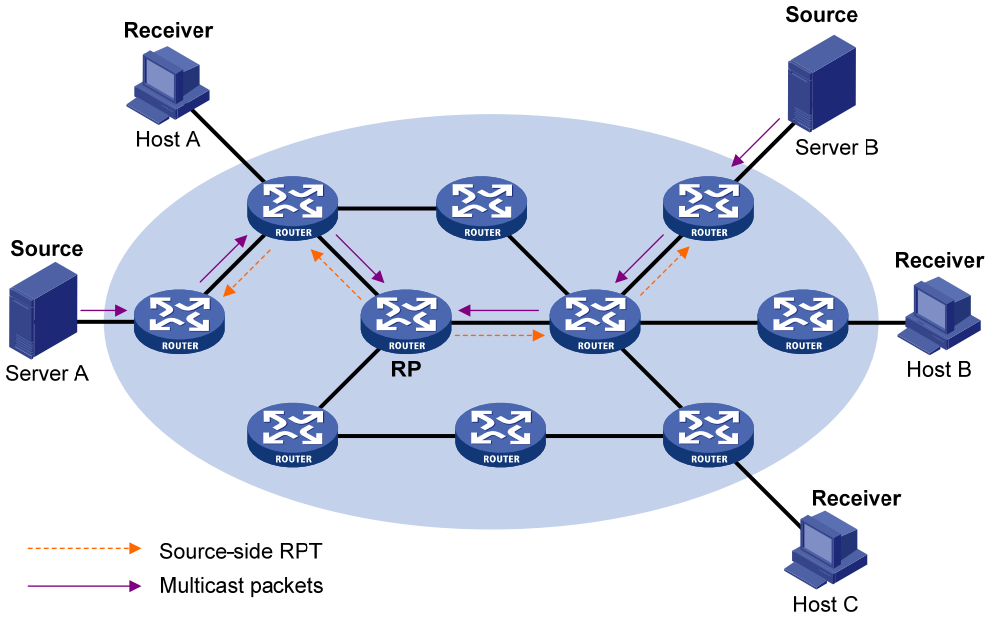
(1) 当接收者加入一个组播组 G 时，先通过 IGMP 报文通知与其直连的路由器；

(2) 该路由器掌握了组播组 G 的接收者的信息后，向该组所对应的 RP 方向逐跳发送加入报文；

(3) 从直连接收者的路由器到 RP 所经过的路由器就形成了接收者侧 RPT 的分支，这些路由器都在其转发表中生成了（*，G）表项。

当某接收者对组播组 G 的信息不再感兴趣时，与其直连的路由器会逆着接收者侧 RPT 向该组的 RP 方向逐跳发送剪枝报文；上游节点收到该报文后在其出接口列表中删除与下游节点相连的接口，并检查自己是否拥有该组播组的接收者，如果没有则继续向其上游转发该剪枝报文。

图1-9 组播源侧 RPT 构建示意图



组播源侧RPT的构建过程则相对简单，如 图 1-9 所示，其构建过程如下：

(1) 组播源发向组播组 G 的组播数据在途径的每个网段，都被该网段的 DF 无条件地向 RP 转发；

(2) 从直连组播源的路由器到 RP 所经过的路由器就形成了组播源侧 RPT 的分支，这些路由器都在其转发表中生成了（*，G）表项。

当双向 RPT 构建完成之后，由组播源发出的组播数据将依次沿着组播源侧 RPT 和接收者侧 RPT，经由 RP 转发至接收者。

说明

当接收者和组播源位于 RP 同一侧时，组播源侧 RPT 与接收者侧 RPT 有可能在到达 RP 之前就已汇合。在这种情况下，由该组播源发往该接收者的组播数据将在此汇合点直接被转发给该接收者，而不必经由 RP。

1.1.4 管理域机制简介

1. 两种域机制的划分

一般情况下，在一个 PIM-SM/双向 PIM 域内只能有一个 BSR，并由该 BSR 负责在整个 PIM-SM/双向 PIM 域内宣告 RP-Set 信息，所有组播组的信息都在此 BSR 管理的网络范围内进行转发，我们称之为非管理域机制。

考虑到管理的精细化，可以将整个 PIM-SM/双向 PIM 域划分为一个 Global 域（Global-scope Zone）和多个管理域（Admin-scope Zone），一方面可以有效分担单一 BSR 的管理压力，另一方面可以使用私有组地址为特定区域提供专门的服务。相应地，我们称之为管理域机制。

管理域与组播组相对应，针对不同组播组划分相应的管理域。管理域的边界由 ZBR（Zone Border Router，区域边界路由器）构成，每个管理域各维护一个 BSR，为特定范围的组播组服务，属于此

范围的组播协议报文（如断言报文、BSR 自举报文等）无法通过管理域边界。不同管理域所服务的组播组范围可以重叠，该范围内的组播组只在本管理域内有效，相当于私有组地址。而不属于任何管理域服务范围的组播组则一律属于 Global 域的服务范围，Global 域中维护一个 BSR，为剩余的所有组播组服务。

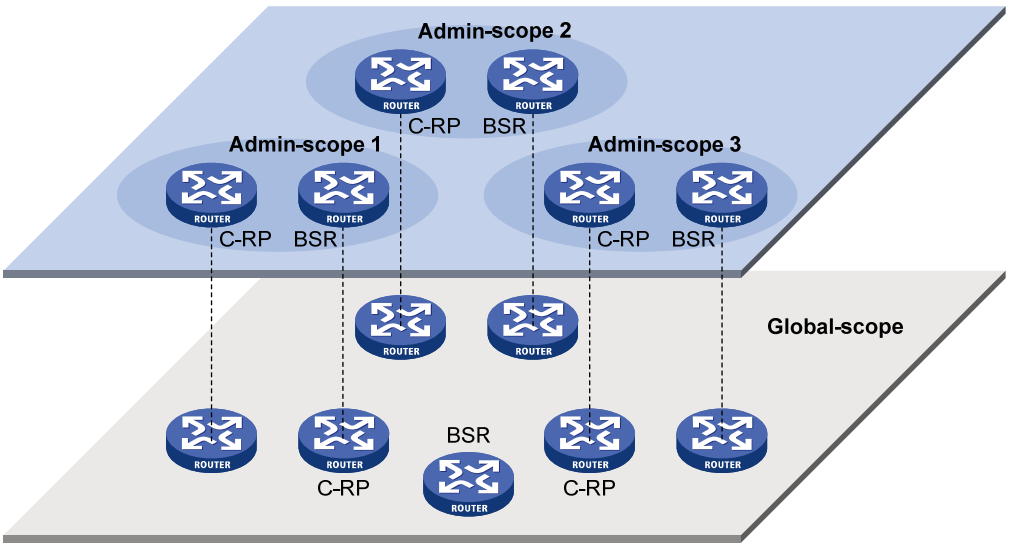
2. 管理域与Global域的关系

每个管理域以及 Global 域都有独立的 C-RP 和 BSR 设备，这些设备仅在其所属的域有效，也就是说 BSR 机制与 RP 选举在各管理域之间是隔离的；每个管理域都有自己的边界，各管理域所服务组播组范围内的组播信息不能进、出该边界。为了更清晰地理解管理域和 Global 域之间的关系，可以从以下两个角度进行考虑：

(1)　地域空间角度

管理域是针对特定范围组播组的逻辑管理区域，属于此范围的组播报文只能在本管理域的域内或域外传播，无法跨过管理域的边界。

图1-10 地域空间上管理域与 Global 域的关系

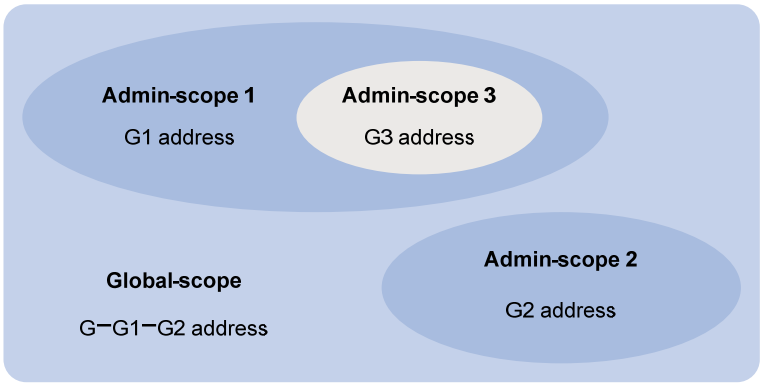


如 图 1-10 所示，对于同一地址范围内的组播组而言，各管理域在地域上必须相互独立、相互隔离，即同一路由器不能从属于多个管理域，各管理域所包含的路由器也互不相同。而Global域则包含了PIM-SM/双向PIM域内的所有路由器，不属于任何管理域服务范围的组播报文，可以在整个PIM-SM/双向PIM域范围内传播。

(2)　组地址范围角度

每个管理域为特定的组播组提供服务，这些组播组地址之间通常没有交集，但是也可能存在相互交叉和重叠关系。

图1-11 组地址范围上管理域与 Global 域的关系



如 图 1-11 所示，管理域 1 与管理域 2 所对应的组地址范围无交集，而管理域 3 的组地址是管理域 1 组地址的子集；Global域所对应的组地址范围是除各管理域组地址外的其它所有组地址，即G−G1−G2。也就是说，Global域和所有管理域之间就组地址范围来说是互补关系。

1.1.5 PIM-SSM简介

SSM（Source-Specific Multicast，指定信源组播）模型和 ASM（Any-Source Multicast，任意信源组播）模型是两个完全对等的模型。目前，ASM 模型包括 PIM-DM 和 PIM-SM 两种模式，SSM 模型能够借助 PIM-SM 的部分技术来实现，也称为 PIM-SSM。

SSM 模型为指定源组播提供了解决方案，通过 IGMPv3 来维护主机与路由器之间的关系。在实际应用中，通常采用 IGMPv3 以及 PIM-SM 的一部分技术来实现 SSM 模型。由于接收者已经通过其它渠道（如广告咨询等）知道了组播源的具体位置，因此在 SSM 模型中无需 RP，无需构建 RPT，无需组播源注册过程，也无需通过 MSDP（Multicast Source Discovery Protocol，组播源发现协议）来发现其它 PIM 域内的组播源。

PIM-SSM 的工作机制可以概括如下：

- 邻居发现
- DR 选举
- 构建 SPT

1. 邻居发现

PIM-SSM使用与PIM-SM完全相同的邻居发现机制，具体请参见“1.1.2 1. 邻居发现”一节。

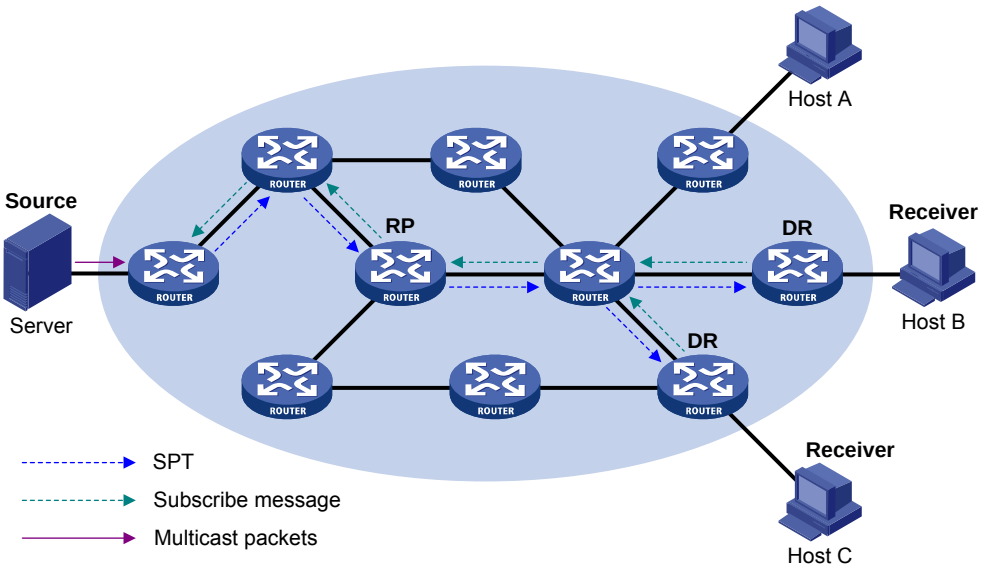
2. DR选举

PIM-SSM使用与PIM-SM完全相同的DR选举机制，具体请参见 “1.1.2 2. DR选举”一节。

3. 构建SPT

构建为 PIM-SM 服务的 RPT，还是构建为 PIM-SSM 服务的 SPT，关键在于接收者准备加入的组播组是否属于 SSM 组地址范围（IANA 保留的 SSM 组地址范围为 232.0.0.0/8）。

图1-12 PIM-SSM 中构建 SPT 示意图



如 图 1-12 所示，Host B和Host C为组播信息的接收者（Receiver），由其借助IGMPv3 的报告报文向DR报告自己对来自组播源S、发往组播组G的信息感兴趣。收到该报告报文的DR先判断该报文中的组地址是否在SSM组地址范围内：

- 如果在 SSM 组地址范围内，则构建 PIM-SSM，并向组播源 S 逐跳发送通道（Channel）的订阅报文（Subscribe Message）。沿途所有路由器上都创建（S，G）表项，从而在网络内构建了一棵以组播源 S 为根、以接收者为叶子的 SPT，该 SPT 就是 PIM-SSM 中的传输通道；
- 如果不在 SSM 组地址范围内，则仍旧按照 PIM-SM 的流程进行后续处理，此时接收者侧 DR 需要向 RP 发送（*，G）加入报文，同时组播源侧 DR 需要进行组播源的注册。

说明

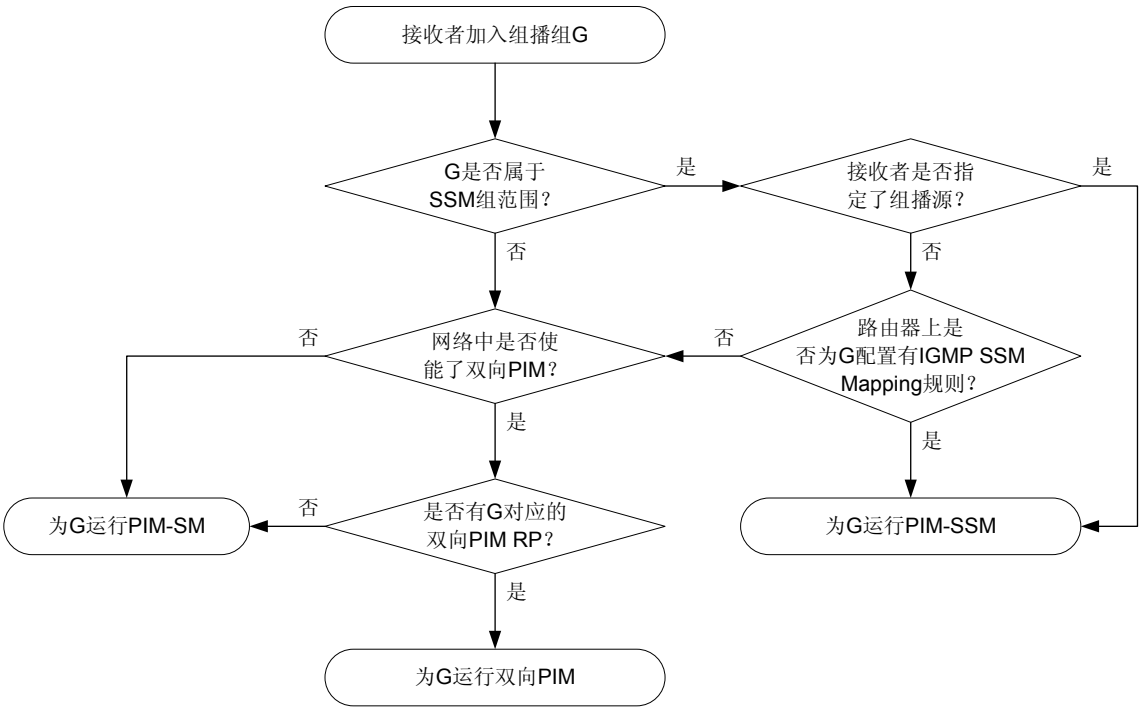
在 PIM-SSM 中，借助“通道”的概念表示组播组，借助“订阅报文”的概念表示加入报文。

1.1.6 各PIM协议运行关系

在一个 PIM 网络中，不允许 PIM-DM 与其它类型的 PIM 协议（PIM-SM、双向 PIM 和 PIM-SSM）同时运行，但允许同时运行 PIM-SM、双向 PIM 和 PIM-SSM。

当网络中同时运行PIM-SM、双向PIM和PIM-SSM时，针对具体的组加入行为运行哪种类型的PIM协议，其判断过程如 图 1-13 所示。

图1-13 各 PIM 协议运行关系示意图



说明

有关 IGMP SSM Mapping 的详细介绍，请参见“IP 组播配置指导”中的“IGMP”。

1.1.7 多实例的PIM

在多实例应用中，组播路由器需要针对不同的实例分别维护 PIM 邻居表、组播路由表、BSR 信息和 RP-Set 信息，并保持各实例间上述信息的相互独立。

当组播路由器收到组播数据报文时，需要区分出该数据报文所属的实例，并根据该实例对应的组播路由表将其转发，或创建与该实例的 PIM 相关的组播路由表项。

1.1.8 协议规范

与 PIM 相关的协议规范有：

- RFC 3973：Protocol Independent Multicast-Dense Mode (PIM-DM): Protocol Specification (Revised)
- RFC 4601：Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification (Revised)
- RFC 5015：Bidirectional Protocol Independent Multicast (BIDIR-PIM)
- RFC 5059：Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)
- RFC 4607：Source-Specific Multicast for IP
- draft-ietf-ssm-overview-05：An Overview of Source-Specific Multicast (SSM)

1.2 配置PIM-DM

1.2.1 PIM-DM配置任务简介

表1-2 PIM-DM 配置任务简介

配置任务	说明	详细配置
使能PIM-DM	必选	1.2.3
使能状态刷新能力	可选	1.2.4
配置状态刷新参数	可选	1.2.5
配置PIM-DM定时器	可选	1.2.6
配置PIM公共特性	可选	1.6

1.2.2 配置准备

在配置 PIM-DM 之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通

在配置 PIM-DM 之前，需准备以下数据：

- 发送状态刷新报文的时间间隔
- 等待接收新状态刷新报文的最小时间
- 状态刷新报文的 TTL 值
- 嫁接报文的重传时间

1.2.3 使能PIM-DM

在接口上使能了 PIM-DM 后，路由器会定期发送 Hello 报文以发现 PIM 邻居，并对收到的来自 PIM 邻居的报文进行处理。在部署 PIM-DM 域时，建议在其所有非边界接口上均使能 PIM-DM。

1. 使能公网实例中的PIM-DM

表1-3 使能公网实例中的 PIM-DM

操作	命令	说明
进入系统视图	**system-view**	-
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能PIM-DM	**pim dm**	必选 缺省情况下，PIM-DM处于关闭状态

2. 使能VPN实例中的PIM-DM

表1-4 使能 VPN 实例中的 PIM-DM

操作	命令	说明
进入系统视图	**system-view**	-
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
将接口与VPN实例进行关联	**ip binding vpn-instance** *vpn-instance-name*	必选 缺省情况下，接口只属于公网，未与任何VPN实例关联
使能PIM-DM	**pim dm**	必选 缺省情况下，PIM-DM处于关闭状态

注意

- 在同一台设备的同一个 VPN 实例中，所有接口上所启用的 PIM 模式必须相同。
- PIM-DM 不能与处于 SSM 组地址范围内的组播组同时使用。

说明

- 有关 **ip vpn-instance**、**route-distinguisher** 和 **ip binding vpn-instance** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。
- 有关 **multicast routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

1.2.4 使能状态刷新能力

为了避免各路由器上被剪枝的接口因为超时而恢复转发，与组播源直连的路由器会周期性地发送（S，G）状态刷新报文，该报文沿着 PIM-DM 域最初的扩散路径逐跳进行转发，从而刷新沿途所有路由器上的剪枝定时器的状态。只有当一个共享网段内的所有 PIM 路由器上都使能了状态刷新能力时，该共享网段才具备状态刷新能力。

请在 PIM-DM 域内的所有路由器上进行如下配置。

表1-5 使能状态刷新能力

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能状态刷新能力	**pim state-refresh-capable**	可选 缺省情况下，状态刷新能力处于使能状态

1.2.5 配置状态刷新参数

在与组播源直连的路由器上，会以一定的时间间隔周期性地发送状态刷新报文，可以通过配置来改变这个时间间隔。

路由器可能在短时间内收到多个状态刷新报文，而其中有些报文可能是重复的。为了避免接收这些重复的报文，可以配置接收新状态刷新报文的等待时间：路由器将丢弃在该时间内收到的状态刷新报文；当该时间超时后，路由器将正常接收新的状态刷新报文，并更新自己的 PIM-DM 状态，同时重置该等待时间。

在收到状态刷新报文时，路由器会将该报文的 TTL 值减 1 后转发给其下游，直至该报文的 TTL 值减为 0，当网络规模很小时，状态刷新报文将在网络中循环传递。因此，为了有效控制刷新报文的传递范围，需要根据网络规模大小配置合适的 TTL 值。

请在 PIM-DM 域内的所有路由器上进行如下配置。

表1-6 配置状态刷新参数

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置发送状态刷新报文的时间间隔	**state-refresh-interval** *interval*	可选 缺省情况下，发送状态刷新报文的时间间隔为60秒
配置接收新状态刷新报文的等待时间	**state-refresh-rate-limit** *interval*	可选 缺省情况下，接收新状态刷新报文的等待时间为30秒
配置状态刷新报文的TTL值	**state-refresh-ttl** *ttl-value*	可选 缺省情况下，状态刷新报文的TTL值为255

1.2.6 配置PIM-DM定时器

嫁接报文是 PIM-DM 中唯一使用确认机制的报文。在 PIM-DM 域中，下游路由器发出嫁接报文后，如果在指定时间内没有收到来自其上游路由器的嫁接应答报文，则会重发嫁接报文，直到被确认。

表1-7 配置 PIM-DM 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置嫁接报文的重传时间	**pim timer graft-retry** *interval*	可选 缺省情况下，嫁接报文的重传时间为3秒

说明

有关PIM-DM其它定时器的相关配置，请参见“1.6.7 配置PIM公共定时器”。

1.3 配置PIM-SM

1.3.1 PIM-SM配置任务简介

表1-8 PIM-SM 配置任务简介

配置任务		说明	详细配置
使能PIM-SM		必选	1.3.3
配置RP	配置静态RP	三者必选其一	1.3.4 1.
	配置C-RP		1.3.4 2.
	使能自动RP侦听		1.3.4 3.
	全局配置C-RP定时器	可选	1.3.4 4.
配置BSR	配置C-BSR	必选	1.3.5 1.
	配置BSR服务边界	可选	1.3.5 2.
	全局配置C-BSR参数	可选	1.3.5 3.
	配置C-BSR定时器	可选	1.3.5 4.
	关闭自举报文语义分片功能	可选	1.3.5 5.
配置管理域	使能管理域机制	可选	1.3.6 1.
	配置管理域边界	可选	1.3.6 2.
	配置管理域和Global域的C-BSR	可选	1.3.6 3.
配置组播源注册		可选	1.3.7
配置禁止SPT切换		可选	1.3.8
配置PIM公共特性		可选	1.6

1.3.2 配置准备

在配置 PIM-SM 之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通

在配置 PIM-SM 之前，需准备以下数据：

- 静态 RP 的 IP 地址及表示其所服务组范围的 ACL 规则
- C-RP 的优先级及表示其所服务组范围的 ACL 规则
- 合法 C-RP 的地址范围及表示其所服务组范围的 ACL 规则
- 发送宣告报文的时间间隔
- C-RP 超时时间
- C-BSR 的优先级
- 哈希掩码长度
- 表示合法 BSR 地址范围的 ACL 规则
- 自举时间间隔
- 自举超时时间
- 表示注册报文过滤规则的 ACL 规则
- 注册抑制时间
- 注册探测时间
- 禁止发起 SPT 切换的 ACL 规则和排序规则

1.3.3 使能PIM-SM

在接口上使能了 PIM-SM 后，路由器会定期发送 Hello 报文以发现 PIM 邻居，并对收到的来自 PIM 邻居的报文进行处理。在部署 PIM-SM 域时，建议在其所有非边界接口上均使能 PIM-SM。

1. 使能公网实例中的PIM-SM

表1-9 使能公网实例中的 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能PIM-SM	**pim sm**	必选 缺省情况下，PIM-SM处于关闭状态

2. 使能VPN实例中的PIM-SM

表1-10 使能 VPN 实例中的 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
将接口与VPN实例进行关联	**ip binding vpn-instance** *vpn-instance-name*	必选 缺省情况下，接口只属于公网，未与任何VPN实例关联
使能PIM-SM	**pim sm**	必选 缺省情况下，PIM-SM处于关闭状态

注意

在同一台设备的同一个 VPN 实例中，所有接口上所启用的 PIM 模式必须相同。

说明

- 有关 **ip vpn-instance**、**route-distinguisher** 和 **ip binding vpn-instance** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。
- 有关 **multicast routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

1.3.4 配置RP

RP 可以通过手工方式静态配置，也可以通过 BSR 机制动态选举。对于大型 PIM 网络，配置静态 RP 将会非常繁琐。所以，通常静态 RP 是作为动态选举 RP 机制的备份手段，以提高网络的健壮性，增强组播网络的运营管理能力。

注意

当 PIM 网络中同时运行 PIM-SM 和双向 PIM 时，请勿使一个 RP 同时为 PIM-SM 和双向 PIM 工作，否则可能引起 PIM 路由表出错。

1. 配置静态RP

当网络内仅有一个动态 RP 时，通过手工配置静态 RP 可以避免因单一节点故障而引起的通信中断，同时也可以避免 C-RP 与 BSR 之间频繁的信息交互而占用带宽。

请在 PIM-SM 域内的所有路由器上进行如下配置。

表1-11 配置静态 RP

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置服务于PIM-SM的静态RP	**static-rp** *rp-address* [*acl-number*] [**preferred**]	必选 缺省情况下，没有配置静态RP

注意

为了让静态 RP 功能正常发挥作用，必须在所有路由器上指定相同的静态 RP 地址。

2. 配置C-RP

在 PIM-SM 域中，可以把有意成为 RP 的路由器配置为 C-RP。BSR 通过接收来自 C-RP 的 C-RP 信息，或者接收来自其它路由器的自动 RP 宣告，收集 C-RP 信息并将其汇总为 RP-Set 信息，然后在全网内扩散。之后，网络内的其它路由器根据 RP-Set 信息计算出特定组播组范围所对应的 RP。建议在骨干网路由器上配置 C-RP。

为了防止 C-RP 欺骗，需要在 BSR 上配置合法的 C-RP 地址范围及其所服务的组播组范围。同时由于每个 C-BSR 都可能成为 BSR，因此必须在 PIM-SM 域内的所有 C-BSR 上都配置相同的过滤策略。

表1-12 配置 C-RP

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置某接口为服务于PIM-SM的C-RP	**c-rp** *interface-type interface-number* [**group-policy** *acl-number* | **priority** *priority* | **holdtime** *hold-interval* | **advertisement-interval** *adv-interval*] *	必选 缺省情况下，没有配置C-RP
配置合法C-RP的地址范围及其所服务的组播组范围	**crp-policy** *acl-number*	可选 缺省情况下，C-RP的地址范围及其所服务的组播组范围不受任何限制

说明

- 在配置 C-RP 时，应在 C-RP 与 PIM-SM 域中的其它设备之间保留较大的通信带宽。
- 一个 RP 可以为多个组播组服务，也可以为所有组播组服务。每个组播组在任意时刻，只能由唯一的一个 RP 为其转发数据，而不能由多个 RP 转发数据。

3. 使能自动RP侦听

自动 RP 宣告（Announce）和发现（Discovery）报文的目的地址分别为组播组地址 224.0.1.39 和 224.0.1.40。在设备上使能了自动 RP 侦听功能后，该设备便能够接收这两种报文，并记录报文中所携带的 RP 信息。

表1-13 使能自动 RP 侦听

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
使能自动RP侦听功能	**auto-rp enable**	必选 缺省情况下，自动RP侦听功能处于关闭状态

4. 全局配置C-RP定时器

为了使 BSR 能够在 PIM-SM 域内分发 RP-Set 信息，C-RP 必须周期性地向 BSR 发送宣告报文，BSR 从该报文中学习 RP-Set 信息，并将该信息与自己的 IP 地址一起封装在自举报文中向域中的所有 PIM 路由器进行宣告。

C-RP 在其宣告报文中封装一个保持时间，BSR 在收到该报文后，从中获得该时间值并启动 C-RP 超时定时器，如果超时后 BSR 仍没有收到来自 C-RP 后续的宣告报文，则认为目前网络中的 C-RP 失效或不可达。

请在已配置为 C-RP 的路由器上进行如下配置。

表1-14 全局配置 C-RP 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置发送宣告报文的时间间隔	**c-rp advertisement-interval** *interval*	可选 缺省情况下，发送宣告报文的时间间隔为60秒
配置C-RP超时时间	**c-rp holdtime** *interval*	可选 缺省情况下，C-RP的超时时间为150秒

说明

有关PIM-SM其它定时器的相关配置，请参见“1.6.7 配置PIM公共定时器”。

1.3.5 配置BSR

在一个 PIM-SM 域中只能有一个 BSR，但需要配置至少一个 C-BSR。任意一台路由器都可以被配置为 C-BSR。在 C-BSR 之间通过自动选举产生 BSR，BSR 负责在 PIM-SM 域中收集并发布 RP 信息。

1. 配置C-BSR

C-BSR 应配置在骨干网的路由器上，在将路由器配置为 C-BSR 时，必须同时指定一个使能了 PIM-SM 的接口。C-BSR 间的自动选举机制简单描述如下：

- 最初，每个 C-BSR 都认为自己是本 PIM-SM 域的 BSR，并使用接口的 IP 地址作为 BSR 地址，发送自举报文；
- 当某 C-BSR 收到其它 C-BSR 发来的自举报文时，首先比较自己与后者的优先级，优先级较高者获胜；在优先级相同的情况下，再比较自己与后者的 BSR 地址，拥有较大 IP 地址者获胜。如果后者获胜，则用后者的 BSR 地址替换自己的 BSR 地址，并不再认为自己是 BSR；否则，保留自己的 BSR 地址，并继续认为自己是 BSR。

通过在路由器上配置合法 BSR 的地址范围，可以对收到的自举报文按照地址范围进行过滤，从而防止某些恶意主机非法伪装成 BSR，以避免合法的 BSR 被恶意取代。必须在 PIM-SM 域内的所有路由器上进行相同的配置。通常针对以下两类情况实施预防措施：

- 某些恶意主机通过伪造自举报文以欺骗路由器，试图更改 RP 映射关系。这种攻击通常发生在边缘路由器上，由于 BSR 处于网络内部，主机在网络外部，因此边缘路由器通过对收到的自举报文进行邻居检查和 RPF 检查，丢弃不符合要求的报文，就可以避免外部网络用户对内部网络 BSR 的攻击；
- 网络中某台路由器被攻击者控制，或者有非法接入的路由器时，攻击者可以将这样的路由器配置为 C-BSR，并使其在竞争中获胜，从而控制网络中 RP 信息的发布权。由于在被配置为 C-BSR 后，路由器会自动向整个网络扩散自举报文，而自举报文是 TTL 值为 1 的组播报文，所以只要其邻居路由器不接收该自举报文，就不会影响整个网络。因此，通过在整个网络的所有路由器上都配置合法 BSR 的地址范围，从而丢弃合法范围之外的自举报文，就可以防止此类攻击。

以上两种预防措施可以部分地保护网络中 BSR 的安全。但是如果某台合法的 BSR 路由器被攻击者控制，还是可能导致问题。

表1-15 配置 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-

操作	命令	说明
配置某接口为C-BSR	**c-bsr** *interface-type interface-number* [*hash-length* [*priority*]]	必选 缺省情况下，没有配置C-BSR
配置合法的BSR地址范围	**bsr-policy** *acl-number*	可选 缺省情况下，BSR的地址范围不受任何限制

- 由于 BSR 与 PIM-SM 域中的其它设备需要交换大量信息，因此应在 C-BSR 与 PIM-SM 域中的其它设备之间保留较大的通信带宽。
- 当 C-BSR 与其它 PIM 路由器通过隧道连接时，如果单播路由中到 C-BSR 下一跳的不是 Tunnel 接口，请在 PIM 路由器上通过配置组播静态路由以保证这一点，否则将影响 RPF 检查。有关组播静态路由的相关配置，请参见“IP 组播配置指导”中的“组播路由与转发”。

2. 配置BSR服务边界

BSR 作为 PIM-SM 域中的管理核心，负责将收集到的 RP-Set 信息以自举报文的形式发向 PIM-SM 域中的所有路由器。

BSR 的服务边界，即 PIM-SM 域的边界。BSR 是针对特定的服务范围而言的，众多的 BSR 服务边界接口将网络划分成不同的 PIM-SM 域，自举报文无法通过 PIM-SM 域的边界，BSR 服务边界之外的路由器也不能参与本 PIM-SM 域内的组播转发。

请在欲配置为 BSR 服务边界的路由器上进行如下配置。

表1-16 配置 BSR 服务边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置BSR的服务边界	**pim bsr-boundary**	必选 缺省情况下，没有配置BSR的服务边界

3. 全局配置C-BSR参数

在一个 PIM-SM 域中，从众多 C-BSR 中选举出唯一的 BSR。PIM-SM 域内的 C-RP 向 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向本 PIM-SM 域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定组播组所对应 RP 的地址。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-17 全局配置 C-BSR 参数

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置哈希掩码长度	**c-bsr hash-length** *hash-length*	可选 缺省情况下，哈希掩码长度为30
配置C-BSR的优先级	**c-bsr priority** *priority*	可选 缺省情况下，C-BSR的优先级为64

说明

哈希掩码长度和 C-BSR 的优先级可以在全局、管理域和 Global 域这三种范围内进行配置：

- 如果在管理域或 Global 域下进行了配置，则取其配置值；
- 如果未在管理域或 Global 域下进行配置，则取相应的全局值。

有关管理域和Global域C-BSR参数的相关配置，请参见“1.3.6 3. 配置管理域和Global域的C-BSR”。

4. 配置C-BSR定时器

当某 C-BSR 竞选成为 BSR 后，它会通过自举报文向其所服务的区域以组播方式发送自己的 IP 地址和 RP-Set 信息。BSR 以自举时间间隔周期性地向网络发送自举报文，收到该报文的 C-BSR 会在自举超时时间内将其保持，此时 BSR 选举过程暂停；当自举超时时间超时后，C-BSR 之间会触发新一轮的 BSR 选举过程。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-18 配置 C-BSR 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置自举时间间隔	**c-bsr interval** *interval*	可选 缺省取值请参见表后的说明
配置自举超时时间	**c-bsr holdtime** *interval*	可选 缺省取值请参见表后的说明

说明

关于自举时间间隔的取值：

- 如果没有进行手工配置，则其取值由如下公式决定：自举时间间隔 =（自举超时时间 − 10）÷ 2。缺省情况下，自举超时时间为 130 秒，则自举时间间隔的缺省值 =（130 − 10）÷ 2 = 60（秒）。
- 如果进行了手工配置，则取配置值。

关于自举超时时间的取值：

- 如果没有进行手工配置，则其取值由如下公式决定：自举超时时间 = 自举时间间隔 × 2 + 10。缺省情况下，自举时间间隔为 60 秒，则自举超时时间的缺省值 = 60 × 2 + 10 = 130（秒）。
- 如果进行了手工配置，则取配置值。

注意

配置时，应保证自举时间间隔小于自举超时时间。

5. 关闭自举报文语义分片功能

BSR 周期性地向所在 PIM-SM 域发送自举报文以通告 RP-Set 信息。当 RP-Set 信息较少时，自举报文被封装在一个 IP 报文中发送出去；而当 RP-Set 信息较多时，自举报文的大小可能超过接口的 MTU（Maximum Transmission Unit，最大传输单元）值，从而触发其在 IP 层的分片。在这种情况下，一个 IP 分片的丢失就会导致整个自举报文都被丢弃。

自举报文语义分片功能可以解决上述问题：当自举报文大于接口 MTU 时，会被分解为多个自举报文分片（Bootstrap Message Fragment，BSMF）。非 BSR 收到自举报文分片后，若发现某组范围对应的 RP 信息都在这一个分片中，便立即更新该组范围对应的 RP-Set；若发现某组范围映射的 RP 信息被分在了多个分片中，则待收齐了这些分片后再更新该组范围对应的 RP-Set。这样，由于不同分片所含组范围对应的 RP 信息不同，因此个别分片的丢失只影响该分片所含组范围对应的 RP 信息，而不会导致整个自举报文都被丢弃。

自举报文语义分片功能是缺省使能的，但由于不支持该功能的设备会将自举报文分片当作完整的自举报文处理，从而导致其学到的 RP-Set 信息不完整，因此当 PIM-SM 域中存在此类设备时，请在已配置为 C-BSR 的路由器上关闭本功能。

表1-19 关闭自举报文语义分片功能

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
关闭自举报文语义分片功能	**undo bsm-fragment enable**	必选 缺省情况下，自举报文语义分片功能处于使能状态

说明

通常，BSR 根据其 BSR 接口的 MTU 值对自举报文进行语义分片；而对由于新学到 PIM 邻居而触发的自举报文发送，则根据发送接口的 MTU 值进行语义分片。

1.3.6 配置管理域

在非管理域机制下，一个 PIM-SM 域中只有唯一的 BSR，整个网络都在该 BSR 的管理范围之内。为了更有针对性地管理，可以将整个 PIM-SM 域划分为多个管理域：每个管理域中各维护一个 BSR，服务于特定范围的组播组；而 Global 域中也维护一个 BSR，为所有剩余的组播组提供服务。

1. 使能管理域机制

在配置管理域各项功能之前，必须先使能管理域机制。

请在 PIM-SM 域内的所有路由器上进行如下配置。

表1-20 使能管理域机制

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
使能管理域机制	**c-bsr admin-scope**	必选 缺省情况下，管理域机制处于关闭状态

2. 配置管理域边界

在管理域机制中，各管理域的边界由 ZBR 构成，每个管理域各维护一个 BSR，服务于特定范围的组播组，属于此范围的组播协议报文（如断言报文、BSR 自举报文等）无法通过管理域边界。

请在欲配置为 ZBR 的路由器上进行如下配置。

表1-21 配置管理域边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置组播转发边界	**multicast boundary** *group-address* { *mask* | *mask-length* }	必选 缺省情况下，没有配置组播转发边界

说明

通过 **multicast boundary** 命令中的参数 *group-address* { *mask* | *mask-length* }可指定管理域所服务的组播组范围，有效的组范围在 239.0.0.0/8 之内。有关 **multicast boundary** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

3. 配置管理域和Global域的C-BSR

在应用了管理域机制的网络中，从众多 C-BSR 中选举出针对不同组播组的 BSR。网络内的 C-RP 只向对应的 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向其所服务管理域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定组播组所对应 RP 的地址。

请为管理域和 Global 域分别配置其各自的 C-BSR。

(1) 配置管理域的 C-BSR

请在欲配置为管理域 C-BSR 的路由器上进行如下配置。

表1-22 配置管理域的 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置管理域的C-BSR	**c-bsr group** *group-address* { *mask* | *mask-length* } [**hash-length** *hash-length* | **priority** *priority*] *	必选 缺省情况下，没有配置管理域的C-BSR

说明

通过 **c-bsr group** 命令中的参数 *group-address* { *mask* | *mask-length* }可指定 C-BSR 所服务的组播组范围，有效的组范围在 239.0.0.0/8 之内。

(2) 配置 Global 域的 C-BSR

请在欲配置为 Global 域 C-BSR 的路由器上进行如下配置。

表1-23 配置 Global 域的 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置Global域的C-BSR	**c-bsr global** [**hash-length** *hash-length* | **priority** *priority*] *	必选 缺省情况下，没有配置Global域的C-BSR

说明

哈希掩码长度和 C-BSR 的优先级可以在全局、管理域和 Global 域这三种范围内进行配置：

- 如果在管理域或 Global 域下进行了配置，则取其配置值；
- 如果未在管理域或 Global 域下进行配置，则取相应的全局值。

有关全局C-BSR参数的相关配置，请参见“1.3.5 3. 全局配置C-BSR参数”。

1.3.7 配置组播源注册

在 PIM-SM 域内，组播源侧 DR 向 RP 发送注册报文，而这些注册报文拥有不同的组播源或组播组地址。为了让 RP 服务于特定的组播组，可以对注册报文进行过滤。如果某个（S，G）表项被过滤规则拒绝，或者过滤规则中没有定义对它的操作，RP 都会向 DR 发送注册停止报文，以停止该组播数据的注册过程。

出于对注册报文在传递过程中完整性的考虑，可以配置根据整个报文来计算校验和。但为了减少往注册报文中封装数据报文的工作量并考虑到互通性，一般情况下不建议配置根据注册报文的全部内容来计算校验和的方式。

当接收者不再通过 RP 接收发往某组播组的数据（即 RP 不再服务于该组播组），或 RP 开始接收组播源沿着 SPT 发来的组播数据时，RP 将向组播源侧 DR 发送注册停止报文，DR 收到该报文后将停止发送封装有组播数据的注册报文并启动注册停止定时器（Register-Stop Timer）。在注册停止定时器超时之前，DR 会向 RP 发送一个空注册报文（Null-Register Message，即不封装组播数据的注册报文）：如果 DR 在注册探测时间（Register_Probe_Time）内收到了来自 RP 的注册停止报文，DR 将刷新其注册停止定时器；否则，DR 将重新开始发送封装有组播数据的注册报文。

注册停止定时器的超时时间是一个随机值，由其它两个时间值决定：注册抑制时间（Register_Suppression_Time）和注册探测时间。其具体取值范围如下：（0.5×注册抑制时间，1.5×注册抑制时间）－注册探测时间。

请在所有已配置为 C-RP 的路由器上配置注册报文的过滤规则和根据注册报文的全部内容来计算校验和；请在所有可能成为组播源侧 DR 的路由器上配置注册抑制时间和注册探测时间。

表1-24 配置组播源注册

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置注册报文的过滤规则	**register-policy** *acl-number*	可选 缺省情况下，没有配置注册报文的过滤规则
配置根据注册报文的全部内容来计算校验和	**register-whole-checksum**	可选 缺省情况下，仅根据注册报文头来计算校验和
配置注册抑制时间	**register-suppression-timeout** *interval*	可选 缺省情况下，注册抑制时间为60秒

操作	命令	说明
配置注册探测时间	**probe-interval** *interval*	可选 缺省情况下，注册探测时间为5秒

1.3.8 配置禁止SPT切换

缺省情况下，当本系列交换机作为 RP 或接收者侧 DR 时，设备会在收到第一个组播数据包后立即发起 SPT 切换，可以通过下面的配置关闭从 RPT 到 SPT 的切换。

表1-25 配置禁止 SPT 切换

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置禁止发起SPT切换的条件	**spt-switch-threshold infinity** [**group-policy** *acl-number* [**order** *order-value*]]	可选 缺省情况下，设备收到第一个组播数据包后便立即向SPT切换

说明

由于某些设备无法将组播报文封装在注册报文中发给 RP，因此在可能成为 RP 的设备上不建议配置永不发起 SPT 切换，以免导致组播报文转发失败。

1.4 配置双向PIM

1.4.1 双向PIM配置任务简介

表1-26 双向 PIM 配置任务简介

配置任务		说明	详细配置
使能PIM-SM		必选	1.4.3
使能双向PIM		必选	1.4.4
配置RP	配置静态RP	三者必选其一	1.4.5 1.
	配置C-RP		1.4.5 2.
	使能自动RP侦听		1.4.5 3.
	全局配置C-RP定时器	可选	1.4.5 4.
配置BSR	配置C-BSR	必选	1.4.6 1.
	配置BSR服务边界	可选	1.4.6 2.

配置任务		说明	详细配置
	全局配置C-BSR参数	可选	1.4.6 3.
	配置C-BSR定时器	可选	1.4.6 4.
	关闭自举报文语义分片功能	可选	1.4.6 5.
配置管理域	使能管理域机制	可选	1.4.7 1.
	配置管理域边界	可选	1.4.7 2.
	配置管理域和Global域的C-BSR	可选	1.4.7 3.
配置PIM公共特性		可选	1.6

1.4.2 配置准备

在配置双向 PIM 之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通

在配置双向 PIM 之前，需准备以下数据：

- 静态 RP 的 IP 地址及表示其所服务组范围的 ACL 规则
- C-RP 的优先级及表示其所服务组范围的 ACL 规则
- 合法 C-RP 的地址范围及表示其所服务组范围的 ACL 规则
- 发送宣告报文的时间间隔
- C-RP 超时时间
- C-BSR 的优先级
- 哈希掩码长度
- 表示合法 BSR 地址范围的 ACL 规则
- 自举时间间隔
- 自举超时时间

1.4.3 使能PIM-SM

由于双向 PIM 是在 PIM-SM 的基础上实现的，因此在使能双向 PIM 之前须先使能 PIM-SM。在部署双向 PIM 域时，建议在其所有非边界接口上均使能 PIM-SM。

1. 使能公网实例中的PIM-SM

表1-27 使能公网实例中的 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
使能PIM-SM	**pim sm**	必选 缺省情况下，PIM-SM处于关闭状态

2. 使能VPN实例中的PIM-SM

表1-28 使能 VPN 实例中的 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
将接口与VPN实例进行关联	**ip binding vpn-instance** *vpn-instance-name*	必选 缺省情况下，接口只属于公网，未与任何VPN实例关联
使能PIM-SM	**pim sm**	必选 缺省情况下，PIM-SM处于关闭状态

注意

在同一台设备的同一个 VPN 实例中，所有接口上所启用的 PIM 模式必须相同。

说明

- 有关 **ip vpn-instance**、**route-distinguisher** 和 **ip binding vpn-instance** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。
- 有关 **multicast routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

1.4.4 使能双向PIM

请在双向 PIM 域内的所有路由器上进行如下配置。

表1-29 使能双向 PIM

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
使能双向PIM	**bidir-pim enable**	必选 缺省情况下，双向PIM处于关闭状态

1.4.5 配置RP

RP 可以通过手工方式静态配置，也可以通过 BSR 机制动态选举。对于大型 PIM 网络，配置静态 RP 将会非常繁琐。所以，通常静态 RP 是作为动态选举 RP 机制的备份手段，以提高网络的健壮性，增强组播网络的运营管理能力。

注意

当 PIM 网络中同时运行 PIM-SM 和双向 PIM 时，请勿使一个 RP 同时为 PIM-SM 和双向 PIM 工作，否则可能引起 PIM 路由表出错。

1. 配置静态RP

当网络内仅有一个动态 RP 时，通过手工配置静态 RP 可以避免因单一节点故障而引起的通信中断，同时也可以避免 C-RP 与 BSR 之间频繁的信息交互而占用带宽。

请在双向 PIM 域内的所有路由器上进行如下配置。

表1-30 配置静态 RP

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置服务于双向PIM的静态RP	**static-rp** *rp-address* [*acl-number*] [**preferred**] **bidir**	必选 缺省情况下，没有配置静态RP

注意

为了让静态 RP 功能正常发挥作用，必须在所有路由器上指定相同的静态 RP 地址。

说明

双向 PIM 允许将静态 RP 的 IP 地址指定为一个实际不存在的 IP 地址。譬如，一条链路两端接口的 IP 地址分别为 10.1.1.1/24 和 10.1.1.2/24，可以将静态 RP 的 IP 地址指定为同网段但实际不存在的一个地址，如 10.1.1.100/24，该链路就成为了 RPL。

2. 配置C-RP

在双向 PIM 域中，可以把有意成为 RP 的路由器配置为 C-RP。BSR 通过接收来自 C-RP 的 C-RP 信息，或者接收来自其它路由器的自动 RP 宣告，收集 C-RP 信息并将其汇总为 RP-Set 信息，然后在全网内扩散。之后，网络内的其它路由器根据 RP-Set 信息计算出特定组播组范围所对应的 RP。建议在骨干网路由器上配置 C-RP。

为了防止 C-RP 欺骗，需要在 BSR 上配置合法的 C-RP 地址范围及其所服务的组播组范围。同时由于每个 C-BSR 都可能成为 BSR，因此必须在双向 PIM 域内的所有 C-BSR 上都配置相同的过滤策略。

表1-31 配置 C-RP

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置某接口为服务于双向PIM的C-RP	**c-rp** *interface-type interface-number* [**group-policy** *acl-number* | **priority** *priority* | **holdtime** *hold-interval* | **advertisement-interval** *adv-interval*] * **bidir**	必选 缺省情况下，没有配置C-RP

说明

- 在配置 C-RP 时，应在 C-RP 与双向 PIM 域中的其它设备之间保留较大的通信带宽。
- 一个 RP 可以为多个组播组服务，也可以为所有组播组服务。每个组播组在任意时刻，只能由唯一的一个 RP 为其转发数据，而不能由多个 RP 转发数据。

3. 使能自动RP侦听

自动 RP 宣告和发现报文的目的地址分别为组播组地址 224.0.1.39 和 224.0.1.40。在设备上使能了自动 RP 侦听功能后，该设备便能够接收这两种报文，并记录报文中所携带的 RP 信息。

表1-32 使能自动 RP 侦听

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-

操作	命令	说明
使能自动RP侦听功能	**auto-rp enable**	必选 缺省情况下，自动RP侦听功能处于关闭状态

4. 全局配置C-RP定时器

为了使 BSR 能够在双向 PIM 域内分发 RP-Set 信息，C-RP 必须周期性地向 BSR 发送宣告报文，BSR 从该报文中学习 RP-Set 信息，并将该信息与自己的 IP 地址一起封装在自举报文中向域中的所有 PIM 路由器进行宣告。

C-RP 在其宣告报文中封装一个保持时间，BSR 在收到该报文后，从中获得该时间值并启动 C-RP 超时定时器，如果超时后 BSR 仍没有收到来自 C-RP 后续的宣告报文，则认为目前网络中的 C-RP 失效或不可达。

请在已配置为 C-RP 的路由器上进行如下配置。

表1-33 全局配置 C-RP 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置发送宣告报文的时间间隔	**c-rp advertisement-interval** *interval*	可选 缺省情况下，发送宣告报文的时间间隔为60秒
配置C-RP超时时间	**c-rp holdtime** *interval*	可选 缺省情况下，C-RP的超时时间为150秒

说明

有关双向PIM其它定时器的相关配置，请参见“1.6.7 配置PIM公共定时器”。

1.4.6 配置BSR

在一个双向 PIM 域中只能有一个 BSR，但需要配置至少一个 C-BSR。任意一台路由器都可以被配置为 C-BSR。在 C-BSR 之间通过自动选举产生 BSR，BSR 负责在双向 PIM 域中收集并发布 RP 信息。

1. 配置C-BSR

C-BSR 应配置在骨干网的路由器上，在将路由器配置为 C-BSR 时，必须同时指定一个使能了 PIM-SM 的接口。C-BSR 间的自动选举机制简单描述如下：

- 最初，每个 C-BSR 都认为自己是本双向 PIM 域的 BSR，并使用接口的 IP 地址作为 BSR 地址，发送自举报文；

- 当某 C-BSR 收到其它 C-BSR 发来的自举报文时，首先比较自己与后者的优先级，优先级较高者获胜；在优先级相同的情况下，再比较自己与后者的 BSR 地址，拥有较大 IP 地址者获胜。如果后者获胜，则用后者的 BSR 地址替换自己的 BSR 地址，并不再认为自己是 BSR；否则，保留自己的 BSR 地址，并继续认为自己是 BSR。

通过在路由器上配置合法 BSR 的地址范围，可以对收到的自举报文按照地址范围进行过滤，从而防止某些恶意主机非法伪装成 BSR，以避免合法的 BSR 被恶意取代。必须在双向 PIM 域内的所有路由器上进行相同的配置。通常针对以下两类情况实施预防措施：

- 某些恶意主机通过伪造自举报文以欺骗路由器，试图更改 RP 映射关系。这种攻击通常发生在边缘路由器上，由于 BSR 处于网络内部，主机在网络外部，因此边缘路由器通过对收到的自举报文进行邻居检查和 RPF 检查，丢弃不符合要求的报文，就可以避免外部网络用户对内部网络 BSR 的攻击；
- 网络中某台路由器被攻击者控制，或者有非法接入的路由器时，攻击者可以将这样的路由器配置为 C-BSR，并使其在竞争中获胜，从而控制网络中 RP 信息的发布权。由于在被配置为 C-BSR 后，路由器会自动向整个网络扩散自举报文，而自举报文是 TTL 值为 1 的组播报文，所以只要其邻居路由器不接收该自举报文，就不会影响整个网络。因此，通过在整个网络的所有路由器上都配置合法 BSR 的地址范围，从而丢弃合法范围之外的自举报文，就可以防止此类攻击。

以上两种预防措施可以部分地保护网络中 BSR 的安全。但是如果某台合法的 BSR 路由器被攻击者控制，还是可能导致问题。

表1-34 配置 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置某接口为C-BSR	**c-bsr** *interface-type interface-number* [*hash-length* [*priority*]]	必选 缺省情况下，没有配置C-BSR
配置合法的BSR地址范围	**bsr-policy** *acl-number*	可选 缺省情况下，BSR的地址范围不受任何限制

- 由于 BSR 与双向 PIM 域中的其它设备需要交换大量信息，因此应在 C-BSR 与双向 PIM 域中的其它设备之间保留较大的通信带宽。
- 当 C-BSR 与其它 PIM 路由器通过隧道连接时，如果单播路由中到 C-BSR 下一跳的不是 Tunnel 接口，请在 PIM 路由器上通过配置组播静态路由以保证这一点，否则将影响 RPF 检查。有关组播静态路由的相关配置，请参见“IP 组播配置指导”中的“组播路由与转发”。

2. 配置BSR服务边界

BSR 作为双向 PIM 域中的管理核心，负责将收集到的 RP-Set 信息以自举报文的形式发向双向 PIM 域中的所有路由器。

BSR 的服务边界，即双向 PIM 域的边界。BSR 是针对特定的服务范围而言的，众多的 BSR 服务边界接口将网络划分成不同的双向 PIM 域，自举报文无法通过双向 PIM 域的边界，BSR 服务边界之外的路由器也不能参与本双向 PIM 域内的组播转发。

请在欲配置为 BSR 服务边界的路由器上进行如下配置。

表1-35 配置 BSR 服务边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置BSR的服务边界	**pim bsr-boundary**	必选 缺省情况下，没有配置BSR的服务边界

3. 全局配置C-BSR参数

在一个双向 PIM 域中，从众多 C-BSR 中选举出唯一的 BSR。双向 PIM 域内的 C-RP 向 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向本双向 PIM 域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定组播组所对应 RP 的地址。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-36 全局配置 C-BSR 参数

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置哈希掩码长度	**c-bsr hash-length** *hash-length*	可选 缺省情况下，哈希掩码长度为30
配置C-BSR的优先级	**c-bsr priority** *priority*	可选 缺省情况下，C-BSR的优先级为64

说明

哈希掩码长度和 C-BSR 的优先级可以在全局、管理域和 Global 域这三种范围内进行配置：

- 如果在管理域或 Global 域下进行了配置，则取其配置值；
- 如果未在管理域或 Global 域下进行配置，则取相应的全局值。

有关管理域和Global域C-BSR参数的相关配置，请参见“1.4.7 3. 配置管理域和Global域的C-BSR”。

4. 配置C-BSR定时器

当某 C-BSR 竞选成为 BSR 后，它会通过自举报文向其所服务的区域以组播方式发送自己的 IP 地址和 RP-Set 信息。BSR 以自举时间间隔周期性地向网络发送自举报文，收到该报文的 C-BSR 会在自举超时时间内将其保持，此时 BSR 选举过程暂停；当自举超时时间超时后，C-BSR 之间会触发新一轮的 BSR 选举过程。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-37 配置 C-BSR 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置自举时间间隔	**c-bsr interval** *interval*	可选 缺省取值请参见表后的说明
配置自举超时时间	**c-bsr holdtime** *interval*	可选 缺省取值请参见表后的说明

说明

关于自举时间间隔的取值：

- 如果没有进行手工配置，则其取值由如下公式决定：自举时间间隔 =（自举超时时间 − 10）÷ 2。缺省情况下，自举超时时间为 130 秒，则自举时间间隔的缺省值 =（130 − 10）÷ 2 = 60（秒）。
- 如果进行了手工配置，则取配置值。

关于自举超时时间的取值：

- 如果没有进行手工配置，则其取值由如下公式决定：自举超时时间 = 自举时间间隔 × 2 + 10。缺省情况下，自举时间间隔为 60 秒，则自举超时时间的缺省值 = 60 × 2 + 10 = 130（秒）。
- 如果进行了手工配置，则取配置值。

注意

配置时，应保证自举时间间隔小于自举超时时间。

5. 关闭自举报文语义分片功能

BSR 周期性地向所在双向 PIM 域发送自举报文以通告 RP-Set 信息。当 RP-Set 信息较少时，自举报文被封装在一个 IP 报文中发送出去；而当 RP-Set 信息较多时，自举报文的大小可能超过接口的 MTU 值，从而触发其在 IP 层的分片。在这种情况下，一个 IP 分片的丢失就会导致整个自举报文都被丢弃。

自举报文语义分片功能可以解决上述问题：当自举报文大于接口 MTU 时，会被分解为多个自举报文分片。非 BSR 收到自举报文分片后，若发现某组范围对应的 RP 信息都在这一个分片中，便立即更新该组范围对应的 RP-Set；若发现某组范围映射的 RP 信息被分在了多个分片中，则待收齐了这些分片后再更新该组范围对应的 RP-Set。这样，由于不同分片所含组范围对应的 RP 信息不同，因此个别分片的丢失只影响该分片所含组范围对应的 RP 信息，而不会导致整个自举报文都被丢弃。

自举报文语义分片功能是缺省使能的，但由于不支持该功能的设备会将自举报文分片当作完整的自举报文处理，从而导致其学到的 RP-Set 信息不完整，因此当 PIM-SM 域中存在此类设备时，请在已配置为 C-BSR 的路由器上关闭本功能。

表1-38 关闭自举报文语义分片功能

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
关闭自举报文语义分片功能	**undo bsm-fragment enable**	必选 缺省情况下，自举报文语义分片功能处于使能状态

说明

通常，BSR 根据其 BSR 接口的 MTU 值对自举报文进行语义分片；而对由于新学到 PIM 邻居而触发的自举报文发送，则根据发送接口的 MTU 值进行语义分片。

1.4.7 配置管理域

在非管理域机制下，一个双向 PIM 域中只有唯一的 BSR，整个网络都在该 BSR 的管理范围之内。为了更有针对性地管理，可以将整个双向 PIM 域划分为多个管理域：每个管理域中各维护一个 BSR，服务于特定范围的组播组；而 Global 域中也维护一个 BSR，为所有剩余的组播组提供服务。

1. 使能管理域机制

在配置管理域各项功能之前，必须先使能管理域机制。

请在双向 PIM 域内的所有路由器上进行如下配置。

表1-39 使能管理域机制

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
使能管理域机制	**c-bsr admin-scope**	必选 缺省情况下，管理域机制处于关闭状态

2. 配置管理域边界

在管理域机制中，各管理域的边界由 ZBR 构成，每个管理域各维护一个 BSR，服务于特定范围的组播组，属于此范围的组播协议报文（如断言报文、BSR 自举报文等）无法通过管理域边界。

请在欲配置为 ZBR 的路由器上进行如下配置。

表1-40 配置管理域边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置组播转发边界	**multicast boundary** *group-address* { *mask* | *mask-length* }	必选 缺省情况下，没有配置组播转发边界

说明

通过 **multicast boundary** 命令中的参数 *group-address* { *mask* | *mask-length* }可指定管理域所服务的组播组范围，有效的组范围在 239.0.0.0/8 之内。有关 **multicast boundary** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

3. 配置管理域和Global域的C-BSR

在应用了管理域机制的网络中，从众多 C-BSR 中选举出针对不同组播组的 BSR。网络内的 C-RP 只向对应的 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向其所服务管理域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定组播组所对应 RP 的地址。

请为管理域和 Global 域分别配置其各自的 C-BSR。

(1) 配置管理域的 C-BSR

请在欲配置为管理域 C-BSR 的路由器上进行如下配置。

表1-41 配置管理域的 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置管理域的C-BSR	**c-bsr group** *group-address* { *mask* | *mask-length* } [**hash-length** *hash-length* | **priority** *priority*] *	必选 缺省情况下，没有配置管理域的C-BSR

说明

通过 **c-bsr group** 命令中的参数 *group-address* { *mask* | *mask-length* }可指定 C-BSR 所服务的组播组范围，有效的组范围在 239.0.0.0/8 之内。

(2) 配置 Global 域的 C-BSR

请在欲配置为 Global 域 C-BSR 的路由器上进行如下配置。

表1-42 配置 Global 域的 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置Global域的C-BSR	**c-bsr global** [**hash-length** *hash-length* | **priority** *priority*] *	必选 缺省情况下，没有配置Global域的C-BSR

说明

哈希掩码长度和 C-BSR 的优先级可以在全局、管理域和 Global 域这三种范围内进行配置：

- 如果在管理域或 Global 域下进行了配置，则取其配置值；
- 如果未在管理域或 Global 域下进行配置，则取相应的全局值。

有关全局C-BSR参数的相关配置，请参见“1.4.6 3. 全局配置C-BSR参数”。

1.5 配置PIM-SSM

说明

PIM-SSM 模型需要 IGMPv3 的支持，因此应确保连接有接收者的 PIM 路由器上使能了 IGMPv3。

1.5.1 PIM-SSM配置任务简介

表1-43 PIM-SSM 配置任务简介

配置任务	说明	详细配置
使能PIM-SM	必选	1.5.3
配置SSM组播组范围	可选	1.5.4
配置PIM公共特性	可选	1.6

1.5.2 配置准备

在配置 PIM-SSM 之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通

在配置 PIM-SSM 之前，需准备以下数据：

- SSM 组播组范围

1.5.3 使能PIM-SM

由于 PIM-SSM 是通过 PIM-SM 的部分子集功能实现的，因此在配置 PIM-SSM 之前须先使能 PIM-SM。在部署 PIM-SSM 域时，建议在其所有非边界接口上均使能 PIM-SM。

1. 使能公网实例中的PIM-SM

表1-44 使能公网实例中的 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能PIM-SM	**pim sm**	必选 缺省情况下，PIM-SM处于关闭状态

2. 使能VPN实例中的PIM-SM

表1-45 使能 VPN 实例中的 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
将接口与VPN实例进行关联	**ip binding vpn-instance** *vpn-instance-name*	必选 缺省情况下，接口只属于公网，未与任何VPN实例关联
使能PIM-SM	**pim sm**	必选 缺省情况下，PIM-SM处于关闭状态

注意

在同一台设备的同一个 VPN 实例中，所有接口上所启用的 PIM 模式必须相同。

说明

- 有关 **ip vpn-instance**、**route-distinguisher** 和 **ip binding vpn-instance** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。
- 有关 **multicast routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

1.5.4 配置SSM组播组范围

在把来自组播源的信息传递给接收者的过程中，是采用 PIM-SSM 模型还是 PIM-SM 模型，这取决于接收者订阅通道（S，G）中的组播组是否在 SSM 组播组范围之内，所有使能了 PIM-SM 的接口将会认为属于该范围内的组播组采用了 PIM-SSM 模型。

请在 PIM-SSM 域内的所有路由器上进行如下配置。

表1-46 配置 SSM 组播组范围

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置SSM组播组的范围	**ssm-policy** *acl-number*	可选 缺省情况下，SSM组播组的范围为232.0.0.0/8

注意

- 应确保域内所有路由器上配置的 SSM 组播组地址范围都一致，否则组播信息将无法通过 SSM 模型进行传输。
- 如果某组播组属于 SSM 组播组范围，但该组成员使用 IGMPv1 或 IGMPv2 发送加入报文，则设备不会触发（*，G）加入报文。

1.6 配置PIM公共特性

说明

对于本节中那些既可在 PIM 视图、又可在接口视图下进行配置的功能或参数来说:

- 在 PIM 视图下所作的配置对所有接口生效，在接口视图下所作的配置仅对当前接口生效;
- 如果在这两个视图下对某功能或参数均进行了配置，则不论配置先后，接口视图下的配置将被优先采用。

1.6.1 PIM公共特性配置任务简介

表1-47 PIM 公共特性配置任务简介

配置任务	说明	详细配置
配置组播数据过滤器	可选	1.6.3
配置Hello报文过滤器	可选	1.6.4
配置Hello报文选项	可选	1.6.5
配置剪枝延迟时间	可选	1.6.6
配置PIM公共定时器	可选	1.6.7
配置加入/剪枝报文规格	可选	1.6.8
配置PIM与BFD联动	可选	1.6.9
配置PIM报文的DSCP优先级	可选	1.6.10

1.6.2 配置准备

在配置 PIM 公共特性之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 配置 PIM-DM（或 PIM-SM、PIM-SSM）

在配置 PIM 公共特性之前，需准备以下数据：

- 表示组播数据过滤规则的 ACL 规则
- 表示合法 Hello 报文的源地址范围的 ACL 规则
- 竞选 DR 的优先级（全局值/接口值）
- 保持 PIM 邻居可达状态的时间（全局值/接口值）
- 发送剪枝报文的延迟时间（全局值/接口值）
- 剪枝否决时间（全局值/接口值）
- 剪枝延迟时间
- 发送 Hello 报文的时间间隔（全局值/接口值）
- 发送 Hello 报文的最大延迟时间（接口值）
- 保持断言状态的时间（全局值/接口值）
- 发送加入/剪枝报文的时间间隔（全局值/接口值）
- 保持加入/剪枝状态的时间（全局值/接口值）
- 组播源生存时间
- 加入/剪枝报文的最大长度
- 加入/剪枝报文中（S，G）表项的最大数量
- PIM 报文的 DSCP 优先级

1.6.3 配置组播数据过滤器

无论在 PIM-DM 还是 PIM-SM 域内，各路由器都可以对流经自己的组播数据进行检查，通过比较是否符合过滤规则，从而决定是否继续转发组播数据。也就是说 PIM 域内的路由器能够成为组播数据的过滤器。过滤器的存在一方面有助于实现信息流量控制，另一方面可以在安全性方面限定下游接收者能够获得的信息。

表1-48 配置组播数据过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置组播数据过滤器	**source-policy** *acl-number*	必选 缺省情况下，没有配置组播数据过滤器

说明

- 通常，过滤器的位置离组播源距离越近，过滤影响越明显。
- 过滤器不仅过滤独立的组播数据，还过滤封装在注册报文中的组播数据。

1.6.4 配置Hello报文过滤器

随着 PIM 协议的推广和应用，对其安全性的要求也越来越高。建立正确的 PIM 邻居是 PIM 协议安全应用的前提。如果在接口上指定了合法 Hello 报文的源地址范围，便能够保证 PIM 邻居的正确建立，从而有效防止各种 PIM 协议报文攻击，提高设备对 PIM 协议报文处理的安全性。

表1-49 配置 Hello 报文过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置合法Hello报文的源地址范围	**pim neighbor-policy** *acl-number*	必选 缺省情况下，Hello报文的源地址范围不受任何限制

说明

当 Hello 报文过滤器的配置生效后，对于之前已建立的 PIM 邻居，若由于其 Hello 报文被过滤而导致无法收到后续的 Hello 报文，将会在老化超时后被自动删除。

1.6.5 配置Hello报文选项

无论在 PIM-DM 域还是在 PIM-SM 域内，各路由器之间发送的 Hello 报文都包含很多可供配置的选项，对各选项的介绍如下：

- DR_Priority（仅用于 PIM-SM）：表示竞选 DR 的优先级，优先级高的设备被选举为 DR。可以在与组播源或接收者直连的共享网络中的所有路由器上都配置此参数。
- Holdtime：表示保持 PIM 邻居可达状态的时间，若超时后仍没有收到 Hello 报文，则认为 PIM 邻居失效或不可达。
- LAN_Prune_Delay：表示在共享网络上传递剪枝报文的延迟时间，该选项由三部分组成：LAN-delay（发送剪枝报文的延迟时间）、Override-interval（剪枝否决时间）和禁止加入报文抑制能力。当共享网段内各 PIM 路由器的 LAN-delay 或 Override-interval 不同时，取其中最大的值；当要禁止加入报文抑制能力时，须在共享网段内的所有 PIM 路由器上都禁止该能力。

LAN-delay 表示路由器从收到下游路由器发来的剪枝报文到继续向上游路由器发送剪枝报文的延迟时间，Override-interval 则表示允许下游路由器否决剪枝动作的时间。路由器在收到下游路由器发来的剪枝报文后并不立即执行剪枝动作，而是仍将当前的转发状态保持 LAN-delay＋Override-interval 时间。如果下游路由器需要继续接收组播数据，则必须在 Override-interval 时间内向上游路由器发送加入报文以否决这个剪枝动作，这就称为剪枝否决；如果 Override-interval 时间超时后未收到任何加入报文，上游路由器就会在 LAN-delay＋Override-interval 时间超时后执行剪枝动作。

PIM 路由器发送 Hello 报文时，会生成一个随机数 Generation ID 并携带在该报文中。一台 PIM 路由器的 Generation ID 一般不会变化，除非该路由器的状态发生了改变（如接口刚使能了 PIM 或设备进行了重启），此时，当其开始或重新开始发送 Hello 报文时，会生成一个新的 Generation ID。这样，如果 PIM 路由器发现其上游邻居发来的 Hello 报文中 Generation ID 发生了改变，就会认为该邻居的状态已经丢失或其上游邻居已经改变，从而触发发送加入报文以进行状态刷新。

在禁止加入报文抑制能力(即使能邻居跟踪)时，应在共享网段的所有 PIM 路由器上都禁止该能力，否则上游路由器无法跟踪每台下游路由器的加入报文。

1. 全局配置Hello报文选项

表1-50 全局配置 Hello 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置竞选DR的优先级	**hello-option dr-priority** *priority*	可选 缺省情况下，竞选DR的优先级为1
配置保持PIM邻居可达状态的时间	**hello-option holdtime** *interval*	可选 缺省情况下，保持PIM邻居可达状态的时间为105秒
配置发送剪枝报文的延迟时间	**hello-option lan-delay** *interval*	可选 缺省情况下，发送剪枝报文的延迟时间为500毫秒

操作	命令	说明
配置剪枝否决时间	**hello-option override-interval** *interval*	可选 缺省情况下，剪枝否决时间为2500毫秒
禁止加入报文抑制能力	**hello-option neighbor-tracking**	必选 缺省情况下，加入报文抑制能力处于使能状态

2. 在接口上配置Hello报文选项

表1-51 在接口上配置 Hello 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置竞选DR的优先级	**pim hello-option dr-priority** *priority*	可选 缺省情况下，竞选DR的优先级为1
配置保持PIM邻居可达状态的时间	**pim hello-option holdtime** *interval*	可选 缺省情况下，保持PIM邻居可达状态的时间为105秒
配置发送剪枝报文的延迟时间	**pim hello-option lan-delay** *interval*	可选 缺省情况下，发送剪枝报文的延迟时间为500毫秒
配置剪枝否决时间	**pim hello-option override-interval** *interval*	可选 缺省情况下，剪枝否决时间为2500毫秒
禁止加入报文抑制能力	**pim hello-option neighbor-tracking**	必选 缺省情况下，加入报文抑制能力处于使能状态
配置不接受无Generation ID的Hello报文	**pim require-genid**	必选 缺省情况下，接受无Generation ID的Hello报文

1.6.6 配置剪枝延迟时间

剪枝延迟时间（Prune Delay）的作用是使共享网段中的上游路由器在收到下游路由器发来的剪枝报文后不立即执行剪枝动作，而将当前的转发状态继续保持 Prune Delay 时间：如果在该时间内收到了下游路由器发来的加入报文，则否决这个剪枝动作；否则，便执行这个剪枝动作。

表1-52 配置剪枝延迟时间

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-

操作	命令	说明
配置剪枝延迟时间	**prune delay** *interval*	可选 缺省情况下，未配置剪枝延迟时间

1.6.7 配置PIM公共定时器

PIM 路由器通过周期性地发送 Hello 报文，以发现 PIM 邻居，并维护各路由器之间的 PIM 邻居关系。

为了避免多个 PIM 路由器同时发送 Hello 报文而导致冲突，当 PIM 路由器在收到 Hello 报文时，将延迟一段时间再发送 Hello 报文，该时间值为小于“触发 Hello 报文的最大延迟时间”的一个随机值。

PIM 路由器通过周期性地向其上游路由器发送加入/剪枝报文以更新状态，在该报文中携带有保持时间，上游路由器为被剪枝的下游接口设置保持加入/剪枝状态定时器。

在断言中落选的路由器将会剪掉其下游的转发接口，并把这种断言状态保持一段时间。超时后，落选的路由器会重新恢复转发组播数据。

当路由器没有收到来自组播源 S 的后续组播数据时，不会立即删除（S，G）表项，而是将其维持一段时间后再删除，这段时间就称为组播源的生存时间。

1. 全局配置PIM公共定时器

表1-53 全局配置 PIM 公共定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置发送Hello报文的时间间隔	**timer hello** *interval*	可选 缺省情况下，发送Hello报文的时间间隔为30秒
配置发送加入/剪枝报文的时间间隔	**timer join-prune** *interval*	可选 缺省情况下，发送加入/剪枝报文的时间间隔为60秒
配置保持加入/剪枝状态的时间	**holdtime join-prune** *interval*	可选 缺省情况下，保持加入/剪枝状态的时间为210秒
配置保持断言状态的时间	**holdtime assert** *interval*	可选 缺省情况下，保持断言状态的时间为180秒
配置组播源生存时间	**source-lifetime** *interval*	可选 缺省情况下，组播源的生存时间为210秒

2. 在接口上配置PIM公共定时器

表1-54 在接口上配置 PIM 公共定时器

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入接口视图	**interface** *interface-type interface-number*	-
配置发送Hello报文的时间间隔	**pim timer hello** *interval*	可选 缺省情况下，发送Hello报文的时间间隔为30秒
配置触发Hello报文的最大延迟时间	**pim triggered-hello-delay** *interval*	可选 缺省情况下，触发Hello报文的最大延迟时间为5秒
配置发送加入/剪枝报文的时间间隔	**pim timer join-prune** *interval*	可选 缺省情况下，发送加入/剪枝报文的时间间隔为60秒
配置保持加入/剪枝状态的时间	**pim holdtime join-prune** *interval*	可选 缺省情况下，保持加入/剪枝状态的时间为210秒
配置保持断言状态的时间	**pim holdtime assert** *interval*	可选 缺省情况下，保持断言状态的时间为180秒

说明

如果网络没有特殊要求，建议采用缺省值。

1.6.8 配置加入/剪枝报文规格

如果加入/剪枝报文的尺寸较大，则丢失一个报文将导致较多信息的遗失；如果加入/剪枝报文的尺寸较小，则单个报文的丢失所产生的影响也将降低。

通过控制加入/剪枝报文中（S，G）表项的数目，可以有效减少单位时间内发送的（S，G）表项数量。

表1-55 配置加入/剪枝报文规格

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-
配置加入/剪枝报文的最大长度	**jp-pkt-size** *packet-size*	可选 缺省情况下，加入/剪枝报文的最大长度为8100字节
配置加入/剪枝报文中（S，G）表项的最大数量	**jp-queue-size** *queue-size*	可选 缺省情况下，加入/剪枝报文中（S，G）表项的最大数量为1020个

注意

如果 PIM 网络中部署了配置 PIM Snooping 功能的交换机，请在接收者侧的边缘 PIM 设备上配置加入/剪枝报文的最大长度不能大于路径 MTU。

1.6.9 配置PIM与BFD联动

PIM 借助 Hello 报文在共享网段中选举出 DR，使其成为该网段中组播数据的唯一转发者。当 DR 出现故障时，只有待其老化后才会触发新的 DR 选举过程，这个过程通常比较长。为了实现 DR 的快速切换，可以在共享网段的 PIM 邻居之间引入 BFD 机制进行链路状态的快速检测。通过在共享网段内的所有 PIM 路由器上都使能 PIM 与 BFD 联动功能，可以使这些 PIM 邻居快速感知 DR 故障并重新选举 DR。

表1-56 配置 PIM 与 BFD 联动

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能PIM与BFD联动功能	**pim bfd enable**	必选 缺省情况下，PIM与BFD联动功能处于关闭状态

说明

- 只有在接口上先使能了 PIM-DM 或 PIM-SM，本配置才能生效。
- 有关 BFD 的详细介绍，请参见“可靠性配置指导”中的“BFD”。

1.6.10 配置PIM报文的DSCP优先级

在 IPv4 报文头中，包含一个 8bit 的 ToS 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 PIM 报文的 DSCP 优先级进行配置。

表1-57 配置 PIM 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例PIM视图	**pim** [**vpn-instance** *vpn-instance-name*]	-

操作	命令	说明
配置PIM报文的DSCP优先级	**dscp** *dscp-value*	可选 缺省情况下，PIM报文的DSCP优先级为48

1.7 PIM显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 PIM 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 PIM 统计信息。

表1-58 PIM 显示和维护

操作	命令
查看PIM-SM域中的BSR信息，以及本地配置并生效的C-RP信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **bsr-info** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看PIM所使用的单播路由信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **claimed-route** [*source-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看PIM控制报文的数量	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **control-message counters** [**message-type** { **probe** | **register** | **register-stop** } | [**interface** *interface-type interface-number* | **message-type** { **assert** | **bsr** | **crp** | **graft** | **graft-ack** | **hello** | **join-prune** | **state-refresh** }] *] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看双向PIM的DF信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **df-info** [*rp-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看尚未确认的PIM-DM嫁接信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **grafts** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上的PIM信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **interface** [*interface-type interface-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看待发送的加入/剪枝报文信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **join-prune mode** { **sm** [**flags** *flag-value*] | **ssm** } [**interface** *interface-type interface-number* | **neighbor** *neighbor-address*] * [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看PIM邻居信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **neighbor** [**interface** *interface-type interface-number* | *neighbor-address* | **verbose**] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看PIM路由表的内容	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **routing-table** [*group-address* [**mask** { *mask-length* | *mask* }] | *source-address* [**mask** { *mask-length* | *mask* }] | **incoming-interface** [*interface-type interface-number* | **register**] | **outgoing-interface** { **include** | **exclude** | **match** } { *interface-type interface-number* | **register** } | **mode** *mode-type* | **flags** *flag-value* | **fsm**] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看RP的信息	**display pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **rp-info** [*group-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
重置PIM控制报文计数器	**reset pim** [**all-instance** | **vpn-instance** *vpn-instance-name*] **control-message counters** [**interface** *interface-type interface-number*]

1.8 PIM典型配置举例

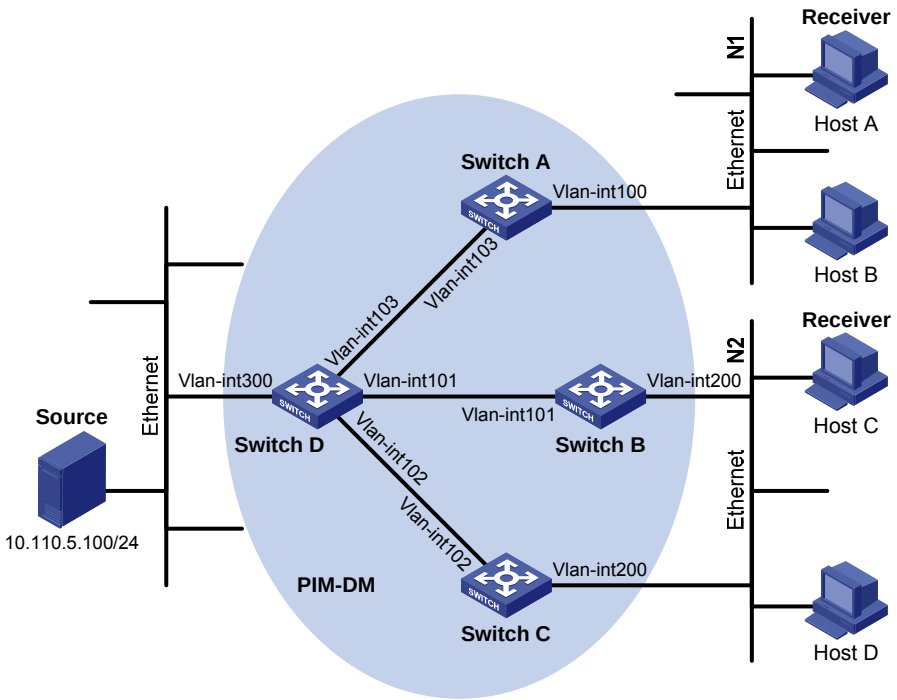
1.8.1 PIM-DM典型配置举例

1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者群体组成末梢网络，每个末梢网络中都存在至少一个接收者，整个 PIM 域采用 DM 方式。
- Host A 和 Host C 为两个末梢网络中的组播信息接收者；Switch D 通过 Vlan-interface300 接口与组播源（Source）所在的网络连接；Switch A 通过 Vlan-interface100 接口连接末梢网络 N1，通过 Vlan-interface103 接口连接 Switch D；Switch B 和 Switch C 通过各自的 Vlan-interface200 接口连接末梢网络 N2，分别通过 Vlan-interface101 和 Vlan-interface102 接口连接 Switch D。
- Switch A 与末梢网络 N1 之间运行 IGMPv2；Switch B 和 Switch C 与末梢网络 N2 之间也运行 IGMPv2。

2. 组网图

图1-14 PIM-DM 典型配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	10.110.1.1/24	Switch D	Vlan-int300	10.110.5.1/24
	Vlan-int103	192.168.1.1/24		Vlan-int103	192.168.1.2/24
Switch B	Vlan-int200	10.110.2.1/24		Vlan-int101	192.168.2.2/24
	Vlan-int101	192.168.2.1/24		Vlan-int102	192.168.3.2/24
Switch C	Vlan-int200	10.110.2.2/24			
	Vlan-int102	192.168.3.1/24			

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-14 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 PIM-DM 域内的各交换机之间采用 OSPF 协议进行互连，确保 PIM-DM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-DM 和 IGMP

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-DM，并在其连接末梢网络的接口 Vlan-interface100 上使能 IGMP。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim dm
[SwitchA-Vlan-interface103] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，配置过程略。

在 Switch D 上使能 IP 组播路由，并在各接口上使能 PIM-DM。

```
<SwitchD> system-view
[SwitchD] multicast routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] pim dm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim dm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 101
[SwitchD-Vlan-interface101] pim dm
[SwitchD-Vlan-interface101] quit
[SwitchD] interface vlan-interface 102
[SwitchD-Vlan-interface102] pim dm
[SwitchD-Vlan-interface102] quit
```

(3) 检验配置效果

通过使用 **display pim interface** 命令可以查看交换机接口上 PIM 的配置和运行情况。例如：

查看 Switch D 上 PIM 的配置信息。

```
[SwitchD] display pim interface
 VPN-Instance: public net
 Interface           NbrCnt HelloInt   DR-Pri     DR-Address
 Vlan300             0      30         1          10.110.5.1     (local)
 Vlan103             1      30         1          192.168.1.2    (local)
 Vlan101             1      30         1          192.168.2.2    (local)
 Vlan102             1      30         1          192.168.3.2    (local)
```

通过使用 **display pim neighbor** 命令可以查看交换机之间的 PIM 邻居关系。例如：

查看 Switch D 上 PIM 的邻居关系信息。

```
[SwitchD] display pim neighbor
 VPN-Instance: public net
 Total Number of Neighbors = 3

 Neighbor        Interface          Uptime   Expires  Dr-Priority
 192.168.1.1     Vlan103            00:02:22 00:01:27 1
 192.168.2.1     Vlan101            00:00:22 00:01:29 3
 192.168.3.1     Vlan102            00:00:23 00:01:31 5
```

假如 Host A 需要接收组播组 G（225.1.1.1）的信息，当组播源 S（10.110.5.100/24）向组播组 G 发送组播数据时，通过扩散生成 SPT，SPT 路径中各交换机（Switch A 和 Switch D）上都存在（S，G）表项，Host A 向 Switch A 发送 IGMP 报告以加入组播组 G，在 Switch A 上生成（*，G）表项。

通过使用 **display pim routing-table** 命令可以查看交换机的 PIM 路由表信息。例如：

查看 Switch A 上的 PIM 路由表信息。

```
[SwitchA] display pim routing-table
 VPN-Instance: public net
 Total 1 (*, G) entry; 1 (S, G) entry

 (*, 225.1.1.1)
     Protocol: pim-dm, Flag: WC
     UpTime: 00:04:25
     Upstream interface: NULL
         Upstream neighbor: NULL
         RPF prime neighbor: NULL
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: igmp, UpTime: 00:04:25, Expires: never

 (10.110.5.100, 225.1.1.1)
     Protocol: pim-dm, Flag: ACT
     UpTime: 00:06:14
     Upstream interface: Vlan-interface103
         Upstream neighbor: 192.168.1.2
         RPF prime neighbor: 192.168.1.2
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: pim-dm, UpTime: 00:04:25, Expires: never
```

查看 Switch D 上的 PIM 路由表信息。

```
[SwitchD] display pim routing-table
 VPN-Instance: public net
 Total 0 (*, G) entry; 1 (S, G) entry
 (10.110.5.100, 225.1.1.1)
     Protocol: pim-dm, Flag: LOC ACT
     UpTime: 00:03:27
     Upstream interface: Vlan-interface300
         Upstream neighbor: NULL
```

```
    RPF prime neighbor: NULL
Downstream interface(s) information:
Total number of downstreams: 2
    1: Vlan-interface103
        Protocol: pim-dm, UpTime: 00:03:27, Expires: never
    2: Vlan-interface102
        Protocol: pim-dm, UpTime: 00:03:27, Expires: never
```

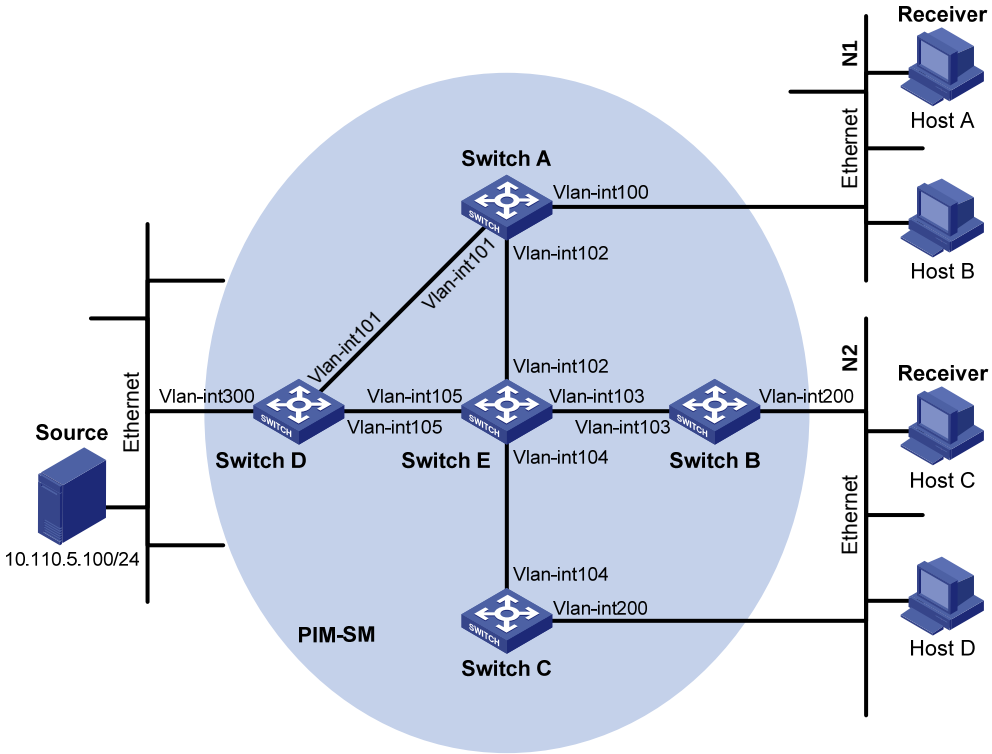
1.8.2 PIM-SM非管理域配置举例

1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者群体组成末梢网络，每个末梢网络中都存在至少一个接收者，整个 PIM 域采用 SM 非管理域方式。
- Host A 和 Host C 为两个末梢网络中的组播信息接收者；Switch D 通过 Vlan-interface300 接口与组播源（Source）所在网络连接；Switch A 通过 Vlan-interface100 接口连接末梢网络 N1，通过 Vlan-interface101 接口和 Vlan-interface102 接口分别连接 Switch D 和 Switch E；Switch B 和 Switch C 通过各自的 Vlan-interface200 接口连接末梢网络 N2，分别通过 Vlan-interface103 和 Vlan-interface104 接口连接 Switch E。
- 将 Switch D 的 Vlan-interface105 接口和 Switch E 的 Vlan-interface102 接口都配置为 C-BSR 和 C-RP，其中 Switch E 上 C-BSR 的优先级较高。C-RP 所服务的组播组范围为 225.1.1.0/24，通过改变哈希掩码长度使此范围内的组地址间隔映射到这两个 C-RP 上。
- Switch A 与末梢网络 N1 之间运行 IGMPv2；Switch B 和 Switch C 与末梢网络 N2 之间也运行 IGMPv2。

2. 组网图

图1-15 PIM-SM 非管理域配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	10.110.1.1/24	Switch D	Vlan-int300	10.110.5.1/24
	Vlan-int101	192.168.1.1/24		Vlan-int101	192.168.1.2/24
	Vlan-int102	192.168.9.1/24		Vlan-int105	192.168.4.2/24
Switch B	Vlan-int200	10.110.2.1/24	Switch E	Vlan-int104	192.168.3.2/24
	Vlan-int103	192.168.2.1/24		Vlan-int103	192.168.2.2/24
Switch C	Vlan-int200	10.110.2.2/24		Vlan-int102	192.168.9.2/24
	Vlan-int104	192.168.3.1/24		Vlan-int105	192.168.4.1/24

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-15 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 PIM-SM 域内的各交换机之间采用 OSPF 协议进行互连，确保 PIM-SM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-SM 和 IGMP

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在其连接末梢网络的接口 Vlan-interface100 上使能 IGMP。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
```

```
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim sm
[SwitchA-Vlan-interface102] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，Switch D 和 Switch E 除了不需要在相应接口上使能 IGMP 外，其它的配置也与 Switch A 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP

在 Switch D 上配置 RP 通告的服务范围，以及 C-BSR 和 C-RP 的位置，并指定哈希掩码长度为 32，C-BSR 的优先级为 10。

```
<SwitchD> system-view
[SwitchD] acl number 2005
[SwitchD-acl-basic-2005] rule permit source 225.1.1.0 0.0.0.255
[SwitchD-acl-basic-2005] quit
[SwitchD] pim
[SwitchD-pim] c-bsr vlan-interface 105 32 10
[SwitchD-pim] c-rp vlan-interface 105 group-policy 2005
[SwitchD-pim] quit
```

在 Switch E 上配置 RP 通告的服务范围，以及 C-BSR 和 C-RP 的位置，并指定哈希掩码长度为 32，C-BSR 的优先级为 20。

```
<SwitchE> system-view
[SwitchE] acl number 2005
[SwitchE-acl-basic-2005] rule permit source 225.1.1.0 0.0.0.255
[SwitchE-acl-basic-2005] quit
[SwitchE] pim
[SwitchE-pim] c-bsr vlan-interface 102 32 20
[SwitchE-pim] c-rp vlan-interface 102 group-policy 2005
[SwitchE-pim] quit
```

(4) 检验配置效果

通过使用 **display pim interface** 命令可以查看交换机接口上 PIM 的配置和运行情况。例如：

查看 Switch A 上 PIM 的配置信息。

```
[SwitchA] display pim interface
 VPN-Instance: public net
 Interface           NbrCnt HelloInt   DR-Pri     DR-Address
 Vlan100             0      30         1          10.110.1.1     (local)
 Vlan101             1      30         1          192.168.1.2
 Vlan102             1      30         1          192.168.9.2
```

通过使用 **display pim bsr-info** 命令可以查看交换机上 BSR 选举的信息，以及本地配置并生效的 C-RP 信息。例如：

查看 Switch A 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchA] display pim bsr-info
 VPN-Instance: public net
 Elected BSR Address: 192.168.9.2
     Priority: 20
     Hash mask length: 32
```

```
     State: Accept Preferred
     Scope: Not scoped
     Uptime: 00:40:40
     Expires: 00:01:42
```

查看 Switch D 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchD] display pim bsr-info
 VPN-Instance: public net
 Elected BSR Address: 192.168.9.2
     Priority: 20
     Hash mask length: 32
     State: Accept Preferred
     Scope: Not scoped
     Uptime: 00:05:26
     Expires: 00:01:45
 Candidate BSR Address: 192.168.4.2
     Priority: 10
     Hash mask length: 32
     State: Candidate
     Scope: Not scoped

 Candidate RP: 192.168.4.2(Vlan-interface105)
     Priority: 192
     HoldTime: 150
     Advertisement Interval: 60
     Next advertisement scheduled at: 00:00:34
```

查看 Switch E 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchE] display pim bsr-info
 VPN-Instance: public net
 Elected BSR Address: 192.168.9.2
     Priority: 20
     Hash mask length: 32
     State: Elected
     Scope: Not scoped
     Uptime: 00:01:18
     Next BSR message scheduled at: 00:01:52
 Candidate BSR Address: 192.168.9.2
     Priority: 20
     Hash mask length: 32
     State: Elected
     Scope: Not scoped

Candidate RP: 192.168.9.2(Vlan-interface102)
     Priority: 192
     HoldTime: 150
     Advertisement Interval: 60
     Next advertisement scheduled at: 00:00:48
```

通过使用 **display pim rp-info** 命令可以查看交换机上获取的 RP 信息。例如：

查看 Switch A 上的 RP 信息。

```
[SwitchA] display pim rp-info
 VPN-Instance: public net
 PIM-SM BSR RP information:
 Group/MaskLen: 225.1.1.0/24
     RP: 192.168.4.2
     Priority: 192
     HoldTime: 150
     Uptime: 00:51:45
     Expires: 00:02:22

     RP: 192.168.9.2
     Priority: 192
     HoldTime: 150
     Uptime: 00:51:45
     Expires: 00:02:22
```

假如 Host A 需要接收组播组 G（225.1.1.0）的信息，由于根据哈希算法得出 G 对应的 RP 为 Switch E，因此 Switch A 和 Switch E 之间会生成 RPT。当组播源 S（10.110.5.100/24）向 RP 发起注册后，Switch D 和 Switch E 之间会生成 SPT。当 Switch A 收到组播数据后立即执行从 RPT 到 SPT 的切换。RPT 路径中的交换机（Switch A 和 Switch E）上存在（*，G）表项，而 SPT 路径中的交换机（Switch A 和 Switch D）上存在（S，G）表项，通过使用 **display pim routing-table** 命令可以查看交换机的 PIM 路由表信息。例如：

查看 Switch A 上的 PIM 路由表信息。

```
[SwitchA] display pim routing-table
 VPN-Instance: public net
 Total 1 (*, G) entry; 1 (S, G) entry

 (*, 225.1.1.0)
     RP: 192.168.9.2
     Protocol: pim-sm, Flag: WC
     UpTime: 00:13:46
     Upstream interface: Vlan-interface102
         Upstream neighbor: 192.168.9.2
         RPF prime neighbor: 192.168.9.2
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: igmp, UpTime: 00:13:46, Expires: 00:03:06

 (10.110.5.100, 225.1.1.0)
     RP: 192.168.9.2
     Protocol: pim-sm, Flag: SPT ACT
     UpTime: 00:00:42
     Upstream interface: Vlan-interface101
         Upstream neighbor: 192.168.1.2
         RPF prime neighbor: 192.168.1.2
     Downstream interface(s) information:
```

```
    Total number of downstreams: 1
        1: Vlan-interface100
            Protocol: pim-sm, UpTime: 00:00:42, Expires: 00:03:06
```

查看 Switch D 上的 PIM 路由表信息。

```
[SwitchD] display pim routing-table
 VPN-Instance: public net
 Total 0 (*, G) entry; 1 (S, G) entry

 (10.110.5.100, 225.1.1.0)
    RP: 192.168.9.2
    Protocol: pim-sm, Flag: SPT LOC ACT
    UpTime: 00:00:42
    Upstream interface: Vlan-interface300
        Upstream neighbor: NULL
        RPF prime neighbor: NULL
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface105
            Protocol: pim-sm, UpTime: 00:00:42, Expires: 00:02:26
```

查看 Switch E 上的 PIM 路由表信息。

```
[SwitchE] display pim routing-table
 VPN-Instance: public net
 Total 1 (*, G) entry; 0 (S, G) entry

 (*, 225.1.1.0)
    RP: 192.168.9.2 (local)
    Protocol: pim-sm, Flag: WC
    UpTime: 00:13:16
    Upstream interface: Register
        Upstream neighbor: 192.168.4.2
        RPF prime neighbor: 192.168.4.2
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface102
            Protocol: pim-sm, UpTime: 00:13:16, Expires: 00:03:22
```

1.8.3 PIM-SM管理域配置举例

1. 组网需求

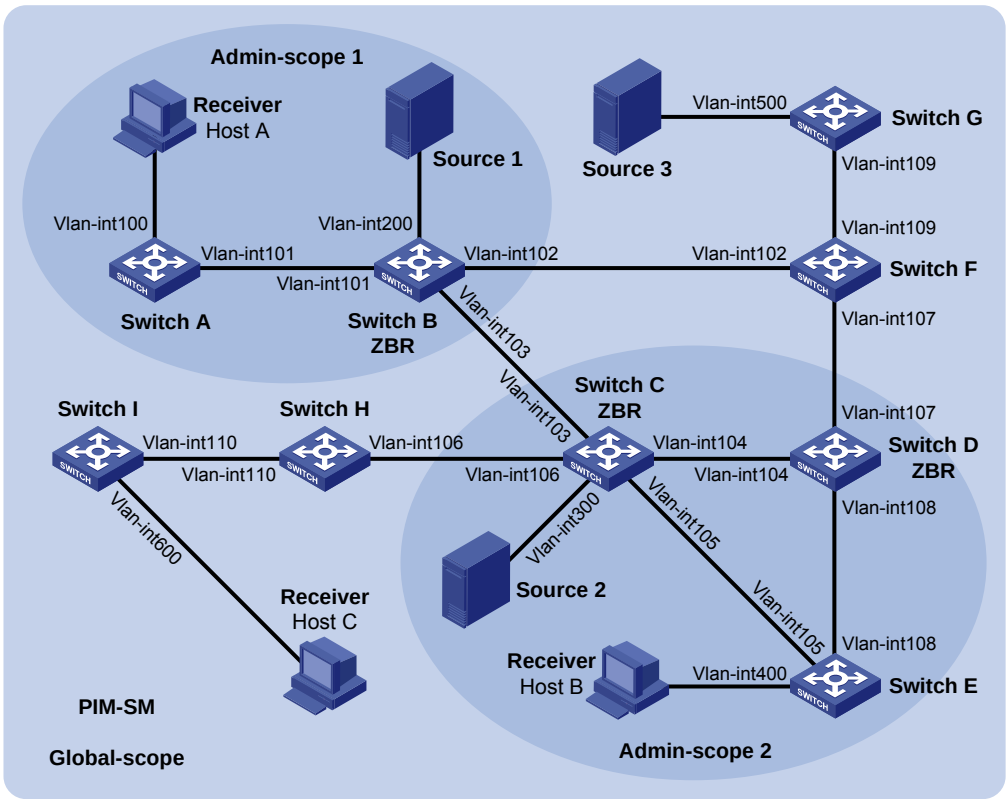
- 接收者通过组播方式接收视频点播信息，整个PIM域采用SM管理域方式，划分为管理域1、管理域 2 和 Global 域，Switch B、Switch C 和 Switch D 为各管理域的 ZBR。
- Source 1 和 Source 2 分别向组播组 239.1.1.1 发送内容不同的组播信息，Host A 和 Host B 则分别只接收来自 Source 1 和 Source 2 的组播信息；Source 3 向组播组 224.1.1.1 发送组播信息，Host C 为其接收者。
- Switch B 的 Vlan-interface101 接口为管理域 1 的 C-BSR 和 C-RP，服务于 239.0.0.0/8；Switch D 的 Vlan-interface104 接口为管理域 2 的 C-BSR 和 C-RP，服务于 239.0.0.0/8；

Switch F 的 Vlan-interface109 为 Global 域的 C-BSR 和 C-RP，服务于 239.0.0.0/8 以外的所有组播组。

- Switch A、Switch E 和 Switch I 分别与各自所连接的接收者之间运行 IGMPv2。

2. 组网图

图1-16 PIM-SM 管理域配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	192.168.1.1/24	Switch D	Vlan-int104	10.110.4.2/24
	Vlan-int101	10.110.1.1/24		Vlan-int108	10.110.7.1/24
Switch B	Vlan-int200	192.168.2.1/24		Vlan-int107	10.110.8.1/24
	Vlan-int101	10.110.1.2/24	Switch E	Vlan-int400	192.168.4.1/24
	Vlan-int103	10.110.2.1/24		Vlan-int105	10.110.5.2/24
	Vlan-int102	10.110.3.1/24		Vlan-int108	10.110.7.2/24
Switch C	Vlan-int300	192.168.3.1/24	Switch F	Vlan-int109	10.110.9.1/24
	Vlan-int104	10.110.4.1/24		Vlan-int107	10.110.8.2/24
	Vlan-int105	10.110.5.1/24		Vlan-int102	10.110.3.2/24
	Vlan-int103	10.110.2.2/24	Switch G	Vlan-int500	192.168.5.1/24
	Vlan-int106	10.110.6.1/24		Vlan-int109	10.110.9.2/24
Switch H	Vlan-int110	10.110.10.1/24	Source 1	-	192.168.2.10/24
	Vlan-int106	10.110.6.2/24	Source 2	-	192.168.3.10/24
Switch I	Vlan-int600	192.168.6.1/24	Source 3	-	192.168.5.10/24
	Vlan-int110	10.110.10.2/24			

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-16 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 PIM-SM 域内的各交换机之间采用 OSPF 协议进行互连，确保 PIM-SM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由和管理域机制，并使能 PIM-SM 和 IGMP

在 Switch A 上使能 IP 组播路由和管理域机制，在各接口上使能 PIM-SM，并在其连接有接收者的接口 Vlan-interface100 上使能 IGMP。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] pim
[SwitchA-pim] c-bsr admin-scope
[SwitchA-pim] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
```

Switch E 和 Switch I 的配置与 Switch A 相似，配置过程略。

在 Switch B 上使能 IP 组播路由和管理域机制，并在各接口上使能 PIM-SM。

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] pim
[SwitchB-pim] c-bsr admin-scope
[SwitchB-pim] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] pim sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim sm
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim sm
[SwitchB-Vlan-interface103] quit
```

Switch C、Switch D、Switch F、Switch G 和 Switch H 的配置与 Switch B 相似，配置过程略。

(3) 配置管理域边界

在 Switch B 上将接口 Vlan-interface102 和 Vlan-interface103 配置为管理域 1 的边界。

```
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] multicast boundary 239.0.0.0 8
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] multicast boundary 239.0.0.0 8
[SwitchB-Vlan-interface103] quit
```

在 Switch C 上将接口 Vlan-interface103 和 Vlan-interface106 配置为管理域 2 的边界。

```
<SwitchC> system-view
[SwitchC] interface vlan-interface 103
```

```
[SwitchC-Vlan-interface103] multicast boundary 239.0.0.0 8
[SwitchC-Vlan-interface103] quit
[SwitchC] interface vlan-interface 106
[SwitchC-Vlan-interface106] multicast boundary 239.0.0.0 8
[SwitchC-Vlan-interface106] quit
```

在 Switch D 上将接口 Vlan-interface107 配置为管理域 2 的边界。

```
<SwitchD> system-view
[SwitchD] interface vlan-interface 107
[SwitchD-Vlan-interface107] multicast boundary 239.0.0.0 8
[SwitchD-Vlan-interface107] quit
```

(4) 配置 C-BSR 和 C-RP

在 Switch B 上配置 RP 通告的服务范围，并将接口 Vlan-interface101 配置为管理域 1 的 C-BSR 和 C-RP。

```
[SwitchB] acl number 2001
[SwitchB-acl-basic-2001] rule permit source 239.0.0.0 0.255.255.255
[SwitchB-acl-basic-2001] quit
[SwitchB] pim
[SwitchB-pim] c-bsr group 239.0.0.0 8
[SwitchB-pim] c-bsr vlan-interface 101
[SwitchB-pim] c-rp vlan-interface 101 group-policy 2001
[SwitchB-pim] quit
```

在 Switch D 上配置 RP 通告的服务范围，并将接口 Vlan-interface104 配置为管理域 2 的 C-BSR 和 C-RP。

```
[SwitchD] acl number 2001
[SwitchD-acl-basic-2001] rule permit source 239.0.0.0 0.255.255.255
[SwitchD-acl-basic-2001] quit
[SwitchD] pim
[SwitchD-pim] c-bsr group 239.0.0.0 8
[SwitchD-pim] c-bsr vlan-interface 104
[SwitchD-pim] c-rp vlan-interface 104 group-policy 2001
[SwitchD-pim] quit
```

在 Switch F 上将接口 Vlan-interface109 配置为 Global 域的 C-BSR 和 C-RP。

```
<SwitchF> system-view
[SwitchF] pim
[SwitchF-pim] c-bsr global
[SwitchF-pim] c-bsr vlan-interface 109
[SwitchF-pim] c-rp vlan-interface 109
[SwitchF-pim] quit
```

(5) 检验配置效果

通过使用 **display pim bsr-info** 命令可以查看交换机上 BSR 选举的信息，以及本地配置并生效的 C-RP 信息。例如：

查看 Switch B 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchB] display pim bsr-info
 VPN-Instance: public net
 Elected BSR Address: 10.110.9.1
     Priority: 64
```

```
    Hash mask length: 30
    State: Accept Preferred
    Scope: Global
    Uptime: 00:01:45
    Expires: 00:01:25
 Elected BSR Address: 10.110.1.2
    Priority: 64
    Hash mask length: 30
    State: Elected
    Scope: 239.0.0.0/8
    Uptime: 00:04:54
    Next BSR message scheduled at: 00:00:06
 Candidate BSR Address: 10.110.1.2
    Priority: 64
    Hash mask length: 30
    State: Elected
    Scope: 239.0.0.0/8

 Candidate RP: 10.110.1.2(Vlan-interface101)
    Priority: 192
    HoldTime: 150
    Advertisement Interval: 60
    Next advertisement scheduled at: 00:00:15
```

查看 Switch D 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchD] display pim bsr-info
 VPN-Instance: public net
 Elected BSR Address: 10.110.9.1
    Priority: 64
    Hash mask length: 30
    State: Accept Preferred
    Scope: Global
    Uptime: 00:01:45
    Expires: 00:01:25
 Elected BSR Address: 10.110.4.2
    Priority: 64
    Hash mask length: 30
    State: Elected
    Scope: 239.0.0.0/8
    Uptime: 00:03:48
    Next BSR message scheduled at: 00:01:12
 Candidate BSR Address: 10.110.4.2
    Priority: 64
    Hash mask length: 30
    State: Elected
    Scope: 239.0.0.0/8

 Candidate RP: 10.110.4.2(Vlan-interface104)
    Priority: 192
```

```
     HoldTime: 150
     Advertisement Interval: 60
     Next advertisement scheduled at: 00:00:10
```

查看 Switch F 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchF] display pim bsr-info
 VPN-Instance: public net
 Elected BSR Address: 10.110.9.1
     Priority: 64
     Hash mask length: 30
     State: Elected
     Scope: Global
     Uptime: 00:11:11
     Next BSR message scheduled at: 00:00:49
 Candidate BSR Address: 10.110.9.1
     Priority: 64
     Hash mask length: 30
     State: Elected
     Scope: Global

 Candidate RP: 10.110.9.1(Vlan-interface109)
     Priority: 192
     HoldTime: 150
     Advertisement Interval: 60
     Next advertisement scheduled at: 00:00:55
```

通过使用 **display pim rp-info** 命令可以查看交换机上获取的 RP 信息。例如：

查看 Switch B 上的 RP 信息。

```
[SwitchB] display pim rp-info
 VPN-Instance: public net
 PIM-SM BSR RP information:
 Group/MaskLen: 224.0.0.0/4
     RP: 10.110.9.1
     Priority: 192
     HoldTime: 150
     Uptime: 00:03:39
     Expires: 00:01:51

 Group/MaskLen: 239.0.0.0/8
     RP: 10.110.1.2 (local)
     Priority: 192
     HoldTime: 150
     Uptime: 00:07:44
     Expires: 00:01:51
```

查看 Switch D 上的 RP 信息。

```
[SwitchD] display pim rp-info
 VPN-Instance: public net
 PIM-SM BSR RP information:
 Group/MaskLen: 224.0.0.0/4
```

```
    RP: 10.110.9.1
    Priority: 192
    HoldTime: 150
    Uptime: 00:03:42
    Expires: 00:01:48

 Group/MaskLen: 239.0.0.0/8
    RP: 10.110.4.2 (local)
    Priority: 192
    HoldTime: 150
    Uptime: 00:06:54
    Expires: 00:02:41
```

查看 Switch F 上的 RP 信息。

```
[SwitchF] display pim rp-info
 VPN-Instance: public net
 PIM-SM BSR RP information:
 Group/MaskLen: 224.0.0.0/4
    RP: 10.110.9.1 (local)
    Priority: 192
    HoldTime: 150
    Uptime: 00:00:32
    Expires: 00:01:58
```

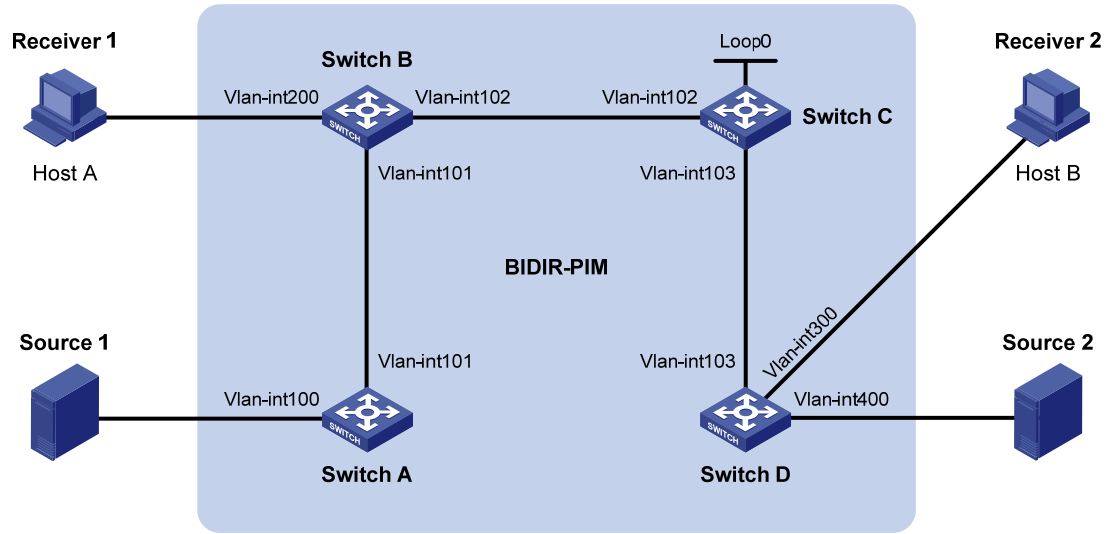
1.8.4 双向PIM典型配置举例

1. 组网需求

- 整个 PIM 域采用 BIDIR 方式，Source 1 和 Source 2 都向组播组 225.1.1.1 发送组播信息，Host A 和 Host B 为组播信息的接收者。
- 将 Switch C 的 Vlan-interface102 接口配置为 C-BSR，Loopback0 接口配置为服务于双向 PIM 的 C-RP。
- Switch B 和 Switch D 分别与各自所连接的接收者之间运行 IGMPv2。

2. 组网图

图1-17 双向 PIM 典型配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	192.168.1.1/24	Switch D	Vlan-int300	192.168.3.1/24
	Vlan-int101	10.110.1.1/24		Vlan-int400	192.168.4.1/24
Switch B	Vlan-int200	192.168.2.1/24		Vlan-int103	10.110.3.2/24
	Vlan-int101	10.110.1.2/24	Source 1	-	192.168.1.100/24
	Vlan-int102	10.110.2.1/24	Source 2	-	192.168.4.100/24
Switch C	Vlan-int102	10.110.2.2/24	Receiver 1	-	192.168.2.100/24
	Vlan-int103	10.110.3.1/24	Receiver 2	-	192.168.3.100/24
	Loop0	1.1.1.1/32			

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-17 配置各接口的IP地址和子网掩码，具体配置过程略。

配置双向 PIM 域内的各交换机之间采用 OSPF 协议进行互连，确保双向 PIM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-SM、双向 PIM 和 IGMP

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-SM，并使能双向 PIM。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] pim
[SwitchA-pim] bidir-pim enable
[SwitchA-pim] quit
```

在 Switch B 上使能 IP 组播路由，在各接口上使能 PIM-SM，在其连接有接收者的接口 Vlan-interface200 上使能 IGMP，并使能双向 PIM。

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] igmp enable
[SwitchB-Vlan-interface200] pim sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim sm
[SwitchB-Vlan-interface102] quit
[SwitchB] pim
[SwitchB-pim] bidir-pim enable
[SwitchB-pim] quit
```

在 Switch C 上使能 IP 组播路由，在各接口上使能 PIM-SM，并使能双向 PIM。

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] pim sm
[SwitchC-Vlan-interface103] quit
[SwitchC] interface loopback 0
[SwitchC-LoopBack0] pim sm
[SwitchC-LoopBack0] quit
[SwitchC] pim
[SwitchC-pim] bidir-pim enable
```

在 Switch D 上使能 IP 组播路由，在各接口上使能 PIM-SM，在其连接有接收者的接口 Vlan-interface300 上使能 IGMP，并使能双向 PIM。

```
<SwitchD> system-view
[SwitchD] multicast routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] igmp enable
[SwitchD-Vlan-interface300] pim sm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] pim sm
[SwitchD-Vlan-interface400] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim sm
[SwitchD-Vlan-interface103] quit
[SwitchD] pim
[SwitchD-pim] bidir-pim enable
[SwitchD-pim] quit
```

(3) 配置 C-BSR 和 C-RP

在 Switch C 上将接口 Vlan-interface102 配置为 C-BSR，并将接口 Loopback0 配置为服务于双向 PIM 的 C-RP。

```
[SwitchC-pim] c-bsr vlan-interface 102
[SwitchC-pim] c-rp loopback 0 bidir
[SwitchC-pim] quit
```

(4) 检验配置效果

通过使用 **display pim df-info** 命令可以查看交换机上双向 PIM 的 DF 信息。例如：

查看 Switch A 上双向 PIM 的 DF 信息。

```
[SwitchA] display pim df-info
 VPN-Instance: public net

 RP Address: 1.1.1.1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Vlan100             Win     100        2          01:08:50  192.168.1.1 (local)
  Vlan101             Lose    100        1          01:07:49  10.110.1.2
```

查看 Switch B 上双向 PIM 的 DF 信息。

```
[SwitchB] display pim df-info
 VPN-Instance: public net

 RP Address: 1.1.1.1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Vlan200             Win     100        1          01:24:09  192.168.2.1 (local)
  Vlan101             Win     100        1          01:24:09  10.110.1.2 (local)
  Vlan102             Lose    0          0          01:23:12  10.110.2.2
```

查看 Switch C 上双向 PIM 的 DF 信息。

```
[SwitchC] display pim df-info
 VPN-Instance: public net

 RP Address: 1.1.1.1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Loop0               -       -          -          -         -
  Vlan102             Win     0          0          01:06:07  10.110.2.2 (local)
  Vlan103             Win     0          0          01:06:07  10.110.3.1 (local)
```

查看 Switch D 上双向 PIM 的 DF 信息。

```
[SwitchD] display pim df-info
 VPN-Instance: public net

 RP Address: 1.1.1.1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Vlan300             Win     100        1          01:19:53  192.168.3.1 (local)
  Vlan400             Win     100        1          00:39:34  192.168.4.1 (local)
  Vlan103             Lose    0          0          01:21:40  10.110.3.1
```

通过使用 **display multicast forwarding-table df-info** 命令可以查看交换机上组播转发表的 DF 信息，有关 **display multicast forwarding-table df-info** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。例如：

查看 Switch A 上组播转发表的 DF 信息。

```
[SwitchA] display multicast forwarding-table df-info
Multicast DF information of VPN-Instance: public net
Total 1 RP

Total 1 RP matched

00001. RP Address: 1.1.1.1
     MID: 0, Flags: 0x2100000:0
     Uptime: 00:08:32
     RPF interface: Vlan-interface101
     List of 1 DF interfaces:
       1: Vlan-interface100
```

查看 Switch B 上组播转发表的 DF 信息。

```
[SwitchB] display multicast forwarding-table df-info
Multicast DF information of VPN-Instance: public net
Total 1 RP

Total 1 RP matched

00001. RP Address: 1.1.1.1
     MID: 0, Flags: 0x2100000:0
     Uptime: 00:06:24
     RPF interface: Vlan-interface102
     List of 2 DF interfaces:
       1: Vlan-interface101
       2: Vlan-interface200
```

查看 Switch C 上组播转发表的 DF 信息。

```
[SwitchC] display multicast forwarding-table df-info
Multicast DF information of VPN-Instance: public net
Total 1 RP

Total 1 RP matched

00001. RP Address: 1.1.1.1
     MID: 0, Flags: 0x2100000:0
     Uptime: 00:07:21
     RPF interface: LoopBack0
     List of 2 DF interfaces:
       1: Vlan-interface102
       2: Vlan-interface103
```

查看 Switch D 上组播转发表的 DF 信息。

```
[SwitchD] display multicast forwarding-table df-info
Multicast DF information of VPN-Instance: public net
Total 1 RP

Total 1 RP matched

00001. RP Address: 1.1.1.1
```

```
MID: 0, Flags: 0x2100000:0
Uptime: 00:05:12
RPF interface: Vlan-interface103
List of 2 DF interfaces:
  1: Vlan-interface300
  2: Vlan-interface400
```

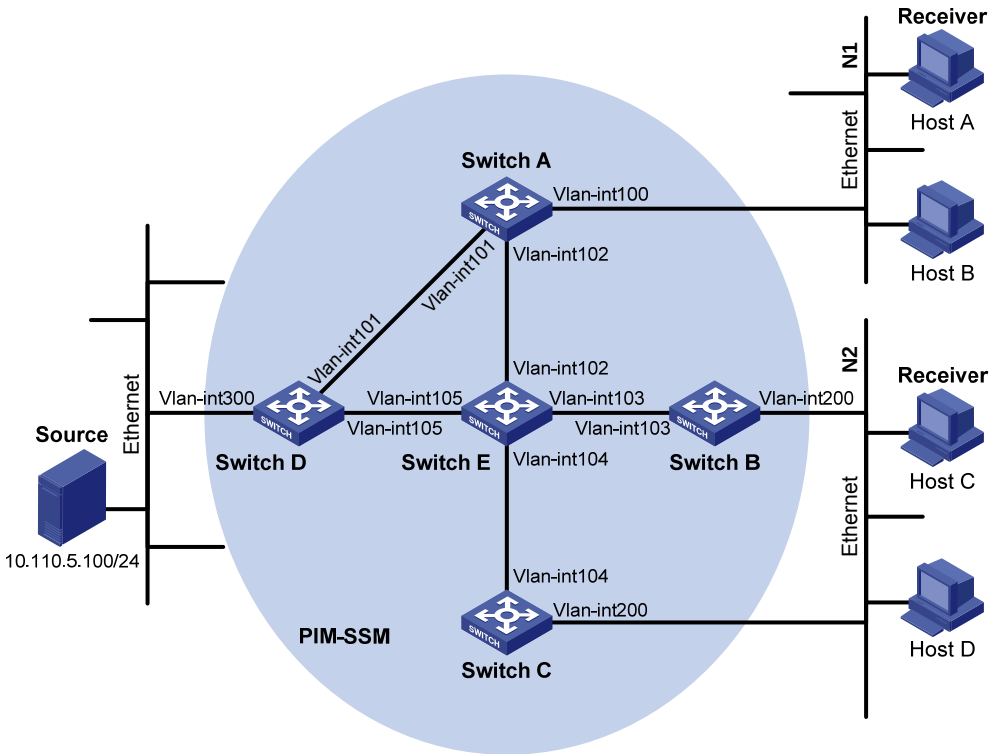
1.8.5 PIM-SSM典型配置举例

1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者群体组成末梢网络，每个末梢网络中都存在至少一个接收者，整个 PIM 域采用 SSM 方式。
- Host A 和 Host C 为两个末梢网络中的组播信息接收者；Switch D 通过 Vlan-interface300 接口与组播源（Source）所在网络连接；Switch A 通过 Vlan-interface100 接口连接末梢网络 N1，通过 Vlan-interface101 接口和 Vlan-interface102 接口分别连接 Switch D 和 Switch E；Switch B 和 Switch C 通过各自的 Vlan-interface200 接口连接末梢网络 N2，分别通过 Vlan-interface103 和 Vlan-interface104 接口连接 Switch E；Switch E 连接 Switch A、Switch B、Switch C 和 Switch D。
- SSM 组播组的范围是 232.1.1.0/24。
- Switch A 与末梢网络 N1 之间运行 IGMPv3；Switch B 和 Switch C 与末梢网络 N2 之间也运行 IGMPv3。

2. 组网图

图1-18 PIM-SSM 典型配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	10.110.1.1/24	Switch D	Vlan-int300	10.110.5.1/24

	Vlan-int101	192.168.1.1/24		Vlan-int101	192.168.1.2/24
	Vlan-int102	192.168.9.1/24		Vlan-int105	192.168.4.2/24
Switch B	Vlan-int200	10.110.2.1/24	Switch E	Vlan-int104	192.168.3.2/24
	Vlan-int103	192.168.2.1/24		Vlan-int103	192.168.2.2/24
Switch C	Vlan-int200	10.110.2.2/24		Vlan-int102	192.168.9.2/24
	Vlan-int104	192.168.3.1/24		Vlan-int105	192.168.4.1/24

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-18 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 PIM-SSM 域内的各交换机之间采用 OSPF 协议进行互连，确保 PIM-SSM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-SM 和 IGMP

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在其连接末梢网络的接口 Vlan-interface100 上使能 IGMP，且配置其版本为 3。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] igmp version 3
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim sm
[SwitchA-Vlan-interface102] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，Switch D 和 Switch E 除了不需要在相应接口上使能 IGMP 外，其它的配置也与 Switch A 相似，配置过程略。

(3) 配置 SSM 组播组的地址范围

在 Switch A 上配置 SSM 组播组地址范围为 232.1.1.0/24。

```
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule permit source 232.1.1.0 0.0.0.255
[SwitchA-acl-basic-2000] quit
[SwitchA] pim
[SwitchA-pim] ssm-policy 2000
[SwitchA-pim] quit
```

Switch B、Switch C、Switch D 和 Switch E 的配置与 Switch A 相似，配置过程略。

(4) 检验配置效果

通过使用 **display pim interface** 命令可以查看交换机接口上 PIM 的配置和运行情况。例如：

查看 Switch A 上 PIM 的配置信息。

```
[SwitchA] display pim interface
 VPN-Instance: public net
 Interface           NbrCnt HelloInt   DR-Pri     DR-Address
```

```
 Vlan100             0      30         1          10.110.1.1      (local)
 Vlan101             1      30         1          192.168.1.2
 Vlan102             1      30         1          192.168.9.2
```

假如 Host A 需要接收指定组播源 S（10.110.5.100/24）发往组播组 G（232.1.1.1）的信息，Switch A 会向组播源方向构造 SPT，SPT 路径中的交换机（Switch A 和 Switch D）上生成（S，G）表项，而 SPT 路径之外的交换机（Switch E）上没有组播路由项，通过使用 **display pim routing-table** 命令可以查看交换机的 PIM 路由表信息。例如：

查看 Switch A 上的 PIM 路由表信息。

```
[SwitchA] display pim routing-table
 VPN-Instance: public net
 Total 0 (*, G) entry; 1 (S, G) entry

 (10.110.5.100, 232.1.1.1)
     Protocol: pim-ssm, Flag:
     UpTime: 00:13:25
     Upstream interface: Vlan-interface101
         Upstream neighbor: 192.168.1.2
         RPF prime neighbor: 192.168.1.2
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: igmp, UpTime: 00:13:25, Expires: 00:03:25
```

查看 Switch D 上的 PIM 路由表信息。

```
[SwitchD] display pim routing-table
 VPN-Instance: public net
 Total 0 (*, G) entry; 1 (S, G) entry

 (10.110.5.100, 232.1.1.1)
     Protocol: pim-ssm, Flag: LOC
     UpTime: 00:12:05
     Upstream interface: Vlan-interface300
         Upstream neighbor: NULL
         RPF prime neighbor: NULL
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface105
             Protocol: pim-ssm, UpTime: 00:12:05, Expires: 00:03:25
```

1.9 常见配置错误举例

1.9.1 无法正确建立组播分发树

1. 故障现象

网络中各路由器（包括直连组播源或接收者的路由器）上都没有组播转发项，也就是说无法正确建立组播分发树，客户端无法接收组播数据。

2. 分析

- 当全网运行 PIM-DM 时，组播数据由直连组播源的第一跳路由器扩散到直连客户端的最后一跳路由器。无论组播数据扩散到哪一台路由器，只有该路由器存在到达组播源的路由，才会创建（S，G）表项。反之，如果没有到达组播源的路由或者到达组播源的 RPF 接口没有使能 PIM-DM，则该路由器无法创建（S，G）表项。
- 当全网运行 PIM-SM 时，路由器在准备加入 SPT 时，只有存在到达组播源的路由，才会创建（S，G）表项。反之，如果没有到达组播源的路由或者到达组播源的 RPF 接口没有使能 PIM-SM，则该路由器无法创建（S，G）表项。
- 对于某个 RPF 检查对象，在现存的单播路由表中查找到达该对象的最优路由，该路由的出接口作为 RPF 接口，下一跳作为 RPF 邻居。RPF 接口完全依赖于现存的单播路由，并且与 PIM 本身无关。RPF 接口必须使能 PIM，而且 RPF 邻居也必须是 PIM 邻居。如果 RPF 接口或 RPF 邻居所在路由器上没有使能 PIM，必然导致组播分发树无法正确建立，导致组播数据转发异常。
- Hello 报文并不携带 PIM 的模式信息，所以运行 PIM 的路由器无法掌握自己的 PIM 邻居运行的是何种模式的 PIM。如果 RPF 接口和 RPF 邻居所在路由器的对应接口没有使能相同模式的 PIM，必然导致组播分发树无法正确建立，导致组播数据转发异常。
- 全网必须运行相同模式的 PIM。否则，组播分发树必然无法正确建立，导致组播数据转发异常。

3. 处理过程

(1) 检查单播路由。使用命令 **display ip routing-table** 检查从接收者主机是否有到达组播源的单播路由项。

(2) 检查接口上是否使能 PIM，尤其是 RPF 接口上是否使能 PIM。通过命令 **display pim interface** 观察接口上的 PIM 信息。若接口上没有使能 PIM，请使用 **pim dm** 或 **pim sm** 命令使能 PIM-DM 或者 PIM-SM。

(3) 检查 RPF 邻居是否是 PIM 邻居。通过命令 **display pim neighbor** 观察 PIM 邻居信息。

(4) 检查直连组播源或接收者的路由器接口上是否使能了 PIM 和 IGMP。

(5) 检查 PIM 模式是否一致。通过命令 **display pim interface verbose** 检查 RPF 接口和 RPF 邻居所在路由器的对应接口上是否使能了相同模式的 PIM。

(6) 检查全部网络中各路由器上的 PIM 模式是否一致。确保全网所有路由器配置相同模式的 PIM，即要么全部配置为 PIM-SM，要么全部配置为 PIM-DM。如果配置为 PIM-SM，则还需要检查 BSR 以及 RP 的配置是否正确。

1.9.2 组播数据异常终止在中间路由器

1. 故障现象

组播数据可以到达中间路由器，但无法到达最后一跳路由器。中间路由器某接口上收到组播数据，但 PIM 路由表中没有创建相应的（S，G）表项。

2. 分析

- 命令 **multicast boundary** 用于在接口上设置组播转发边界，如果组播数据无法通过该边界，PIM 是无法创建路由项的。

- 此外，**source-policy** 命令用于过滤接收到的组播数据报文。如果组播数据报文无法通过该命令的 ACL 规则，PIM 也是无法创建路由项的。

3. 处理过程

(1) 检查组播转发边界的配置。通过命令 **display current-configuration** 查看组播转发边界上的设置，使用 **multicast boundary** 命令更改组播转发边界的设置，使组播数据能够通过该边界。

(2) 检查组播过滤器配置。通过命令 **display current-configuration** 查看组播过滤器的配置，更改 **source-policy** 命令的 ACL 规则，使组播数据的源/组地址通过 ACL 过滤。

1.9.3 PIM-SM中RP无法加入SPT

1. 故障现象

共享树无法正确建立，或者 RP 无法加入到达组播源的最短路径树。

2. 分析

- RP 是 PIM-SM 的核心，为特定的组服务。网络中可以同时存在多个 RP。必须保证所有路由器的 RP 信息完全一致，并且对于某个特定的组映射到相同的 RP。否则必然导致组播数据转发异常。
- 如果使用静态 RP 机制，必须在全网所有路由器上配置完全相同的静态 RP 命令。否则必然导致组播数据转发异常。

3. 处理过程

(1) 检查是否有到达 RP 的路由。通过命令 **display ip routing-table** 查看各路由器上是否有到达 RP 的路由。

(2) 检查动态 RP 信息。通过命令 **display pim rp-info** 查看各路由器上的 RP 信息是否一致。

(3) 检查静态 RP 的配置。通过命令 **display pim rp-info** 查看全网所有路由器上是否配置了完全相同的静态 RP。

1.9.4 PIM-SM中无法建立RPT或无法进行源注册

1. 故障现象

C-RP 无法向 BSR 单播通告报文，BSR 没有发布包含 C-RP 的自举报文，BSR 上没有到达各 C-RP 的单播路由，共享树无法正确建立，或者 DR 无法向 RP 进行源注册。

2. 分析

- 由于 C-RP 周期性地向 BSR 单播宣告报文，如果 C-RP 没有到达 BSR 的单播路由就无法发送宣告报文，BSR 就收不到 C-RP 宣告报文，也就不会发布包含该 C-RP 的自举报文。
- 另外，如果 BSR 没有到达 C-RP 的单播路由，就会丢弃该宣告报文，也就不会发布包含该 C-RP 的自举报文。
- RP 是 PIM-SM 域的核心。必须保证全网所有路由器的 RP 信息完全一致，并且对于某个特定的组 G 映射到相同的 RP，并且存在到达 RP 的单播路由。

3. 处理过程

(1) 检查是否有到达各 C-RP、BSR 的路由。通过命令 **display ip routing-table** 查看各路由器上是否有到达 RP 和 BSR 的路由，及 RP 和 BSR 之间的路由。确保 C-RP 上存在到达 BSR 的单播路由，BSR 上存在到达 C-RP 的单播路由，全网络所有路由器上存在到达 RP 的单播路由。

(2) 检查 RP 和 BSR 信息。PIM-SM 协议需要有 RP 和 BSR 的支持，首先使用命令 **display pim bsr-info** 查看是否有 BSR 信息，使用 **display pim rp-info** 命令查看 RP 信息是否正确。

(3) 检查 PIM 邻居关系。通过命令 **display pim neighbor** 来查看是否正确建立了邻居关系。

目 录

1 MSDP配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- 本文中所提到的 DR（Designated Router，指定路由器）、BSR（Bootstrap Router，自举路由器）、C-BSR（Candidate-BSR，候选 BSR）、RP（Rendezvous Point，汇集点）、C-RP（Candidate-RP，候选 RP）、SPT（Shortest Path Tree，最短路径树）和 RPT（Rendezvous Point Tree，共享树）等概念的详细介绍，请参见“IP 组播配置指导”中的“PIM”。
- MSDP 功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

1.1 MSDP简介

1.1.1 MSDP概述

MSDP 是 Multicast Source Discovery Protocol（组播源发现协议）的简称，是为了解决多个 PIM-SM（Protocol Independent Multicast Sparse Mode，协议无关组播—稀疏模式）域之间的互连而开发的一种域间组播解决方案，用来发现其它 PIM-SM 域内的组播源信息。

在基本的 PIM-SM 模式下，组播源只向本 PIM-SM 域内的 RP 注册，且各域的组播源信息是相互隔离的，因此 RP 仅知道本域内的组播源信息，只能在本域内建立组播分发树，将本域内组播源发出的组播数据分发给本地用户。如果能够有一种机制，将其它域内的组播源信息传递给本域内的 RP，则本域内的 RP 就可以向其它域内的组播源发起加入过程并建立组播分发树，从而实现组播数据的跨域传输。

基于这一设想，MSDP 通过在网络中选取适当的路由器建立 MSDP 对等体关系，以连通各 PIM-SM 域的 RP。通过在各 MSDP 对等体之间交互 SA（Source Active，信源有效）消息来共享组播源信息。

注意

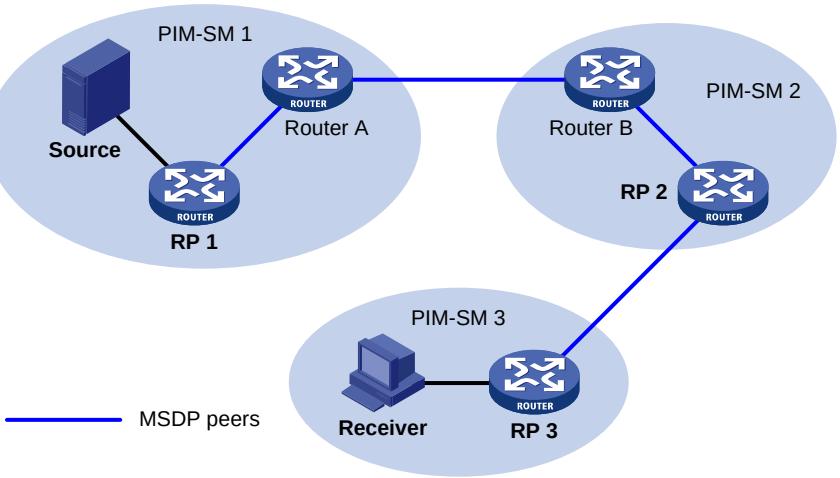
- MSDP 的适用前提：域内组播路由协议必须是 PIM-SM。
- MSDP 仅对 ASM（Any-Source Multicast，任意信源组播）模型有意义。

1.1.2 MSDP原理

1. MSDP对等体

通过在网络中配置一对或多对 MSDP 对等体，形成彼此相连的一张“MSDP 连通图”，以连通各个 PIM-SM 域的 RP。通过这些 MSDP 对等体之间的接力，可以把某 RP 发出的 SA 消息传递给其它所有的 RP。

图1-1 MSDP 对等体的位置



如 图 1-1 所示，MSDP对等体可以创建在任意的PIM-SM路由器上，在不同角色的PIM-SM路由器上所创建的MSDP对等体的功能有所不同：

(1) 在 RP 上创建的 MSDP 对等体

- 源端 MSDP 对等体：即离组播源（Source）最近的 MSDP 对等体（通常也就是源端 RP，如 RP 1）。源端 RP 创建 SA 消息并发送给远端 MSDP 对等体，通告在本 RP 上注册的组播源信息。源端 MSDP 对等体必须配置在 RP 上，否则将无法向外发布组播源信息。
- 接收者端 MSDP 对等体：即离接收者（Receiver）最近的 MSDP 对等体（如 RP 3）。接收者端 MSDP 对等体在收到 SA 消息后，根据该消息中所包含的组播源信息，跨域加入以该组播源为根的 SPT；当来自该组播源的组播数据到达后，再沿 RPT 向本地接收者转发。
- 中间 MSDP 对等体：即拥有多个远端 MSDP 对等体的 MSDP 对等体（如 RP 2）。中间 MSDP 对等体把从一个远端 MSDP 对等体收到的 SA 消息转发给其它远端 MSDP 对等体，其作用相当于传输组播源信息的中转站。

(2) 在普通的 PIM-SM 路由器（非 RP）上创建的 MSDP 对等体

如 Router A 和 Router B，其作用仅限于将收到的 SA 消息转发出去。

说明

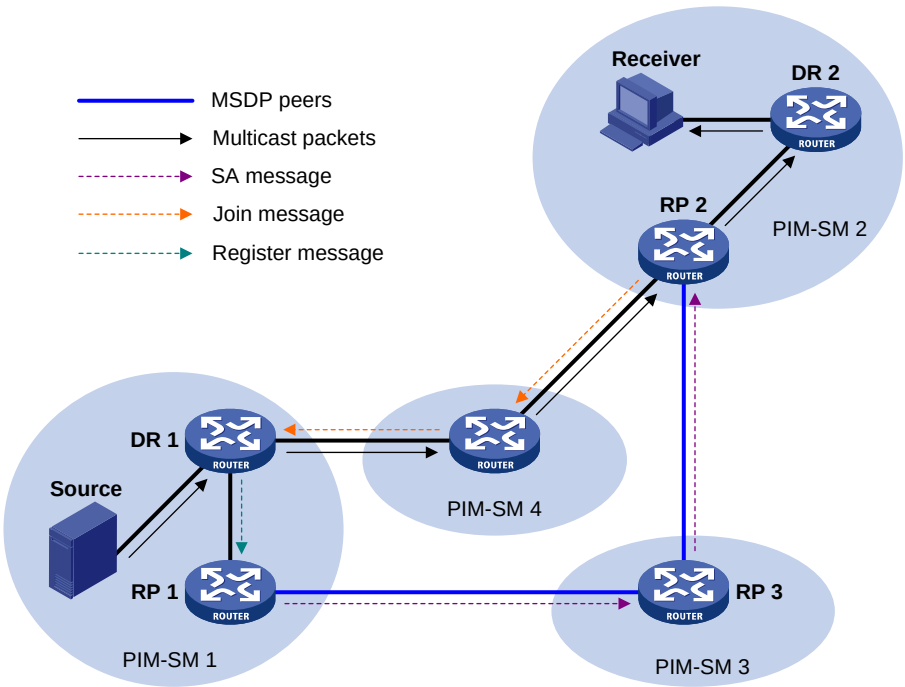
对于通过 BSR 机制动态选举 RP 的 PIM-SM 网络来说，RP 是由 C-RP 选举产生的。为了增强其网络的健壮性，一个 PIM-SM 域内往往存在不止一个 C-RP。由于无法预计 RP 选举的结果，为了保证选举获胜的 C-RP 能始终位于“MSDP 连通图”上，需要在所有的 C-RP 之间建立 MSDP 对等体关系。而选举落败的 C-RP 在“MSDP 连通图”上所担当的角色相当于普通的 PIM-SM 路由器。

---

2. MSDP实现域间组播

如 图 1-2 所示，PIM-SM 1 域内存在激活的组播源（Source），RP 1 通过组播源注册过程了解到了该组播源的存在。如果PIM-SM 2 和PIM-SM 3 域也希望知道该组播源的具体位置，进而能够从该组播源获取组播数据，则需要在RP 1 与RP 3、RP 2 与RP 3 之间分别建立MSDP对等体关系。

图1-2 MSDP 实现域间组播示意图



借助 MSDP 对等体进行 PIM-SM 域间组播的工作过程如下：

(1) 当 PIM-SM 1 域内的组播源向组播组 G 发送第一个组播数据包时，DR 1 将该组播数据封装在注册消息（Register Message）中，并发给 RP 1。RP 1 因此获知了该组播源的相关信息。

(2) RP 1 作为源端 RP，创建 SA 消息，并周期性地向其它 MSDP 对等体发送。SA 消息中包含组播源的地址 S、组播组的地址 G 以及创建该 SA 消息的源端 RP（即 RP 1）的地址。

(3) MSDP 对等体对收到的 SA 消息进行 RPF（Reverse Path Forwarding，逆向路径转发）检查，以及各种转发策略的过滤，从而只接受和转发来自正确路径并通过过滤的 SA 消息，以避免 SA 消息传递环路；另外，可以在 MSDP 对等体之间配置 MSDP 全连接组（Mesh Group），以避免 SA 消息在 MSDP 对等体之间的泛滥。

(4) SA 消息在 MSDP 对等体之间转发，最终该组播源的相关信息将传遍所有建立了 MSDP 对等体关系的 PIM-SM 域（即 PIM-SM 2 和 PIM-SM 3）。

(5) PIM-SM 2 中的 RP 2 在收到该 SA 消息后，检查本域内是否有组播组 G 的接收者（Receiver）存在：

- 如果有接收者，RP 2 与接收者之间维护组播组 G 的 RPT。RP 2 创建（S，G）表项，向源端的 DR 1 逐跳发送（S，G）加入消息（Join Message），从而跨越各 PIM-SM 域直接加入以该组播源为根的 SPT。组播数据沿 SPT 到达 RP 2 后，再沿 RPT 向接收者转发。当接收者端的 DR 2 收到来自组播源的组播数据后，可以自行决定是否发起从 RPT 向 SPT 的切换；
- 如果没有接收者，RP 2 不会创建（S，G）表项，也不加入以该组播源为根的 SPT。

说明

- MSDP 全连接组：要求所有组成员之间两两建立 MSDP 对等体关系，且所有组成员均使用相同的组名称。
- 在使用 MSDP 进行域间组播时，RP 在收到组播源的信息后就不再需要依赖其它 PIM-SM 域内的 RP，此时接收者可以跨越各 PIM-SM 域内的 RP，而直接加入基于组播源的 SPT。

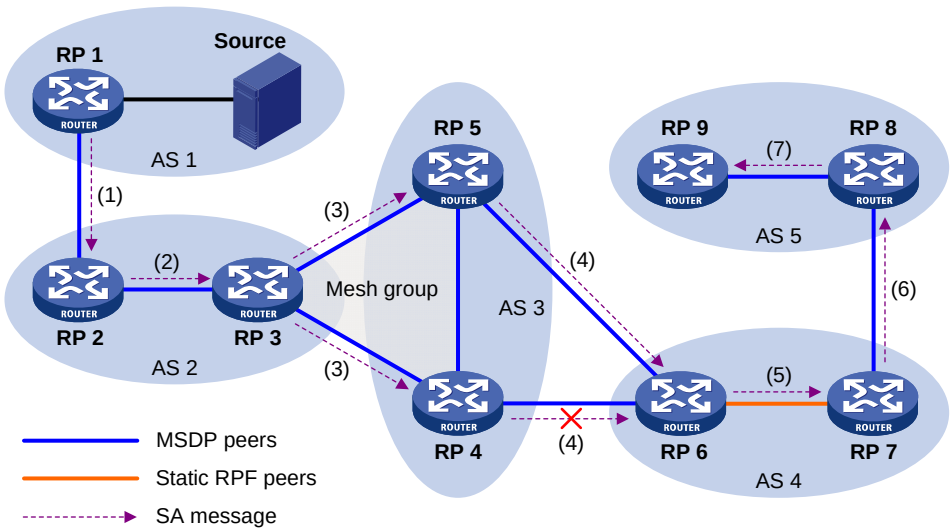
3. SA消息的RPF检查规则

如 图 1-3 所示，网络中有五个自治系统AS 1～AS 5，AS内部使用IGP互联，AS之间使用BGP或MBGP互联。每个AS中包含至少一个PIM-SM域，且每个PIM-SM域中包含至少一个RP。各RP之间建立起MSDP对等体关系，其中RP 3、RP 4 和RP 5 之间建立MSDP全连接组，并在RP 7 上将RP 6 配置为其静态RPF对等体。

说明

设备对于来自静态 RPF 对等体的 SA 消息不进行 RPF 检查，直接接受并向其它对等体转发。

图1-3 SA 消息的 RPF 检查规则



对照 图 1-3，这些MSDP对等体将按照如下RPF检查规则处理收到的SA消息：

(1) 当 RP 2 收到 RP 1 发来的 SA 消息时：由于 SA 消息中所携带的源端 RP 的地址与 MSDP 对等体的地址相同，说明发出 SA 消息的 MSDP 对等体就是创建该 SA 消息的 RP，于是 RP 2 接受该 SA 消息并向其它对等体（RP 3）转发。

(2) 当 RP 3 收到 RP 2 发来的 SA 消息时：由于 SA 消息来自同一个 AS 的 MSDP 对等体（RP 2），且该对等体是到源端 RP 最佳路径上的下一跳，于是 RP 3 接受该 SA 消息并向其它对等体（RP 4 和 RP 5）转发。

(3) 当 RP 4 和 RP 5 分别收到 RP 3 发来的 SA 消息时：由于 SA 消息来自同一个全连接组的 MSDP 对等体（RP 3），于是 RP 4 和 RP 5 均接受该 SA 消息并不再向本组其它成员转发，而只向本组之外的其它 MSDP 对等体（RP 6）转发。

(4) 当 RP 6 收到 RP 4 和 RP 5（假设 RP 5 的 IP 地址较大）发来的 SA 消息时：尽管同处 AS 3 的 RP 4 和 RP 5 都与 RP 6 建立了 MSDP 对等体关系，但 RP 6 只接受 IP 地址较高的 MSDP 对等体（RP 5）发来的 SA 消息。

(5) 当 RP 7 收到 RP 6 发来的 SA 消息时：由于 SA 消息来自其静态 RPF 对等体（RP 6），于是 RP 7 接受该 SA 消息并向其它对等体（RP 8）转发。

(6) 当 RP 8 收到 RP 7 发来的 SA 消息时：属于不同 AS 的 MSDP 对等体之间存在 BGP 或 MBGP 路由。由于 SA 消息来自不同 AS 的 MSDP 对等体（RP 7），且该对等体是到源端 RP 的 BGP 或 MBGP 路由的下一跳，于是 RP 8 接受该 SA 消息并向其它对等体（RP 9）转发。

(7) 当 RP 9 收到 RP 8 发来的 SA 消息时：由于只有一个 MSDP 对等体（RP 8），于是 RP 9 接受该 SA 消息。

对于由其它路径到来的 SA 消息，MSDP 对等体将不接受也不转发。

4. MSDP实现域内Anycast RP

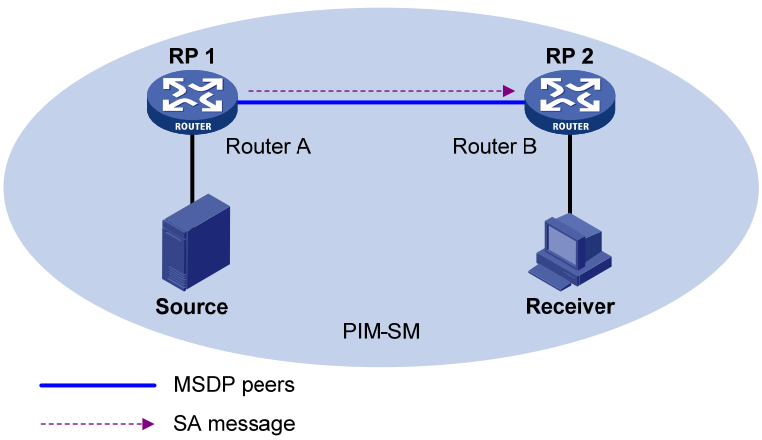
Anycast RP（任播 RP）是指在同一个 PIM-SM 域内设置两个或多个具有相同地址的 RP，并在这些 RP 之间建立 MSDP 对等体关系，以实现域内各 RP 之间的负载分担和冗余备份。

如 图 1-4 所示，在一个PIM-SM域内，组播源（Source）向组播组G发送组播数据，接收者（Receiver）是组播组G的成员。分别在Router A和Router B上配置相同的IP地址（称为Anycast RP地址，通常使用私有地址），同时将这些接口配置为C-RP，并在Router A和Router B之间建立MSDP对等体关系。

说明

通常在设备的逻辑接口（如 Loopback 接口）上配置 Anycast RP 地址。

图1-4 MSDP 实现域内 Anycast RP 示意图



Anycast RP 的工作过程如下：

(1) 组播源选择距离最近的 RP 进行注册。如：Source 向 RP 1 注册，注册消息中封装有 Source 发出的组播数据。当该注册消息到达 RP 1 后，进行解封装。

(2) 接收者向距离最近的 RP 发送加入消息，加入以该 RP 为根的 RPT。如：Receiver 加入以 RP 2 为根的 RPT。

(3) RP 之间通过发送 SA 消息，共享注册的组播源信息。如：RP 1 创建一个 SA 消息，发送给 RP 2，该 SA 消息中封装有 Source 发出的组播数据。当该 SA 消息到达 RP 2 后，进行解封装。

(4) 接收者沿 RPT 收到组播数据后，直接加入以该组播源为根的 SPT。如：RP 2 沿 RPT 将组播数据向下转发。当 Receiver 收到来自 Source 的组播数据后，直接加入以 Source 为根的 SPT。

Anycast RP 的意义如下：

- RP 路径最优：组播源向距离最近的 RP 进行注册，建立路径最优的 SPT；接收者向距离最近的 RP 发起加入，建立路径最优的 RPT。
- RP 间的负载分担：每个 RP 上只需维护 PIM-SM 域内的部分源/组信息、转发部分的组播数据，从而实现了 RP 间的负载分担。
- RP 间的冗余备份：当某 RP 失效后，原先在该 RP 上注册或加入的组播源或接收者会自动选择就近的 RP 进行注册或加入操作，从而实现了 RP 间的冗余备份。

注意

- 必须为 Anycast RP 地址配置 32 位的子网掩码（即 255.255.255.255），也即将其配置为一个主机地址。
- MSDP 对等体的地址不能与 Anycast RP 地址相同。

1.1.3 多实例的MSDP

属于同一实例的组播路由器各接口之间可以建立 MSDP 对等体。通过在 MSDP 对等体之间交互 SA 消息，可以实现跨域的 VPN 组播。

应用多实例的组播路由器，为其所支持的每一个实例都独立维护了一套 MSDP 机制，包括：SA 缓存、对等体连接、定时器、发送缓存和 PIM 交互的缓冲区。同时，保证不同实例之间信息隔离。所以，只有属于同一实例的 MSDP 和 PIM-SM 信息才可以交互。

1.1.4 协议规范

与 MSDP 相关的协议规范有：

- RFC 3618：Multicast Source Discovery Protocol (MSDP)
- RFC 3446：Anycast Rendevous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP)

1.2 MSDP配置任务简介

表1-1 MSDP 配置任务简介

配置任务		说明	详细配置
配置MSDP基本功能	使能MSDP	必选	1.3.2
	创建MSDP对等体连接	必选	1.3.3
	配置静态RPF对等体	可选	1.3.4
配置MSDP对等体连接	配置MSDP对等体描述信息	可选	1.4.2
	配置MSDP全连接组	可选	1.4.3
	配置MSDP对等体连接控制	可选	1.4.4
配置SA消息	配置SA消息内容	可选	1.5.2
	配置SA请求消息	可选	1.5.3
	配置SA消息过滤规则	可选	1.5.4
	配置SA消息缓存	可选	1.5.5

1.3 配置MSDP基本功能

说明

本节的所有配置都是在 PIM-SM 域内的 RP 上进行的，这些 RP 将成为 MSDP 对等体的一端。

1.3.1 配置准备

在配置 MSDP 基本功能之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 配置 PIM-SM，实现域内组播

在配置 MSDP 基本功能之前，需准备以下数据：

- MSDP 对等体的 IP 地址
- 表示 RP 地址过滤策略的地址前缀列表

1.3.2 使能MSDP

在配置 MSDP 各功能之前，必须先使能 MSDP。

1. 使能公网实例中的MSDP

表1-2 使能公网实例中的 MSDP

操作	命令	说明
进入系统视图	**system-view**	-
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
使能MSDP，并进入公网实例MSDP视图	**msdp**	必选 缺省情况下，MSDP处于关闭状态

2. 使能VPN实例中的MSDP

表1-3 使能 VPN 实例中的 MSDP

操作	命令	说明
进入系统视图	**system-view**	-
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态
退回系统视图	**quit**	-
使能MSDP，并进入VPN实例MSDP视图	**msdp vpn-instance** *vpn-instance-name*	必选 缺省情况下，MSDP处于关闭状态

说明

- 有关 **ip vpn-instance** 和 **route-distinguisher** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。
- 有关 **multicast routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

1.3.3 创建MSDP对等体连接

MSDP 对等体使用地址对来标识，即本端 MSDP 对等体地址和远端 MSDP 对等体地址。需要在互为对等体的两端都创建 MSDP 对等体连接。

表1-4 创建 MSDP 对等体连接

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
创建MSDP对等体连接	**peer** *peer-address* **connect-interface** *interface-type interface-number*	必选 缺省情况下，没有创建MSDP对等体连接

说明

如果某接口同时作为 MSDP 对等体和 BGP/MBGP 对等体中的一端，则建议为 MSDP 对等体配置与 BGP 或 MBGP 对等体相同的 IP 地址。

1.3.4 配置静态RPF对等体

通过配置静态 RPF 对等体可以免除对收到的 SA 消息进行 RPF 检查。

表1-5 配置静态 RPF 对等体

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
配置静态RPF对等体	**static-rpf-peer** *peer-address* [**rp-policy** *ip-prefix-name*]	必选 缺省情况下，没有配置静态RPF对等体

说明

如果在一台路由器上只配置了一个 MSDP 对等体，则该 MSDP 对等体将被当作静态 RPF 对等体。

1.4 配置MSDP对等体连接

1.4.1 配置准备

在配置 MSDP 对等体连接之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 配置 MSDP 基本功能

在配置 MSDP 对等体连接之前，需准备以下数据：

- MSDP 对等体的描述信息

- MSDP 全连接组的名称
- 建立 MSDP 对等体连接的重试周期
- MSDP 对等体建立 TCP 连接时进行 MD5 认证的密码

1.4.2 配置MSDP对等体描述信息

管理员可以通过MSDP对等体的描述信息方便地区分不同的MSDP对等体，从而更好地管理MSDP对等体。

表1-6 配置 MSDP 对等体描述信息

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
配置MSDP对等体的描述信息	**peer** *peer-address* **description** *text*	必选 缺省情况下，MSDP对等体没有描述信息

1.4.3 配置MSDP全连接组

一个自治系统内可能包含多个 MSDP 对等体，为了避免这些 MSDP 对等体之间泛滥 SA 消息，可以使用 MSDP 全连接组来优化数据流量。

构成全连接组的 MSDP 对等体，一方面将来自组外并通过了 RPF 检查的 SA 消息转发给组内的其它成员；另一方面，对来自组内成员的 SA 消息不经 RPF 检查就接受，也不在组内进行重复转发。这种操作既避免了 SA 消息的泛滥，同时还由于不需要在 MSDP 对等体之间运行 BGP 或 MBGP，所以也就简化了对等体 RPF 检查机制。

通过为多个 MSDP 对等体配置相同的全连接组名称，可以建立 MSDP 全连接组。

表1-7 配置 MSDP 全连接组

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
把MSDP对等体加入全连接组	**peer** *peer-address* **mesh-group** *name*	必选 缺省情况下，MSDP对等体不属于任何全连接组

说明

- 在配置 MSDP 全连接组之前，应使各路由器之间保持两两互连。
- 如果在同一 MSDP 对等体上多次配置加入全连接组，最后一个配置有效。

1.4.4 配置MSDP对等体连接控制

MSDP 对等体之间使用 TCP 进行连接（端口号为 639），用户可以手工关闭或重建 MSDP 对等体连接，灵活控制 MSDP 对等体之间的会话。当关闭了 MSDP 对等体连接后，MSDP 对等体之间不再传递 SA 消息，TCP 连接关闭，并不再重试建立连接，但配置信息会被保留。

当新创建了 MSDP 对等体、或重新启动了被关闭的 MSDP 对等体连接、或发生故障的 MSDP 对等体尝试恢复工作时，需要在 MSDP 对等体之间建立 TCP 连接。用户可以灵活地调整建立 MSDP 对等体连接的重试周期。

为了提高 MSDP 的安全性，可以配置 MSDP 对等体在建立 TCP 连接时进行 MD5 认证。该认证并不能对 MSDP 报文进行认证，它只是为 TCP 连接设置 MD5 认证密码，并由 TCP 完成认证。如果认证失败，则无法建立 TCP 连接。

表1-8 配置 MSDP 对等体连接控制

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
手工关闭MSDP对等体连接	**shutdown** *peer-address*	可选 缺省情况下，MSDP对等体处于连接状态
配置建立MSDP对等体连接的重试周期	**timer retry** *interval*	可选 缺省情况下，建立MSDP对等体连接的重试周期为30秒
配置MSDP对等体建立TCP连接时进行MD5认证	**peer** *peer-address* **password** { **cipher** | **simple** } *password*	可选 缺省情况下，MSDP对等体建立TCP连接时不进行MD5认证

注意

参与 MD5 认证的两端 MSDP 对等体必须配置相同的认证方式和密码，否则将由于不能通过认证而无法建立 TCP 连接。

1.5 配置SA消息

1.5.1 配置准备

在配置 SA 消息传递之前，需完成以下任务：

- 配置任一单播路由协议，实现域内网络层互通
- 配置 MSDP 基本功能

在配置 SA 消息传递之前，需准备以下数据：

- 表示 SA 请求消息过滤规则的 ACL 规则
- 表示 SA 消息创建规则的 ACL 规则

- 表示接收或转发 SA 消息的过滤规则的 ACL 规则
- 封装在 SA 消息中组播数据报文的 TTL（Time to Live，生存时间）阈值
- 可缓存从指定 MSDP 对等体学到的（S，G）表项的最大数量

1.5.2 配置SA消息内容

某些组播源发送组播数据的时间间隔较长，超出了（S，G）表项的超时时间。在这种情况下，源端 DR 只能将组播数据逐个封装在注册消息中，发送给源端 RP。源端 RP 使用 SA 消息将（S，G）信息传输给远端 RP。然后，远端 RP 向源端 DR 发起加入过程，并创建 SPT。由于（S，G）表项已超时，远端用户将永远无法收到该组播源发出的组播数据。

当在源端 RP 上使能了在 SA 消息中封装组播数据报文的功能后，源端 RP 会将组播数据报文封装在 SA 消息中发送出去。远端 RP 收到该 SA 消息后解封装，并将组播数据报文沿 RPT 传输给本域内的用户。

MSDP 对等体之间传递 SA 消息，当路由器对收到的 SA 消息进行 RPF 检查时，如果发现对端 RP 的地址与本地 RP 的地址相同，就会丢弃该 SA 消息。但在 Anycast RP 应用中，要求在同一个 PIM-SM 域内的两台或多台路由器上配置 IP 地址相同的 RP，并在这些路由器之间建立 MSDP 对等体关系，因此必须为 SA 消息指定一个与实际 RP 的地址不同的逻辑 RP 地址（即逻辑接口上的 RP 地址），以通过 RPF 检查。

表1-9 配置 SA 消息内容

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
使能在SA消息中封装组播数据报文	**encap-data-enable**	可选 缺省情况下，在SA消息中只包含（S，G）表项，不封装组播数据报文
配置接口地址为SA消息的RP地址	**originating-rp** *interface-type interface-number*	可选 缺省情况下，SA消息的RP地址为PIM的RP地址

1.5.3 配置SA请求消息

缺省情况下，当一个新接收者加入时，路由器不会主动向其 MSDP 对等体发送 SA 请求消息，而是等待其 MSDP 对等体在下一个周期发来的 SA 消息，这将延迟接收者获取组播源信息的时间。为了尽快让新接收者了解到当前活跃的组播源信息，需要主动向 MSDP 对等体发送 SA 请求消息。

表1-10 配置 SA 请求消息

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-

操作	命令	说明
使能发送SA请求消息	**peer** *peer-address* **request-sa-enable**	可选 缺省情况下，路由器收到新的组加入消息时，不向其MSDP对等体发送SA请求消息，而是等待下一周期SA消息的到来
配置SA请求消息的过滤规则	**peer** *peer-address* **sa-request-policy** [**acl** *acl-number*]	可选 缺省情况下，不对SA请求消息进行过滤

注意

在使能发送 SA 请求消息功能之前，必须首先关闭 SA 消息缓存机制，否则设备不会向外发送 SA 请求消息。

1.5.4 配置SA消息过滤规则

通过配置 SA 消息的创建规则，路由器可以在创建 SA 消息时，对其通告的（S，G）表项进行过滤，从而实现在创建 SA 消息时对组播源消息传播的控制。

通过配置接收或转发 SA 消息的过滤规则，路由器可以在接收或转发 SA 消息时，对其通告的（S，G）转发项进行过滤，从而实现在接收和转发 SA 消息时，对组播源消息传播的控制。

通过配置封装在 SA 消息中组播数据报文的 TTL 阈值，可以对组播数据报文在 SA 消息中的封装以及传输范围进行限制：

- 路由器在创建封装有组播数据报文的 SA 消息之前，先检查该组播数据报文 IP 头的 TTL 值：如果小于阈值，则不创建该 SA 消息；如果大于或等于阈值，则将组播数据报文封装在 SA 消息中并转发出去。
- 路由器在收到封装有组播数据报文的SA消息之后，先将该组播数据报文IP头的TTL值减1，再检查此时的 TTL 值：如果小于阈值，则不再向其指定的 MSDP 对等体转发；如果大于或等于阈值，则重新将组播数据报文封装在 SA 消息中并转发出去。

表1-11 配置 SA 消息过滤规则

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
配置SA消息的创建规则	**import-source** [**acl** *acl-number*]	必选 缺省情况下，在创建SA消息时，对其通告的（S，G）表项不作限制
配置接收或转发SA消息的过滤规则	**peer** *peer-address* **sa-policy** { **import** | **export** } [**acl** *acl-number*]	必选 缺省情况下，不对接收或转发的SA消息进行过滤

操作	命令	说明
配置封装在SA消息中组播数据报文的TTL阈值	**peer** *peer-address* **minimum-ttl** *ttl-value*	可选 缺省情况下，封装在SA消息中组播数据报文的TTL阈值为0

1.5.5 配置SA消息缓存

为了减少获取组播信息的延迟时间，可以在路由器上使能 SA 消息缓存机制，即在本地缓存 SA 消息中所包含的（S，G）表项。缓存的（S，G）表项越多，所占用的内存空间越大。

在使能了 SA 消息缓存机制后，当收到一个新的组加入消息（*，G）时，路由器首先查找 SA 缓存：

- 如果缓存中没有对应的（S，G），便等候其 MSDP 对等体在下一个周期发来的 SA 消息；
- 如果缓存中有对应的（S，G），则直接加入以 S 为根的 SPT。

为了有效防止路由器受到 DoS（Denial of Service，拒绝服务）攻击，可以在路由器上配置可缓存（S，G）表项的最大数量。

表1-12 配置 SA 消息缓存

操作	命令	说明
进入系统视图	**system-view**	-
进入公网实例或VPN实例MSDP视图	**msdp** [**vpn-instance** *vpn-instance-name*]	-
使能SA消息缓存机制	**cache-sa-enable**	可选 在缺省情况下，SA消息缓存机制处于使能状态，即在设备收到SA消息后缓存其中包含的（S，G）表项
配置可缓存从指定MSDP对等体学到的（S，G）表项的最大数量	**peer** *peer-address* **sa-cache-maximum** *sa-limit*	可选 缺省情况下，可缓存从任一MSDP对等体学到的（S，G）表项的最大数量为8192

1.6 MSDP显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 MSDP 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 MSDP 的统计信息。

表1-13 MSDP 显示和维护

操作	命令
查看MSDP对等体的简要信息	**display msdp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **brief** [**state** { **connect** | **down** | **listen** | **shutdown** | **up** }] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看MSDP对等体的详细状态信息	**display msdp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **peer-status** [*peer-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
查看SA缓存中的（S，G）表项信息	**display msdp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **sa-cache** [*group-address* | *source-address* | *as-number*] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看SA缓存中（S，G）表项的数量	**display msdp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **sa-count** [*as-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
重置与MSDP对等体的TCP连接	**reset msdp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **peer** [*peer-address*]
清除SA缓存中的（S，G）表项	**reset msdp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **sa-cache** [*group-address*]
清除MSDP对等体的统计信息	**reset msdp** [**all-instance** | **vpn-instance** *vpn-instance-name*] **statistics** [*peer-address*]

1.7 MSDP典型配置举例

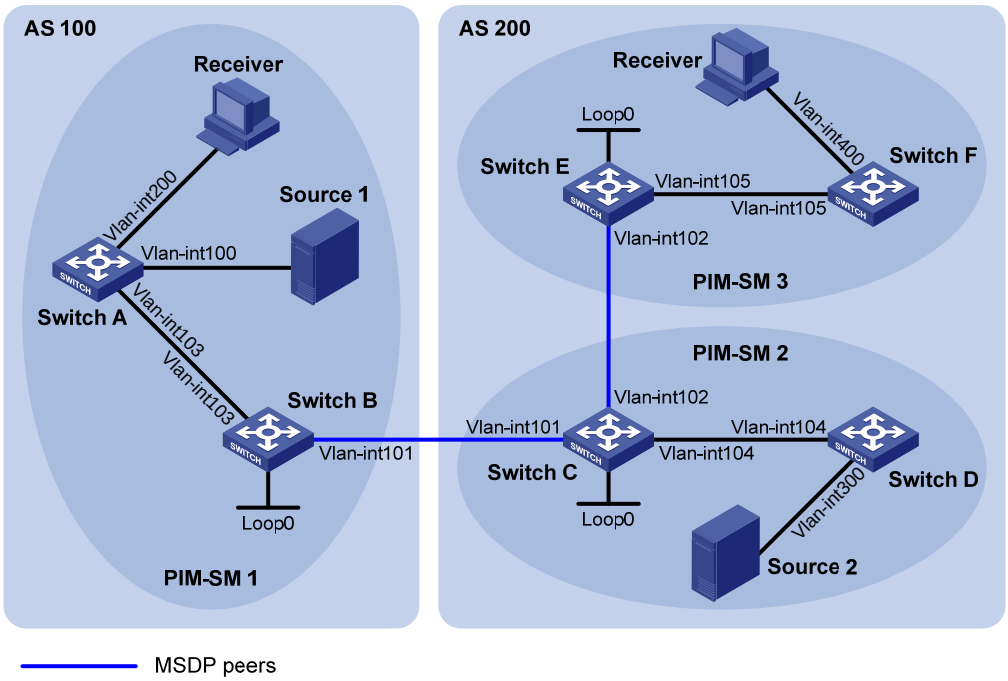
1.7.1 PIM-SM域间组播配置举例

1. 组网需求

- 网络中存在两个自治系统：AS 100 和 AS 200，各 AS 内部采用 OSPF 协议、AS 之间则采用 BGP 协议以保证单播路由的畅通；
- PIM-SM 1 属于 AS 100，PIM-SM 2 和 PIM-SM 3 属于 AS 200，每个 PIM-SM 域分别拥有至少一个组播源（Source）或接收者（Receiver）；
- 将 Switch B、Switch C 和 Switch E 各自的 Loopback0 接口分别配置为各自 PIM-SM 域的 C-BSR 和 C-RP；
- 通过在各 PIM-SM 域的 RP 之间建立 MSDP 对等体，从而实现各 PIM-SM 域之间组播源信息的共享。

2. 组网图

图1-5 PIM-SM 域间组播配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int103	10.110.1.2/24	Switch D	Vlan-int104	10.110.4.2/24
	Vlan-int100	10.110.2.1/24		Vlan-int300	10.110.5.1/24
	Vlan-int200	10.110.3.1/24	Switch E	Vlan-int105	10.110.6.1/24
Switch B	Vlan-int103	10.110.1.1/24		Vlan-int102	192.168.3.2/24
	Vlan-int101	192.168.1.1/24		Loop0	3.3.3.3/32
	Loop0	1.1.1.1/32	Switch F	Vlan-int105	10.110.6.2/24
Switch C	Vlan-int104	10.110.4.1/24		Vlan-int400	10.110.7.1/24
	Vlan-int102	192.168.3.1/24	Source 1	-	10.110.2.100/24
	Vlan-int101	192.168.1.2/24	Source 2	-	10.110.5.100/24
	Loop0	2.2.2.2/32			

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-5 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 AS 内的各交换机之间采用 OSPF 协议进行互连，确保各 AS 内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，使能 PIM-SM 和 IGMP，并配置 BSR 的服务边界

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在主机侧接口 Vlan-interface200 上使能 IGMP。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim sm
[SwitchA-Vlan-interface103] quit
[SwitchA] interface vlan-interface 100
```

```
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] igmp enable
[SwitchA-Vlan-interface200] pim sm
[SwitchA-Vlan-interface200] quit
```

Switch B、Switch C、Switch D、Switch E 和 Switch F 上的配置与 Switch A 相似，配置过程略。

在 Switch B 上配置 BSR 的服务边界。

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim bsr-boundary
[SwitchB-Vlan-interface101] quit
```

Switch C 和 Switch E 上的配置与 Switch B 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP 的位置

在 Switch B 上将 Loopback0 接口配置为 C-BSR 和 C-RP。

```
[SwitchB] pim
[SwitchB-pim] c-bsr loopback 0
[SwitchB-pim] c-rp loopback 0
[SwitchB-pim] quit
```

Switch C 和 Switch E 上的配置与 Switch B 相似，配置过程略。

(4) 配置 BGP 协议，将 BGP 与 OSPF 互相引入

在 Switch B 上配置 EBGP 对等体，并引入 OSPF 路由。

```
[SwitchB] bgp 100
[SwitchB-bgp] router-id 1.1.1.1
[SwitchB-bgp] peer 192.168.1.2 as-number 200
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] quit
```

在 Switch C 上配置 EBGP 对等体，并引入 OSPF 路由。

```
[SwitchC] bgp 200
[SwitchC-bgp] router-id 2.2.2.2
[SwitchC-bgp] peer 192.168.1.1 as-number 100
[SwitchC-bgp] import-route ospf 1
[SwitchC-bgp] quit
```

在 Switch B 的 OSPF 中引入 BGP。

```
[SwitchB] ospf 1
[SwitchB-ospf-1] import-route bgp
[SwitchB-ospf-1] quit
```

在 Switch C 的 OSPF 中引入 BGP。

```
[SwitchC] ospf 1
[SwitchC-ospf-1] import-route bgp
[SwitchC-ospf-1] quit
```

(5) 配置 MSDP 对等体

在 Switch B 上配置 MSDP 对等体。

```
[SwitchB] msdp
[SwitchB-msdp] peer 192.168.1.2 connect-interface vlan-interface 101
[SwitchB-msdp] quit
```

在 Switch C 上配置 MSDP 对等体。

```
[SwitchC] msdp
[SwitchC-msdp] peer 192.168.1.1 connect-interface vlan-interface 101
[SwitchC-msdp] peer 192.168.3.2 connect-interface vlan-interface 102
[SwitchC-msdp] quit
```

在 Switch E 上配置 MSDP 对等体。

```
[SwitchE] msdp
[SwitchE-msdp] peer 192.168.3.1 connect-interface vlan-interface 102
[SwitchE-msdp] quit
```

(6) 检验配置效果

通过使用 **display bgp peer** 命令可以查看交换机之间 BGP 对等体的关系。例如：

查看 Switch B 上 BGP 对等体关系的信息。

```
[SwitchB] display bgp peer

 BGP local router ID : 1.1.1.1
 Local AS number : 100
 Total number of peers : 1                 Peers in established state : 1

  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  192.168.1.2            200       24       21    0       6 00:13:09 Established
```

查看 Switch C 上 BGP 对等体关系的信息。

```
[SwitchC] display bgp peer

 BGP local router ID : 2.2.2.2
 Local AS number : 200
 Total number of peers : 1                 Peers in established state : 1

  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  192.168.1.1            100       18       16    0       1 00:12:04 Established
```

通过使用 **display bgp routing-table** 命令可以查看交换机上的 BGP 路由表。例如：

查看 Switch C 上 BGP 路由表的信息。

```
[SwitchC] display bgp routing-table

 Total Number of Routes: 5

 BGP Local router ID is 2.2.2.2
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >  1.1.1.1/32         192.168.1.1     0                     0       100?
* >i 2.2.2.2/32         0.0.0.0         0                     0       ?
```

```
* >  192.168.1.0        0.0.0.0         0                    0       ?
* >  192.168.1.1/32     0.0.0.0         0                    0       ?
* >  192.168.1.2/32     0.0.0.0         0                    0       ?
```

当 PIM-SM 1 和 PIM-SM 2 域内的组播源 Source 1 和 Source 2 发送组播信息时，PIM-SM 1 和 PIM-SM 3 域内的接收者能收到该组播信息。通过使用 **display msdp brief** 命令可以查看交换机之间 MSDP 对等体建立情况。例如：

查看 Switch B 上 MSDP 对等体建立情况的简要信息。

```
[SwitchB] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0

  Peer's Address     State     Up/Down time    AS     SA Count   Reset Count
  192.168.1.2        Up        00:12:27       200    13         0
```

查看 Switch C 上 MSDP 对等体建立情况的简要信息。

```
[SwitchC] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  2            2            0            0            0            0

  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  192.168.3.2       Up       00:15:32        200    8          0
  192.168.1.1       Up       00:06:39        100    13         0
```

查看 Switch E 上 MSDP 对等体建立情况的简要信息。

```
[SwitchE] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0

  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  192.168.3.1       Up       01:07:08        200    8          0
```

查看 Switch B 上 MSDP 对等体的详细信息。

```
[SwitchB] display msdp peer-status
MSDP Peer Information of VPN-Instance: public net
  MSDP Peer 192.168.1.2, AS 200
  Description:
  Information about connection status:
    State: Up
    Up/down time: 00:15:47
    Resets: 0
    Connection interface: Vlan-interface101 (192.168.1.1)
    Number of sent/received messages: 16/16
    Number of discarded output messages: 0
    Elapsed time since last connection or counters clear: 00:17:51
  Information about (Source, Group)-based SA filtering policy:
    Import policy: none
    Export policy: none
```

```
Information about SA-Requests:
  Policy to accept SA-Request messages: none
  Sending SA-Requests status: disable
Minimum TTL to forward SA with encapsulated data: 0
SAs learned from this peer: 0, SA-cache maximum for the peer: none
Input queue size: 0, Output queue size: 0
Counters for MSDP message:
  Count of RPF check failure: 0
  Incoming/outgoing SA messages: 0/0
  Incoming/outgoing SA requests: 0/0
  Incoming/outgoing SA responses: 0/0
  Incoming/outgoing data packets: 0/0
```

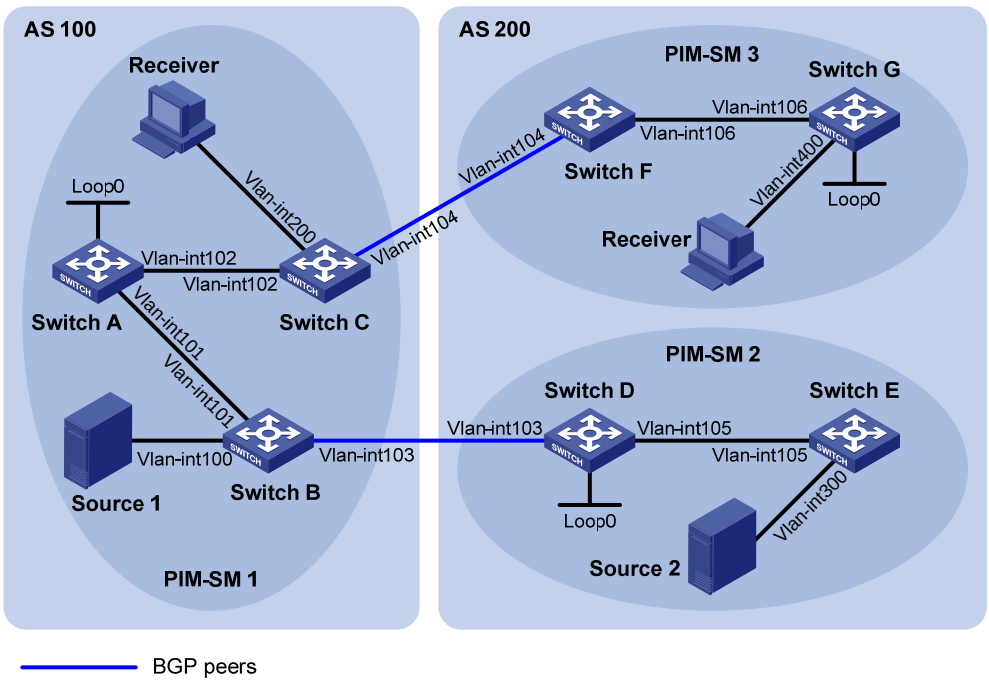
1.7.2 借助静态RPF对等体的AS间组播配置举例

1. 组网需求

- 网络中存在两个自治系统：AS 100 和 AS 200，各 AS 内部采用 OSPF 协议、AS 之间则采用 BGP 协议以保证单播路由的畅通；
- PIM-SM 1 属于 AS 100，PIM-SM 2 和 PIM-SM 3 属于 AS 200，每个 PIM-SM 域分别拥有至少一个组播源（Source）或接收者（Receiver）；
- 将 Switch A、Switch D 和 Switch G 各自的 Loopback0 接口分别配置为各自 PIM-SM 域的 C-BSR 和 C-RP；
- 根据 RPF 规则，设备接受来自其静态 RPF 对等体且被过滤策略所允许的 SA 消息。通过在各 PIM-SM 域的 RP 之间建立 MSDP 对等体，并在各 MSDP 对等体之间建立静态 RPF 对等体，从而在不改变单播拓扑的基础上实现各 PIM-SM 域之间组播源信息的共享。

2. 组网图

图1-6 借助静态 RPF 对等体的 AS 间组播配置组网图



设备	接口	IP地址	设备	接口	IP地址
Source 1	-	192.168.1.100/24	Switch D	Vlan-int105	10.110.5.1/24
Source 2	-	192.168.3.100/24		Vlan-int103	10.110.3.2/24
Switch A	Vlan-int101	10.110.1.1/24		Loop0	2.2.2.2/32
	Vlan-int102	10.110.2.1/24	Switch E	Vlan-int105	10.110.5.2/24
	Loop0	1.1.1.1/32		Vlan-int300	192.168.3.1/24
Switch B	Vlan-int101	10.110.1.2/24	Switch F	Vlan-int106	10.110.6.1/24
	Vlan-int100	192.168.1.1/24		Vlan-int104	10.110.4.2/24
	Vlan-int103	10.110.3.1/24	Switch G	Vlan-int106	10.110.6.2/24
Switch C	Vlan-int102	10.110.2.2/24		Vlan-int400	192.168.4.1/24
	Vlan-int200	192.168.2.1/24		Loop0	3.3.3.3/32
	Vlan-int104	10.110.4.1/24			

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-6 配置各接口的IP地址和掩码，具体配置过程略。

配置 AS 内的各交换机之间采用 OSPF 协议进行互连，确保 AS 内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，使能 PIM-SM 和 IGMP，并配置 BSR 的服务边界

在 Switch C 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在主机侧接口 Vlan-interface200 上使能 IGMP。

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] igmp enable
[SwitchC-Vlan-interface200] pim sm
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim sm
[SwitchC-Vlan-interface104] quit
```

Switch A、Switch B、Switch D、Switch E、Switch F 和 Switch G 上的配置与 Switch C 相似，配置过程略。

在 Switch B 上配置 BSR 的服务边界。

```
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim bsr-boundary
[SwitchB-Vlan-interface103] quit
```

Switch C、Switch D 和 Switch F 上的配置与 Switch B 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP 的位置

在 Switch A 上将 Loopback0 接口配置为 C-BSR 和 C-RP。

```
[SwitchA] pim
[SwitchA-pim] c-bsr loopback 0
[SwitchA-pim] c-rp loopback 0
[SwitchA-pim] quit
```

Switch D 和 Switch G 上的配置与 Switch A 相似，配置过程略。

(4) 配置 BGP 协议，将 BGP 与 OSPF 互相引入

在 Switch B 上配置 EBGP 对等体，并引入 OSPF 路由。

```
[SwitchB] bgp 100
[SwitchB-bgp] router-id 1.1.1.2
[SwitchB-bgp] peer 10.110.3.2 as-number 200
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] quit
```

在 Switch D 上配置 EBGP 对等体，并引入 OSPF 路由。

```
[SwitchD] bgp 200
[SwitchD-bgp] router-id 2.2.2.2
[SwitchD-bgp] peer 10.110.3.1 as-number 100
[SwitchD-bgp] import-route ospf 1
[SwitchD-bgp] quit
```

在 Switch C 上配置 EBGP 对等体，并引入 OSPF 路由。

```
[SwitchC] bgp 100
[SwitchC-bgp] router-id 1.1.1.3
[SwitchC-bgp] peer 10.110.4.2 as-number 200
[SwitchC-bgp] import-route ospf 1
[SwitchC-bgp] quit
```

在 Switch F 上配置 EBGP 对等体，并引入 OSPF 路由。

```
[SwitchF] bgp 200
[SwitchF-bgp] router-id 3.3.3.1
[SwitchF-bgp] peer 10.110.4.1 as-number 100
[SwitchF-bgp] import-route ospf 1
[SwitchF-bgp] quit
```

在 Switch B 的 OSPF 中引入 BGP。

```
[SwitchB] ospf 1
[SwitchB-ospf-1] import-route bgp
[SwitchB-ospf-1] quit
```

在 Switch D 的 OSPF 中引入 BGP。

```
[SwitchD] ospf 1
[SwitchD-ospf-1] import-route bgp
[SwitchD-ospf-1] quit
```

在 Switch C 的 OSPF 中引入 BGP。

```
[SwitchC] ospf 1
[SwitchC-ospf-1] import-route bgp
[SwitchC-ospf-1] quit
```

在 Switch F 的 OSPF 中引入 BGP。

```
[SwitchF] ospf 1
[SwitchF-ospf-1] import-route bgp
[SwitchF-ospf-1] quit
```

(5) 配置 MSDP 对等体及静态 RPF 对等体

配置 Switch D 和 Switch G 作为 Switch A 的 MSDP 对等体及静态 RPF 对等体。

```
[SwitchA] ip ip-prefix list-dg permit 10.110.0.0 16 greater-equal 16 less-equal 32
[SwitchA] msdp
```

```
[SwitchA-msdp] peer 10.110.3.2 connect-interface vlan-interface 101
[SwitchA-msdp] peer 10.110.6.2 connect-interface vlan-interface 102
[SwitchA-msdp] static-rpf-peer 10.110.3.2 rp-policy list-dg
[SwitchA-msdp] static-rpf-peer 10.110.6.2 rp-policy list-dg
[SwitchA-msdp] quit
```

配置 Switch A 作为 Switch D 的 MSDP 对等体及静态 RPF 对等体。

```
[SwitchD] ip ip-prefix list-a permit 10.110.0.0 16 greater-equal 16 less-equal 32
[SwitchD] msdp
[SwitchD-msdp] peer 10.110.1.1 connect-interface vlan-interface 103
[SwitchD-msdp] static-rpf-peer 10.110.1.1 rp-policy list-a
[SwitchD-msdp] quit
```

配置 Switch A 作为 Switch G 的 MSDP 对等体及静态 RPF 对等体。

```
[SwitchG] ip ip-prefix list-a permit 10.110.0.0 16 greater-equal 16 less-equal 32
[SwitchG] msdp
[SwitchG-msdp] peer 10.110.2.1 connect-interface vlan-interface 106
[SwitchG-msdp] static-rpf-peer 10.110.2.1 rp-policy list-a
[SwitchG-msdp] quit
```

(6) 检验配置效果

通过使用 **display bgp peer** 命令可以查看交换机之间 BGP 对等体建立情况。Switch A 上无任何信息输出，说明 Switch A 与 Switch D、Switch A 与 Switch G 之间均未建立 BGP 对等体关系。

当 PIM-SM 1 和 PIM-SM 2 域内的组播源 Source 1 和 Source 2 发送组播信息时，PIM-SM 1 和 PIM-SM 3 域内的接收者能收到该组播信息，通过使用 **display msdp brief** 命令可以查看交换机之间 MSDP 对等体的建立情况。例如：

查看 Switch A 上 MSDP 对等体的简要信息。

```
[SwitchA] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  2            2            0            0            0            0

  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  10.110.3.2        Up       01:07:08        ?      8          0
  10.110.6.2        Up       00:16:39        ?      13         0
```

查看 Switch D 上 MSDP 对等体的简要信息。

```
[SwitchD] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0

  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  10.110.1.1        Up       01:07:09        ?      8          0
```

查看 Switch G 上 MSDP 对等体的简要信息。

```
[SwitchG] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0
```

```
Peer's Address     State     Up/Down time     AS      SA Count     Reset Count
10.110.2.1         Up        00:16:40         ?       13           0
```

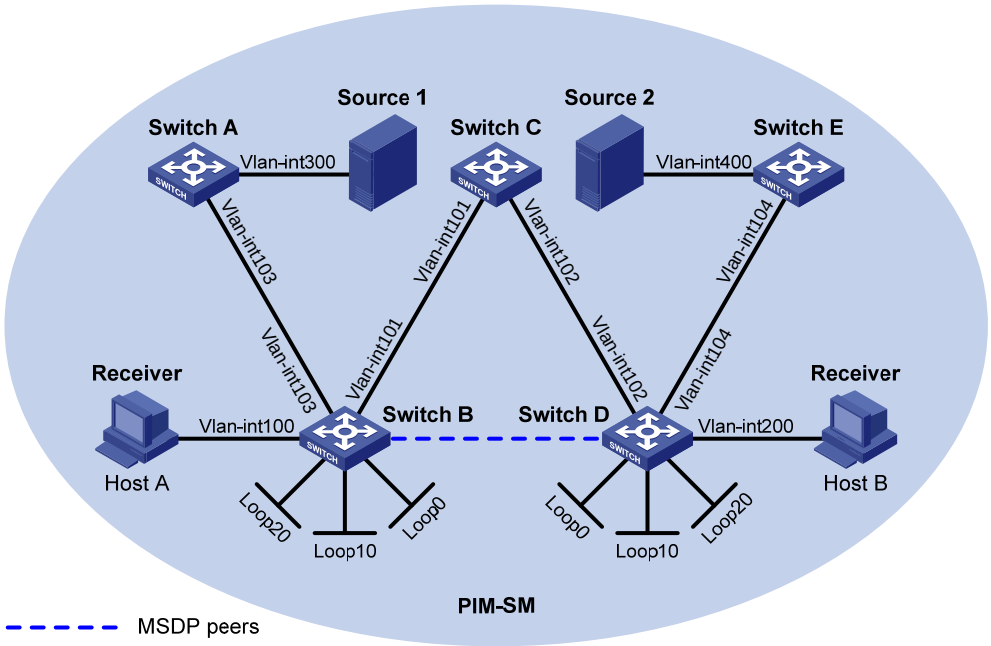
1.7.3 Anycast RP应用配置举例

1. 组网需求

- PIM-SM 域内拥有多个组播源（Source）和接收者（Receiver），并在域内运行 OSPF 协议以提供单播路由；
- 通过配置 Anycast RP，使接收者侧 DR 能够向拓扑距离最近的 RP 发起加入，组播源侧 DR 也向拓扑距离最近的 RP 发起注册；
- 将 Switch B 和 Switch D 各自的 Loopback10 接口配置为 C-BSR、Loopback20 接口配置为 C-RP；
- Switch B 的 Router ID 为 1.1.1.1，Switch D 的 Router ID 为 2.2.2.2，在 Switch B 和 Switch D 之间建立 MSDP 对等体关系。

2. 组网图

图1-7 Anycast RP 应用配置组网图



设备	接口	IP地址	设备	接口	IP地址
Source 1	-	10.110.5.100/24	Switch C	Vlan-int101	192.168.1.2/24
Source 2	-	10.110.6.100/24		Vlan-int102	192.168.2.2/24
Switch A	Vlan-int300	10.110.5.1/24	Switch D	Vlan-int200	10.110.3.1/24
	Vlan-int103	10.110.2.2/24		Vlan-int104	10.110.4.1/24
Switch B	Vlan-int100	10.110.1.1/24		Vlan-int102	192.168.2.1/24
	Vlan-int103	10.110.2.1/24		Loop0	2.2.2.2/32
	Vlan-int101	192.168.1.1/24		Loop10	4.4.4.4/32
	Loop0	1.1.1.1/32		Loop20	10.1.1.1/32
	Loop10	3.3.3.3/32	Switch E	Vlan-int400	10.110.6.1/24
	Loop20	10.1.1.1/32		Vlan-int104	10.110.4.2/24

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-7 配置各接口的IP地址和子网掩码，具体配置过程略。

配置 PIM-SM 域内的各交换机之间采用 OSPF 协议进行互连，确保 PIM-SM 域内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，并使能 PIM-SM 和 IGMP

在 Switch B 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在主机侧接口 Vlan-interface100 上使能 IGMP。

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] igmp enable
[SwitchB-Vlan-interface100] pim sm
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim sm
[SwitchB-Vlan-interface103] quit
[SwitchB] interface Vlan-interface 101
[SwitchB-Vlan-interface101] pim sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface loopback 0
[SwitchB-LoopBack0] pim sm
[SwitchB-LoopBack0] quit
[SwitchB] interface loopback 10
[SwitchB-LoopBack10] pim sm
[SwitchB-LoopBack10] quit
[SwitchB] interface loopback 20
[SwitchB-LoopBack20] pim sm
[SwitchB-LoopBack20] quit
```

Switch A、Switch C、Switch D 和 Switch E 上的配置与 Switch B 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP 的位置

在 Switch B 上将 Loopback10 配置为 C-BSR，将 Loopback20 配置为 C-RP。

```
[SwitchB] pim
[SwitchB-pim] c-bsr loopback 10
[SwitchB-pim] c-rp loopback 20
[SwitchB-pim] quit
```

Switch D 上的配置与 Switch B 相似，配置过程略。

(4) 配置 MSDP 对等体

在 Switch B 的 Loopback0 接口上配置 MSDP 对等体。

```
[SwitchB] msdp
[SwitchB-msdp] originating-rp loopback 0
[SwitchB-msdp] peer 2.2.2.2 connect-interface loopback 0
[SwitchB-msdp] quit
```

在 Switch D 的 Loopback0 接口上配置 MSDP 对等体。

```
[SwitchD] msdp
[SwitchD-msdp] originating-rp loopback 0
[SwitchD-msdp] peer 1.1.1.1 connect-interface loopback 0
[SwitchD-msdp] quit
```

(5) 检验配置效果

通过使用 **display msdp brief** 命令可以查看交换机之间 MSDP 对等体建立情况。

查看 Switch B 上 MSDP 对等体的简要信息。

```
[SwitchB] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0

  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  2.2.2.2           Up       00:10:17        ?      0          0
```

查看 Switch D 上 MSDP 对等体的简要信息。

```
[SwitchD] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0

  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  1.1.1.1           Up       00:10:18        ?      0          0
```

通过使用 **display pim routing-table** 命令可以查看交换机上的 PIM 路由。

当 Source 1（10.110.5.100/24）开始向组播组 G（225.1.1.1）发送组播信息时，Host A 加入组播组 G。通过比较 Switch B 与 Switch D 上 PIM 路由的显示信息，可知当前的有效 RP 为 Switch B：Source 1 向 Switch B 注册，Host A 向 Switch B 加入。

查看 Switch B 上的 PIM 路由的信息。

```
[SwitchB] display pim routing-table
 VPN-Instance: public net
 Total 1 (*, G) entry; 1 (S, G) entry

 (*, 225.1.1.1)
     RP: 10.1.1.1 (local)
     Protocol: pim-sm, Flag: WC
     UpTime: 00:15:04
     Upstream interface: Register
         Upstream neighbor: NULL
         RPF prime neighbor: NULL
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: igmp, UpTime: 00:15:04, Expires: -

 (10.110.5.100, 225.1.1.1)
     RP: 10.1.1.1 (local)
     Protocol: pim-sm, Flag: SPT 2MSDP ACT
```

```
    UpTime: 00:46:28
    Upstream interface: Vlan-interface103
        Upstream neighbor: 10.110.2.2
        RPF prime neighbor: 10.110.2.2
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface100
            Protocol: pim-sm, UpTime:  - , Expires:  -
```

查看 Switch D 上的 PIM 路由的信息。

```
[SwitchD] display pim routing-table
```

Switch D 上没有信息输出。

Host A 离开组播组 G，Source 1 也停止向组播组 G 发送组播数据。当 Source 2（10.110.6.100/24）开始向组播组 G 发送组播信息时，Host B 加入组播组 G。通过比较 Switch B 与 Switch D 上 PIM 路由的显示信息，可知当前的有效 RP 为 Switch D：Source 2 向 Switch D 注册，Host B 向 Switch D 加入。

查看 Switch B 上的 PIM 路由的信息。

```
[SwitchB] display pim routing-table
```

Switch B 上没有信息输出。

查看 Switch D 上的 PIM 路由的信息。

```
[SwitchD] display pim routing-table
 VPN-Instance: public net
 Total 1 (*, G) entry; 1 (S, G) entry

 (*, 225.1.1.1)
     RP: 10.1.1.1 (local)
     Protocol: pim-sm, Flag: WC
     UpTime: 00:12:07
     Upstream interface: Register
         Upstream neighbor: NULL
         RPF prime neighbor: NULL
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface200
             Protocol: igmp, UpTime: 00:12:07, Expires: -

 (10.110.6.100, 225.1.1.1)
     RP: 10.1.1.1 (local)
     Protocol: pim-sm, Flag: SPT 2MSDP ACT
     UpTime: 00:40:22
     Upstream interface: Vlan-interface104
         Upstream neighbor: 10.110.4.2
         RPF prime neighbor: 10.110.4.2
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface200
             Protocol: pim-sm, UpTime:  - , Expires:  -
```

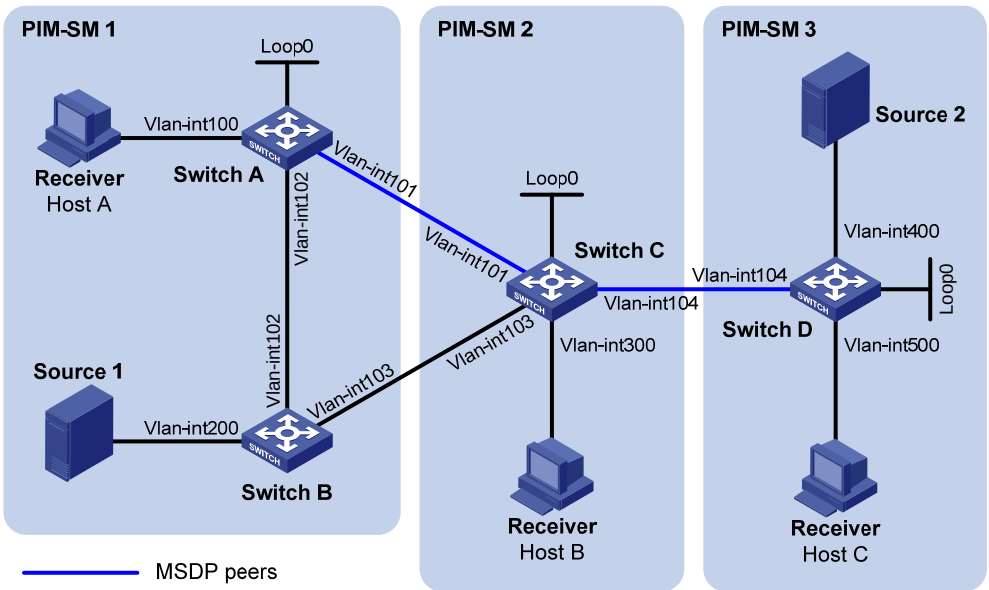
1.7.4 SA消息过滤机制配置举例

1. 组网需求

- 网络中存在三个 PIM-SM 域，各域内部以及域之间均运行 OSPF 协议以提供单播路由；
- 将 Switch A、Switch C 和 Switch D 各自的 Loopback0 接口分别配置为各自 PIM-SM 域的 C-BSR 和 C-RP；
- 分别在 Switch A 与 Switch C、Switch C 与 Switch D 之间建立 MSDP 对等体关系；
- 组播源 Source 1 向组播组 225.1.1.0/30 和 226.1.1.0/30 发送组播数据，组播源 Source 2 向组播组 227.1.1.0/30 发送组播数据；
- 通过配置 SA 消息过滤规则，使接收者 Host A 和 Host B 只能接收发往组播组 225.1.1.0/30 和 226.1.1.0/30 的组播数据，而 Host C 则只能接收发往组播组 226.1.1.0/30 和 227.1.1.0/30 的组播数据。

2. 组网图

图1-8 SA 消息过滤机制配置组网图



设备	接口	IP地址	设备	接口	IP地址
Source 1	-	10.110.3.100/24	Switch C	Vlan-int300	10.110.4.1/24
Source 2	-	10.110.6.100/24		Vlan-int104	10.110.5.1/24
Switch A	Vlan-int100	10.110.1.1/24		Vlan-int101	192.168.1.2/24
	Vlan-int102	10.110.2.1/24		Vlan-int103	192.168.2.2/24
	Vlan-int101	192.168.1.1/24		Loop0	2.2.2.2/32
	Loop0	1.1.1.1/32	Switch D	Vlan-int400	10.110.6.1/24
Switch B	Vlan-int200	10.110.3.1/24		Vlan-int500	10.110.7.1/24
	Vlan-int102	10.110.2.2/24		Vlan-int104	10.110.5.2/24
	Vlan-int103	192.168.2.1/24		Loop0	3.3.3.3/32

3. 配置步骤

(1) 配置 IP 地址和单播路由协议

请按照 图 1-8 配置各接口的IP地址和掩码，具体配置过程略。

配置各交换机之间采用 OSPF 协议进行互连，确保 PIM-SM 域内部以及各 PIM-SM 域之间在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IP 组播路由，使能 PIM-SM 和 IGMP，并配置 BSR 的服务边界

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在主机侧接口 Vlan-interface100 上使能 IGMP。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim sm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface loopback 0
[SwitchA-LoopBack0] pim sm
[SwitchA-LoopBack0] quit
```

Switch B、Switch C 和 Switch D 上的配置与 Switch A 相似，配置过程略。

在 Switch C 上配置 BSR 的服务边界。

```
[SwitchC] interface vlan-interface 101
[SwitchC-Vlan-interface101] pim bsr-boundary
[SwitchC-Vlan-interface101] quit
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] pim bsr-boundary
[SwitchC-Vlan-interface103] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim bsr-boundary
[SwitchC-Vlan-interface104] quit
```

Switch A、Switch B 和 Switch D 上的配置与 Switch C 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP 的位置

在 Switch A 上将 Loopback0 接口配置为 C-BSR 和 C-RP。

```
[SwitchA] pim
[SwitchA-pim] c-bsr loopback 0
[SwitchA-pim] c-rp loopback 0
[SwitchA-pim] quit
```

Switch C 和 Switch D 上的配置与 Switch A 相似，配置过程略。

(4) 配置 MSDP 对等体

在 Switch A 上配置 MSDP 对等体。

```
[SwitchA] msdp
[SwitchA-msdp] peer 192.168.1.2 connect-interface vlan-interface 101
[SwitchA-msdp] quit
```

在 Switch C 上配置 MSDP 对等体。

```
[SwitchC] msdp
[SwitchC-msdp] peer 192.168.1.1 connect-interface vlan-interface 101
[SwitchC-msdp] peer 10.110.5.2 connect-interface vlan-interface 104
[SwitchC-msdp] quit
```

在 Switch D 上配置 MSDP 对等体。

```
[SwitchD] msdp
[SwitchD-msdp] peer 10.110.5.1 connect-interface vlan-interface 104
[SwitchD-msdp] quit
```

(5) 配置 SA 消息过滤规则

在 Switch C 上配置不向 Switch D 转发有关（Source 1，225.1.1.0/30）的 SA 消息。

```
[SwitchC] acl number 3001
[SwitchC-acl-adv-3001] rule deny ip source 10.110.3.100 0 destination 225.1.1.0 0.0.0.3
[SwitchC-acl-adv-3001] rule permit ip source any destination any
[SwitchC-acl-adv-3001] quit
[SwitchC] msdp
[SwitchC-msdp] peer 10.110.5.2 sa-policy export acl 3001
[SwitchC-msdp] quit
```

在 Switch D 上配置不创建有关 Source 2 的 SA 消息。

```
[SwitchD] acl number 2001
[SwitchD-acl-basic-2001] rule deny source 10.110.6.100 0
[SwitchD-acl-basic-2001] quit
[SwitchD] msdp
[SwitchD-msdp] import-source acl 2001
[SwitchD-msdp] quit
```

(6) 检验配置效果

通过使用 **display msdp sa-cache** 命令可以查看交换机上 SA 缓存中的（S，G）表项信息。例如：

查看 Switch C 上 SA 缓存中的（S，G）表项信息。

```
[SwitchC] display msdp sa-cache
MSDP Source-Active Cache Information of VPN-Instance: public net
 MSDP Total Source-Active Cache - 8 entries
 MSDP matched 8 entries

(Source, Group)              Origin RP       Pro  AS     Uptime   Expires
(10.110.3.100, 225.1.1.0)    1.1.1.1         ?    ?      02:03:30 00:05:31
(10.110.3.100, 225.1.1.1)    1.1.1.1         ?    ?      02:03:30 00:05:31
(10.110.3.100, 225.1.1.2)    1.1.1.1         ?    ?      02:03:30 00:05:31
(10.110.3.100, 225.1.1.3)    1.1.1.1         ?    ?      02:03:30 00:05:31
(10.110.3.100, 226.1.1.0)    1.1.1.1         ?    ?      02:03:30 00:05:31
(10.110.3.100, 226.1.1.1)    1.1.1.1         ?    ?      02:03:30 00:05:31
(10.110.3.100, 226.1.1.2)    1.1.1.1         ?    ?      02:03:30 00:05:31
(10.110.3.100, 226.1.1.3)    1.1.1.1         ?    ?      02:03:30 00:05:31
```

查看 Switch D 上 SA 缓存中的（S，G）表项信息。

```
[SwitchD] display msdp sa-cache
MSDP Source-Active Cache Information of VPN-Instance: public net
 MSDP Total Source-Active Cache - 4 entries
 MSDP matched 4 entries
```

```
(Source, Group)              Origin RP       Pro  AS     Uptime   Expires
(10.110.3.100, 226.1.1.0)    1.1.1.1         ?    ?      00:32:53 00:05:07
(10.110.3.100, 226.1.1.1)    1.1.1.1         ?    ?      00:32:53 00:05:07
(10.110.3.100, 226.1.1.2)    1.1.1.1         ?    ?      00:32:53 00:05:07
(10.110.3.100, 226.1.1.3)    1.1.1.1         ?    ?      00:32:53 00:05:07
```

1.8 常见配置错误举例

1.8.1 MSDP对等体一直处于down状态

1. 故障现象

配置了 MSDP 对等体，但其状态一直显示为 down。

2. 分析

- 所配置的本地接口地址与 MSDP 对等体地址之间，建立起基于 TCP 连接的 MSDP 对等体关系；
- 如果本地接口地址与对端路由器上所配置的 MSDP 对等体地址不一致，TCP 连接就不会建立起来；
- 如果两个 MSDP 对等体之间没有路由，TCP 连接也不会建立起来。

3. 处理过程

(1) 检查各路由器之间的路由是否通达。通过命令 **display ip routing-table** 查看各路由器之间单播路由是否正确。

(2) 检查将成为 MSDP 对等体的两个路由器间是否存在到达对方的单播路由。

(3) 检查 MSDP 对等体之间的接口地址是否匹配。通过命令 **display current-configuration** 查看本地接口地址是否与对端 MSDP 对等体地址一致，应确保一致。

1.8.2 路由器SA缓存中没有SA表项

1. 故障现象

MSDP 没有将（S，G）转发项通过 SA 消息发送出去。

2. 分析

- 命令 **import-source** 用于控制将本域的（S，G）表项通过 SA 消息发送给 MSDP 对等体。如果没有指定参数 *acl-number* 则表示默认过滤掉所有的（S，G）表项，即不通告本域的所有（S，G）表项；
- 未配置 **import-source** 命令时，系统将发送本域的所有（S，G）信息。如果 MSDP 没有将本域的（S，G）表项通过 SA 消息发送出去，应检查 **import-source** 命令的配置是否正确。

3. 处理过程

(1) 检查各路由器之间的路由是否通达。通过命令 **display ip routing-table** 查看各路由器之间单播路由是否正确。

(2) 检查将成为 MSDP 对等体的两个路由器间是否存在到达对方的单播路由。

(3) 检查命令 **import-source** 及其参数 *acl-number* 的配置情况，确保 ACL 规则能够过滤合适的（S，G）信息。

1.8.3 Anycast RP应用中RP间互通信息异常

1. 故障现象

在 Anycast RP 的应用中，各 RP 之间没有互相交换其本地注册的（S，G）信息。

2. 分析

- 在 Anycast RP 应用中，通过将同一 PIM-SM 域内的 RP 配置为 MSDP 对等体，可以实现 RP 间的负载分担；
- MSDP 对等体地址不能与 Anycast RP 地址相同，且 C-BSR 和 C-RP 必须配置在不同的设备或接口上；
- 在使用 **originating-rp** 命令进行配置后，MSDP 将利用该命令所指定的接口地址替换 SA 消息中的 RP 地址；
- 当 MSDP 对等体对收到的 SA 消息进行 RPF 检查时，如果发现 RP 地址与本地配置的 RP 地址相同，将拒绝接收 SA 消息。

3. 处理过程

(1) 检查各路由器之间的路由是否通达。通过命令 **display ip routing-table** 查看各路由器之间单播路由是否正确。

(2) 检查将成为 MSDP 对等体的两个路由器间是否存在到达对方的单播路由。

(3) 检查 **originating-rp** 命令的配置情况。在 Anycast RP 的应用环境中，一定要配置 **originating-rp** 命令，而且 **originating-rp** 命令所指定的接口地址要与建立 MSDP 对等体连接的本地接口地址相同。

(4) 检查所配置的 C-BSR 地址是否与 Anycast RP 的地址不同，应确保两个地址不同。

目 录

1 MBGP配置

说明

在以下路由协议的介绍中所指的路由器及路由器图标，代表了一般意义下的路由器以及运行了路由协议的以太网交换机。

1.1 MBGP简介

为了提供对多种网络层协议的支持，IETF（Internet Engineering Task Force，互联网工程任务组）对 BGP-4 进行了扩展，形成 MP-BGP（Multiprotocol Border Gateway Protocol，多协议边界网关协议），使 BGP 能够为多种路由应用提供路由信息。

组播的网络拓扑和单播拓扑有可能不同，需要通过 MP-BGP 扩展使得 BGP 能够将单播 NLRI（Network Layer Reachability Information，网络层可达性信息）和组播 NLRI 分开运载，其中组播 NLRI 专用于执行 RPF（Reverse Path Forwarding，逆向路径转发）功能。这使得在单播路由表和组播路由表中，对于相同的前缀有不同的路径选择，维护了一致的单播转发并使域间组播正常工作。

目前的 MP-BGP 标准是 RFC 2858（Multiprotocol Extensions for BGP-4，BGP-4 的多协议扩展）。MP-BGP 在组播上的应用简称为 MBGP（Multicast BGP，组播 BGP）。

说明

- 本章主要介绍 MP-BGP 应用于组播的配置事项，即 MBGP 配置事项，有关 BGP 的详细内容以及配置事项，请参见“三层技术-IP 路由配置指导”中的“BGP”。
- 有关 RPF 检查的详细内容以及配置事项，请参见“IP 组播配置指导”中的“组播路由与转发”。

1.2 协议规范

与 MBGP 相关的协议规范有：

- RFC 2858：Multiprotocol Extensions for BGP-4
- RFC 3392：Capabilities Advertisement with BGP-4
- draft-ietf-idmr-bgp-mcast-attr-00：BGP Attributes for Multicast Tree Construction
- RFC4271，A Border Gateway Protocol 4 (BGP-4)
- RFC5291，Outbound Route Filtering Capability for BGP-4
- RFC5292，Address-Prefix-Based Outbound Route Filter for BGP-4

1.3 MBGP配置任务简介

表1-1 MBGP 配置任务简介

配置任务		说明	详细配置
配置MBGP基本功能		必选	1.4.2
控制路由信息的发布与接收	配置MBGP引入其他路由	必选	1.5.2
	配置MBGP引入缺省路由	可选	1.5.3
	配置MBGP路由聚合	可选	1.5.4
	配置向IPv4 MBGP对等体/IPv4 MBGP对等体组发送缺省路由	可选	1.5.5
	配置MBGP路由信息的发布策略	可选	1.5.6
	配置MBGP路由信息的接收策略	可选	1.5.7
	配置MBGP路由衰减	可选	1.5.8
配置MBGP的路由属性	配置MBGP路由管理优先级	可选	1.6.2
	配置本地优先级缺省值	可选	1.6.3
	配置MED属性	可选	1.6.4
	配置发布路由时将自身地址做为下一跳	可选	1.6.5
	配置AS_PATH属性	可选	1.6.6
调整和优化MBGP网络	配置MBGP软复位	可选	1.7.2
	配置BGP ORF能力	可选	1.7.3
	配置MBGP负载分担路由条数	可选	1.7.4
配置大型MBGP网络	配置IPv4 MBGP对等体组	可选	1.8.2
	配置MBGP团体	可选	1.8.3
	配置MBGP路由反射器	可选	1.8.4

1.4 配置MBGP基本功能

1.4.1 配置准备

在配置 MBGP 之前，需完成以下任务：相邻节点的网络层互通。

1.4.2 配置MBGP基本功能

表1-2 配置 MBGP 基本功能

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入BGP视图	**bgp** *as-number*	-
指定对等体/对等体组的AS号	**peer** { *group-name* | *ip-address* } **as-number** *as-number*	必选 缺省情况下，对等体/对等体组无AS号
创建并进入MBGP地址族视图	**ipv4-family multicast**	必选
激活在IPv4单播视图下创建的对等体/对等体组	**peer** { *group-name* | *ip-address* } **enable**	必选 缺省情况下，对等体/对等体组在BGP IPv4组播地址族视图下处于非激活状态
为从IPv4 MBGP对等体/IPv4 MBGP对等体组接收的路由分配首选值	**peer** { *group-name* | *ip-address* } **preferred-value** *value*	可选 缺省情况下，从IPv4 MBGP对等体/IPv4 MBGP对等体组接收的路由的首选值为0

1.5 控制路由信息的发布与接收

1.5.1 配置准备

在控制 MBGP 路由信息的发布与接收之前，需完成以下任务：配置 MBGP 基本功能。

1.5.2 配置MBGP引入其他路由

MBGP 可以向邻居 AS 发送本地 AS 内部网络的路由信息，但 MBGP 不是自己去发现 AS 内部的路由信息，而是引入 IGP 的路由信息到 MBGP 路由表中，并发布给 MBGP 对等体。在引入 IGP 路由时，还可以针对不同的路由协议来对路由信息进行过滤。

表1-3 配置 MBGP 引入其他路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
引入其它协议路由信息并通告	**import-route** *protocol* [{ *process-id* | **all-processes** } [**allow-direct** | **med** *med-value* | **route-policy** *route-policy-name*] *]	二者必选其一 缺省情况下，MBGP不引入、不通告其它协议的路由，不发布任何网段路由
将网段路由发布到MBGP路由表中	**network** *ip-address* [*mask* | *mask-length*] [**short-cut** | **route-policy** *route-policy-name*]	目前只有OSPF协议支持配置**allow-direct**参数

说明

- 通过 **import-route** 命令引入到 MBGP 路由表中的路由的 ORIGIN 属性为 Incomplete。
- 使用 **network** 命令发布到 MBGP 路由表中的网段路由的 ORIGIN 属性为 IGP。
- 要发布的网段路由必须存在于本地的 IP 路由表中，使用路由策略可以更为灵活的控制所发布的路由。

1.5.3 配置MBGP引入缺省路由

MBGP 不能通过 **import-route** 命令从其它协议引入缺省路由，如果要引入其它协议的缺省路由，必须要配置下列命令。

表1-4 配置 MBGP 引入其它协议的缺省路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
引入其它协议路由信息并通告	**import-route** *protocol* [{ *process-id* | **all-processes** } [**allow-direct** | **med** *med-value* | **route-policy** *route-policy-name*] *]	必选 缺省情况下，MBGP不引入其它协议的路由 目前只有OSPF协议支持配置**allow-direct**参数
允许将其它协议的缺省路由引入到MBGP路由表中	**default-route imported**	必选 缺省情况下，MBGP不引入其它协议的缺省路由

1.5.4 配置MBGP路由聚合

在中型或大型 MBGP 网络中，在向组播对等体发布路由信息时，需要配置路由聚合，减小对等体路由表中的路由数量。MBGP 支持以下两种聚合方式：

自动聚合：对 MBGP 引入的 IGP 子网路由进行聚合。配置自动聚合后，MBGP 将不再发布从 IGP 引入的子网路由，而是发布聚合后的自然网段的路由。缺省路由和用 **network** 命令引入的路由不能进行自动聚合。

手动聚合：对 MBGP 本地路由进行聚合。手动聚合的优先级高于自动聚合的优先级。

表1-5 配置 MBGP 路由聚合

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-

操作		命令	说明
进入MBGP地址族视图		**ipv4-family multicast**	-
配置MBGP路由聚合	配置对引入的子网路由进行自动聚合	**summary automatic**	二者必选其一 缺省情况下，不进行路由聚合 可以根据需求选择路由聚合方式；当二者同时配置时，手动路由聚合生效
	配置手动路由聚合	**aggregate** *ip-address* { *mask* | *mask-length* } [**as-set** | **attribute-policy** *route-policy-name* | **detail-suppressed** | **origin-policy** *route-policy-name* | **suppress-policy** *route-policy-name*] *	

1.5.5 配置向IPv4 MBGP对等体/IPv4 MBGP对等体组发送缺省路由

表1-6 配置向 IPv4 MBGP/IPv4 MBGP 组发送缺省路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
向IPv4 MBGP对等体/IPv4 MBGP对等体组发送缺省路由	**peer** { *group-name* | *ip-address* } **default-route-advertise** [**route-policy** *route-policy-name*]	必选 缺省情况下，不向IPv4 MBGP对等体/IPv4 MBGP对等体组发送缺省路由

说明

执行 **peer default-route-advertise** 命令后，不论本地路由表中是否存在缺省路由，都将向指定 IPv4 MBGP 对等体/IPv4 MBGP 对等体组发布一条下一跳地址为本地地址的缺省路由。

1.5.6 配置MBGP路由信息的发布策略

用户可以根据需求选择过滤策略，同时配置几种过滤策略时，按照如下顺序执行：

- **filter-policy export**
- **peer filter-policy export**
- **peer as-path-acl export**
- **peer ip-prefix export**
- **peer route-policy export**

只有通过前面的过滤策略，才能继续执行后面的过滤策略；只有通过所有配置的过滤策略后，路由信息才能被发布。

表1-7 配置 MBGP 路由信息的发布策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
配置对发布的路由信息进行过滤	**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* } **export** [**direct** | **isis** *process-id* | **ospf** *process-id* | **rip** *process-id* | **static**]	至少选其一 缺省情况下，发布路由信息时不对路由进行过滤
对发布给IPv4 MBGP对等体/IPv4 MBGP对等体组的路由指定路由策略	**peer** { *group-name* | *peer-address* } **route-policy** *route-policy-name* **export**	
为IPv4 MBGP对等体/IPv4 MBGP对等体组设置基于ACL的过滤策略	**peer** { *group-name* | *ip-address* } **filter-policy** *acl-number* **export**	
为IPv4 MBGP对等体/IPv4 MBGP对等体组设置基于AS路径过滤列表的MBGP路由过滤策略	**peer** { *group-name* | *ip-address* } **as-path-acl** *as-path-acl-number* **export**	
为IPv4 MBGP对等体/IPv4 MBGP对等体组设置基于IP前缀列表的路由过滤策略	**peer** { *group-name* | *ip-address* } **ip-prefix** *ip-prefix-name* **export**	

1.5.7 配置MBGP路由信息的接收策略

通过配置 MBGP 路由信息的接收策略，可以对从 IPv4 MBGP 对等体/IPv4 MBGP 对等体组接收的路由进行过滤，只有满足指定条件的路由才能被接收并加到 MBGP 路由表中。

用户可以根据需求选择过滤策略，同时配置几种过滤策略时，按照如下顺序执行：

- **filter-policy import**
- **peer filter-policy import**
- **peer as-path-acl import**
- **peer ip-prefix import**
- **peer route-policy import**

只有通过前面的过滤策略，才能继续执行后面的过滤策略；只有通过所有配置的过滤策略后，路由信息才能被接收。

表1-8 配置 MBGP 路由信息的接收策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** as-number	-
进入MBGP地址族视图	**ipv4-family multicast**	-
对接收的路由信息进行过滤	**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* } **import**	至少选其一

操作	命令	说明
对来自IPv4 MBGP对等体/IPv4 MBGP对等体组的路由指定路由策略	**peer** { *group-name* | *ip-address* } **route-policy** *policy-name* **import**	缺省情况下，不对接收的路由信息进行过滤
为IPv4 MBGP对等体/IPv4 MBGP对等体组设置基于ACL的过滤策略	**peer** { *group-name* | *ip-address* } **filter-policy** *acl-number* **import**	
为IPv4 MBGP对等体/IPv4 MBGP对等体组设置基于AS路径过滤列表的MBGP路由过滤策略	**peer** { *group-name* | *ip-address* } **as-path-acl** *as-path-acl-number* **import**	
为IPv4 MBGP对等体/IPv4 MBGP对等体组设置基于IP前缀列表的路由过滤策略	**peer** { *group-name* | *ip-address* } **ip-prefix** *ip-prefix-name* **import**	
配置允许从IPv4 MBGP对等体/IPv4 MBGP对等体组接收的最大路由数	**peer** { *group-name* | *ip-address* } **route-limit** limit [*percentage*]	可选 缺省情况下，允许从IPv4 MBGP对等体/IPv4 MBGP对等体组接收的最大路由数无限制

注意

对等体组的成员可以与所在的组使用不同的入方向路由策略，即接收路由时，各对等体可以选择自己的策略。

1.5.8 配置MBGP路由衰减

通过配置 MBGP 衰减，可以抑制不稳定的路由信息，不将这类路由加入到 MBGP 路由表中，也不将这类路由向其他 IPv4 MBGP 对等体发布。

表1-9 配置 MBGP 路由衰减

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
配置MBGP路由衰减参数	**dampening** [*half-life-reachable half-life-unreachable reuse suppress ceiling* | **route-policy** *route-policy-name*] *	必选 缺省情况下，没有配置MBGP路由衰减

1.6 配置MBGP的路由属性

MBGP 具有很多路由属性，利用这些属性可以改变 MBGP 的选路策略。

1.6.1 配置准备

在配置 MBGP 的路由属性之前，需完成以下任务：配置 MBGP 基本功能。

1.6.2 配置MBGP路由管理优先级

应用路由策略，可以为匹配过滤条件的特定路由配置优先级。对于那些没有匹配的路由，使用缺省优先级。

表1-10 配置 MBGP 路由管理优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
配置MBGP路由的管理优先级	**preference** { *external-preference internal-preference local-preference* | **route-policy** *route-policy-name* }	可选 缺省情况下，MBGP EBGP路由的管理优先级为255，MBGP IBGP路由的管理优先级为255，本地产生的MBGP路由的管理优先级为130

1.6.3 配置本地优先级缺省值

表1-11 配置本地优先级缺省值

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
配置本地优先级的缺省值	**default local-preference** *value*	可选 缺省情况下，本地优先级的缺省值为100

1.6.4 配置MED属性

在其它条件相同的情况下，MED 较小的路由被优选作为自治系统的外部路由。

表1-12 配置 MED 属性

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-

操作		命令	说明
配置MED属性	配置系统MED的缺省值	**default med** *med-value*	可选 缺省情况下，MED的缺省值为0
	配置允许比较来自不同AS邻居的路由路径的MED属性值	**compare-different-as-med**	可选 缺省情况下，不允许比较来自不同AS邻居的路由路径的MED属性值
	配置根据路由来自的AS进行分组对MED排序优选	**bestroute compare-med**	可选 缺省情况下，不根据路由来自的AS进行分组对MED排序优选
	配置允许比较联盟对等体的路由按MED值进行优选	**bestroute med-confederation**	可选 缺省情况下，比较联盟对等体的路由时不考虑MED值

1.6.5 配置发布路由时将自身地址作为下一跳

在某些组网环境中，本地路由器向 MBGP IBGP 对等体/MBGP IBGP 对等体组发布路由时，为保证 MBGP IBGP 邻居能够找到正确的下一跳，可以配置 **peer next-hop-local** 命令将自身地址作为下一跳地址。如果配置了负载分担，则不论是否配置了 **peer next-hop-local** 命令，本地路由器向 MBGP IBGP 对等体/MBGP IBGP 对等体组发布路由时都先将下一跳地址改变为自身地址。

在两个 MBGP 连接在同一网段的广播网络中，缺省情况下，向 MBGP EBGP 对等体/MBGP EBGP 对等体组发布路由时，不将自身地址作为下一跳；只有配置了 **peer next-hop-local** 命令，才将自身地址作为下一跳。

表1-13 配置发布路由时将自身地址作为下一跳

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-
配置发布路由时将自身地址作为下一跳	**peer** { *group-name* | *ip-address* } **next-hop-local**	可选 缺省情况下，向MBGP EBGP对等体/对等体组发布路由时，将自身地址作为下一跳；向MBGP IBGP对等体/对等体组发布路由时，不将自身地址作为下一跳

1.6.6 配置AS_PATH属性

通常情况下，MBGP 会检查对等体发来的路由的 AS_PATH 属性，如果其中已存在本地 AS 号，则 MBGP 会忽略此路由，以免形成路由环路。

表1-14 配置 AS_PATH 属性

操作		命令	说明
进入系统视图		**system-view**	-
进入BGP视图		**bgp** *as-number*	-
进入MBGP地址族视图		**ipv4-family multicast**	-
配置AS_PATH属性	配置允许本地AS号重复出现的次数	**peer** { *group-name* | *ip-address* } **allow-as-loop** [*number*]	可选 缺省情况下，不允许本地AS重复
	禁止路由器将AS_PATH当作选路算法中的一个因素	**bestroute as-path-neglect**	可选 缺省情况下，路由器可以将AS_PATH当作选路算法中的一个因素
	配置发送MBGP更新报文时AS_PATH属性中不携带私有AS号	**peer** { *group-name* | *ip-address* } **public-as-only**	可选 缺省情况下，发送MBGP更新报文时，允许携带私有自治系统号

1.7 调整和优化MBGP网络

1.7.1 配置准备

在调整和优化 MBGP 网络之前，需完成以下任务：配置 MBGP 基本功能。

1.7.2 配置MBGP软复位

MBGP 的选路策略改变后，为了使新的策略生效，必须复位 MBGP 连接，但这样会造成短暂的 MBGP 连接中断。

通过使能 Route-Refresh 功能，当策略改变后，系统可以在不中断 MBGP 连接的情况下，自动对 MBGP 路由表进行动态刷新。

如果对等体不支持 Route-Refresh 功能，则可以将从对等体接收的所有路由更新保存在本地，当选路策略发生改变后，在不中断连接的情况下重新刷新 MBGP 路由表，并应用新的策略。

1. 通过Route-Refresh实现MBGP软复位

在对等体支持并使能 Route-Refresh 功能的情况下，如果 MBGP 的路由策略发生了变化，本地路由器会向 MBGP 对等体发布 Router-Refresh 消息，收到此消息的对等体会将其路由信息重新发给本地路由器。这样，在不中断 MBGP 连接的情况下，就可以对 MBGP 路由表进行动态更新，并应用新的选路策略。

表1-15 通过 Route-Refresh 实现 MBGP 软复位

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-

操作	命令	说明
使能BGP路由刷新功能	**peer** { *group-name* | *ip-address* } **capability-advertise route-refresh**	可选 缺省情况下，路由刷新功能处于使能状态

2. 通过将所有路由更新保存在本地实现MBGP软复位

当对等体不支持 Route-Refresh 功能时，可通过配置 **peer keep-all-routes** 命令实现软复位功能。用户也可以通过执行 **refresh bgp ipv4 multicast** 命令对保存在本地的所有路由重新过一遍策略。

表1-16 通过将所有路由更新保存在本地实现 MBGP 软复位

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
禁止BGP路由刷新和多协议扩展功能	**peer** { *group-name* | *ip-address* } **capability-advertise conventional**	可选 缺省情况下，BGP路由刷新和多协议扩展功能处于使能状态
进入MBGP地址族视图	**ipv4-family multicast**	-
保存所有来自IPv4 MBGP对等体/IPv4 MBGP对等体组的原始路由信息，即使这些路由没有通过已配置的入口策略	**peer** { *group-name* | *ip-address* } **keep-all-routes**	必选 缺省情况下，不保存IPv4 MBGP对等体/IPv4 MBGP对等体组的原始路由信息
退回用户视图	**return**	-
手工对MBGP连接进行软复位	**refresh bgp ipv4 multicast** { **all** | *ip-address* | **group** *group-name* | **external** | **internal** } { **export** | **import** }	可选

1.7.3 使能MBGP ORF能力功能

BGP ORF 特性是将本地入口策略通过 Route-refresh 报文发送给邻居，当邻居需要向 BGP 对等体发送 Update 更新报文时，通过本地的路由策略后还需要进行 ORF 策略的过滤，只有通过 ORF 策略的路由信息才会发给 BGP 对等体，以达到减少 BGP 邻居间 Update 更新报文的交互，节省网络资源的目的。

使能 BGP ORF 能力后，本地和 BGP 对等体会通过 Open 报文协商 ORF 能力（即收发的报文里是否允许携带 ORF 信息，如果允许携带，是否可以携带非标准的 ORF 信息），当协商完毕并成功建立邻居关系后，可以通过特殊的 Route-refresh 报文交互 ORF 信息。

ORF能力协商成功需要两端的配置来保证，关于两端参数的选择请参见 表 1-18。

表1-17 配置 BGP ORF 能力

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
使能BGP路由刷新功能	**peer** { *group-name* | *ip-address* } **capability-advertise route-refresh**	可选 缺省情况下，BGP路由刷新功能处于使能状态 如果该功能当前处于未使能状态，则必须配置该命令 本命令的详细介绍请参见“三层技术-IP路由命令参考”中的“BGP”
使能BGP ORF非标准功能	**peer** { *group-name* | *ipv6-address* } **capability-advertise orf non-standard**	可选 缺省情况下，BGP ORF能力支持RFC5291和RFC5292的标准能力 如果该功能当前处于未使能状态，则必须配置该命令 本命令的详细介绍请参见”IP路由命令参考”中的“BGP”
进入MBGP地址族视图	**ipv4-family multicast**	-
使能BGP ORF地址前缀能力协商功能	**peer** { *group-name* | *ip-address* } **capability-advertise orf ip-prefix** { **both** | **receive** | **send** }	必选 缺省情况下，BGP 不支持ORF 地址前缀的能力协商

表1-18 both、send、receive 参数选择以及配置效果描述表

本地选择参数	对端选择参数	协商成功后
send	**receive**	本端的ORF发送能力，对端的ORF接收能力
	both	
receive	**send**	本端的ORF接收能力，对端的ORF发送能力
	both	
both	**both**	双向的ORF发送和接收能力

1.7.4 配置MBGP负载分担路由条数

表1-19 配置 MBGP 负载分担路由条数

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-

操作	命令	说明
配置进行MBGP负载分担的最大路由条数	**balance** *number*	必选 没有进行MBGP负载分担

1.8 配置大型MBGP网络

1.8.1 配置准备

在配置大型 MBGP 网络之前，需完成以下任务：相邻节点的网络层互通。

1.8.2 配置IPv4 MBGP对等体组

在大型网络中，对等体的数目众多，配置和维护极为不便。使用对等体组可以降低管理的难度，还可以提高路由发布效率。可以在 MBGP 中使能 IPv4 单播对等体组，从而达到共同管理组内 IPv4 MBGP 对等体的目的。MBGP 中不能单独创建对等体组。

表1-20 配置 IPv4 MBGP 对等体组

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	-
创建BGP对等体组	**group** *group-name* [**external** | **internal**]	必选 缺省情况下，没有创建BGP对等体组
将对等体加入已存在的对等体组	**peer** *ip-address* **group** *group-name* [**as-number** *as-number*]	必选 缺省情况下，对等体不属于任何对等体组
进入MBGP地址族视图	**ipv4-family multicast**	-
使能IPv4单播对等体组	**peer** *group-name* **enable**	必选
将IPv4 MBGP对等体加入对等体组中	**peer** *ip-address* **group** *group-name*	必选 缺省情况下，IPv4 MBGP对等体不属于任何对等体组

注意

- 配置 IPv4 MBGP 对等体组的前提是，在 BGP-IPv4 组播地址族下存在已经使能的 IPv4 单播对等体组。
- 将 IPv4 MBGP 对等体加入对等体组，需要先在 BGP 视图下将相应的对等体加入对等体组。

1.8.3 配置MBGP团体

对等体组可以使一组对等体共享相同的策略，而利用团体可以使多个 AS 中的一组 BGP 路由器共享相同的策略。团体是一个路由属性，在 BGP 对等体之间传播，它并不受到 AS 范围的限制。

BGP 路由器在将带有团体属性的路由发布给其它对等体之前，可以改变此路由原有的团体属性。除了使用公认的团体属性外，用户还可以使用团体属性列表自定义扩展团体属性，以便更为灵活地控制路由策略。

表1-21 配置 MBGP 团体

操作		命令	说明
进入系统视图		**system-view**	-
进入BGP视图		**bgp** *as-number*	-
进入MBGP地址族视图		**ipv4-family multicast**	-
配置向MBGP对等体/MBGP对等体组发布团体属性	配置发布团体属性	**peer** { *group-name* | *ip-address* } **advertise-community**	必选 缺省情况下，不将团体属性和扩展团体属性发布给任何IPv4 MBGP对等体/IPv4 MBGP对等体组
	配置发布扩展团体属性	**peer** { *group-name* | *ip-address* } **advertise-ext-community**	
对发布给IPv4 MBGP对等体/IPv4 MBGP对等体组的路由指定路由策略		**peer** { *group-name* | *ip-address* } **route-policy** *route-policy-name* **export**	必选 缺省情况下，不指定IPv4 MBGP对等体/IPv4 MBGP对等体组的路由策略

注意

- 配置 MBGP 团体时，必须使用路由策略来定义具体的团体属性，然后在发布路由信息时应用此路由策略。
- 关于路由策略的配置，请参见“三层技术-IP 路由命令参考”中的“路由策略”。

1.8.4 配置MBGP路由反射器

在 AS 内部，为保证 MBGP IBGP 对等体之间的连通性，需要在 MBGP IBGP 对等体之间建立全连接关系。当 MBGP IBGP 对等体数目很多时，在 MBGP IBGP 对等体之间建立全连接的开销很大。使用路由反射器，可以解决这个问题。

表1-22 配置 MBGP 路由反射器

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入MBGP地址族视图	**ipv4-family multicast**	-

操作	命令	说明
将本设备配置为路由反射器，并指定IPv4 MBGP对等体（组）作为路由反射器的客户	**peer** { *group-name* | *peer-address* } **reflect-client**	必选 缺省情况下，没有配置路由反射器及其客户
配置允许客户到客户的路由反射	**reflect between-clients**	可选 缺省情况下，允许客户到客户的路由反射
配置路由反射器的集群ID	**reflector cluster-id** *cluster-id*	可选 缺省情况下，每个路由反射器是使用自己的Router ID作为集群ID

注意

- 通常情况下，路由反射器的客户之间不要求是全连接的，路由缺省通过反射器从一个客户反射到其它客户；如果客户之间是全连接的，可以禁止客户间的反射，以便减少开销。
- 通常，一个集群里只有一个路由反射器。此时是由反射器的路由器 ID 来识别该集群的。设置多个路由反射器可提高网络的稳定性。如果一个集群中配有多个路由反射器，请使用相关命令为所有的路由反射器配置同样的集群 ID，以避免路由循环。

1.9 MBGP显示和维护

1.9.1 MBGP配置显示

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 MBGP 的运行情况，通过查看显示信息验证配置的效果。

表1-23 MBGP 配置显示

操作	命令
显示IPv4 BGP组播路由表中的路由信息	**display ip multicast routing-table** [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定目的地址的组播路由信息	**display ip multicast routing-table** *ip-address* [*mask-length* | *mask*] [**longer-match**] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv4 MBGP对等体组信息	**display bgp multicast group** [*group-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示MBGP发布的路由信息	**display bgp multicast network** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示AS路径信息	**display bgp multicast paths** [*as-regular-expression* | **|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv4 MBGP对等体/IPv4 MBGP对等体组的信息	**display bgp multicast peer** [[*ip-address*] **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
收到的邻居ORF信息中的前缀信息	**display bgp multicast peer** *ip-address* **received ip-prefix** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示MBGP路由信息	**display bgp multicast routing-table** [*ip-address* [{ *mask* | *mask-length* } [**longer-prefixes**]]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定AS路径过滤列表的MBGP路由信息	**display bgp multicast routing-table as-path-acl** *as-path-acl-number* [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示CIDR的MBGP路由信息	**display bgp multicast routing-table cidr** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定MBGP团体的MBGP路由信息	**display bgp multicast routing-table community**[*aa:nn&<1-13>*] [**no-advertise** | **no-export** | **no-export-subconfed**] * [**whole-match**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定MBGP团体列表的MBGP路由	**display bgp multicast routing-table community-list** { { *basic-community-list-number* | *comm-list-name* } [**whole-match**] | *adv-community-list-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示MBGP衰减的MBGP路由信息	**display bgp multicast routing-table dampened** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示MBGP衰减的配置参数	**display bgp multicast routing-table dampening parameter** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示源AS不一致的路由	**display bgp multicast routing-table different-origin-as** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示MBGP路由振荡统计信息	**display bgp multicast routing-table flap-info** [**regular-expression** *as-regular-expression* | [**as-path-acl** *as-path-acl-number* | *ip-address* [{ *mask* | *mask-length* } [**longer-match**]]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]]
显示向指定的IPv4 MBGP对等体发送或者从IPv4 MBGP对等体收到的路由信息	**display bgp multicast routing-table peer** *ip-address* { **advertised-routes** | **received-routes** } [*network-address* [*mask* | *mask-length*] | **statistic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示与指定的AS路径正则表达式相匹配的IPv4 MBGP路由信息	**display bgp multicast routing-table regular-expression** *as-regular-expression*
显示MBGP的路由统计信息	**display bgp multicast routing-table statistic** [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.9.2 复位MBGP连接

当 MBGP 路由策略或协议发生变化后，如果需要通过复位 MBGP 连接使新的配置生效，请在用户视图下进行下列配置。

表1-24 复位 MBGP 连接

操作	命令
复位指定的MBGP连接	**reset bgp ipv4 multicast** { **all** | *as-number* | *ip-address* | **group** *group-name* | **external** | **internal** }

1.9.3 清除MBGP信息

在用户视图下，执行 **reset** 命令可以清除 MBGP 相关统计信息。

表1-25 清除 MBGP 信息

操作	命令
清除路由的衰减信息并释放被抑制的路由	**reset bgp ipv4 multicast dampening** [*ip-address* [*mask* | *mask-length*]]
清除路由的振荡统计信息	**reset bgp ipv4 multicast flap-info** [**regexp** *as-path-regexp* | **as-path-acl** *as-path-acl-number* | *ip-address* [*mask* | *mask-length*]]

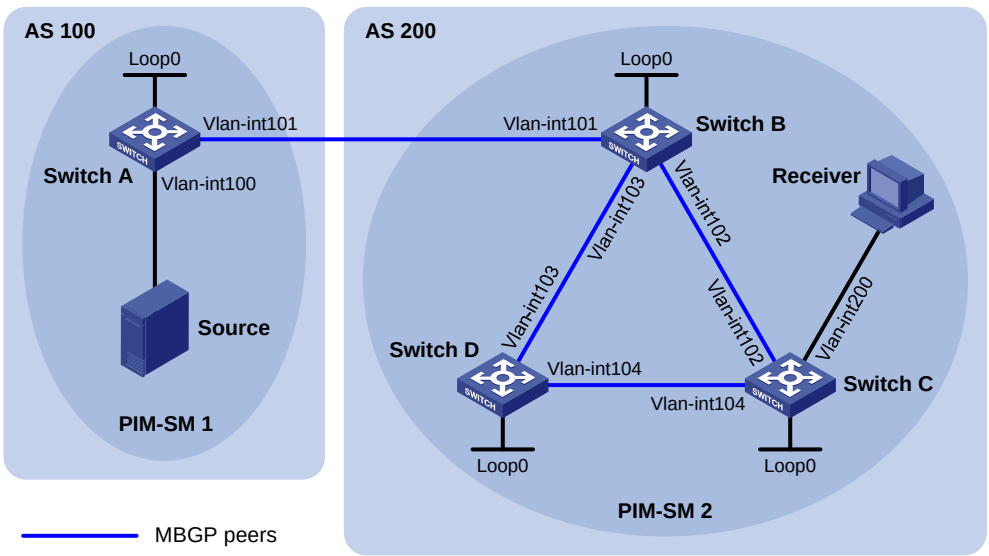
1.10 MBGP典型配置举例

1. 组网需求

- 网络中存在两个自治系统：PIM-SM 1 属于 AS 100，PIM-SM 2 属于 AS 200。各 AS 内部采用 OSPF 进行互联，AS 之间采用 MBGP 交换组播路由信息；
- 组播源属于 AS 100 内的 PIM-SM 1，接收者则属于 AS 200 内的 PIM-SM 2；
- 将 Switch A 和 Switch B 各自的 Loopback0 接口分别配置为各自 PIM-SM 域的 C-BSR 和 C-RP；
- 在 Switch A 与 Switch B 之间通过 MBGP 建立 MSDP 对等体关系。

2. 组网图

图1-1 MBGP 典型配置组网图



设备	接口	IP地址	设备	接口	IP地址
Source	-	10.110.1.100/24	Switch C	Vlan-int200	10.110.2.1/24
Switch A	Vlan-int100	10.110.1.1/24		Vlan-int102	192.168.2.2/24
	Vlan-int101	192.168.1.1/24		Vlan-int104	192.168.4.1/24
	Loop0	1.1.1.1/32		Loop0	3.3.3.3/32
Switch B	Vlan-int101	192.168.1.2/24	Switch D	Vlan-int103	192.168.3.2/24

	Vlan-int102	192.168.2.1/24		Vlan-int104	192.168.4.2/24
	Vlan-int103	192.168.3.1/24		Loop0	4.4.4.4/32
	Loop0	2.2.2.2/32			

3. 配置步骤

(1) 配置各交换机接口的 IP 地址和单播路由协议

- 请按照 图 1-1 配置各接口的IP地址和子网掩码，具体配置过程略。
- 配置 AS200 内的各交换机之间采用 OSPF 路由协议交换路由信息（AS 内各路由器创建的 OSPF 进程号为 1），确保各 AS 内部在网络层互通，能学到彼此 Loopback 接口的路由，具体配置过程略。

(2) 使能 IP 组播路由，使能 PIM-SM 和 IGMP，并配置 BSR 的服务边界

在 Switch A 上使能 IP 组播路由，在各接口上使能 PIM-SM。

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
```

Switch B 和 Switch D 上的配置与 Switch A 相似，配置过程略。

在 Switch C 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在主机侧接口 Vlan-interface200 上使能 IGMP。

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim sm
[SwitchC-Vlan-interface104] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] pim sm
[SwitchC-Vlan-interface200] igmp enable
[SwitchC-Vlan-interface200] quit
```

在 Switch A 上配置 BSR 的服务边界。

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim bsr-boundary
[SwitchA-Vlan-interface101] quit
```

在 Switch B 上配置 BSR 的服务边界。

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim bsr-boundary
[SwitchB-Vlan-interface101] quit
```

(3) 配置 Loopback0 接口和 C-BSR、C-RP 的位置

在 Switch A 上配置 Loopback0 接口和 C-BSR、C-RP 的位置。

```
[SwitchA] interface loopback 0
[SwitchA-LoopBack0] ip address 1.1.1.1 32
[SwitchA-LoopBack0] pim sm
[SwitchA-LoopBack0] quit
[SwitchA] pim
[SwitchA-pim] c-bsr loopback 0
[SwitchA-pim] c-rp loopback 0
[SwitchA-pim] quit
```

在 Switch B 上配置 Loopback0 接口和 C-BSR、C-RP 的位置。

```
[SwitchB] interface loopback 0
[SwitchB-LoopBack0] ip address 2.2.2.2 32
[SwitchB-LoopBack0] pim sm
[SwitchB-LoopBack0] quit
[SwitchB] pim
[SwitchB-pim] c-bsr loopback 0
[SwitchB-pim] c-rp loopback 0
[SwitchB-pim] quit
```

(4) 配置 BGP 协议，配置 MBGP 对等体，并互相引入路由

在 Switch A 上配置 EBGP 邻接关系、配置 MBGP 对等体。

```
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 192.168.1.2 as-number 200
[SwitchA-bgp] import-route direct
[SwitchA-bgp] ipv4-family multicast
[SwitchA-bgp-af-mul] peer 192.168.1.2 enable
[SwitchA-bgp-af-mul] import-route direct
[SwitchA-bgp-af-mul] quit
[SwitchA-bgp] quit
```

在 Switch B 上配置 EBGP 邻接关系、配置 MBGP 对等体，并引入 OSPF 路由。

```
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 192.168.1.1 as-number 100
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] ipv4-family multicast
[SwitchB-bgp-af-mul] peer 192.168.1.1 enable
[SwitchB-bgp-af-mul] import-route ospf 1
[SwitchB-bgp-af-mul] quit
[SwitchB-bgp] quit
```

(5) 配置 MSDP 对等体

在 Switch A 上配置 MSDP 对等体。

```
[SwitchA] msdp
[SwitchA-msdp] peer 192.168.1.2 connect-interface vlan-interface 101
[SwitchA-msdp] quit
```

在 Switch B 上配置 MSDP 对等体。

```
[SwitchB] msdp
[SwitchB-msdp] peer 192.168.1.1 connect-interface vlan-interface 101
[SwitchB-msdp] quit
```

(6) 检验配置效果

通过使用 **display bgp multicast peer** 命令可以查看交换机之间 MBGP 对等体的关系。例如：

查看 Switch B 上 MBGP 对等体关系的信息。

```
[SwitchB] display bgp multicast peer

 BGP local router ID : 2.2.2.2
 Local AS number : 200
 Total number of peers : 3                 Peers in established state : 3

  Peer            AS  MsgRcvd  MsgSent  OutQ PrefRcv Up/Down  State

  192.168.1.1    100       56       56     0       0 00:40:54 Established
```

通过使用 **display msdp brief** 命令可以查看交换机之间 MSDP 对等体建立情况。例如：

查看 Switch B 上 MSDP 对等体建立情况的简要信息。

```
[SwitchB] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0
  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  192.168.1.1       Up       00:07:17        100    1          0
```

目　录

1 组播VPN配置

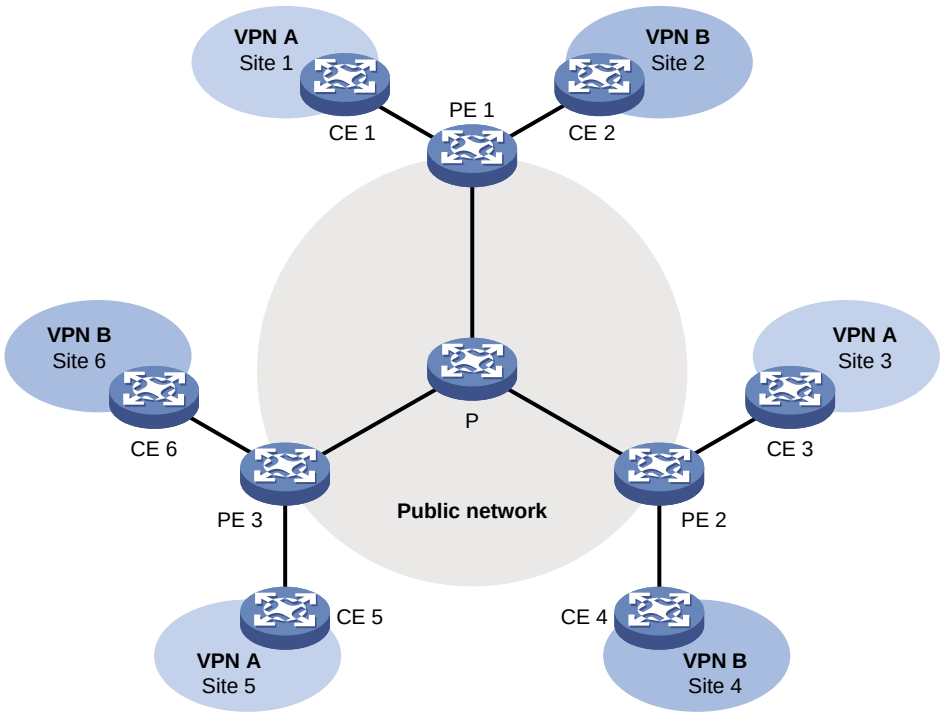
说明

- 本系列交换机中 S5500-28SC-HI 和 S5500-52SC-HI 不支持组播 VPN 功能。
- 本文所指的路由器代表运行了路由协议的三层设备。
- 有关 MPLS L3VPN 的详细介绍和相关配置，请参见“MPLS 配置指导”中的“MPLS L3VPN”。
- 有关 BGP 的详细介绍和相关配置，请参见“三层技术-IP 路由配置指导”中的“BGP”。

1.1 组播VPN简介

组播 VPN（Virtual Private Network，虚拟专用网络）是一项在 VPN 网络中实现组播传输的技术。

图1-1 典型的 VPN 组网



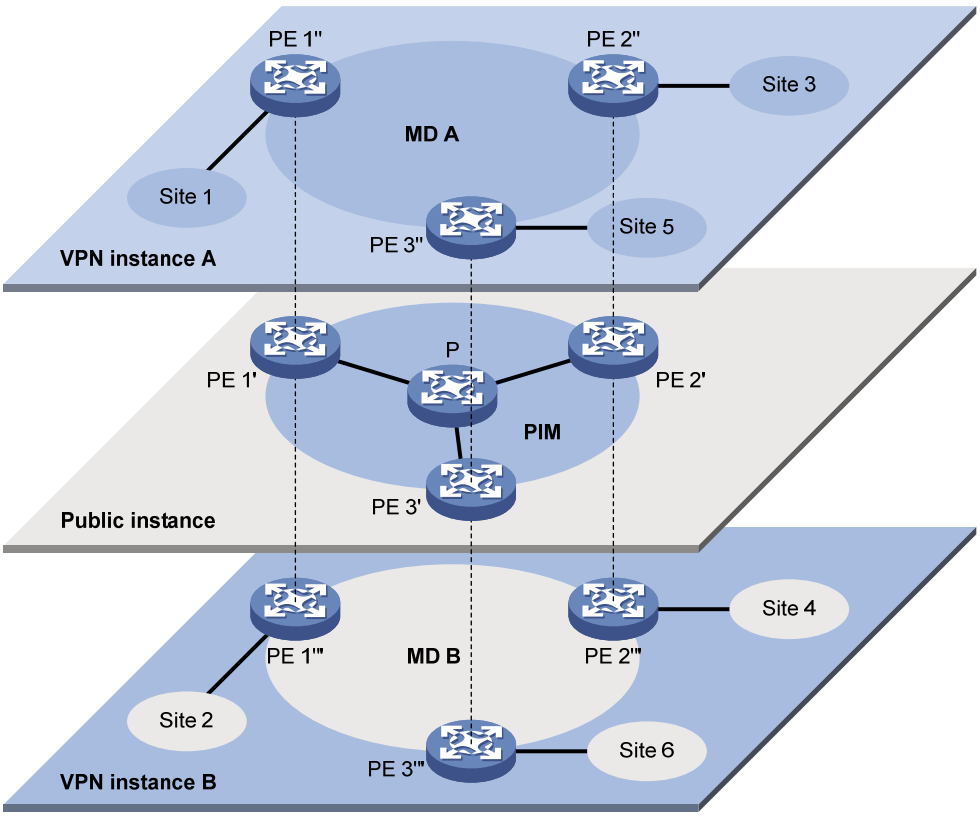
一个VPN网络由运营商的公共网络和用户的各个Site（站点）组成，各Site之间彼此相互孤立，只有借助公共网络才能实现互通。如 图 1-1 所示，由Site 1、Site 3 和Site 5 组成VPN A网络，由Site 2、Site 4 和Site 6 组成VPN B网络，其中包括以下三种类型的设备：

- P（Provider）：公共网络核心设备，不与 CE 直接相连。
- PE（Provider Edge）：公共网络边缘设备，与 CE 直接相连，负责 VPN 路由的处理。
- CE（Customer Edge）：用户网络边缘设备，可以是路由器或交换机，也可以是一台主机，负责用户网络路由的发布。

当 图 1-1 所示的VPN网络中运行组播VPN时，该网络中将同时承载着三个相互独立的组播业务：公网实例、VPN实例A和VPN实例B。公共网络边缘的PE组播设备支持多实例，相当于多台独立运行的组播设备。各实例之间形成彼此隔离的平面，每个实例对应一个平面，如 图 1-2 所示。

举例来说，图 1-1 中的PE 1 上运行着三个实例：公网实例、VPN实例A和VPN实例B，可以把这三个不同的实例想象成三台独立运行的虚拟设备，分别是PE 1’、PE 1”和PE 1”’，每台虚拟设备则分别对应着一个平面，对应关系如 图 1-2 所示。

图1-2 VPN 中基于多实例的组播



以 VPN 实例 A 为例，组播 VPN 是指：当 VPN A 中的组播源向某组播组发送组播数据时，在网络中所有可能的接收者中，只有属于 VPN A（即 Site 1、Site 3 或 Site 5 中）的组播组成员才能收到该组播源发来的组播数据。组播数据在各 Site 以及公网中均以组播方式进行传输。

由此分析，实现组播 VPN 所需具备的网络条件如下：

(1) 在每个 Site 内支持基于 VPN 实例的组播。

(2) 在公共网络内支持基于公网实例的组播。

(3) PE 设备支持多实例组播：

- 通过 VPN 实例连接 Site，支持基于 VPN 实例的组播；
- 通过公网实例连接公共网络，支持基于公网实例的组播；
- 支持公网实例与 VPN 实例之间的信息交流和数据转换。

Comware 利用 MD（Multicast Domain，组播域）方案来实现组播 VPN，简称为 MD VPN。该方案的最大优点就是仅需要 PE 设备支持多实例，而无需升级 CE 和 P 设备，且无需修改 CE 和 P 设备上原有的 PIM 配置——也就是说，该方案对于 CE 和 P 是透明的。

1.1.1 MD VPN概述

说明

有关 PIM（Protocol Independent Multicast，协议无关组播）、BSR（Bootstrap Router，自举路由器）、C-BSR（Candidate-BSR，候选 BSR）、RP（Rendezvous Point，汇集点）、DF（Designated Forwarder，指定转发者）、C-RP（Candidate-RP，候选 RP）、SPT（Shortest Path Tree，最短路径树）和 RPT（Rendezvous Point Tree，共享树）等概念的介绍，请参见“IP 组播配置指导”中的“PIM”。

1. MD VPN基本概念

MD VPN中所涉及到的基本概念如 表 1-1 所示。

表1-1 MD VPN 中的基本概念

概念	全称及中文解释	说明
MD	Multicast Domain，组播域	由各PE上能交互组播报文的相同VPN实例所构成的集合称为MD，不同的VPN实例属于不同的MD
MDT	Multicast Distribution Tree，组播分发树	建立在属于同一VPN内所有PE间的组播分发树，类型为Share-MDT
MT	Multicast Tunnel，组播隧道	在MD内将各PE连接到一起的通道称为MT，它用来传输MD内部的私网数据
MTI	Multicast Tunnel Interface，组播隧道接口	MTI是MT的入/出口，相当于MD的入/出口，PE通过MTI连接到MT上。MTI只收发组播报文，而不收发单播报文。MTI在VPN实例配置了Share-Group地址并绑定MTI后自动创建。
Share-Group	共享组	每个MD在公网上分配一个独立的组播地址，称为Share-Group地址。它是MD在公网上的唯一标志，用来在公网上建立MD所对应的Share-MDT
Share-MDT	Share-Multicast Distribution Tree，共享组播分发树	以Share-Group为组地址的MDT，称为Share-MDT。VPN使用Share-Group唯一标识一棵Share-MDT。Share-MDT是在配置完成后自动生成的，在公网中将会一直存在，而不论公网或私网中有没有实际的组播业务

2. MD VPN实现机制

MD VPN 实现机制的要点列举如下：

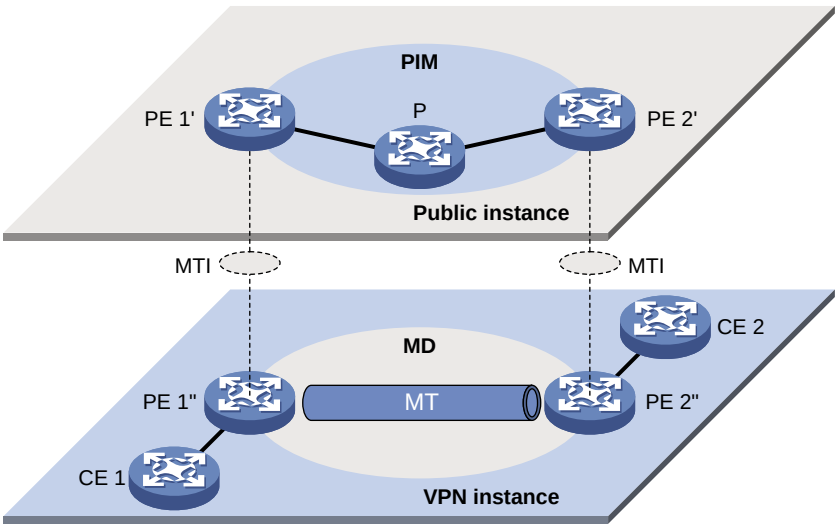
(1) 运营商构建的公共网络支持组播功能。PE 同时支持公网实例和多个 VPN 实例，每个实例各自运行相互间独立的 PIM。PE 与 CE 之间通过 VPN 实例进行私网组播传输；PE 与 P 之间则通过公网实例进行公网组播传输。

(2) MD在逻辑上表示某一特定VPN的私网组播数据在公网中的传播范围，在实际中则标识了网络中支持该VPN实例的所有PE设备。不同的VPN实例对应不同的MD。如 图 1-2 所示，其中每个VPN实例平面的中央椭圆区域表示一个MD，服务于某个特定的VPN，在该VPN中传输的所有私网组播数据，都在此MD内传输。

(3) 在 MD 内部，私网数据通过 MT 进行传输。MT 传输过程为：本地 PE 将私网组播报文封装成公网组播数据报文，并在公网内进行组播转发，远端 PE 收到该报文后通过解封装将其还原成私网报文。

(4) 本地PE将私网数据通过MTI发出，而远端PE则从MTI接收私网数据。如 图 1-3 所示，可以将MD比作一个私网数据的传输池，MTI则是MD的入/出口。本地PE将私网数据从入口（MTI）投入传输池，传输池自动将私网数据复制并传输到MD的所有出口（MTI），任何有需求的远端PE都可以从各自的出口（MTI）“打捞”私网数据。

图1-3 公网实例 PIM 与 VPN 实例 MD 对应关系示意图



(5) 一个 VPN 实例唯一指定一个 Share-Group 地址。私网数据信息对于公网来说是透明的，不论私网组播报文属于哪个组播组、是协议报文还是数据报文，PE 都统一将其封装为普通的公网组播数据报文，并以 Share-Group 作为其所属的公网组播组。之后，PE 将封装好的公网组播数据报文发送到公网中。

(6) 一个 Share-Group 唯一对应一个 MD，并利用公网资源唯一创建一棵 Share-MDT 以进行数据转发。在该 VPN 中传输的所有私网组播报文，无论从哪个 PE 进入公网，都经由此 Share-MDT 转发。

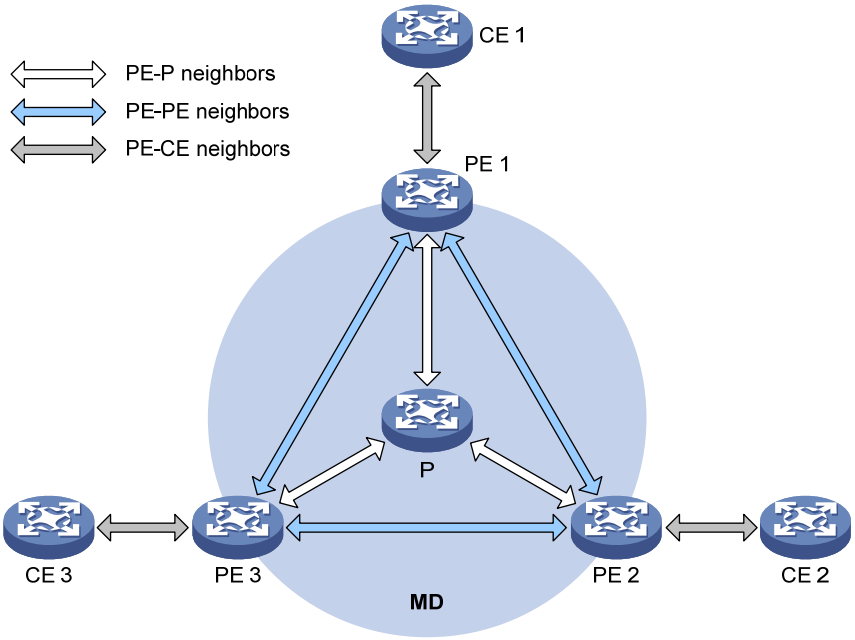
说明

一个 VPN 唯一确定一个 MD，而一个 MD 也只能服务于一个 VPN，这种关系就称为“一一对应”。VPN、MD、MTI 和 Share-Group 地址两两之间分别一一对应。

---

3. MD VPN中的PIM邻居关系

图1-4 MD VPN 中的 PIM 邻居关系示意图



PIM邻居关系建立在直接相连且属于同一网段的两台或多台设备之间。如 图 1-4 所示，在MD VPN中存在如下三种PIM邻居关系：

- PE-P 邻居关系：指 PE 上公网实例接口与链路对端 P 上的接口之间所建立的 PIM 邻居关系。
- PE-PE 邻居关系：指 PE 上的 VPN 实例通过 MTI 收到远端 PE 上的 VPN 实例发来的 PIM Hello 报文后所建立的邻居关系。
- PE-CE 邻居关系：指 PE 上绑定 VPN 实例的接口与链路对端 CE 上的接口之间所建立的 PIM 邻居关系。

1.1.2 协议规范

与组播 VPN 相关的协议规范有：

- draft-rosen-vpn-mcast-08：Multicast in MPLS/BGP IP VPNs

1.2 MD VPN实现原理

本节介绍 MD VPN 技术的实现原理，包括组播分发树的构建和工作过程，以及跨 AS 的实现方式。

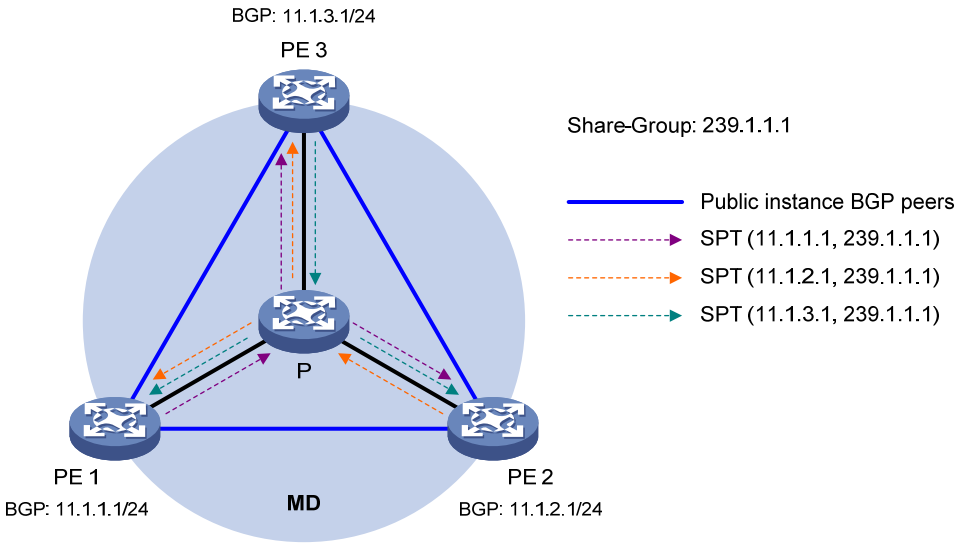
对于 VPN 实例来说，公网传输是透明的，私网数据的传输在 PE 上的 MTI 处完成了无缝连接：VPN 实例只知道将私网数据从 MTI 发出，然后远端就能从 MTI 接收。其实中间经历了复杂的公网传输过程，即 MDT 传输过程。

1.2.1 创建Share-MDT

公网中运行的组播路由协议可以是 PIM-DM、PIM-SM、双向 PIM 或 PIM-SSM 中的一种。在这四种情况下，创建 Share-MDT 的过程是有区别的。

1. 在PIM-DM网络中创建Share-MDT

图1-5 在 PIM-DM 网络中创建 Share-MDT

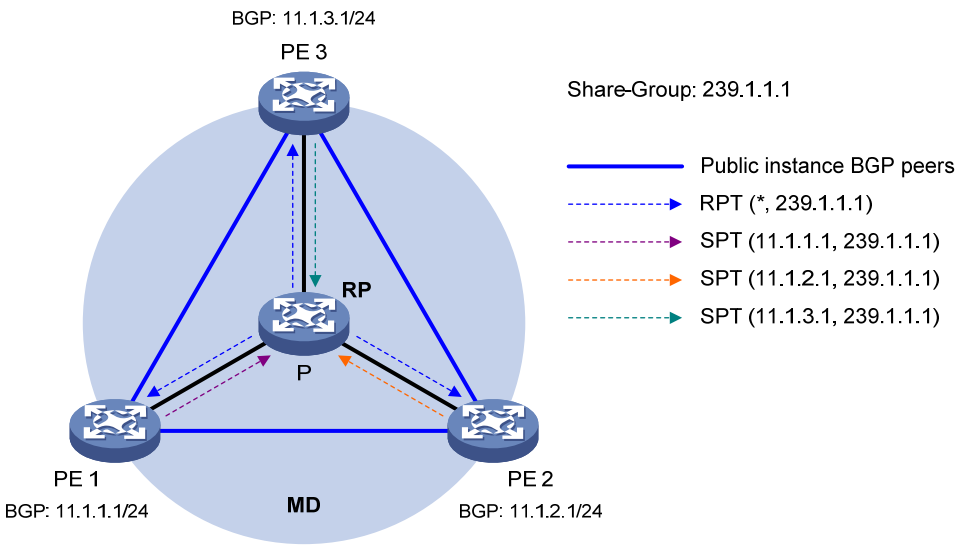


如 图 1-5 所示，公网中运行PIM-DM，且PE 1、PE 2 和PE 3 都支持VPN实例A。Share-MDT的创建过程如下：PE 1 的公网实例以BGP接口地址（即建立BGP对等体时所使用的接口地址）为组播源地址、Share-Group地址为组播组地址、其它支持VPN实例A的PE为组播组成员，在整个公网范围内发起扩散—剪枝过程，在公网沿途的各设备上分别创建（11.1.1.1，239.1.1.1）表项，从而形成一棵以PE 1 为根、PE 2 和PE 3 为叶的SPT。

同时，PE 2 和 PE 3 也各自发起类似的扩散—剪枝过程，最终在 MD 中形成三棵相互独立的 SPT。在 PIM-DM 网络中，由这三棵相互独立的 SPT 共同组成了一棵 Share-MDT。

2. 在PIM-SM网络中创建Share-MDT

图1-6 在 PIM-SM 网络中创建 Share-MDT



如 图 1-6 所示，公网中运行PIM-SM，且PE 1、PE 2 和PE 3 都支持VPN实例A。Share-MDT的创建过程如下：

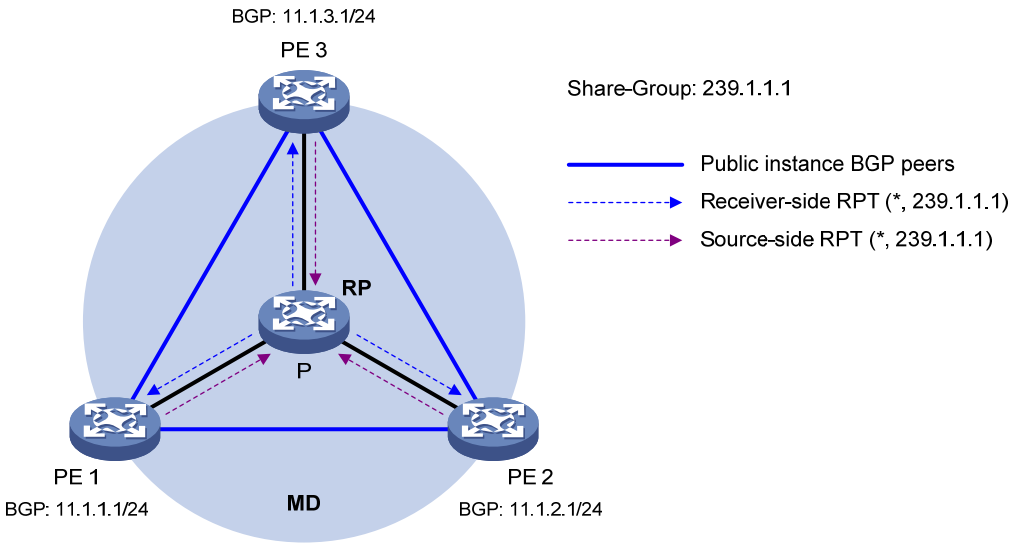
(1) PE 1 的公网实例向公网 RP 发起加入（Join），以 Share-Group 地址为组播组地址，在公网沿途的各设备上分别创建（*，239.1.1.1）表项。同时，PE 2 和 PE 3 也各自发起类似的加入过程，最终在 MD 中形成一棵以公网 RP 为根，PE 1、PE 2 和 PE 3 为叶的 RPT。

(2) 先由 PE 1 的公网实例向公网 RP 发起注册（Register），再由公网 RP 向 PE 1 发起加入，并以 BGP 接口地址为组播源地址、Share-Group 地址为组播组地址，在公网沿途的各设备上分别创建（11.1.1.1，239.1.1.1）表项。同时，PE 2 和 PE 3 也各自发起类似的注册过程，最终在 MD 中形成三棵相互独立的、连接 PE 与 RP 的 SPT。

在 PIM-SM 网络中，由 RPT（*，239.1.1.1）和这三棵相互独立的 SPT 共同组成了一棵 Share-MDT。

3. 在双向PIM网络中创建Share-MDT

图1-7 在双向 PIM 网络中创建 Share-MDT



如 图 1-7 所示，公网中运行双向PIM，且PE 1、PE 2 和PE 3 都支持VPN实例A。Share-MDT的创建过程如下：

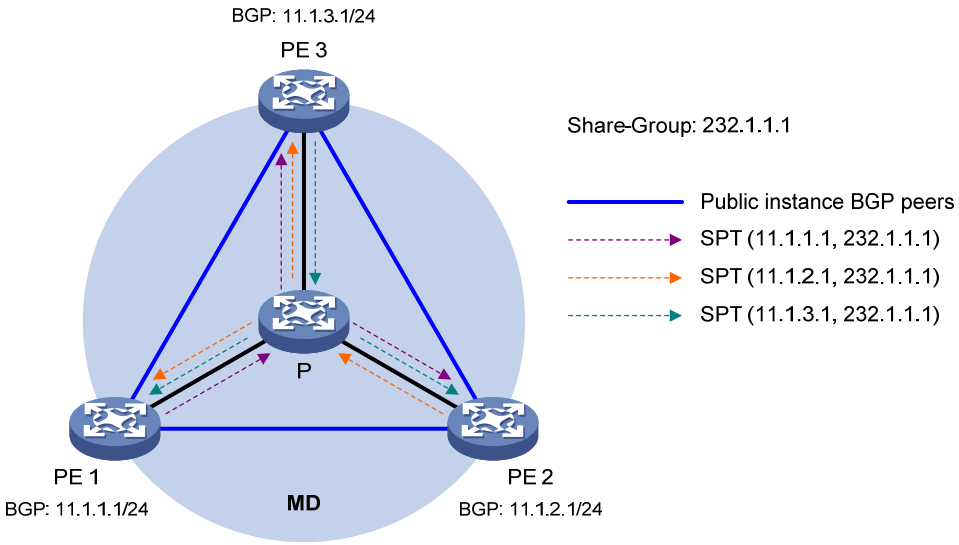
(1) PE 1 的公网实例向公网 RP 发起加入，以 Share-Group 地址为组播组地址，在公网沿途的各设备上分别创建（*，239.1.1.1）表项。同时，PE 2 和 PE 3 也各自发起类似的加入过程，最终在 MD 中形成一棵以公网 RP 为根，PE 1、PE 2 和 PE 3 为叶的接收者侧 RPT。

(2) PE 1 的公网实例以 Share-Group 地址为组播组地址发送组播数据，在公网途径的每个网段都被该网段的 DF 无条件向 RP 转发，并在沿途各设备上分别创建（*，239.1.1.1）表项。同时，PE 2 和 PE 3 也各自发起类似的创建过程，最终在 MD 中形成一棵以 PE 1、PE 2 和 PE 3 为根、公网 RP 为叶的组播源侧 RPT。

在双向 PIM 网络中，由接收者侧 RPT 和组播源侧 RPT 共同构成的双向 RPT（*，239.1.1.1）就成为了一棵 Share-MDT。

4. 在PIM-SSM网络中创建Share-MDT

图1-8 在 PIM-SSM 网络中创建 Share-MDT



如 图 1-8 所示，公网中运行PIM-SSM，且PE 1、PE 2 和PE 3 都支持VPN实例A。Share-MDT的创建过程如下：PE 1 的公网实例分别向PE 2 和PE 3 的公网实例通知其本地的BGP MDT信息（包含BGP接口地址和Share-Group地址等信息），并互换BGP MDT路由信息；PE 2 和PE 3 收到来自PE 1 的BGP MDT信息后，分别向PE 1 的BGP接口地址逐跳发送订阅报文（Subscribe Message），在公网沿途的各设备上分别创建（11.1.1.1，232.1.1.1）表项，从而形成一棵以PE 1 为根、PE 2 和PE 3 为叶的SPT。

同时，分别以PE 2和PE 3为根的SPT形成过程与此类似，最终在MD中形成三棵相互独立的SPT。在 PIM-SSM 网络中，由这三棵相互独立的 SPT 共同组成了一棵 Share-MDT。

说明

在 PIM-SSM 中，借助“订阅报文”的概念表示加入报文。

5. Share-MDT的特点

综前，无论公网中运行的是 PIM-DM、PIM-SM、双向 PIM 还是 PIM-SSM，Share-MDT 都具有以下特点：

- 网络中所有支持 VPN 实例 A 的 PE 都加入该 Share-MDT。
- 所有属于 VPN A 的私网组播报文（包括协议报文和数据报文）在进入公网后，均沿该 Share-MDT 向各 PE 转发，无论 PE 所连接的 Site 中是否存在接收者。

1.2.2 基于Share-MDT的传输

Share-MDT 构建完成后，就可以进行组播报文的传输。其中，组播协议报文和组播数据报文的传输过程是不同的，下面分别进行介绍。

1. 组播协议报文的传输

当私网组播协议报文需要跨越公网进行传输时，这些报文在本地 PE 上被封装为普通的公网组播数据并沿 Share-MDT 进行传输，在远端 PE 上被解封装，从而继续进行正常的协议交互过程，最终建立一棵跨越公网的组播分发树。组播协议报文的作用可分为以下几种情况：

(1) 当 VPN 网络中运行 PIM-DM 或 PIM-SSM 时：

- 各 MTI 之间通过交互 Hello 报文，建立 PIM 邻居。
- 通过跨越公网发起扩散—剪枝（PIM-DM）或加入（PIM-SSM）以创建 SPT。

(2) 当 VPN 网络中运行 PIM-SM 时：

- 各 MTI 之间通过交互 Hello 报文，建立 PIM 邻居。
- 如果接收者与私网 RP 属于不同的 Site，通过跨越公网发起加入以创建 RPT。
- 如果组播源与私网 RP 属于不同的 Site，通过跨越公网发起注册以创建 SPT。

(3) 当 VPN 网络中运行双向 PIM 时：

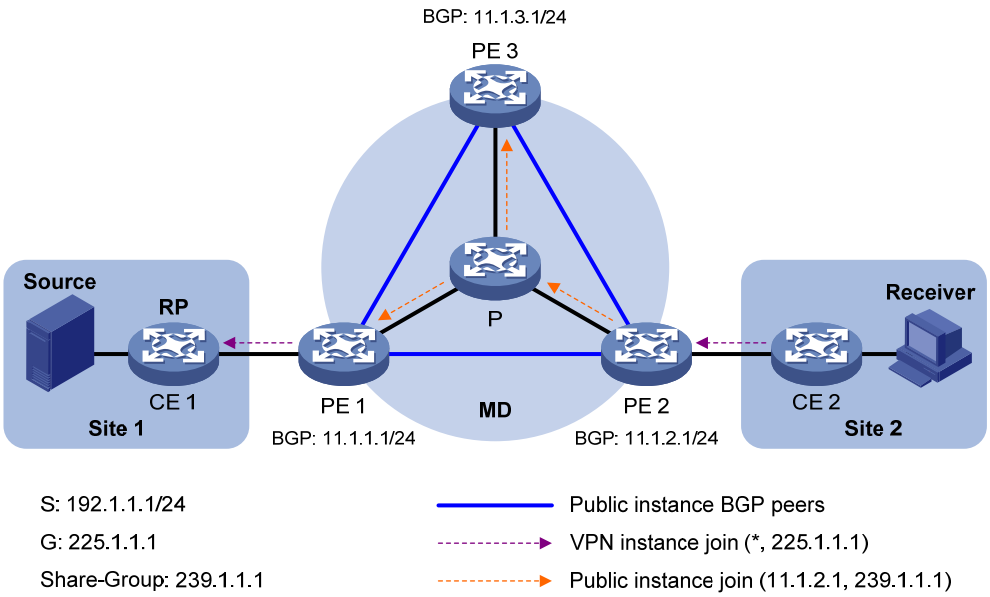
- 各 MTI 之间通过交互 Hello 报文，建立 PIM 邻居。
- 如果接收者与私网 RP 属于不同的 Site，通过跨越公网发起加入以创建接收者侧 RPT。
- 如果组播源与私网 RP 属于不同的 Site，通过跨越公网发送组播数据以创建组播源侧 RPT。

说明

- 在属于同一 VPN 的所有接口（包括 PE 上绑定 VPN 实例的接口和 MTI）上必须运行相同模式的 PIM 协议。
- 下面以公网和 VPN 网络中均运行 PIM-SM、私网接收者跨越公网发起加入为例，介绍基于 Share-MDT 的组播协议报文的传输过程。

如 图 1-9 所示，公网和VPN网络中分别运行PIM-SM，属于Site 2 的私网组播组G（225.1.1.1）的接收者（Receiver）与CE 2 相连；属于Site 1 的CE 1 作为G的RP；用于公网组播数据转发的Share-Group地址为 239.1.1.1。

图1-9 组播协议报文的传输过程



组播协议报文的交互过程如下：

(1) Receiver 向 CE 2 发送 IGMP 报告以加入组播组 G。CE 2 在本地创建（*，225.1.1.1）表项，同时向私网 RP（CE 1）发起加入。

(2) PE 2 上的 VPN 实例收到 CE 2 发来的加入消息（Join Message）后，在本地创建（*，225.1.1.1）表项，并指定上游接口为 MTI，然后 PE 2 将对该加入消息做进一步处理（详见第(3)步）。这时，PE 2 上的 VPN 实例将认为加入消息已从 MTI 发出。

(3) PE 2 对加入消息以 GRE（Generic Routing Encapsulation，通用路由封装）方式进行封装，以 PE 2 的 BGP 接口地址为组播源地址、Share-Group 地址为组播组地址，转换成普通的公网组播数据报文（11.1.2.1，239.1.1.1），然后交由 PE 2 上的公网实例向公网进行转发。

(4) 组播数据报文（11.1.2.1，239.1.1.1）沿 Share-MDT 传输给各 PE 上的公网实例。各 PE 对其进行解封装，还原为发往私网 RP 的加入消息。然后，各 PE 分别检查该加入消息，如果发现私网 RP 在与其直连的 Site 中，则交由其上的 VPN 实例处理，否则丢弃该加入消息。

(5) PE 1 上的 VPN 实例收到加入消息后，认为是从 MTI 获得的。在本地创建（*，225.1.1.1）表项，并指定下游接口为 MTI、上游接口为朝向 CE 1 的接口。同时向私网 RP 发送加入消息。

(6) CE 1 收到 PE 1 上的 VPN 实例发来的加入消息后，在本地更新或创建（*，225.1.1.1）表项。至此跨越公网的 RPT 创建完成。

说明

有关 GRE 的详细介绍，请参见“三层技术-IP 业务配置指导”中的“GRE”。

---

2. 组播数据报文的传输

当组播分发树创建完成后，组播源通过组播分发树将私网组播数据发送给各 Site 中的接收者。私网组播数据在本地 PE 上被封装为普通的公网组播数据并沿 Share-MDT 进行传输，在远端 PE 上被解封装，从而继续在私网内进行组播数据的传输。私网组播数据跨越公网的传输可分为以下几种情况：

(1) 当 VPN 网络中运行 PIM-DM 或 PIM-SSM 时，组播源通过私网 SPT 跨越公网向接收者发送私网组播数据。

(2) 当 VPN 网络中运行 PIM-SM 时：

- 在 SPT 切换前，如果组播源与私网 RP 属于不同的 Site，组播源先通过其与私网 RP 间的私网 SPT 跨越公网向私网 RP 发送私网组播数据；如果私网 RP 与接收者也属于不同的 Site，私网 RP 再通过私网 RPT 跨越公网向接收者继续转发私网组播数据。
- 在 SPT 切换后，如果组播源与接收者属于不同的 Site，组播源通过私网 SPT 跨越公网向接收者直接发送私网组播数据。

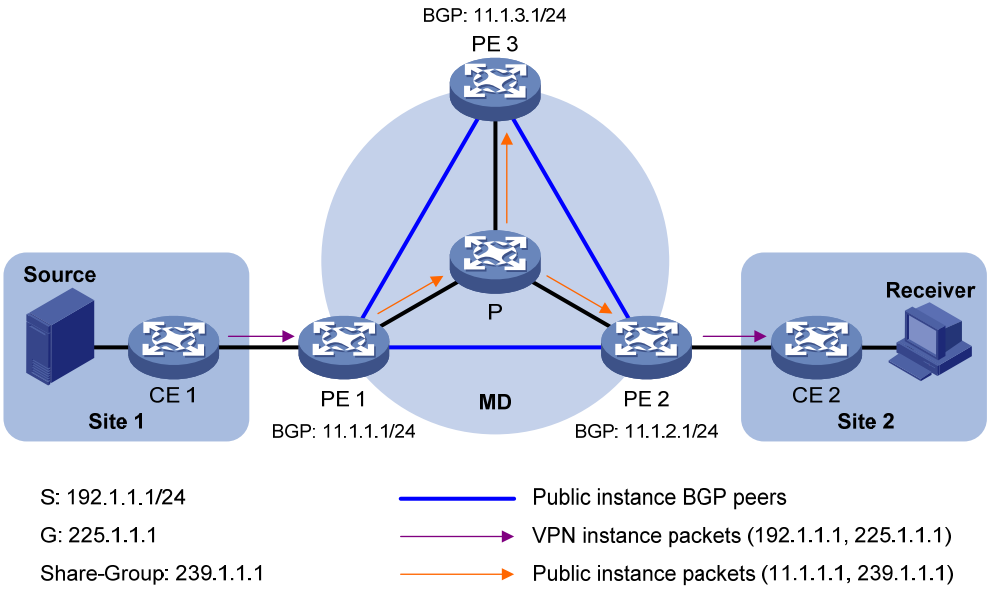
(3) 当 VPN 网络中运行双向 PIM 时，如果组播源与私网 RP 属于不同的 Site，组播源先通过其与私网 RP 间的私网组播源侧 RPT 跨越公网向私网 RP 发送私网组播数据；如果私网 RP 与接收者也属于不同的 Site，私网 RP 再通过私网接收者侧 RPT 跨越公网向接收者继续转发私网组播数据。

说明

- 有关 SPT 切换的详细介绍，请参见“IP 组播配置指导”中的“PIM”。
- 下面以公网和 VPN 网络中均运行 PIM-DM、沿私网 SPT 跨越公网传输私网组播数据为例，介绍基于 Share-MDT 的组播数据报文的传输过程。

如 图 1-10 所示，公网和VPN网络中分别运行PIM-DM，属于Site 2 的私网组播组G（225.1.1.1）的接收者（Receiver）与CE 2 相连；属于Site 1 的组播源（Source）向G发送组播数据；用于公网组播数据转发的Share-Group地址为 239.1.1.1。

图1-10 组播数据报文的传输过程



私网组播数据跨越公网进行传输的过程如下：

(1) Source 发送私网组播数据（192.1.1.1，225.1.1.1）到 CE 1。

(2) CE 1 沿 SPT 将私网组播数据转发给 PE 1，PE 1 上的 VPN 实例查找转发表项。如果对应转发表项的出接口列表中包含 MTI，则 PE 1 将对该私网组播数据做进一步处理（详见第(3)步）。这时，PE 1 上的 VPN 实例将认为私网组播数据已从 MTI 发出。

(3) PE 1 对该私网组播数据以 GRE 方式进行封装，以 PE 1 的 BGP 接口地址为组播源地址、Share-Group 地址为组播组地址，转换成普通的公网组播数据报文（11.1.1.1，239.1.1.1），然后交由 PE 1 上的公网实例向公网进行转发。

(4) 组播数据报文（11.1.1.1，239.1.1.1）沿 Share-MDT 传输给各 PE 上的公网实例。各 PE 对其进行解封装，还原为私网组播数据，然后交由相应的 VPN 实例处理。如果该 PE 上存在 SPT 的下游接口，则沿 SPT 转发该私网组播数据，否则将其丢弃。

(5) PE 2 上的 VPN 实例查找转发表项，最终将私网组播数据送达 Receiver。至此跨越公网的私网组播数据传输完成。

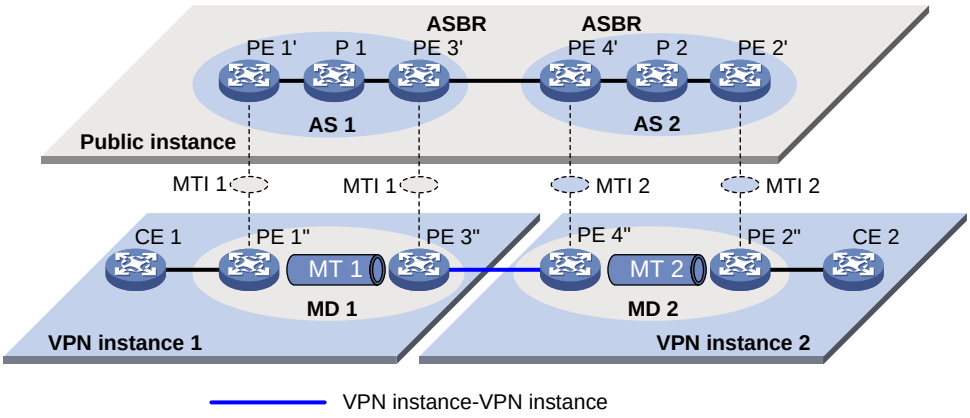
1.2.3 跨AS的MD VPN

当整个 VPN 跨越多个 AS（Autonomous System，自治系统）时，需要连通分布在不同 AS 内的各 VPN 节点。跨 AS 的 MD VPN 有以下两种实现方式：

1. VPN实例－VPN实例连接方式

如图 1-11 所示，VPN跨越了AS 1 和AS 2 两个自治系统，PE 3 和PE 4 分别是AS 1 和AS 2 的ASBR。PE 3 和PE 4 通过各自的VPN实例相连，并互把对方视为CE设备。

图1-11 VPN 实例－VPN 实例连接方式示意图



采用 VPN 实例－VPN 实例方式时，需在每个 AS 内各建立一个独立的 MD，在各 MD 之间实现私网组播数据跨 AS 的传输。

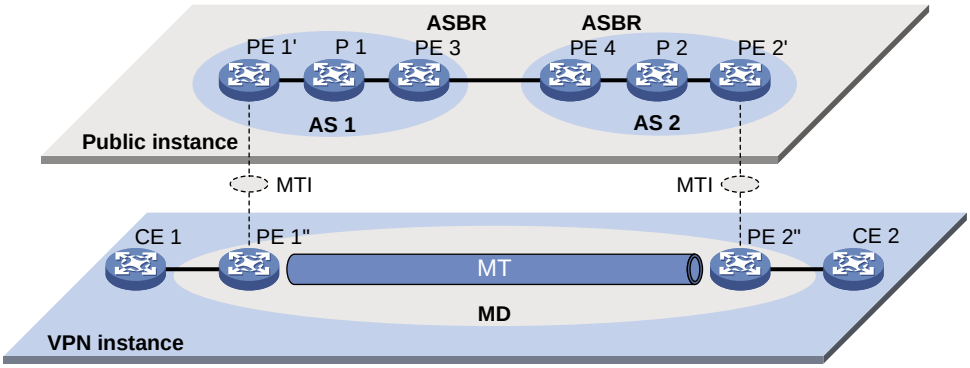
说明

由于 ASBR 之间转发的只是私网组播数据，因此各 AS 内部运行的公网 PIM 模式可以不同，但属于同一VPN的所有接口（包括ASBR上绑定VPN实例的接口）上必须运行统一的PIM模式（PIM-DM、PIM-SM、双向 PIM 或 PIM-SSM）。

2. Multi-hop EBGP连接方式

如图 1-12所示，VPN跨越了AS 1和AS 2两个自治系统，PE 3和PE 4分别是AS 1和AS 2的ASBR。PE 3 和PE 4 通过各自的公网实例相连，并互把对方视为P设备。

图1-12 Multi-hop EBGP 连接方式



采用 Multi-hop EBGP 方式时，只需在所有 AS 内统一建立一个 MD 即可，在该 MD 内部实现公网组播数据跨 AS 的传输。

1.3 组播VPN配置任务简介

表1-2 组播 VPN 配置任务简介

配置任务		说明	详细配置
配置MD VPN	使能VPN实例中的IP组播路由	必选	1.4.2
	配置Share-Group并绑定MTI	必选	1.4.3
配置BGP MDT	配置BGP MDT对等体/对等体组	必选	1.5.2
	配置BGP MDT路由反射器	可选	1.5.3

1.4 配置MD VPN

1.4.1 配置准备

在配置 MD VPN 之前，需完成以下任务：

- 在公网中配置任一单播路由协议
- 在公网中配置 MPLS L3VPN
- 在公网中配置 PIM-DM、PIM-SM、双向 PIM 或 PIM-SSM

在配置 MD VPN 之前，需准备以下数据：

- VPN 实例的名称和 RD（Route Distinguisher，路由标识符）
- Share-Group 地址和 MTI 接口编号

1.4.2 使能VPN实例中的IP组播路由

在配置 MD VPN 的各项功能之前，必须先使能 VPN 实例中的 IP 组播路由。

请在 PE 上进行本配置。

表1-3 使能 VPN 实例中的 IP 组播路由

操作	命令	说明
进入系统视图	**system-view**	-
创建VPN实例，并进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
配置VPN实例的RD	**route-distinguisher** *route-distinguisher*	必选 缺省情况下，VPN实例没有RD
使能IP组播路由	**multicast routing-enable**	必选 缺省情况下，IP组播路由处于关闭状态

说明

- 有关 **ip vpn-instance** 和 **route-distinguisher** 命令的详细介绍，请参见“MPLS 命令参考”中的“MPLS L3VPN”。
- 有关 **multicast routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“组播路由与转发”。

1.4.3 配置Share-Group并绑定MTI

通过在 PE 上启动多实例，可以为多个 VPN 服务。不同 PE 上相同的 VPN 实例需要配置相同的 Share-Group 地址。配置 Share-Group 并绑定 MTI 时，系统会自动创建 MTI、将 Share-Group 绑定到 MTI，并将 MTI 绑定到本 VPN 实例。

请在 PE 上进行本配置。

表1-4 配置 Share-Group 并绑定 MTI

操作	命令	说明
进入系统视图	**system-view**	-
进入VPN实例视图	**ip vpn-instance** *vpn-instance-name*	-
指定Share-Group并为本VPN实例绑定MTI	**multicast-domain share-group** *group-address* **binding mtunnel** *mtunnel-number*	必选 缺省情况下，没有指定Share-Group，也没有为本VPN实例绑定MTI

说明

- 当使用命令 **peer connect-interface** 配置了 BGP 对等体之后，MTI 接口将自动获取 **connect-interface** 的地址为其 IP 地址，且该地址不能在 VPN 中再被使用，否则 MTI 接口将无法获取 IP 地址；在一台设备上配置多个 BGP 对等体时应指定相同的 **connect-interface**，否则 MTI 接口也无法获取 IP 地址。有关 **peer connect-interface** 命令的详细介绍，请参见“三层技术-IP 路由命令参考”中的“BGP”。
- 在本系列交换机上必须使用命令 **service-loopback group** 配置 **multicast-tunnel** 类型的业务环回组，MTI 接口才会被激活。有关 **service-loopback group** 命令的详细介绍，请参见“二层技术-以太网交换命令参考”中的“业务环回组”。
- 只有当 VPN 实例内至少一个接口上使能了 PIM 协议，MTI 接口上的 PIM 协议才能被使能；当 VPN 实例内的所有接口上都关闭了 PIM 协议后，MTI 接口上的 PIM 协议也将被同时关闭。

1.5 配置BGP MDT

当在公网中运行 PIM-SSM 时，需要进行 BGP MDT 的配置。

1.5.1 配置准备

在配置 BGP MDT 之前，需完成以下任务：

- 在公网中配置 MPLS L3VPN
- 在公网中配置 BGP 基本功能

在配置 BGP MDT 之前，需准备以下数据：

- 路由反射器的集群 ID

1.5.2 配置BGP MDT对等体/对等体组

通过配置 BGP MDT 对等体/对等体组，使各 PE 间可通过交换 BGP MDT 路由信息以获取远端 PE 的地址来建立 Share-MDT。

请在 PE 上进行本配置。

表1-5 配置 BGP MDT 对等体/对等体组

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入BGP-MDT子地址族视图	**ipv4-family mdt**	-
激活BGP MDT对等体/对等体组	**peer** { *group-name* | *ip-address* } **enable**	必选 缺省情况下，BGP MDT对等体/对等体组处于未激活状态
向BGP MDT对等体组中添加BGP MDT对等体	**peer** *ip-address* **group** *group-name*	可选 缺省情况下，BGP MDT对等体不属于任何BGP MDT对等体组

说明

BGP MDT 对等体/对等体组是指在 BGP-MDT 子地址族视图下创建的对等体/对等体组。

1.5.3 配置BGP MDT路由反射器

为保证位于同一 AS 内的 BGP MDT 对等体间的连通性，需在对等体之间建立全连接关系，而当对等体的数目很多时，建立全连接的开销很大，使用路由反射器则可以解决这个问题。

在配置了路由反射器之后，由其它路由器作为客户机与路由反射器之间建立 BGP MDT 连接，路由反射器在客户机之间传递（反射）BGP MDT 路由信息，从而使各客户机之间无需建立全连接。但如果客户机之间已经是全连接的，则可以通过禁止客户机间的路由反射以减少开销。

路由反射器及其客户机共同组成了一个集群。通常，一个集群中只有一个路由反射器，并通过其 Router ID 来识别该集群。为了增强网络的可靠性，可在一个集群中配置多个路由反射器，此时应为每个路由反射器配置相同的集群 ID，以避免产生路由环路。

请在 PE 上进行本配置。

表1-6 配置 BGP MDT 路由反射器

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入BGP-MDT子地址族视图	**ipv4-family mdt**	-
配置本设备作为路由反射器，并指定其路由反射客户	**peer** { *group-name* | *ip-address* } **reflect-client**	必选 缺省情况下，没有配置路由反射器及其客户
禁止客户到客户的路由反射	**undo reflect between-clients**	可选 缺省情况下，允许客户到客户的路由反射
配置路由反射器的集群ID	**reflector cluster-id** { *cluster-id* | *ip-address* }	可选 缺省情况下，路由反射器的Router ID就是其集群ID

1.6 组播VPN显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后组播 VPN 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以复位 BGP MDT 的连接。

表1-7 组播 VPN 显示和维护

操作	命令
查看MD中指定VPN实例的Share-Group组信息	**display multicast-domain vpn-instance** *vpn-instance-name* **share-group** { **local** | **remote** } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看BGP MDT对等体组的信息	**display bgp mdt group** [*group-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看BGP MDT对等体的信息	**display bgp mdt peer** [[*ip-address*] **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
查看BGP MDT路由信息	**display bgp mdt** { **all** | **route-distinguisher** *route-distinguisher* } **routing-table** [*ip-address* [*mask* | *mask-length*]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
复位BGP MDT连接	**reset bgp mdt** { *as-number* | *ip-address* | **all** | **external** | **group** *group-name* | **internal** }

1.7 组播VPN典型配置举例

1.7.1 单AS内MD VPN配置举例

1. 组网需求

组网需求如 表 1-8 所示。

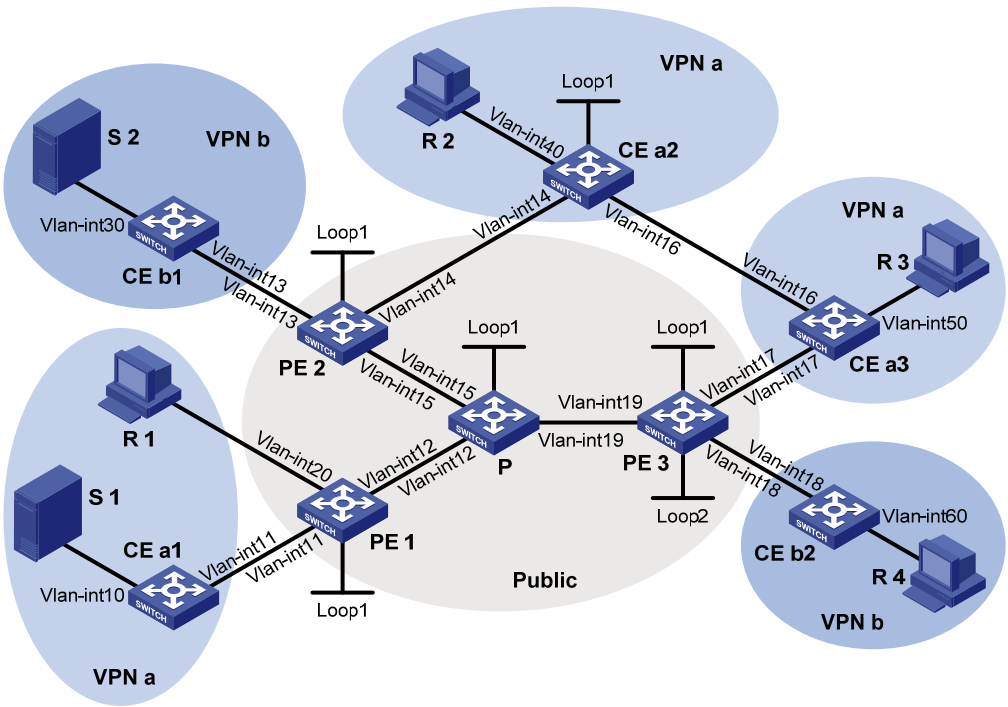
表1-8 单 AS 内的 MD VPN 配置组网需求

项目	组网需求
组播源和接收者	• VPN a 中的组播源为 S 1，接收者为 R 1、R 2 和 R 3 • VPN b 中的组播源为 S 2，接收者为 R 4 • VPN a 中的 Share-Group 地址为 239.1.1.1 • VPN b 中的 Share-Group 地址为 239.2.2.2
PE 上各接口所属的VPN实例	• PE 1：Vlan-interface11 和 Vlan-interface20 属于 VPN 实例 a，Vlan-interface12 和 Loopback1 属于公网实例 • PE 2：Vlan-interface13 属于 VPN 实例 b，Vlan-interface14 属于 VPN 实例 a，Vlan-interface15 和 Loopback1 属于公网实例 • PE 3：Vlan-interface17 属于 VPN 实例 a，Vlan-interface18 和 Loopback2 属于 VPN 实例 b，Vlan-interface19 和 Loopback1 接口属于公网实例
单播路由协议和MPLS	• 在公网中配置 OSPF，在各 PE 与 CE 之间配置 RIP • 在 PE 1、PE 2 和 PE 3 各自的 Loopback1 接口之间建立 BGP 对等体连接并传递所有私网路由 • 在公网中配置 MPLS
IP组播路由功能	• 在 P 上使能 IP 组播路由 • 在 PE 1、PE 2 和 PE 3 的公网实例中均使能 IP 组播路由 • 在 PE 1、PE 2 和 PE 3 的 VPN 实例 a 中均使能 IP 组播路由 • 在 PE 2 和 PE 3 的 VPN 实例 b 中均使能 IP 组播路由 • 在 CE a1、CE a2、CE a3、CE b1 和 CE b2 上均使能 IP 组播路由

项目	组网需求
IGMP功能	• 在 PE 1 的 Vlan-interface20 接口上使能 IGMPv2 • 在 CE a2 的 Vlan-interface40、CE a3 的 Vlan-interface50 和 CE b2 的 Vlan-interface60 接口上分别使能 IGMPv2
PIM功能	• 在 P 的所有接口上均使能 PIM-SM • 在 PE 1、PE 2 和 PE 3 的所有公网和私网接口上均使能 PIM-SM • 在 CE a1、CE a2、CE a3、CE b1 和 CE b2 的所有接口上均使能 PIM-SM • P 的 Loopback1 接口为公网的 C-BSR 和 C-RP（服务于所有组播组） • CE a2 的 Loopback1 接口为 VPN a 的 C-BSR 和 C-RP（服务于所有组播组） • PE 3 的 Loopback2 接口为 VPN b 的 C-BSR 和 C-RP（服务于所有组播组）

2. 组网图

图1-13 单 AS 内 MD VPN 配置组网图



设备	接口	IP地址	设备	接口	IP地址
S 1	-	10.110.7.2/24	PE 3	Vlan-int19	192.168.8.1/24
S 2	-	10.110.8.2/24		Vlan-int17	10.110.5.1/24
R 1	-	10.110.1.2/24		Vlan-int18	10.110.6.1/24
R 2	-	10.110.9.2/24		Loop1	1.1.1.3/32
R 3	-	10.110.10.2/24		Loop2	33.33.33.33/32
R 4	-	10.110.11.2/24	CE a1	Vlan-int10	10.110.7.1/24
P	Vlan-int12	192.168.6.2/24		Vlan-int11	10.110.2.2/24
	Vlan-int15	192.168.7.2/24	CE a2	Vlan-int40	10.110.9.1/24
	Vlan-int19	192.168.8.2/24		Vlan-int14	10.110.4.2/24
	Loop1	2.2.2.2/32		Vlan-int16	10.110.12.1/24
PE 1	Vlan-int12	192.168.6.1/24		Loop1	22.22.22.22/32

	Vlan-int20	10.110.1.1/24	CE a3	Vlan-int50	10.110.10.1/24
	Vlan-int11	10.110.2.1/24		Vlan-int17	10.110.5.2/24
	Loop1	1.1.1.1/32		Vlan-int16	10.110.12.2/24
PE 2	Vlan-int15	192.168.7.1/24	CE b1	Vlan-int30	10.110.8.1/24
	Vlan-int13	10.110.3.1/24		Vlan-int13	10.110.3.2/24
	Vlan-int14	10.110.4.1/24	CE b2	Vlan-int60	10.110.11.1/24
	Loop1	1.1.1.2/32		Vlan-int18	10.110.6.2/24

3. 配置步骤

(1) 配置 PE 1

配置 Router ID，使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<PE1> system-view
[PE1] router id 1.1.1.1
[PE1] multicast routing-enable
[PE1] mpls lsr-id 1.1.1.1
[PE1] mpls
[PE1-mpls] quit
[PE1] mpls ldp
[PE1-mpls-ldp] quit
```

创建 VPN 实例 a，为其配置 RD 和 VPN Target。

```
[PE1] ip vpn-instance a
[PE1-vpn-instance-a] route-distinguisher 100:1
[PE1-vpn-instance-a] vpn-target 100:1 export-extcommunity
[PE1-vpn-instance-a] vpn-target 100:1 import-extcommunity
```

使能 VPN 实例 a 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE1-vpn-instance-a] multicast routing-enable
[PE1-vpn-instance-a] multicast-domain share-group 239.1.1.1 binding mtunnel 0
[PE1-vpn-instance-a] quit
```

创建业务环回组 1，并指定其业务类型为 Multicast tunnel 类型。

```
[PE1] service-loopback group 1 type multicast-tunnel
```

选择一个未被使用的端口，关闭该端口上的生成树协议、LLDP 功能和 NDP 功能，并将其加入业务环回组 1。

```
[PE1] interface gigabitethernet1/0/3
[PE1-GigabitEthernet1/0/3] undo stp enable
[PE1-GigabitEthernet1/0/3] undo ndp enable
[PE1-GigabitEthernet1/0/3] undo lldp enable
[PE1-GigabitEthernet1/0/3] port service-loopback group 1
[PE1-GigabitEthernet1/0/3] quit
```

在公网接口 Vlan-interface12 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[PE1] interface vlan-interface 12
[PE1-Vlan-interface12] ip address 192.168.6.1 24
[PE1-Vlan-interface12] pim sm
[PE1-Vlan-interface12] mpls
[PE1-Vlan-interface12] mpls ldp
[PE1-Vlan-interface12] quit
```

将接口 Vlan-interface20 与 VPN 实例 a 进行关联，配置 IP 地址，并使能 IGMP 和 PIM-SM。

```
[PE1] interface vlan-interface 20
[PE1-Vlan-interface20] ip binding vpn-instance a
[PE1-Vlan-interface20] ip address 10.110.1.1 24
[PE1-Vlan-interface20] igmp enable
[PE1-Vlan-interface20] pim sm
[PE1-Vlan-interface20] quit
```

将接口 Vlan-interface11 与 VPN 实例 a 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE1] interface vlan-interface 11
[PE1-Vlan-interface11] ip binding vpn-instance a
[PE1-Vlan-interface11] ip address 10.110.2.1 24
[PE1-Vlan-interface11] pim sm
[PE1-Vlan-interface11] quit
```

配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。

```
[PE1] interface loopback 1
[PE1-LoopBack1] ip address 1.1.1.1 32
[PE1-LoopBack1] pim sm
[PE1-LoopBack1] quit
```

配置 BGP 协议。

```
[PE1] bgp 100
[PE1-bgp] group vpn-g internal
[PE1-bgp] peer vpn-g connect-interface loopback 1
[PE1-bgp] peer 1.1.1.2 group vpn-g
[PE1-bgp] peer 1.1.1.3 group vpn-g
[PE1–bgp] ipv4-family vpn-instance a
[PE1-bgp-a] import-route rip 2
[PE1-bgp-a] import-route direct
[PE1-bgp-a] quit
[PE1–bgp] ipv4-family vpnv4
[PE1–bgp-af-vpnv4] peer vpn-g enable
[PE1-bgp-af-vpnv4] peer 1.1.1.2 group vpn-g
[PE1–bgp-af-vpnv4] peer 1.1.1.3 group vpn-g
[PE1–bgp-af-vpnv4] quit
[PE1–bgp] quit
```

当 PE 1 上配置了 BGP 对等体之后，MTI0 接口将自动获得 IP 地址，该地址与配置 BGP 对等体时所指定的 Loopback 接口的 IP 地址相同；MTI0 接口上运行的 PIM 协议类型也与其所属的 VPN 实例 a 中运行的 PIM 协议类型相同。

配置 OSPF 协议。

```
[PE1] ospf 1
[PE1-ospf-1] area 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] network 1.1.1.1 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] network 192.168.0.0 0.0.255.255
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] quit
```

配置 RIP 协议。

```
[PE1] rip 2 vpn-instance a
[PE1-rip-2] network 10.0.0.0
[PE1-rip-2] import-route bgp
```

```
[PE1-rip-2] return
```

(2) 配置 PE 2

配置 Router ID，使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<PE2> system-view
[PE2] router id 1.1.1.2
[PE2] multicast routing-enable
[PE2] mpls lsr-id 1.1.1.2
[PE2] mpls
[PE2-mpls] quit
[PE2] mpls ldp
[PE2-mpls-ldp] quit
```

创建 VPN 实例 b，为其配置 RD 和 VPN Target。

```
[PE2] ip vpn-instance b
[PE2-vpn-instance-b] route-distinguisher 200:1
[PE2-vpn-instance-b] vpn-target 200:1 export-extcommunity
[PE2-vpn-instance-b] vpn-target 200:1 import-extcommunity
```

使能 VPN 实例 b 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE2-vpn-instance-b] multicast routing-enable
[PE2-vpn-instance-b] multicast-domain share-group 239.2.2.2 binding mtunnel 1
[PE2-vpn-instance-b] quit
```

创建 VPN 实例 a，为其配置 RD 和 VPN Target。

```
[PE2] ip vpn-instance a
[PE2-vpn-instance-a] route-distinguisher 100:1
[PE2-vpn-instance-a] vpn-target 100:1 export-extcommunity
[PE2-vpn-instance-a] vpn-target 100:1 import-extcommunity
```

使能 VPN 实例 a 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE2-vpn-instance-a] multicast routing-enable
[PE2-vpn-instance-a] multicast-domain share-group 239.1.1.1 binding mtunnel 0
[PE2-vpn-instance-a] quit
```

创建业务环回组 1，并指定其业务类型为 Multicast tunnel 类型。

```
[PE2] service-loopback group 1 type multicast-tunnel
```

选择一个未被使用的端口，关闭该端口上的生成树协议、LLDP 功能和 NDP 功能，并将其加入业务环回组 1。

```
[PE2] interface gigabitethernet1/0/3
[PE2-GigabitEthernet1/0/3] undo stp enable
[PE2-GigabitEthernet1/0/3] undo ndp enable
[PE2-GigabitEthernet1/0/3] undo lldp enable
[PE2-GigabitEthernet1/0/3] port service-loopback group 1
[PE2-GigabitEthernet1/0/3] quit
```

在公网接口 Vlan-interface15 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[PE2] interface vlan-interface 15
[PE2-Vlan-interface15] ip address 192.168.7.1 24
[PE2-Vlan-interface15] pim sm
[PE2-Vlan-interface15] mpls
[PE2-Vlan-interface15] mpls ldp
[PE2-Vlan-interface15] quit
```

将接口 Vlan-interface13 与 VPN 实例 b 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE2] interface vlan-interface 13
[PE2-Vlan-interface13] ip binding vpn-instance b
[PE2-Vlan-interface13] ip address 10.110.3.1 24
[PE2-Vlan-interface13] pim sm
[PE2-Vlan-interface13] quit
```

将接口 Vlan-interface14 与 VPN 实例 a 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE2] interface vlan-interface 14
[PE2-Vlan-interface14] ip binding vpn-instance a
[PE2-Vlan-interface14] ip address 10.110.4.1 24
[PE2-Vlan-interface14] pim sm
[PE2-Vlan-interface14] quit
```

配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。

```
[PE2] interface loopback 1
[PE2-LoopBack1] ip address 1.1.1.2 32
[PE2-LoopBack1] pim sm
[PE2-LoopBack1] quit
```

配置 BGP 协议。

```
[PE2] bgp 100
[PE2-bgp] group vpn-g internal
[PE2-bgp] peer vpn-g connect-interface loopback 1
[PE2-bgp] peer 1.1.1.1 group vpn-g
[PE2-bgp] peer 1.1.1.3 group vpn-g
[PE2–bgp] ipv4-family vpn-instance a
[PE2-bgp-a] import-route rip 2
[PE2-bgp-a] import-route direct
[PE2-bgp-a] quit
[PE2–bgp] ipv4-family vpn-instance b
[PE2-bgp-b] import-route rip 3
[PE2-bgp-b] import-route direct
[PE2-bgp-b] quit
[PE2–bgp] ipv4-family vpnv4
[PE2–bgp-af-vpnv4] peer vpn-g enable
[PE2-bgp-af-vpnv4] peer 1.1.1.1 group vpn-g
[PE2–bgp-af-vpnv4] peer 1.1.1.3 group vpn-g
[PE2–bgp-af-vpnv4] quit
[PE2–bgp] quit
```

当 PE 2 上配置了 BGP 对等体之后，MTI0 和 MTI1 接口将自动获得各自的 IP 地址，该地址与配置 BGP 对等体时所指定的 Loopback 接口的 IP 地址相同；MTI0 和 MTI1 接口上运行的 PIM 协议类型也分别与其所属的 VPN 实例 a 和 b 中运行的 PIM 协议类型相同。

配置 OSPF 协议。

```
[PE2] ospf 1
[PE2-ospf-1] area 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] network 1.1.1.2 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] network 192.168.0.0 0.0.255.255
[PE2-ospf-1-area-0.0.0.0] quit
[PE2-ospf-1] quit
```

配置 RIP 协议。

```
[PE2] rip 2 vpn-instance a
[PE2-rip-2] network 10.0.0.0
[PE2-rip-2] import-route bgp
[PE2-rip-2] quit
[PE2] rip 3 vpn-instance b
[PE2-rip-3] network 10.0.0.0
[PE2-rip-3] import-route bgp
[PE2-rip-3] return
```

(3) 配置 PE 3

配置 Router ID，使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<PE3> system-view
[PE3] router id 1.1.1.3
[PE3] multicast routing-enable
[PE3] mpls lsr-id 1.1.1.3
[PE3] mpls
[PE3-mpls] quit
[PE3] mpls ldp
[PE3-mpls-ldp] quit
```

创建 VPN 实例 a，为其配置 RD 和 VPN Target。

```
[PE3] ip vpn-instance a
[PE3-vpn-instance-a] route-distinguisher 100:1
[PE3-vpn-instance-a] vpn-target 100:1 export-extcommunity
[PE3-vpn-instance-a] vpn-target 100:1 import-extcommunity
```

使能 VPN 实例 a 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE3-vpn-instance-a] multicast routing-enable
[PE3-vpn-instance-a] multicast-domain share-group 239.1.1.1 binding mtunnel 0
[PE3-vpn-instance-a] quit
```

创建 VPN 实例 b，为其配置 RD 和 VPN Target。

```
[PE3] ip vpn-instance b
[PE3-vpn-instance-b] route-distinguisher 200:1
[PE3-vpn-instance-b] vpn-target 200:1 export-extcommunity
[PE3-vpn-instance-b] vpn-target 200:1 import-extcommunity
```

使能 VPN 实例 b 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE3-vpn-instance-b] multicast routing-enable
[PE3-vpn-instance-b] multicast-domain share-group 239.2.2.2 binding mtunnel 1
[PE3-vpn-instance-b] quit
```

创建业务环回组 1，并指定其业务类型为 Multicast tunnel 类型。

```
[PE3] service-loopback group 1 type multicast-tunnel
```

选择一个未被使用的端口，关闭该端口上的生成树协议、LLDP 功能和 NDP 功能，并将其加入业务环回组 1。

```
[PE3] interface gigabitethernet1/0/3
[PE3-GigabitEthernet1/0/3] undo stp enable
[PE3-GigabitEthernet1/0/3] undo ndp enable
[PE3-GigabitEthernet1/0/3] undo lldp enable
[PE3-GigabitEthernet1/0/3] port service-loopback group 1
```

```
[PE3-GigabitEthernet1/0/3] quit
```

在公网接口 Vlan-interface19 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[PE3] interface vlan-interface 19
[PE3-Vlan-interface19] ip address 192.168.8.1 24
[PE3-Vlan-interface19] pim sm
[PE3-Vlan-interface19] mpls
[PE3-Vlan-interface19] mpls ldp
[PE3-Vlan-interface19] quit
```

将接口 Vlan-interface17 与 VPN 实例 a 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE3] interface vlan-interface 17
[PE3-Vlan-interface17] ip binding vpn-instance a
[PE3-Vlan-interface17] ip address 10.110.5.1 24
[PE3-Vlan-interface17] pim sm
[PE3-Vlan-interface17] quit
```

将接口 Vlan-interface18 与 VPN 实例 b 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE3] interface vlan-interface 18
[PE3-Vlan-interface18] ip binding vpn-instance b
[PE3-Vlan-interface18] ip address 10.110.6.1 24
[PE3-Vlan-interface18] pim sm
[PE3-Vlan-interface18] quit
```

配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。

```
[PE3] interface loopback 1
[PE3-LoopBack1] ip address 1.1.1.3 32
[PE3-LoopBack1] pim sm
[PE3-LoopBack1] quit
```

将接口 Loopback2 与 VPN 实例 b 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE3] interface loopback 2
[PE3-LoopBack2] ip binding vpn-instance b
[PE3-LoopBack2] ip address 33.33.33.33 32
[PE3-LoopBack2] pim sm
[PE3-LoopBack2] quit
```

配置 Loopback2 接口为 VPN b 的 C-BSR 和 C-RP。

```
[PE3] pim vpn-instance b
[PE3-pim-b] c-bsr loopback 2
[PE3-pim-b] c-rp loopback 2
[PE3-pim-b] quit
```

配置 BGP 协议。

```
[PE3] bgp 100
[PE3-bgp] group vpn-g internal
[PE3-bgp] peer vpn-g connect-interface loopback 1
[PE3-bgp] peer 1.1.1.1 group vpn-g
[PE3-bgp] peer 1.1.1.2 group vpn-g
[PE3–bgp] ipv4-family vpn-instance a
[PE3-bgp-a] import-route rip 2
[PE3-bgp-a] import-route direct
[PE3-bgp-a] quit
```

```
[PE3–bgp] ipv4-family vpn-instance b
[PE3-bgp-b] import-route rip 3
[PE3-bgp-b] import-route direct
[PE3-bgp-b] quit
[PE3–bgp] ipv4-family vpnv4
[PE3–bgp-af-vpnv4] peer vpn-g enable
[PE3-bgp-af-vpnv4] peer 1.1.1.1 group vpn-g
[PE3–bgp-af-vpnv4] peer 1.1.1.2 group vpn-g
[PE3–bgp-af-vpnv4] quit
[PE3–bgp] quit
```

当 PE 3 上配置了 BGP 对等体之后，MTI0 和 MTI1 接口将自动获得各自的 IP 地址，该地址与配置 BGP 对等体时所指定的 Loopback 接口的 IP 地址相同；MTI0 和 MTI1 接口上运行的 PIM 协议类型也分别与其所属的 VPN 实例 a 和 b 中运行的 PIM 协议类型相同。

配置 OSPF 协议。

```
[PE3] ospf 1
[PE3-ospf-1] area 0.0.0.0
[PE3-ospf-1-area-0.0.0.0] network 1.1.1.3 0.0.0.0
[PE3-ospf-1-area-0.0.0.0] network 192.168.0.0 0.0.255.255
[PE3-ospf-1-area-0.0.0.0] quit
[PE3-ospf-1] quit
```

配置 RIP 协议。

```
[PE3] rip 2 vpn-instance a
[PE3-rip-2] network 10.0.0.0
[PE3-rip-2] import-route bgp
[PE3-rip-2] quit
[PE3] rip 3 vpn-instance b
[PE3-rip-3] network 10.0.0.0
[PE3-rip-3] network 33.0.0.0
[PE3-rip-3] import-route bgp
[PE3-rip-3] return
```

(4) 配置 P

使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<P> system-view
[P] multicast routing-enable
[P] mpls lsr-id 2.2.2.2
[P] mpls
[P-mpls] quit
[P] mpls ldp
[P-mpls-ldp] quit
```

在公网接口 Vlan-interface12 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[P] interface vlan-interface 12
[P-Vlan-interface12] ip address 192.168.6.2 24
[P-Vlan-interface12] pim sm
[P-Vlan-interface12] mpls
[P-Vlan-interface12] mpls ldp
[P-Vlan-interface12] quit
```

在公网接口 Vlan-interface15 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[P] interface vlan-interface 15
[P-Vlan-interface15] ip address 192.168.7.2 24
[P-Vlan-interface15] pim sm
[P-Vlan-interface15] mpls
[P-Vlan-interface15] mpls ldp
[P-Vlan-interface15] quit
```

在公网接口 Vlan-interface19 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[P] interface vlan-interface 19
[P-Vlan-interface19] ip address 192.168.8.2 24
[P-Vlan-interface19] pim sm
[P-Vlan-interface19] mpls
[P-Vlan-interface19] mpls ldp
[P-Vlan-interface19] quit
```

配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。

```
[P] interface loopback 1
[P-LoopBack1] ip address 2.2.2.2 32
[P-LoopBack1] pim sm
[P-LoopBack1] quit
```

配置 Loopback1 接口为公网实例的 C-BSR 和 C-RP。

```
[P] pim
[P-pim] c-bsr loopback 1
[P-pim] c-rp loopback 1
[P-pim] quit
```

配置 OSPF 协议。

```
[P] ospf 1
[P-ospf-1] area 0.0.0.0
[P-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[P-ospf-1-area-0.0.0.0] network 192.168.0.0 0.0.255.255
```

(5) 配置 CE a1

使能 IP 组播路由。

```
<CEa1> system-view
[CEa1] multicast routing-enable
```

在接口 Vlan-interface10 上配置 IP 地址，并使能 PIM-SM。

```
[CEa1] interface vlan-interface 10
[CEa1-Vlan-interface10] ip address 10.110.7.1 24
[CEa1-Vlan-interface10] pim sm
[CEa1-Vlan-interface10] quit
```

在接口 Vlan-interface11 上配置 IP 地址，并使能 PIM-SM。

```
[CEa1] interface vlan-interface 11
[CEa1-Vlan-interface11] ip address 10.110.2.2 24
[CEa1-Vlan-interface11] pim sm
[CEa1-Vlan-interface11] quit
```

配置 RIP 协议。

```
[CEa1] rip 2
[CEa1-rip-2] network 10.0.0.0
```

(6) 配置 CE b1

使能 IP 组播路由。

```
<CEb1> system-view
[CEb1] multicast routing-enable
```

在接口 Vlan-interface30 上配置 IP 地址，并使能 PIM-SM。

```
[CEb1] interface vlan-interface 30
[CEb1-Vlan-interface30] ip address 10.110.8.1 24
[CEb1-Vlan-interface30] pim sm
[CEb1-Vlan-interface30] quit
```

在接口 Vlan-interface13 上配置 IP 地址，并使能 PIM-SM。

```
[CEb1] interface vlan-interface 13
[CEb1-Vlan-interface13] ip address 10.110.3.2 24
[CEb1-Vlan-interface13] pim sm
[CEb1-Vlan-interface13] quit
```

配置 RIP 协议。

```
[CEb1] rip 3
[CEb1-rip-3] network 10.0.0.0
```

(7) 配置 CE a2

使能 IP 组播路由。

```
<CEa2> system-view
[CEa2] multicast routing-enable
```

在接口 Vlan-interface40 上配置 IP 地址，使能 IGMP 和 PIM-SM。

```
[CEa2] interface vlan-interface 40
[CEa2-Vlan-interface40] ip address 10.110.9.1 24
[CEa2-Vlan-interface40] igmp enable
[CEa2-Vlan-interface40] pim sm
[CEa2-Vlan-interface40] quit
```

在接口 Vlan-interface14 上配置 IP 地址，并使能 PIM-SM。

```
[CEa2] interface vlan-interface 14
[CEa2-Vlan-interface14] ip address 10.110.4.2 24
[CEa2-Vlan-interface14] pim sm
[CEa2-Vlan-interface14] quit
```

在接口 Vlan-interface16 上配置 IP 地址，并使能 PIM-SM。

```
[CEa2] interface vlan-interface 16
[CEa2-Vlan-interface16] ip address 10.110.12.1 24
[CEa2-Vlan-interface16] pim sm
[CEa2-Vlan-interface16] quit
```

配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。

```
[CEa2] interface loopback 1
[CEa2-LoopBack1] ip address 22.22.22.22 32
[CEa2-LoopBack1] pim sm
[CEa2-LoopBack1] quit
```

配置 Loopback1 接口为 VPN a 的 BSR 和 RP。

```
[CEa2] pim
[CEa2-pim] c-bsr loopback 1
```

```
[CEa2-pim] c-rp loopback 1
[CEa2-pim] quit
```
配置 RIP 协议。
```
[CEa2] rip 2
[CEa2-rip-2] network 10.0.0.0
[CEa2-rip-2] network 22.0.0.0
```
(8) 配置 CE a3

使能 IP 组播路由。
```
<CEa3> system-view
[CEa3] multicast routing-enable
```
在接口 Vlan-interface50 上配置 IP 地址，使能 IGMP 和 PIM-SM。
```
[CEa3] interface vlan-interface 50
[CEa3-Vlan-interface50] ip address 10.110.10.1 24
[CEa3-Vlan-interface50] igmp enable
[CEa3-Vlan-interface50] pim sm
[CEa3-Vlan-interface50] quit
```
在接口 Vlan-interface17 上配置 IP 地址，并使能 PIM-SM。
```
[CEa3] interface vlan-interface 17
[CEa3-Vlan-interface17] ip address 10.110.5.2 24
[CEa3-Vlan-interface17] pim sm
[CEa3-Vlan-interface17] quit
```
在接口 Vlan-interface16 上配置 IP 地址，并使能 PIM-SM。
```
[CEa3] interface vlan-interface 16
[CEa3-Vlan-interface16] ip address 10.110.12.2 24
[CEa3-Vlan-interface16] pim sm
[CEa3-Vlan-interface16] quit
```
配置 RIP 协议。
```
[CEa3] rip 2
[CEa3-rip-2] network 10.0.0.0
```
(9) 配置 CE b2

使能 IP 组播路由。
```
<CEb2> system-view
[CEb2] multicast routing-enable
```
在接口 Vlan-interface60 上配置 IP 地址，使能 IGMP 和 PIM-SM。
```
[CEb2] interface vlan-interface 60
[CEb2-Vlan-interface60] ip address 10.110.11.1 24
[CEb2-Vlan-interface60] igmp enable
[CEb2-Vlan-interface60] pim sm
[CEb2-Vlan-interface60] quit
```
在接口 Vlan-interface18 上配置 IP 地址，并使能 PIM-SM。
```
[CEb2] interface vlan-interface 18
[CEb2-Vlan-interface18] ip address 10.110.6.2 24
[CEb2-Vlan-interface18] pim sm
[CEb2-Vlan-interface18] quit
```
配置 RIP 协议。

```
[CEb2] rip 3
[CEb2-rip-3] network 10.0.0.0
```

4. 检验配置效果

通过使用 **display multicast-domain vpn-instance share-group** 命令可以查看指定 VPN 实例的 Share-Group 组信息。

查看 PE 1 上 VPN 实例 a 中的本地 Share-Group 组信息。

```
<PE1> display multicast-domain vpn-instance a share-group local
MD local share-group information for VPN-Instance: a
  Share-group: 239.1.1.1
  MTunnel address: 1.1.1.1
```

查看 PE 2 上 VPN 实例 a 中的本地 Share-Group 组信息。

```
<PE2> display multicast-domain vpn-instance a share-group local
MD local share-group information for VPN-Instance: a
  Share-group: 239.1.1.1
  MTunnel address: 1.1.1.2
```

查看 PE 2 上 VPN 实例 b 中的本地 Share-Group 组信息。

```
<PE2> display multicast-domain vpn-instance b share-group local
MD local share-group information for VPN-Instance: b
  Share-group: 239.2.2.2
  MTunnel address: 1.1.1.2
```

查看 PE 3 上 VPN 实例 a 中的本地 Share-Group 组信息。

```
<PE3> display multicast-domain vpn-instance a share-group local
MD local share-group information for VPN-Instance: a
  Share-group: 239.1.1.1
  MTunnel address: 1.1.1.3
```

查看 PE 3 上 VPN 实例 b 中的本地 Share-Group 组信息。

```
<PE3> display multicast-domain vpn-instance b share-group local
MD local share-group information for VPN-Instance: b
  Share-group: 239.2.2.2
  MTunnel address: 1.1.1.3
```

1.7.2 跨AS的MD VPN配置举例

1. 组网需求

组网需求如 表 1-9 所示。

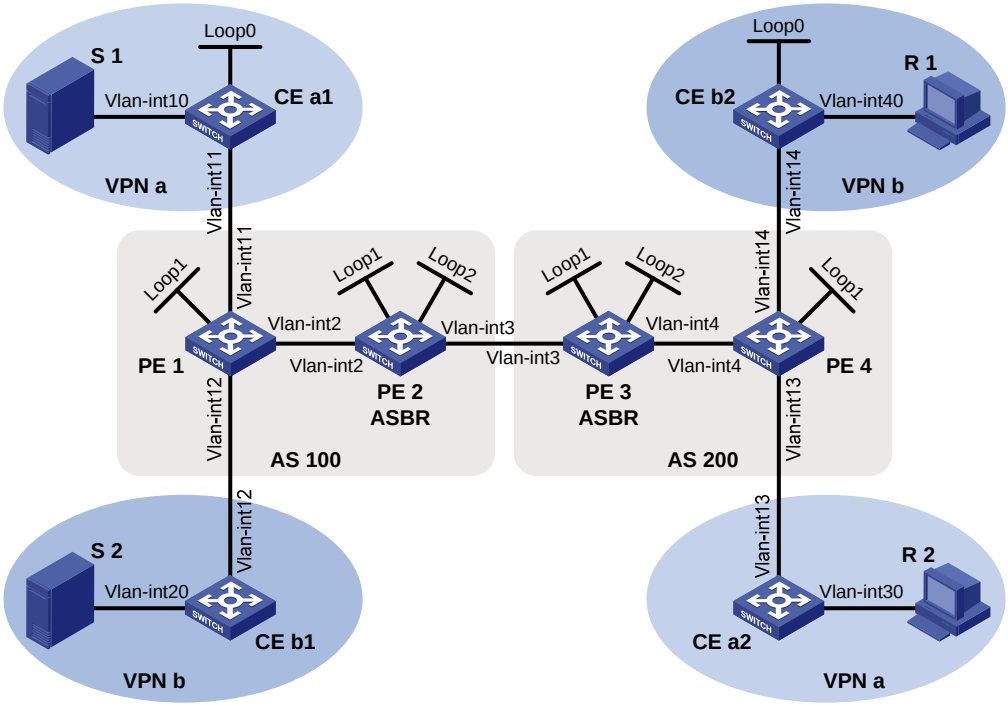
表1-9 跨 AS 的 MD VPN 配置组网需求

项目	组网需求
组播源和接收者	• VPN a 的组播源为 S 1，接收者为 R 2 • VPN b 的组播源为 S 2，接收者为 R 1 • VPN a 中的 Share-Group 地址为 239.1.1.1 • VPN b 中的 Share-Group 地址为 239.4.4.4

项目	组网需求
PE上各接口所属的VPN实例	• PE 1：Vlan-interface11 属于 VPN 实例 a，Vlan-interface12 属于 VPN 实例 b，Vlan-interface2 和 Loopback1 属于公网实例 • PE 2：Vlan-interface2、Vlan-interface3、Loopback1 和 Loopback2 属于公网实例 • PE 3：Vlan-interface3、Vlan-interface4、Loopback1 和 Loopback2 属于公网实例 • PE 4：Vlan-interface13 属于 VPN 实例 a，Vlan-interface14 属于 VPN 实例 b，Vlan-interface4 和 Loopback1 属于公网实例
单播路由协议和MPLS	• 在 AS 100 和 AS 200 中分别配置 OSPF，在各 PE 与 CE 之间也配置 OSPF • 在 PE 1、PE 2、PE 3 和 PE 4 各自的 Loopback1 接口之间建立 BGP 对等体连接并传递所有私网路由 • 在 AS 100 和 AS 200 中分别配置 MPLS
IP组播路由功能	• 在 PE 1、PE 2、PE 3 和 PE 4 的公网实例中均使能 IP 组播路由 • 在 PE 1 和 PE 4 的 VPN 实例 a 中均使能 IP 组播路由 • 在 PE 1 和 PE 4 的 VPN 实例 b 中均使能 IP 组播路由 • 在 CE a1、CE a2、CE b1、CE b2 上均使能 IP 组播路由
IGMP功能	• 在 CE a2 的 Vlan-interface30 接口上使能 IGMPv2 • 在 CE b2 的 Vlan-interface40 接口上使能 IGMPv2
PIM功能	• 在 PE 2 和 PE 3 的所有公网接口上均使能 PIM-SM • 在 PE 1 和 PE 4 的所有公网和私网接口上均使能 PIM-SM • 在 CE a1、CE a2、CE b1 和 CE b2 的所有接口上均使能 PIM-SM • PE 2 和 PE 3 的 Loopback2 接口为各自所在 AS 的 C-BSR 和 C-RP（服务于所有组播组） • CE a1 的 Loopback0 接口为 VPN a 的 C-BSR 和 C-RP（服务于所有组播组） • CE b1 的 Loopback0 接口为 VPN b 的 C-BSR 和 C-RP（服务于所有组播组）
MSDP功能	• 在 PE 2 和 PE 3 的 Loopback1 接口之间建立 MSDP 对等体

2. 组网图

图1-14 跨 AS 的 MD VPN 配置组网图



设备	接口	IP地址	设备	接口	IP地址
S 1	-	10.11.5.2/24	R 1	-	10.11.8.2/24
S 2	-	10.11.6.2/24	R 2	-	10.11.7.2/24
PE 1	Vlan-int2	10.10.1.1/24	PE 3	Vlan-int4	10.10.2.1/24
	Vlan-int11	10.11.1.1/24		Vlan-int3	192.168.1.2/24
	Vlan-int12	10.11.2.1/24		Loop1	1.1.1.3/32
	Loop1	1.1.1.1/32		Loop2	22.22.22.22/32
PE 2	Vlan-int2	10.10.1.2/24	PE 4	Vlan-int4	10.10.2.2/24
	Vlan-int3	192.168.1.1/24		Vlan-int13	10.11.3.1/24
	Loop1	1.1.1.2/32		Vlan-int14	10.11.4.1/32
	Loop2	11.11.11.11/32		Loop2	1.1.1.4/32
CE a1	Vlan-int10	10.11.5.1/24	CE b1	Vlan-int20	10.11.6.1/24
	Vlan-int11	10.11.1.2/24		Vlan-int12	10.11.2.2/24
	Loop0	2.2.2.2/32	CE b2	Vlan-int40	10.11.8.1/24
CE a2	Vlan-int30	10.11.7.1/24		Vlan-int14	10.11.4.2/24
	Vlan-int13	10.11.3.2/24		Loop0	3.3.3.3/32

3. 配置步骤

(1) 配置 PE 1

配置 Router ID，使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<PE1> system-view
[PE1] router id 1.1.1.1
[PE1] multicast routing-enable
[PE1] mpls lsr-id 1.1.1.1
[PE1] mpls
[PE1-mpls] quit
[PE1] mpls ldp
[PE1-mpls-ldp] quit
```

创建 VPN 实例 a，为其配置 RD 和 VPN Target；使能 VPN 实例 a 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE1] ip vpn-instance a
[PE1-vpn-instance-a] route-distinguisher 100:1
[PE1-vpn-instance-a] vpn-target 100:1 export-extcommunity
[PE1-vpn-instance-a] vpn-target 100:1 import-extcommunity
[PE1-vpn-instance-a] multicast routing-enable
[PE1-vpn-instance-a] multicast-domain share-group 239.1.1.1 binding mtunnel 0
[PE1-vpn-instance-a] quit
```

创建 VPN 实例 b，为其配置 RD 和 VPN Target；使能 VPN 实例 b 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE1] ip vpn-instance b
[PE1-vpn-instance-b] route-distinguisher 200:1
[PE1-vpn-instance-b] vpn-target 200:1 export-extcommunity
[PE1-vpn-instance-b] vpn-target 200:1 import-extcommunity
[PE1-vpn-instance-b] multicast routing-enable
[PE1-vpn-instance-b] multicast-domain share-group 239.4.4.4 binding mtunnel 1
[PE1-vpn-instance-b] quit
```

创建业务环回组 1，并指定其业务类型为 Multicast tunnel 类型。

```
[PE1] service-loopback group 1 type multicast-tunnel
```

选择一个未被使用的端口，关闭该端口上的生成树协议、LLDP 功能和 NDP 功能，并将其加入业务环回组 1。

```
[PE1] interface gigabitethernet1/0/3
[PE1-GigabitEthernet1/0/3] undo stp enable
[PE1-GigabitEthernet1/0/3] undo ndp enable
[PE1-GigabitEthernet1/0/3] undo lldp enable
[PE1-GigabitEthernet1/0/3] port service-loopback group 1
[PE1-GigabitEthernet1/0/3] quit
```

在公网接口 Vlan-interface2 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[PE1] interface vlan-interface 2
[PE1-Vlan-interface2] ip address 10.10.1.1 24
[PE1-Vlan-interface2] pim sm
[PE1-Vlan-interface2] mpls
[PE1-Vlan-interface2] mpls ldp
[PE1-Vlan-interface2] quit
```

将接口 Vlan-interface11 与 VPN 实例 a 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE1] interface vlan-interface 11
[PE1-Vlan-interface11] ip binding vpn-instance a
[PE1-Vlan-interface11] ip address 10.11.1.1 24
[PE1-Vlan-interface11] pim sm
[PE1-Vlan-interface11] quit
```

将接口 Vlan-interface12 与 VPN 实例 b 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE1] interface vlan-interface 12
[PE1-Vlan-interface12] ip binding vpn-instance b
[PE1-Vlan-interface12] ip address 10.11.2.1 24
[PE1-Vlan-interface12] pim sm
```

```
[PE1-Vlan-interface12] quit
```
配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。
```
[PE1] interface loopback 1
[PE1-LoopBack1] ip address 1.1.1.1 32
[PE1-LoopBack1] pim sm
[PE1-LoopBack1] quit
```
配置 BGP 协议。
```
[PE1] bgp 100
[PE1-bgp] group pe1-pe2 internal
[PE1-bgp] peer pe1-pe2 label-route-capability
[PE1-bgp] peer pe1-pe2 connect-interface loopback 1
[PE1-bgp] peer 1.1.1.2 group pe1-pe2
[PE1-bgp] group pe1-pe4 external
[PE1-bgp] peer pe1-pe4 as-number 200
[PE1-bgp] peer pe1-pe4 ebgp-max-hop 255
[PE1-bgp] peer 1.1.1.4 group pe1-pe4
[PE1-bgp] peer pe1-pe4 connect-interface loopback 1
[PE1–bgp] ipv4-family vpn-instance a
[PE1-bgp-a] import-route ospf 2
[PE1-bgp-a] import-route direct
[PE1-bgp-a] quit
[PE1–bgp] ipv4-family vpn-instance b
[PE1-bgp-b] import-route ospf 3
[PE1-bgp-b] import-route direct
[PE1-bgp-b] quit
[PE1–bgp] ipv4-family vpnv4
[PE1–bgp-af-vpnv4] peer 1.1.1.4 enable
[PE1–bgp-af-vpnv4] quit
[PE1–bgp] quit
```
当 PE 1 上配置了 BGP 对等体之后，MTI0 和 MTI1 接口将分别自动获得其 IP 地址，该地址与配置 BGP 对等体时所指定的 Loopback 接口的 IP 地址相同；MTI0 接口上运行的 PIM 协议类型与 VPN 实例 a 中运行的 PIM 协议类型相同；MTI1 接口上运行的 PIM 协议类型与 VPN 实例 b 中运行的 PIM 协议类型相同。

配置 OSPF 协议。
```
[PE1] ospf 1
[PE1-ospf-1] area 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] network 1.1.1.1 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] network 10.10.0.0 0.0.255.255
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] quit
[PE1] ospf 2 vpn-instance a
[PE1-ospf-2] import-route bgp
[PE1-ospf-2] area 0.0.0.0
[PE1-ospf-2-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[PE1-ospf-2-area-0.0.0.0] quit
[PE1-ospf-2] quit
[PE1] ospf 3 vpn-instance b
```

```
[PE1-ospf-3] import-route bgp
[PE1-ospf-3] area 0.0.0.0
[PE1-ospf-3-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[PE1-ospf-3-area-0.0.0.0] quit
[PE1-ospf-3] quit
```

(2) 配置 PE 2

配置 Router ID，使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<PE2> system-view
[PE2] router id 1.1.1.2
[PE2] multicast routing-enable
[PE2] mpls lsr-id 1.1.1.2
[PE2] mpls
[PE2-mpls] quit
[PE2] mpls ldp
[PE2-mpls-ldp] quit
```

在公网接口 Vlan-interface2 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[PE2] interface vlan-interface 2
[PE2-Vlan-interface2] ip address 10.10.1.2 24
[PE2-Vlan-interface2] pim sm
[PE2-Vlan-interface2] mpls
[PE2-Vlan-interface2] mpls ldp
[PE2-Vlan-interface2] quit
```

在公网接口 Vlan-interface3 上配置 IP 地址，使能 PIM-SM 和 MPLS。

```
[PE2] interface vlan-interface 3
[PE2-Vlan-interface3] ip address 192.168.1.1 24
[PE2-Vlan-interface3] pim sm
[PE2-Vlan-interface3] mpls
[PE2-Vlan-interface3] quit
```

配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。

```
[PE2] interface loopback 1
[PE2-LoopBack1] ip address 1.1.1.2 32
[PE2-LoopBack1] pim sm
[PE2-LoopBack1] quit
```

配置 Loopback2 接口的 IP 地址，并使能 PIM-SM。

```
[PE2] interface loopback 2
[PE2-LoopBack2] ip address 11.11.11.11 32
[PE2-LoopBack2] pim sm
[PE2-LoopBack2] quit
```

配置 Loopback2 接口为公网实例的 C-BSR 和 C-RP。

```
[PE2] pim
[PE2-pim] c-bsr loopback 2
[PE2-pim] c-rp loopback 2
[PE2-pim] quit
```

配置 BSR 的服务边界。

```
[PE2] interface vlan-interface 3
[PE2-Vlan-interface3] pim bsr-boundary
```

```
[PE2-Vlan-interface3] quit
```

配置 MSDP 对等体。

```
[PE2] msdp
[PE2-msdp] encap-data-enable
[PE2-msdp] peer 1.1.1.3 connect-interface loopback 1
```

配置静态路由。

```
[PE2] ip route-static 1.1.1.3 32 vlan-interface 3 192.168.1.2
```

配置 BGP 协议。

```
[PE2] bgp 100
[PE2-bgp] import-route ospf 1
[PE2-bgp] group pe2-pe1 internal
[PE2-bgp] peer pe2-pe1 route-policy map2 export
[PE2-bgp] peer pe2-pe1 label-route-capability
[PE2-bgp] peer pe2-pe1 connect-interface loopback 1
[PE2-bgp] peer 1.1.1.1 group pe2-pe1
[PE2-bgp] group pe2-pe3 external
[PE2-bgp] peer pe2-pe3 as-number 200
[PE2-bgp] peer pe2-pe3 route-policy map1 export
[PE2-bgp] peer pe2-pe3 label-route-capability
[PE2-bgp] peer pe2-pe3 connect-interface loopback 1
[PE2-bgp] peer 1.1.1.3 group pe2-pe3
[PE2–bgp] quit
```

配置 OSPF 协议。

```
[PE2] ospf 1
[PE2-ospf-1] area 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] network 1.1.1.2 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] network 11.11.11.11 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] network 10.10.0.0 0.0.255.255
[PE2-ospf-1-area-0.0.0.0] quit
[PE2-ospf-1] quit
```

配置路由策略。

```
[PE2] route-policy map1 permit node 10
[PE2-route-policy] apply mpls-label
[PE2-route-policy] quit
[PE2] route-policy map2 permit node 10
[PE2-route-policy] if-match mpls-label
[PE2-route-policy] apply mpls-label
[PE2-route-policy] quit
```

(3) 配置 PE 3

配置 Router ID，使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<PE3> system-view
[PE3] router id 1.1.1.3
[PE3] multicast routing-enable
[PE3] mpls lsr-id 1.1.1.3
[PE3] mpls
[PE3-mpls] quit
[PE3] mpls ldp
```

```
[PE3-mpls-ldp] quit
```
在公网接口 Vlan-interface4 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。
```
[PE3] interface vlan-interface 4
[PE3-Vlan-interface4] ip address 10.10.2.1 24
[PE3-Vlan-interface4] pim sm
[PE3-Vlan-interface4] mpls
[PE3-Vlan-interface4] mpls ldp
[PE3-Vlan-interface4] quit
```
在公网接口 Vlan-interface3 上配置 IP 地址，使能 PIM-SM 和 MPLS。
```
[PE3] interface vlan-interface 3
[PE3-Vlan-interface3] ip address 192.168.1.2 24
[PE3-Vlan-interface3] pim sm
[PE3-Vlan-interface3] mpls
[PE3-Vlan-interface3] quit
```
配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。
```
[PE3] interface loopback 1
[PE3-LoopBack1] ip address 1.1.1.3 32
[PE3-LoopBack1] pim sm
[PE3-LoopBack1] quit
```
配置 Loopback2 接口的 IP 地址，并使能 PIM-SM。
```
[PE3] interface loopback 2
[PE3-LoopBack2] ip address 22.22.22.22 32
[PE3-LoopBack2] pim sm
[PE3-LoopBack2] quit
```
配置 Loopback2 接口为公网实例的 C-BSR 和 C-RP。
```
[PE3] pim
[PE3-pim] c-bsr loopback 2
[PE3-pim] c-rp loopback 2
[PE3-pim] quit
```
配置 BSR 的服务边界。
```
[PE3] interface vlan-interface 3
[PE3-Vlan-interface3] pim bsr-boundary
[PE3-Vlan-interface3] quit
```
配置 MSDP 对等体。
```
[PE3] msdp
[PE3-msdp] encap-data-enable
[PE3-msdp] peer 1.1.1.2 connect-interface loopback 1
```
配置静态路由。
```
[PE3] ip route-static 1.1.1.2 32 vlan-interface 3 192.168.1.1
```
配置 BGP 协议。
```
[PE3] bgp 200
[PE3-bgp] import-route ospf 1
[PE3-bgp] group pe3-pe4 internal
[PE3-bgp] peer pe3-pe4 route-policy map2 export
[PE3-bgp] peer pe3-pe4 label-route-capability
[PE3-bgp] peer pe3-pe4 connect-interface loopback 1
```

```
[PE3-bgp] peer 1.1.1.4 group pe3-pe4
[PE3-bgp] group pe3-pe2 external
[PE3-bgp] peer pe3-pe2 as-number 100
[PE3-bgp] peer pe3-pe2 route-policy map1 export
[PE3-bgp] peer pe3-pe2 label-route-capability
[PE3-bgp] peer pe3-pe2 connect-interface loopback 1
[PE3-bgp] peer 1.1.1.2 group pe3-pe2
[PE3–bgp] quit
```

配置 OSPF 协议。

```
[PE3] ospf 1
[PE3-ospf-1] area 0.0.0.0
[PE3-ospf-1-area-0.0.0.0] network 1.1.1.3 0.0.0.0
[PE3-ospf-1-area-0.0.0.0] network 22.22.22.22 0.0.0.0
[PE3-ospf-1-area-0.0.0.0] network 10.10.0.0 0.0.255.255
[PE3-ospf-1-area-0.0.0.0] quit
[PE3-ospf-1] quit
```

配置路由策略。

```
[PE3] route-policy map1 permit node 10
[PE3-route-policy] apply mpls-label
[PE3-route-policy] quit
[PE3] route-policy map2 permit node 10
[PE3-route-policy] if-match mpls-label
[PE3-route-policy] apply mpls-label
[PE3-route-policy] quit
```

(4) 配置 PE 4

配置 Router ID，使能公网实例的 IP 组播路由，配置 MPLS LSR ID，并使能 LDP 能力。

```
<PE4> system-view
[PE4] router id 1.1.1.4
[PE4] multicast routing-enable
[PE4] mpls lsr-id 1.1.1.4
[PE4] mpls
[PE4-mpls] quit
[PE4] mpls ldp
[PE4-mpls-ldp] quit
```

创建 VPN 实例 a，为其配置 RD 和 VPN Target；使能 VPN 实例 a 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE4] ip vpn-instance a
[PE4-vpn-instance-a] route-distinguisher 100:1
[PE4-vpn-instance-a] vpn-target 100:1 export-extcommunity
[PE4-vpn-instance-a] vpn-target 100:1 import-extcommunity
[PE4-vpn-instance-a] multicast routing-enable
[PE4-vpn-instance-a] multicast-domain share-group 239.1.1.1 binding mtunnel 0
[PE4-vpn-instance-a] quit
```

创建 VPN 实例 b，为其配置 RD 和 VPN Target；使能 VPN 实例 b 中的 IP 组播路由，配置 Share-Group 地址，绑定 MTI 接口。

```
[PE4] ip vpn-instance b
[PE4-vpn-instance-b] route-distinguisher 200:1
```

```
[PE4-vpn-instance-b] vpn-target 200:1 export-extcommunity
[PE4-vpn-instance-b] vpn-target 200:1 import-extcommunity
[PE4-vpn-instance-b] multicast routing-enable
[PE4-vpn-instance-b] multicast-domain share-group 239.4.4.4 binding mtunnel 1
[PE4-vpn-instance-b] quit
```

创建业务环回组 1，并指定其业务类型为 Multicast tunnel 类型。

```
[PE4] service-loopback group 1 type multicast-tunnel
```

选择一个未被使用的端口，关闭该端口上的生成树协议、LLDP 功能和 NDP 功能，并将其加入业务环回组 1。

```
[PE4] interface gigabitethernet1/0/3
[PE4-GigabitEthernet1/0/3] undo stp enable
[PE4-GigabitEthernet1/0/3] undo ndp enable
[PE4-GigabitEthernet1/0/3] undo lldp enable
[PE4-GigabitEthernet1/0/3] port service-loopback group 1
[PE4-GigabitEthernet1/0/3] quit
```

在公网接口 Vlan-interface4 上配置 IP 地址，使能 PIM-SM 和 LDP 能力。

```
[PE4] interface vlan-interface 4
[PE4-Vlan-interface4] ip address 10.10.2.2 24
[PE4-Vlan-interface4] pim sm
[PE4-Vlan-interface4] mpls
[PE4-Vlan-interface4] mpls ldp
[PE4-Vlan-interface4] quit
```

将接口 Vlan-interface13 与 VPN 实例 a 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE4] interface vlan-interface 13
[PE4-Vlan-interface13] ip binding vpn-instance a
[PE4-Vlan-interface13] ip address 10.11.3.1 24
[PE4-Vlan-interface13] pim sm
[PE4-Vlan-interface13] quit
```

将接口 Vlan-interface14 与 VPN 实例 b 进行关联，配置 IP 地址，并使能 PIM-SM。

```
[PE4] interface vlan-interface 14
[PE4-Vlan-interface14] ip binding vpn-instance b
[PE4-Vlan-interface14] ip address 10.11.4.1 24
[PE4-Vlan-interface14] pim sm
[PE4-Vlan-interface14] quit
```

配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。

```
[PE4] interface loopback 1
[PE4-LoopBack1] ip address 1.1.1.4 32
[PE4-LoopBack1] pim sm
[PE4-LoopBack1] quit
```

配置 BGP 协议。

```
[PE4] bgp 200
[PE4-bgp] group pe4-pe3 internal
[PE4-bgp] peer pe4-pe3 label-route-capability
[PE4-bgp] peer pe4-pe3 connect-interface loopback 1
[PE4-bgp] peer 1.1.1.3 group pe4-pe3
[PE4-bgp] group pe4-pe1 external
```

```
[PE4-bgp] peer pe4-pe1 as-number 100
[PE4-bgp] peer pe4-pe1 ebgp-max-hop 255
[PE4-bgp] peer 1.1.1.1 group pe4-pe1
[PE4-bgp] peer pe4-pe1 connect-interface loopback 1
[PE4–bgp] ipv4-family vpn-instance a
[PE4-bgp-a] import-route ospf 2
[PE4-bgp-a] import-route direct
[PE4-bgp-a] quit
[PE4–bgp] ipv4-family vpn-instance b
[PE4-bgp-b] import-route ospf 3
[PE4-bgp-b] import-route direct
[PE4-bgp-b] quit
[PE4–bgp] ipv4-family vpnv4
[PE4–bgp-af-vpnv4] peer 1.1.1.1 enable
[PE4–bgp-af-vpnv4] quit
[PE4–bgp] quit
```

当 PE 4 上配置了 BGP 对等体之后，MTI0 和 MTI1 接口将分别自动获得其 IP 地址，该地址与配置 BGP 对等体时所指定的 Loopback 接口的 IP 地址相同；MTI0 接口上运行的 PIM 协议类型与 VPN 实例 a 中运行的 PIM 协议类型相同；MTI1 接口上运行的 PIM 协议类型与 VPN 实例 b 中运行的 PIM 协议类型相同。

配置 OSPF 协议。

```
[PE4] ospf 1
[PE4-ospf-1] area 0.0.0.0
[PE4-ospf-1-area-0.0.0.0] network 1.1.1.4 0.0.0.0
[PE4-ospf-1-area-0.0.0.0] network 10.10.0.0 0.0.255.255
[PE4-ospf-1-area-0.0.0.0] quit
[PE4-ospf-1] quit
[PE4] ospf 2 vpn-instance a
[PE4-ospf-2] import-route bgp
[PE4-ospf-2] area 0.0.0.0
[PE4-ospf-2-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[PE4-ospf-2-area-0.0.0.0] quit
[PE4-ospf-2] quit
[PE4] ospf 3 vpn-instance b
[PE4-ospf-3] import-route bgp
[PE4-ospf-3] area 0.0.0.0
[PE4-ospf-3-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[PE4-ospf-3-area-0.0.0.0] quit
[PE4-ospf-3] quit
```

(5) 配置 CE a1

使能 IP 组播路由。

```
<CEa1> system-view
[CEa1] multicast routing-enable
```

在接口 Vlan-interface10 上配置 IP 地址，并使能 PIM-SM。

```
[CEa1] interface vlan-interface 10
[CEa1-Vlan-interface10] ip address 10.11.5.1 24
[CEa1-Vlan-interface10] pim sm
```

```
[CEa1-Vlan-interface10] quit
```
在接口 Vlan-interface11 上配置 IP 地址，并使能 PIM-SM。
```
[CEa1] interface vlan-interface 11
[CEa1-Vlan-interface11] ip address 10.11.1.2 24
[CEa1-Vlan-interface11] pim sm
[CEa1-Vlan-interface11] quit
```
配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。
```
[CEa1] interface loopback 1
[CEa1-LoopBack1] ip address 2.2.2.2 32
[CEa1-LoopBack1] pim sm
[CEa1-LoopBack1] quit
```
配置 Loopback1 接口为 VPN a 的 C-BSR 和 C-RP。
```
[CEa1] pim
[CEa1-pim] c-bsr loopback 1
[CEa1-pim] c-rp loopback 1
[CEa1-pim] quit
```
配置 OSPF 协议。
```
[CEa1] ospf 1
[CEa1-ospf-1] area 0.0.0.0
[CEa1-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[CEa1-ospf-1-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[CEa1-ospf-1-area-0.0.0.0] quit
[CEa1-ospf-1] quit
```
(6) 配置 CE b1

使能 IP 组播路由。
```
<CEb1> system-view
[CEb1] multicast routing-enable
```
在接口 Vlan-interface20 上配置 IP 地址，并使能 PIM-SM。
```
[CEb1] interface vlan-interface 20
[CEb1-Vlan-interface20] ip address 10.11.6.1 24
[CEb1-Vlan-interface20] pim sm
[CEb1-Vlan-interface20] quit
```
在接口 Vlan-interface12 上配置 IP 地址，并使能 PIM-SM。
```
[CEb1] interface vlan-interface 12
[CEb1-Vlan-interface12] ip address 10.11.2.2 24
[CEb1-Vlan-interface12] pim sm
[CEb1-Vlan-interface12] quit
```
配置 OSPF 协议。
```
[CEb1] ospf 1
[CEb1-ospf-1] area 0.0.0.0
[CEb1-ospf-1-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[CEb1-ospf-1-area-0.0.0.0] quit
[CEb1-ospf-1] quit
```
(7) 配置 CE a2

使能 IP 组播路由。

```
<CEa2> system-view
[CEa2] multicast routing-enable
```
在接口 Vlan-interface30 上配置 IP 地址，使能 IGMP 和 PIM-SM。
```
[CEa2] interface vlan-interface 30
[CEa2-Vlan-interface30] ip address 10.11.7.1 24
[CEa2-Vlan-interface30] igmp enable
[CEa2-Vlan-interface30] pim sm
[CEa2-Vlan-interface30] quit
```
在接口 Vlan-interface13 上配置 IP 地址，并使能 PIM-SM。
```
[CEa2] interface vlan-interface 13
[CEa2-Vlan-interface13] ip address 10.11.3.2 24
[CEa2-Vlan-interface13] pim sm
[CEa2-Vlan-interface13] quit
```
配置 OSPF 协议。
```
[CEa2] ospf 1
[CEa2-ospf-1] area 0.0.0.0
[CEa2-ospf-1-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[CEa2-ospf-1-area-0.0.0.0] quit
[CEa2-ospf-1] quit
```
(8) 配置 CE b2

使能 IP 组播路由。
```
<CEb2> system-view
[CEb2] multicast routing-enable
```
在接口 Vlan-interface40 上配置 IP 地址，使能 IGMP 和 PIM-SM。
```
[CEb2] interface vlan-interface 40
[CEb2-Vlan-interface40] ip address 10.11.8.1 24
[CEb2-Vlan-interface40] igmp enable
[CEb2-Vlan-interface40] pim sm
[CEb2-Vlan-interface40] quit
```
在接口 Vlan-interface14 上配置 IP 地址，并使能 PIM-SM。
```
[CEb2] interface vlan-interface 14
[CEb2-Vlan-interface14] ip address 10.11.4.2 24
[CEb2-Vlan-interface14] pim sm
[CEb2-Vlan-interface14] quit
```
配置 Loopback1 接口的 IP 地址，并使能 PIM-SM。
```
[CEb2] interface loopback 1
[CEb2-LoopBack1] ip address 3.3.3.3 32
[CEb2-LoopBack1] pim sm
[CEb2-LoopBack1] quit
```
配置 Loopback1 接口为 VPN b 的 C-BSR 和 C-RP。
```
[CEb2] pim
[CEb2-pim] c-bsr loopback 1
[CEb2-pim] c-rp loopback 1
[CEb2-pim] quit
```
配置 OSPF 协议。
```
[CEb2] ospf 1
```

```
[CEb2-ospf-1] area 0.0.0.0
[CEb2-ospf-1-area-0.0.0.0] network 3.3.3.3 0.0.0.0
[CEb2-ospf-1-area-0.0.0.0] network 10.11.0.0 0.0.255.255
[CEb2-ospf-1-area-0.0.0.0] quit
[CEb2-ospf-1] quit
```

4. 检验配置效果

通过使用 **display multicast-domain vpn-instance share-group** 命令可以查看指定 VPN 实例的 Share-Group 组信息。

查看 PE 1 上 VPN 实例 a 中的本地 Share-Group 组信息。

```
<PE1> display multicast-domain vpn-instance a share-group local
MD local share-group information for VPN-Instance: a
  Share-group: 239.1.1.1
  MTunnel address: 1.1.1.1
```

查看 PE 1 上 VPN 实例 b 中的本地 Share-Group 组信息。

```
<PE1> display multicast-domain vpn-instance b share-group local
MD local share-group information for VPN-Instance: b
  Share-group: 239.4.4.4
  MTunnel address: 1.1.1.1
```

查看 PE 4 上 VPN 实例 a 中的本地 Share-Group 组信息。

```
<PE4> display multicast-domain vpn-instance a share-group local
MD local share-group information for VPN-Instance: a
  Share-group: 239.1.1.1
  MTunnel address: 1.1.1.4
```

查看 PE 4 上 VPN 实例 b 中的本地 Share-Group 组信息。

```
<PE4> display multicast-domain vpn-instance b share-group local
MD local share-group information for VPN-Instance: b
  Share-group: 239.4.4.4
  MTunnel address: 1.1.1.4
```

1.8 常见配置错误举例

1.8.1 无法建立Share-MDT

1. 故障现象

无法正确建立 Share-MDT，不同 PE 设备上相同的 VPN 实例之间无法建立起 PIM 邻居关系。

2. 分析

- 在不同的 PE 设备上，相同的 VPN 实例需要配置相同的 Share-Group 地址，每个 Share-Group 地址唯一标识一个 Share-MDT，如果不同 PE 设备上的相同 VPN 实例没有配置相同的 Share-Group 地址，则该 VPN 实例在不同的 PE 设备上无法建立起 Share-MDT。
- 在不同的 PE 设备上，相同 VPN 实例的各接口必须使能相同类型的 PIM 协议，P 设备上的所有接口必须使能相同类型的 PIM 协议，这样才能正确地建立 Share-MDT，本地 PE 设备和远端 PE 设备的相同 VPN 实例上才能建立起 PIM 邻居关系。否则无法建立 Share-MDT。

- 只有配置了 BGP 和单播路由，MTI 接口才能自动获得 IP 地址；只有 VPN 实例内至少有一个接口上使能了 PIM 协议，MTI 接口上的 PIM 协议才能被使能，从而使不同 PE 设备的相同 VPN 实例之间建立起 PIM 邻居。否则无法建立起 PIM 邻居关系。

3. 处理过程

(1) 检查 Share-Group 地址。使用命令 **display multicast-domain vpn-instance share-group** 检查不同 PE 设备上相同的 VPN 实例是否配置了相同的 Share-Group 地址。

(2) 检查各设备 VPN 实例内是否有至少一个接口上使能了 PIM 协议，不同 PE 设备上属于同一 VPN 实例的各接口上是否使能了相同类型的 PIM 协议，以及 P 设备的各接口上是否使能了相同类型的 PIM 协议。使用命令 **display pim interface verbose** 查看各接口上的 PIM 信息。

(3) 检查单播路由。使用命令 **display ip routing-table** 检查本地 PE 设备的 VPN 实例是否有到达远端 PE 设备的相同 VPN 实例的单播路由项。

(4) 检查是否配置 BGP 对等体。使用命令 **display bgp peer** 查看配置的 BGP 对等体信息。

1.8.2 VPN实例无法正确建立组播路由表

1. 故障现象

VPN 实例无法正确建立起组播路由表。

2. 分析

- 如果 VPN 实例使能的是 PIM-SM，需要有该 VPN 实例的 BSR 信息，否则不能正确建立该 VPN 实例的组播路由表。
- 如果 VPN 实例使能的是 PIM-SM，需要有该 VPN 实例的 RP 信息，如果没有通向 RP 的单播路由，公网实例和 VPN 实例没有正确建立起 PIM 邻居关系，VPN 实例就无法正确建立组播路由表。
- 私网 DR 需要有到达私网 RP 的路由。

3. 处理过程

(1) 使用 **display pim bsr-info** 命令查看公网实例和 VPN 实例是否有 BSR 信息。如果不存在 BSR 信息，则需要查看是否有通向 BSR 的单播路由。

(2) 使用 **display pim rp-info** 命令查看 RP 信息是否正确。如果没有 RP 信息，则需要检查是否有通向 RP 的单播路由。使用 **display pim neighbor** 命令查看公网和私网上是否正确建立起了邻居关系。

(3) 使用 **ping** 命令检查私网 DR 与私网 RP 之间是否通达。

目 录

1 MLD Snooping配置

1.1 MLD Snooping简介

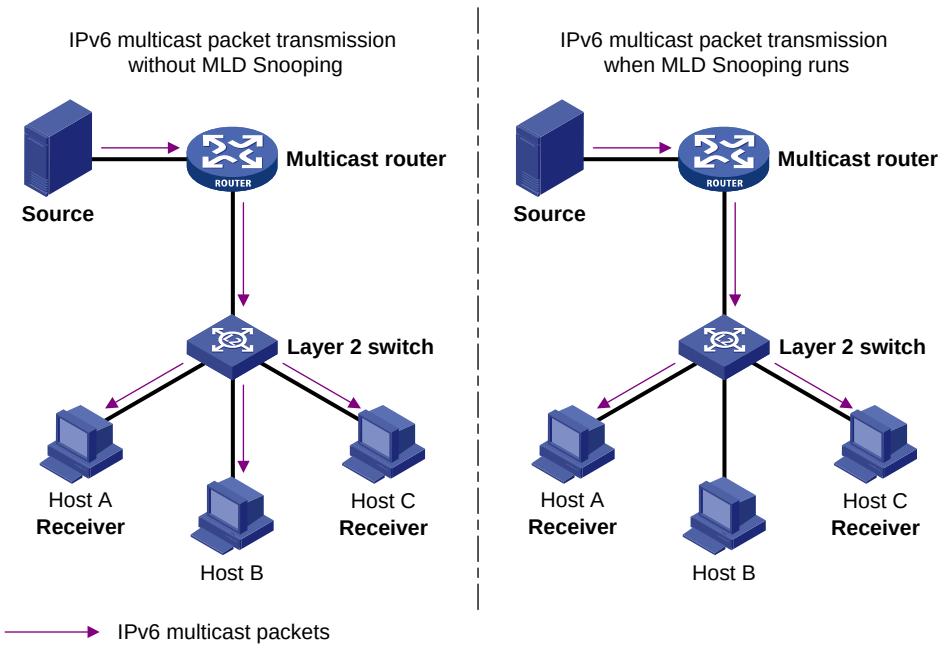
MLD Snooping 是 Multicast Listener Discovery Snooping（组播侦听者发现协议窥探）的简称。它是运行在二层设备上的 IPv6 组播约束机制，用于管理和控制 IPv6 组播组。

1.1.1 MLD Snooping原理

运行 MLD Snooping 的二层设备通过对收到的 MLD 报文进行分析，为端口和 MAC 组播地址建立起映射关系，并根据这样的映射关系转发 IPv6 组播数据。

如 图 1-1 所示，当二层设备没有运行MLD Snooping时，IPv6 组播数据报文在二层被广播；当二层设备运行了MLD Snooping后，已知IPv6 组播组的组播数据报文不会在二层被广播，而在二层被组播给指定的接收者。

图1-1 二层设备运行 MLD Snooping 前后的对比



MLD Snooping 通过二层组播将信息只转发给有需要的接收者，可以带来以下好处：

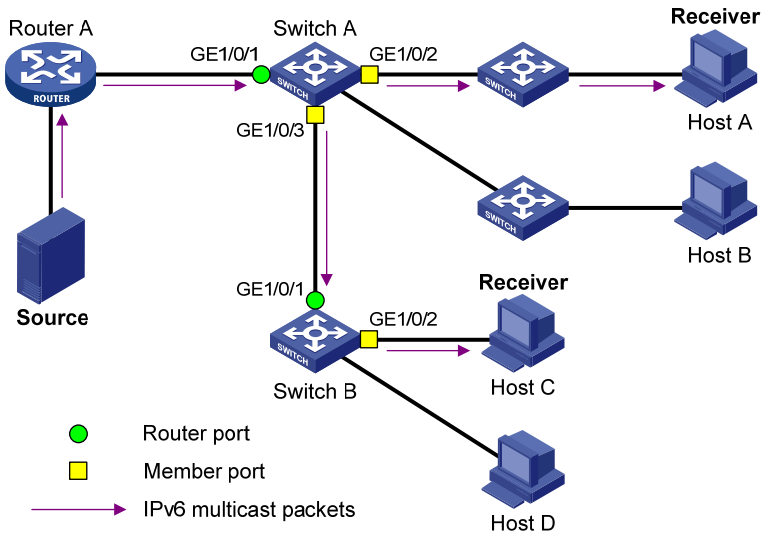
- 减少了二层网络中的广播报文，节约了网络带宽；
- 增强了 IPv6 组播信息的安全性；
- 为实现对每台主机的单独计费带来了方便。

1.1.2 MLD Snooping基本概念

1. MLD Snooping相关端口

如 图 1-2 所示，Router A连接组播源，在Switch A和Switch B上分别运行MLD Snooping，Host A和Host C为接收者主机（即IPv6 组播组成员）。

图1-2 MLD Snooping 相关端口



结合 图 1-2，介绍一下MLD Snooping相关的端口概念：

- 路由器端口（Router Port）：交换机上朝向三层组播设备（DR 或 MLD 查询器）一侧的端口，如 Switch A 和 Switch B 各自的 GigabitEthernet1/0/1 端口。交换机将本设备上的所有路由器端口都记录在路由器端口列表中。
- 成员端口（Member Port）：又称 IPv6 组播组成员端口，表示交换机上朝向 IPv6 组播组成员一侧的端口，如 Switch A 的 GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 端口，以及 Switch B 的 GigabitEthernet1/0/2 端口。交换机将本设备上的所有成员端口都记录在 MLD Snooping 转发表中。

说明

- 本文中提到的路由器端口都是指交换机上朝向组播路由器的端口，而不是指路由器上的端口。
- 如不特别指明，本文中提到的路由器/成员端口均包括动态和静态端口。
- 在运行了 MLD Snooping 的交换机上，所有收到源地址不为 0::0 的 MLD 普遍组查询报文或 IPv6 PIM Hello 报文的端口都将被视为动态路由器端口。有关 IPv6 PIM Hello 报文的详细介绍，请参见“IP 组播配置指导”中的“IPv6 PIM”。

---

2. MLD Snooping动态端口老化定时器

表1-1 MLD Snooping 动态端口老化定时器

定时器	说明	超时前应收到的报文	超时后交换机的动作
动态路由器端口老化定时器	交换机为其每个动态路由器端口都启动一个定时器，其超时时间就是动态路由器端口老化时间	源地址不为0::0的MLD普遍组查询报文或IPv6 PIM Hello报文	将该端口从路由器端口列表中删除
动态成员端口老化定时器	当一个端口动态加入某IPv6组播组时，交换机为该端口启动一个定时器，其超时时间就是动态成员端口老化时间	MLD成员关系报告报文	将该端口从MLD Snooping转发表中删除

说明

MLD Snooping 端口老化机制只针对动态端口，静态端口永不老化。

1.1.3 MLD Snooping工作机制

运行了 MLD Snooping 的交换机对不同 MLD 动作的具体处理方式如下：

注意

本节中所描述的增删端口动作均只针对动态端口，静态端口只能通过相应的配置进行增删，具体步骤请参见“1.4.3 配置静态端口”。

1. 普遍组查询

MLD 查询器定期向本地网段内的所有主机与路由器（FF02::1）发送 MLD 普遍组查询报文，以查询该网段有哪些 IPv6 组播组的成员。

在收到 MLD 普遍组查询报文时，交换机将其通过 VLAN 内除接收端口以外的其它所有端口转发出去，并对该报文的接收端口做如下处理：

- 如果在路由器端口列表中已包含该动态路由器端口，则重置其老化定时器。
- 如果在路由器端口列表中尚未包含该动态路由器端口，则将其添加到路由器端口列表中，并启动其老化定时器。

2. 报告成员关系

以下情况，主机会向 MLD 查询器发送 MLD 成员关系报告报文：

- 当 IPv6 组播组的成员主机收到 MLD 查询报文后，会回复 MLD 成员关系报告报文。
- 如果主机要加入某个 IPv6 组播组，它会主动向 MLD 查询器发送 MLD 成员关系报告报文以声明加入该 IPv6 组播组。

在收到 MLD 成员关系报告报文时，交换机将其通过 VLAN 内的所有路由器端口转发出去，从该报文中解析出主机要加入的 IPv6 组播组地址，并对该报文的接收端口做如下处理：

- 如果不存在该 IPv6 组播组所对应的转发表项，则创建转发表项，将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；

- 如果已存在该 IPv6 组播组所对应的转发表项，但其出端口列表中不包含该端口，则将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；
- 如果已存在该 IPv6 组播组所对应的转发表项，且其出端口列表中已包含该动态成员端口，则重置其老化定时器。

说明

交换机不会将 MLD 成员关系报告报文通过非路由器端口转发出去，因为根据主机上的 MLD 成员关系报告抑制机制，如果非路由器端口下还有该 IPv6 组播组的成员主机，则这些主机在收到该报告报文后便抑制了自身的报告，从而使交换机无法获知这些端口下还有该 IPv6 组播组的成员主机。有关主机上的 MLD 成员关系报告抑制机制的详细介绍，请参见“IP 组播配置指导”中的“MLD”。

---

3. 离开组播组

当主机离开 IPv6 组播组时，会通过发送 MLD 离开组报文，以通知组播路由器自己离开了某个 IPv6 组播组。当交换机从某动态成员端口上收到 MLD 离开组报文时，首先判断要离开的 IPv6 组播组所对应的转发表项是否存在，以及该 IPv6 组播组所对应转发表项的出端口列表中是否包含该接收端口：

- 如果不存在该 IPv6 组播组对应的转发表项，或者该 IPv6 组播组对应转发表项的出端口列表中不包含该端口，交换机不会向任何端口转发该报文，而将其直接丢弃；
- 如果存在该 IPv6 组播组对应的转发表项，且该 IPv6 组播组对应转发表项的出端口列表中包含该端口，交换机会将该报文通过 VLAN 内的所有路由器端口转发出去。同时，由于并不知道该接收端口下是否还有该 IPv6 组播组的其它成员，所以交换机不会立刻把该端口从该 IPv6 组播组所对应转发表项的出端口列表中删除，而是重置其老化定时器。

当 MLD 查询器收到 MLD 离开组报文后，从中解析出主机要离开的 IPv6 组播组的地址，并通过接收端口向该 IPv6 组播组发送 MLD 特定组查询报文。交换机在收到 MLD 特定组查询报文后，将其通过 VLAN 内的所有路由器端口和该 IPv6 组播组的所有成员端口转发出去。对于 MLD 离开组报文的接收端口（假定为动态成员端口），交换机在其老化时间内：

- 如果从该端口收到了主机响应该特定组查询的 MLD 成员关系报告报文，则表示该端口下还有该 IPv6 组播组的成员，于是重置其老化定时器；
- 如果没有从该端口收到主机响应该特定组查询的 MLD 成员关系报告报文，则表示该端口下已没有该 IPv6 组播组的成员，则在其老化时间超时后，将其从该 IPv6 组播组所对应转发表项的出端口列表中删除。

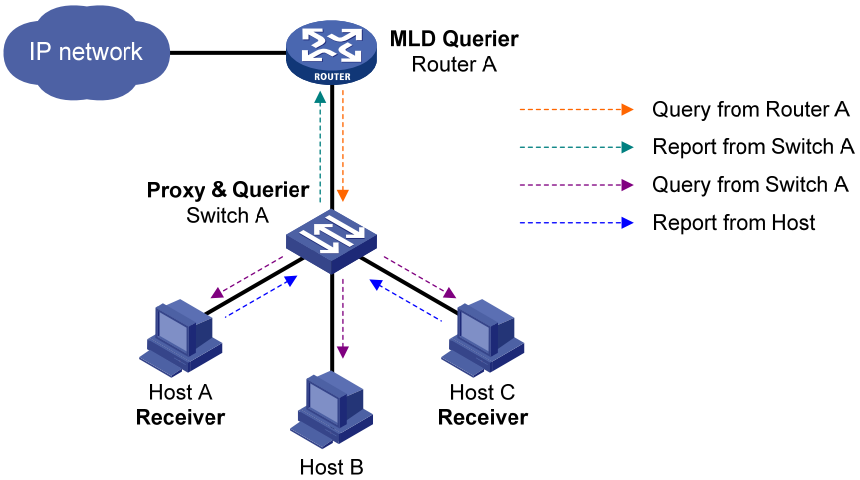
1.1.4 MLD Snooping Proxying

为了减少上游设备收到的 MLD 报告报文和离开报文的数量，可以通过在边缘设备上配置 MLD Snooping Proxying（MLD Snooping 代理）功能，使其能够代理下游主机来向上游设备发送报告报文和离开报文。配置了 MLD Snooping Proxying 功能的设备称为 MLD Snooping 代理设备，在其上游设备看来，它就相当于一台主机。

说明

尽管在其上游设备看来，MLD Snooping 代理设备相当于一台主机，但主机上的 MLD 成员关系报告抑制机制在 MLD Snooping 代理设备上并不会生效。有关主机上的 MLD 成员关系报告抑制机制的详细介绍，请参见“IP 组播配置指导”中的“MLD”。

图1-3 MLD Snooping Proxying 组网图



如 图 1-3 所示，作为MLD Snooping代理设备的Switch A，对其上游的MLD查询器Router A来说相当于一台主机，代理下游主机向Router A发送报告报文和离开报文。

MLD Snooping代理设备对MLD报文的处理方式如 表 1-2 所示。

表1-2 MLD Snooping 代理设备对 MLD 报文的处理方式

MLD 报文类型	处理方式
普遍组查询报文	收到普遍组查询报文后，向本VLAN内除接收端口以外的所有端口转发；同时根据本地维护的组成员关系生成报告报文，并向所有路由器端口发送
特定组查询报文	收到针对某IPv6组播组的特定组查询报文时，若该组对应的转发表项中还有成员端口，则向所有路由器端口回复该组的报告报文
报告报文	从某端口收到某IPv6组播组的报告报文时，若已存在该组对应的转发表项，且其出端口列表中已包含该动态成员端口，则重置其老化定时器；若已存在该组对应的转发表项，但其出端口列表中不包含该端口，则将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；若尚不存在该组对应的转发表项，则创建转发表项，将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器，然后向所有路由器端口发送该组的报告报文
离开报文	从某端口收到某IPv6组播组的离开报文后，向该端口发送针对该组的特定组查询报文。只有当删除某IPv6组播组对应转发表项中的最后一个成员端口时，才会向所有路由器端口发送该组的离开报文

1.1.5 协议规范

与 MLD Snooping 相关的协议规范有：

- RFC 4541：Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches

1.2 MLD Snooping配置任务简介

表1-3 MLD Snooping 配置任务简介

配置任务		说明	详细配置
配置MLD Snooping基本功能	使能MLD Snooping	必选	1.3.2
	配置MLD Snooping版本	可选	1.3.3
	配置MLD Snooping转发表项的最大数量	可选	1.3.4
	配置IPv6静态组播MAC地址表项	可选	1.3.5
配置MLD Snooping端口功能	配置动态端口老化定时器	可选	1.4.2
	配置静态端口	可选	1.4.3
	配置模拟主机加入	可选	1.4.4
	配置端口快速离开	可选	1.4.5
	禁止端口成为动态路由器端口	可选	1.4.6
配置MLD Snooping查询器	使能MLD Snooping查询器	可选	1.5.2
	配置MLD查询和响应	可选	1.5.3
	配置MLD查询报文源IPv6地址	可选	1.5.4
配置MLD Snooping Proxying	使能MLD Snooping Proxying	可选	1.6.2
	配置代理发送MLD报文的源IPv6地址	可选	1.6.3
配置MLD Snooping策略	配置IPv6组播组过滤器	可选	1.7.2
	配置IPv6组播数据报文源端口过滤	可选	1.7.3
	配置丢弃未知IPv6组播数据报文	可选	1.7.4
	配置MLD成员关系报告报文抑制	可选	1.7.5
	配置端口加入的IPv6组播组最大数量	可选	1.7.6
	配置IPv6组播组替换	可选	1.7.7
	配置MLD报文的802.1p优先级	可选	1.7.8
	配置IPv6组播用户控制策略	可选	1.7.9
	配置MLD Snooping主机跟踪功能	可选	1.7.10
	配置MLD Snooping协议中发送的MLD报文的DSCP优先级	可选	1.7.11

说明

对于 MLD Snooping 的相关配置来说:

- MLD-Snooping 视图下的配置对所有 VLAN 都有效，VLAN 视图下的配置只对当前 VLAN 有效。对于某 VLAN 来说，优先采用该 VLAN 视图下的配置，只有当在该 VLAN 视图下没有进行配置时，才采用 MLD-Snooping 视图下的相应配置。
- MLD-Snooping 视图下的配置对所有端口都有效；二层以太网端口视图下的配置只对当前端口有效；二层聚合接口视图下的配置只对当前接口有效；端口组视图下的配置对当前端口组中的所有端口有效。对于某端口来说，优先采用二层以太网端口视图、二层聚合接口视图或端口组视图下的配置，只有当在上述视图下没有进行配置时，才采用 MLD-Snooping 视图下的相应配置。
- 二层聚合接口与其各成员端口上的配置互不影响，且成员端口上的配置只有当该端口退出聚合组后才会生效，二层聚合接口上的配置也不会参与聚合计算。

1.3 配置MLD Snooping基本功能

1.3.1 配置准备

在配置 MLD Snooping 基本功能之前，需完成以下任务：

- 使能 IPv6 转发功能
- 配置相应 VLAN

在配置 MLD Snooping 基本功能之前，需准备以下数据：

- MLD Snooping 的版本

1.3.2 使能MLD Snooping

表1-4 使能 MLD Snooping

操作	命令	说明
进入系统视图	**system-view**	-
全局使能MLD Snooping，并进入MLD-Snooping视图	**mld-snooping**	必选 缺省情况下，MLD Snooping处于关闭状态
退回系统视图	**quit**	-
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能MLD Snooping	**mld-snooping enable**	必选 缺省情况下，VLAN内的MLD Snooping处于关闭状态

说明

- 在VLAN内使能MLD Snooping之前，必须先在系统视图下全局使能MLD Snooping，否则将无法在VLAN内使能MLD Snooping。
- 在VLAN内使能MLD Snooping之后，不允许在该VLAN所对应的VLAN接口上再使能MLD和IPv6 PIM，反之亦然。
- 在指定VLAN内使能了MLD Snooping之后，MLD Snooping功能只在属于该VLAN的端口上生效。

1.3.3 配置MLD Snooping版本

配置MLD Snooping的版本，实际上就是配置MLD Snooping可以处理的MLD报文的版本：

- 当MLD Snooping的版本为1时，MLD Snooping能够对MLDv1的报文进行处理，对MLDv2的报文则不进行处理，而是在VLAN内将其广播；
- 当MLD Snooping的版本为2时，MLD Snooping能够对MLDv1和MLDv2的报文进行处理。

表1-5 配置MLD Snooping版本

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置MLD Snooping的版本	**mld-snooping version** *version-number*	必选 缺省情况下，MLD Snooping的版本为1

注意

当MLD Snooping的版本由版本2切换到版本1时，系统将清除所有通过动态加入的MLD Snooping转发表项；对于在版本2下通过手工配置而静态加入的MLD Snooping转发表项，则分为以下两种情况进行不同的处理：

- 如果配置的仅仅是静态加入IPv6组播组，而没有指定IPv6组播源，则这些转发表项将不会被清除；
- 如果配置的是指定了IPv6组播源的静态加入IPv6组播源组，则这些转发表项将会被清除，并且当再次切换回版本2时，这些转发表项将被重新恢复。

有关静态加入的详细配置，请参见“1.4.3 配置静态端口”。

1.3.4 配置MLD Snooping转发表项的最大数量

用户可以调整MLD Snooping转发表项的最大数量，当设备上维护的表项数量达到最大数量后，将不再创建新的表项，直至有表项被老化或被手工删除。

表1-6 配置 MLD Snooping 转发表项的最大数量

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
配置MLD Snooping转发表项的最大数量	**entry-limit** *limit*	必选 缺省情况下， MLD Snooping转发表项的最大数量为1000

- 在对 MLD Snooping 转发表项的最大数量进行配置时，如果设备上维护的表项数量已超过了配置值，系统将提示用户删除多余表项，但并不会自动删除任何已存在的表项，同时也不再继续创建新的表项。
- IPv6 组播 VLAN 建立的转发表项不受 **entry-limit** 命令的限制，可通过 **multicast-vlan ipv6 entry-limit** 命令进行限制，详情请参见“IP 组播配置指导”中的“IPv6 组播 VLAN 配置”。

1.3.5 配置IPv6 静态组播MAC地址表项

在二层组播中，除了可通过二层 IPv6 组播协议（如 MLD Snooping）动态建立 IPv6 组播 MAC 地址表项外，还可以通过手工方式配置 IPv6 组播 MAC 地址表项，将端口与 IPv6 组播 MAC 地址进行静态绑定，以便灵活控制 IPv6 组播信息送达的目的端口。

1. 系统视图下配置IPv6 静态组播MAC地址表项

表1-7 系统视图下配置 IPv6 静态组播 MAC 地址表项

操作	命令	说明
进入系统视图	**system-view**	-
配置静态组播MAC地址表项	**mac-address multicast** *mac-address* **interface** *interface-list* **vlan** *vlan-id*	必选 缺省情况下，没有配置静态组播MAC地址表项

2. 接口视图下配置IPv6 静态组播MAC地址表项

表1-8 接口视图下配置 IPv6 静态组播 MAC 地址表项

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一 进入以太网端口或二层聚合接口

操作		命令	说明
	进入端口组视图	**port-group manual** *port-group-name*	视图后，下面进行的配置只在当前接口生效；进入端口组视图后，下面进行的配置将在端口组的所有接口生效
配置静态组播MAC地址表项		**mac-address multicast** *mac-address* **vlan** *vlan-id*	必选 缺省情况下，没有配置静态组播MAC地址表项

说明

- 有关 **mac-address multicast** 命令的详细介绍，请参见“IP 组播命令参考”中的“IGMP Snooping”。
- 系统视图下的配置对指定端口有效，而端口视图下的配置只对当前端口（或当前端口组内的所有端口）有效。
- 对于 R5206 之前的版本，可手工配置的 IPv6 组播 MAC 地址表项为除 3333-xxxx-xxxx 以外的任意 IPv6 组播 MAC 地址（组播 MAC 地址就是最高字节的最低比特位为 1 的 MAC 地址），其中 x 代表 0～F 的任意一个十六进制数。对于 R5206 及以上版本，可手工配置的 IPv6 组播 MAC 地址表项为任意合法的组播 MAC 地址。

1.4 配置MLD Snooping端口功能

1.4.1 配置准备

在配置 MLD Snooping 端口功能之前，需完成以下任务：

- 在 VLAN 内使能 MLD Snooping
- 配置相应端口组

在配置 MLD Snooping 端口功能之前，需准备以下数据：

- 动态路由器端口老化时间
- 动态成员端口老化时间
- IPv6 组播组和 IPv6 组播源的地址

1.4.2 配置动态端口老化定时器

对于动态路由器端口，如果在其老化时间超时前没有收到 MLD 普遍组查询报文或者 IPv6 PIM Hello 报文，交换机将把该端口从路由器端口列表中删除。

对于动态成员端口，如果在其老化时间超时前没有收到该 IPv6 组播组的 MLD 成员关系报告报文，交换机将把该端口从该 IPv6 组播组所对应转发表的出端口列表中删除。

如果 IPv6 组播组成员的变动比较频繁，可以把动态成员端口老化时间设置小一些，反之亦然。

1. 全局配置动态端口老化定时器

表1-9 全局配置动态端口老化定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
配置动态路由器端口老化时间	**router-aging-time** *interval*	必选 缺省情况下，动态路由器端口的老化时间为260秒
配置动态成员端口老化时间	**host-aging-time** *interval*	必选 缺省情况下，动态成员端口的老化时间为260秒

2. 在VLAN内配置动态端口老化定时器

表1-10 在 VLAN 内配置动态端口老化定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置动态路由器端口老化时间	**mld-snooping** **router-aging-time** *interval*	必选 缺省情况下，动态路由器端口的老化时间为260秒
配置动态成员端口老化时间	**mld-snooping** **host-aging-time** *interval*	必选 缺省情况下，动态成员端口的老化时间为260秒

1.4.3 配置静态端口

如果某端口所连接的主机需要固定接收发往某 IPv6 组播组的 IPv6 组播数据，可以配置该端口静态加入该 IPv6 组播组，成为静态成员端口。

可以通过将交换机上的端口配置为静态路由器端口，从而使交换机上所有收到的 IPv6 组播数据可以通过该端口被转发出去。

表1-11 配置静态端口

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置静态成员端口		**mld-snooping** **static-group** *ipv6-group-address* [**source-ip** *ipv6-source-address*] **vlan** *vlan-id*	必选 缺省情况下，端口不是静态成员端口

操作	命令	说明
配置静态路由器端口	**mld-snooping static-router-port vlan** *vlan-id*	必选 缺省情况下，端口不是静态路由器端口

- 静态成员端口不会对 MLD 查询器发出的查询报文进行响应；当配置静态成员端口或取消静态成员端口的配置时，端口也不会主动发送 MLD 成员关系报告报文或 MLD 离开组报文。
- 静态成员端口和静态路由器端口都不会老化，只能通过相应的 **undo** 命令删除。

1.4.4 配置模拟主机加入

通常情况下，运行 MLD 的主机会对 MLD 查询器发出的查询报文进行响应。如果主机由于某种原因无法响应，就可能导致组播路由器认为该网段没有该 IPv6 组播组的成员，从而取消相应的转发路径。

为避免这种情况的发生，可以将交换机的端口配置成为 IPv6 组播组成员（即配置模拟主机加入）。当收到 MLD 查询报文时由模拟主机进行响应，从而保证该交换机能够继续收到 IPv6 组播报文。

模拟主机加入功能的实现原理如下：

- 在某端口上使能模拟主机加入功能时，交换机会通过该端口主动发送一个 MLD 成员关系报告报文；
- 在某端口上使能了模拟主机加入功能后，当收到 MLD 普遍组查询报文时，交换机会通过该端口响应一个 MLD 成员关系报告报文；
- 在某端口上关闭模拟主机加入功能时，交换机会通过该端口发送一个 MLD 离开组报文。

表1-12 配置模拟主机加入

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置模拟主机加入IPv6组播组或组播源组		**mld-snooping host-join** *ipv6-group-address* [**source-ip** *ipv6-source-address*] **vlan** *vlan-id*	必选 缺省情况下，没有配置模拟主机加入IPv6组播组或组播源组

说明

- 每配置一次模拟主机加入，即相当于启动了一台独立的主机。例如，当收到 MLD 查询报文时，每条配置所对应的模拟主机将分别进行响应。
- 与静态成员端口不同，配置了模拟主机加入的端口会作为动态成员端口而参与动态成员端口的老化过程。

1.4.5 配置端口快速离开

端口快速离开是指当交换机从某端口收到主机发送的离开某 IPv6 组播组的 MLD 离开组报文时，直接把该端口从对应转发表项的出端口列表中删除。此后，当交换机收到对该 IPv6 组播组的 MLD 特定组查询报文时，交换机将不再向该端口转发。

在交换机上，在只连接有一个接收者的端口上，可以通过使能端口快速离开功能来节约带宽和资源；而在连接有多个接收者的端口上，如果交换机或该端口所在的 VLAN 已使能了丢弃未知 IPv6 组播数据报文功能，则不要再使能端口快速离开功能，否则，一个接收者的离开将导致该端口下属于同一 IPv6 组播组的其它接收者无法收到 IPv6 组播数据。

1. 全局配置端口快速离开

表1-13 全局配置端口快速离开

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
使能端口快速离开功能	**fast-leave** [**vlan** *vlan-list*]	必选 缺省情况下，端口快速离开功能处于关闭状态

2. 在端口上配置端口快速离开

表1-14 在端口上配置端口快速离开

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
使能端口快速离开功能		**mld-snooping fast-leave** [**vlan** *vlan-list*]	必选 缺省情况下，端口快速离开功能处于关闭状态

1.4.6 禁止端口成为动态路由器端口

目前，在 IPv6 组播用户接入网络中存在以下问题：

- 如果交换机收到了某用户主机发来的 MLD 普遍组查询报文或 IPv6 PIM Hello 报文，那么该主机所在的端口就将成为动态路由器端口，从而使 VLAN 内的所有 IPv6 组播报文都会向该端口转发，导致该用户主机收到的 IPv6 组播报文失控。
- 同时，用户主机发送 MLD 普遍组查询报文或 IPv6 PIM Hello 报文，也会影响该接入网络中三层设备上的 IPv6 组播路由协议状态（如影响 MLD 查询器或 DR 的选举），严重时可能导致网络中断。

当禁止某端口成为动态路由器端口后，即使该端口收到了 MLD 普遍组查询报文或 IPv6 PIM Hello 报文，该端口也不会成为动态路由器端口，从而能够有效解决上述问题，提高网络的安全性和对 IPv6 组播用户的控制能力。

表1-15 禁止端口成为动态路由器端口

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
禁止端口成为动态路由器端口		**mld-snooping router-port-deny** [**vlan** *vlan-list*]	必选 缺省情况下，不禁止端口成为动态路由器端口

说明

本配置与静态路由器端口的配置互不影响。

1.5 配置MLD Snooping查询器

1.5.1 配置准备

在配置 MLD Snooping 查询器之前，需完成以下任务：

- 在 VLAN 内使能 MLD Snooping

在配置 MLD Snooping 查询器之前，需准备以下数据：

- 发送 MLD 普遍组查询报文的时间间隔
- 发送 MLD 特定组查询报文的时间间隔
- MLD 普遍组查询的最大响应时间
- MLD 普遍组查询报文的源 IPv6 地址
- MLD 特定组查询报文的源 IPv6 地址

1.5.2 使能MLD Snooping查询器

在运行了 MLD 的 IPv6 组播网络中，会有一台三层组播设备充当 MLD 查询器，负责发送 MLD 查询报文，使三层组播设备能够在网络层建立并维护 IPv6 组播转发表项，从而在网络层正常转发 IPv6 组播数据。

但是，在一个没有三层组播设备的网络中，由于二层设备并不支持 MLD，因此无法实现 MLD 查询器的相关功能。为了解决这个问题，可以在二层设备上使能 MLD Snooping 查询器，使二层设备能够在数据链路层建立并维护 IPv6 组播转发表项，从而在数据链路层正常转发 IPv6 组播数据。

表1-16 使能 MLD Snooping 查询器

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
使能MLD Snooping查询器	**mld-snooping querier**	必选 缺省情况下，MLD Snooping查询器处于关闭状态

注意

尽管 MLD Snooping 查询器并不参与 MLD 查询器的选举，但在运行了 MLD 的 IPv6 组播网络中，配置 MLD Snooping 查询器不但没有实际的意义，反而可能会由于其发送的 MLD 普遍组查询报文的源 IPv6 地址较小而影响 MLD 查询器的选举。

有关 MLD 查询器的详细介绍，请参见“IP 组播配置指导”中的“MLD”。

1.5.3 配置MLD查询和响应

可以根据网络的实际情况来修改发送 MLD 普遍组查询报文的时间间隔。

在收到 MLD 查询报文（包括普遍组查询和特定组查询）后，主机会为其所加入的每个 IPv6 组播组都启动一个定时器，定时器的值在 0 到最大响应时间（该时间值由主机从所收到的 MLD 查询报文的最大响应时间字段获得）中随机选定，当定时器的值减为 0 时，主机就会向该定时器对应的 IPv6 组播组发送 MLD 成员关系报告报文。

合理配置 MLD 查询的最大响应时间，既可以使主机对 MLD 查询报文做出快速响应，又可以减少由于定时器同时超时，造成大量主机同时发送报告报文而引起的网络拥塞：

- 对于 MLD 普遍组查询报文来说，通过配置 MLD 普遍组查询的最大响应时间来填充其最大响应时间字段；
- 对于 MLD 特定组查询报文来说，所配置的发送 MLD 特定组查询报文的时间间隔将被填充到其最大响应时间字段。也就是说，MLD 特定组查询的最大响应时间从数值上与发送 MLD 特定组查询报文的时间间隔相同。

1. 全局配置MLD查询和响应

表1-17 全局配置 MLD 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
配置MLD普遍组查询的最大响应时间	**max-response-time** *interval*	必选 缺省情况下，MLD普遍组查询的最大响应时间为10秒
配置发送MLD特定组查询报文的时间间隔	**last-listener-query-interval** *interval*	必选 缺省情况下，发送MLD特定组查询报文的时间间隔为1秒

2. 在VLAN内配置MLD查询和响应

表1-18 在 VLAN 内配置 MLD 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置发送MLD普遍组查询报文的时间间隔	**mld-snooping query-interval** *interval*	必选 缺省情况下，发送MLD普遍组查询报文的时间间隔为125秒
配置MLD普遍组查询的最大响应时间	**mld-snooping max-response-time** *interval*	必选 缺省情况下，MLD普遍组查询的最大响应时间为10秒
配置发送MLD特定组查询报文的时间间隔	**mld-snooping last-listener-query-interval** *interval*	必选 缺省情况下，发送MLD特定组查询报文的时间间隔为1秒

注意

应确保发送 MLD 普遍组查询报文的时间间隔大于 MLD 普遍组查询的最大响应时间，否则有可能造成对 IPv6 组播组成员的误删。

1.5.4 配置MLD查询报文源IPv6 地址

可以通过此项配置改变 MLD 查询报文的源 IPv6 地址。

表1-19 配置 MLD 查询报文源 IPv6 地址

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置MLD普遍组查询报文源IPv6地址	**mld-snooping general-query source-ip** { *ipv6-address* | **current-interface** }	必选 缺省情况下，MLD普遍组查询报文的源IPv6地址为FE80::02FF:FFFF:FE00:0001
配置MLD特定组查询报文源IPv6地址	**mld-snooping special-query source-ip** { *ipv6-address* | **current-interface** }	必选 缺省情况下，MLD特定组查询报文的源IPv6地址为FE80::02FF:FFFF:FE00:0001

注意

MLD 查询报文源 IPv6 地址的改变可能会影响网段内 MLD 查询器的选举。

1.6 配置MLD Snooping Proxying

1.6.1 配置准备

在配置 MLD Snooping Proxying 之前，需完成以下任务：

- 在 VLAN 内使能 MLD Snooping

在配置 MLD Snooping Proxying 之前，需准备以下数据：

- 代理发送 MLD 报告报文的源 IPv6 地址
- 代理发送 MLD 离开报文的源 IPv6 地址

1.6.2 使能MLD Snooping Proxying

当在有组播需求的 VLAN 内使能了 MLD Snooping 代理功能后，该设备就成为该 VLAN 内的 MLD Snooping 代理设备。

表1-20 使能 MLD Snooping Proxying

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能MLD Snooping代理功能	**mld-snooping proxying enable**	必选 缺省情况下，VLAN内的MLD Snooping代理功能处于关闭状态

1.6.3 配置代理发送MLD报文的源IPv6 地址

通过本配置可以改变代理发送的 MLD 报告报文和离开报文的源 IPv6 地址。

表1-21 配置代理发送 MLD 报文的源 IPv6 地址

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置代理发送MLD报告报文的源IPv6地址	**mld-snooping report source-ip** { *ipv6-address* | **current-interface** }	必选 缺省情况下，代理发送MLD报告报文的源IPv6地址为FE80::02FF:FFFF:FE00:0001
配置代理发送MLD离开报文的源IPv6地址	**mld-snooping done source-ip** { *ipv6-address* | **current-interface** }	必选 缺省情况下，代理发送MLD离开报文的源IPv6地址为FE80::02FF:FFFF:FE00:0001

1.7 配置MLD Snooping策略

1.7.1 配置准备

在配置 MLD Snooping 策略之前，需完成以下任务：

- 在 VLAN 内使能 MLD Snooping

在配置 MLD Snooping 策略之前，需准备以下数据：

- IPv6 组播组过滤的 IPv6 ACL 规则
- 端口加入的 IPv6 组播组最大数量
- MLD 报文的 802.1p 优先级

1.7.2 配置IPv6 组播组过滤器

在使能了 MLD Snooping 的交换机上，通过配置 IPv6 组播组过滤器，可以限制用户对组播节目的点播。

在实际应用中，当用户点播某个组播节目时，主机会发起一个 MLD 成员关系报告报文，该报文到达交换机后，进行 ACL 检查：如果该接收端口可以加入这个 IPv6 组播组，则将其列入到 MLD Snooping 转发表中；否则交换机就丢弃该报文。这样，未通过 ACL 检查的 IPv6 组播数据就不会送到该端口，从而达到控制用户点播组播节目的目的。

1. 全局配置IPv6 组播组过滤器

表1-22 全局配置 IPv6 组播组过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-

操作	命令	说明
配置IPv6组播组过滤器	**group-policy** *acl6-number* [**vlan** *vlan-list*]	必选 缺省情况下，没有配置全局IPv6组播组过滤器，即各VLAN内主机可以加入任意合法的IPv6组播组

2. 在端口上配置IPv6 组播组过滤器

表1-23 在端口上配置 IPv6 组播组过滤器

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置IPv6组播组过滤器		**mld-snooping group-policy** *acl6-number* [**vlan** *vlan-list*]	必选 缺省情况下，端口上没有配置IPv6组播组过滤器，即该端口下的主机可以加入任意合法的IPv6组播组

说明

- 当在 IPv6 组播 VLAN 中配置组播组过滤器时，需在 sub-vlan 上进行配置，在主 VLAN 上配置不会生效。
- 当使用 MLD v2 的用户点播多个组播组时，组播组过滤器功能可能不能正确对用户点播的组播组进行过滤，因为 MLD v2 会将多个组播组的 report 信息在一个报文中发送，从而导致组播组过滤的失效。

1.7.3 配置IPv6 组播数据报文源端口过滤

通过配置 IPv6 组播数据报文源端口过滤功能，可以允许或禁止端口作为组播源端口：

- 使能了该功能后，端口下不能连接组播源，因为该端口将过滤掉所有的 IPv6 组播数据报文（但允许 IPv6 组播协议报文通过），只能连接 IPv6 组播数据接收者；
- 关闭了该功能后，端口下既可以连接组播源，也可以连接 IPv6 组播数据接收者。

1. 全局配置IPv6 组播数据报文源端口过滤

表1-24 全局配置 IPv6 组播数据报文源端口过滤

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-

操作	命令	说明
使能IPv6组播数据报文源端口过滤功能	**source-deny port** *interface-list*	必选 缺省情况下，IPv6组播数据报文源端口过滤功能处于关闭状态

2. 在端口上配置IPv6 组播数据报文源端口过滤

表1-25 在端口上配置 IPv6 组播数据报文源端口过滤

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网端口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
使能IPv6组播数据报文源端口过滤功能		**mld-snooping source-deny**	必选 缺省情况下，IPv6组播数据报文源端口过滤功能处于关闭状态

1.7.4 配置丢弃未知IPv6 组播数据报文

未知 IPv6 组播数据报文是指在 MLD Snooping 转发表中不存在对应转发表项的那些 IPv6 组播数据报文，当交换机收到发往未知组播组的 IPv6 报文时，IPv6 数据报文会在 VLAN 内广播，这样会占用大量的网络带宽，影响转发效率。

此时用户可以在交换机上启动丢弃未知 IPv6 组播数据报文功能，当交换机收到未知 IPv6 组播数据报文时，只向其路由器端口转发，不在 VLAN 内广播。如果交换机没有路由器端口，IPv6 数据报文会被丢弃，不再转发。

表1-26 在 VLAN 内配置丢弃未知 IPv6 组播数据报文

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
使能丢弃未知IPv6组播数据报文功能	**mld-snooping drop-unknown**	必选 缺省情况下，丢弃未知IPv6组播数据报文的功能处于关闭状态，即对未知IPv6组播数据报文进行广播

1.7.5 配置MLD成员关系报告报文抑制

当二层设备收到来自某 IPv6 组播组成员的 MLD 成员关系报告报文时，会将该报文转发给与其直连的三层设备。这样，当二层设备上存在属于某 IPv6 组播组的多个成员时，与其直连的三层设备会收到这些成员发送的相同 MLD 成员关系报告报文。

当使能了 MLD 成员关系报告报文抑制功能后，在一个查询间隔内二层设备只会把收到的某 IPv6 组播组内的第一个 MLD 成员关系报告报文转发给三层设备，而不继续向三层设备转发来自同一组播组的其它 MLD 成员关系报告报文，这样可以减少网络中的报文数量。

表1-27 配置 MLD 成员关系报告报文抑制

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
使能MLD成员关系报告报文抑制功能	**report-aggregation**	必选 缺省情况下，MLD成员关系报告报文抑制功能处于使能状态

注意

在 MLD Snooping 代理设备上，不论是否使能了 MLD 成员关系报告报文抑制功能，只要存在某 IPv6 组播组对应的转发表项，就会将从下游收到的针对该组的报告报文都抑制掉。

1.7.6 配置端口加入的IPv6 组播组最大数量

通过配置端口加入的 IPv6 组播组的最大数量，可以限制用户点播组播节目的数量，从而控制了端口上的数据流量。

表1-28 配置端口加入的 IPv6 组播组最大数量

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置端口加入的IPv6组播组最大数量		**mld-snooping group-limit** *limit* [**vlan** *vlan-list*]	必选 缺省情况下，端口加入的IPv6组播组最大数量为1000

说明

在配置端口加入的 IPv6 组播组最大数量时，如果当前端口上的 IPv6 组播组数量已超过配置值，系统将把该端口相关的所有转发表项从 MLD Snooping 转发表中删除，该端口下的主机都需要重新加入 IPv6 组播组，直至该端口上的 IPv6 组播组数量达到限制值为止。其中，如果该端口已配置为静态成员端口，系统会将静态成员端口的配置重新生效一次；如果在该端口上配置了模拟主机加入，系统在收到模拟主机发来的报告报文之后才会重新建立相应的转发表项。

1.7.7 配置IPv6 组播组替换

由于某些特殊的原因，当前交换机或端口上通过的 IPv6 组播组数目有可能会超过交换机或该端口的限定；另外，在某些特定的应用中，交换机上新加入的 IPv6 组播组需要自动替换已存在的 IPv6 组播组（一个典型的应用就是“频道切换”，即用户通过加入一个新的 IPv6 组播组就能完成离开原 IPv6 组播组并切换到新 IPv6 组播组的动作）。

针对以上情况，可以在交换机或者某些端口上使能 IPv6 组播组替换功能。当交换机或端口上加入的 IPv6 组播组数量已达到限定值时：

- 若使能了 IPv6 组播组替换功能，则新加入的 IPv6 组播组会自动替代已存在的 IPv6 组播组，替代规则是替代 IPv6 地址最小的 IPv6 组播组；
- 若没有使能 IPv6 组播组替换功能，则自动丢弃新的 MLD 成员关系报告报文。

1. 全局配置IPv6 组播组替换

表1-29 全局配置 IPv6 组播组替换

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
使能IPv6组播组替换功能	**overflow-replace** [**vlan** *vlan-list*]	必选 缺省情况下，IPv6组播组替换功能处于关闭状态

2. 在端口上配置IPv6 组播组替换

表1-30 在端口上配置 IPv6 组播组替换

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
使能IPv6组播组替换功能		**mld-snooping overflow-replace** [**vlan** *vlan-list*]	必选 缺省情况下，IPv6组播组替换功能处于关闭状态

注意

在使能IPv6 组播组替换功能之前，必须首先配置端口通过的IPv6 组播组的最大数量（具体配置过程请参见“1.7.6 配置端口加入的IPv6 组播组最大数量”），否则IPv6 组播组替换功能将不会生效。

1.7.8 配置MLD报文的 802.1p优先级

可以通过本配置来改变 MLD 报文的 802.1p 优先级。当交换机的出端口发生拥塞时，交换机通过识别报文的 802.1p 优先级，优先发送优先级较高的报文。

1. 全局配置MLD报文的 802.1p优先级

表1-31 全局配置 MLD 报文的 802.1p 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
配置MLD报文的802.1p优先级	**dot1p-priority** *priority-number*	必选 缺省情况下，MLD报文的802.1p优先级为0

2. 在VLAN内配置MLD报文的 802.1p优先级

表1-32 在 VLAN 内配置 MLD 报文的 802.1p 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
配置MLD报文的802.1p优先级	**mld-snooping dot1p-priority** *priority-number*	必选 缺省情况下，MLD报文的802.1p优先级为0

1.7.9 配置IPv6 组播用户控制策略

IPv6 组播用户控制策略通常配置在接入交换机上，是基于用户权限控制来实现的，即只有通过授权的用户才能收到相应的 IPv6 组播流，从而达到限制用户对 IPv6 组播节目点播的目的。在实际应用中，用户先要通过接入交换机向 RADIUS 服务器发起认证（如 802.1X 认证），当认证通过后再根据用户的点播行为进行策略检查：

- 当用户点播 IPv6 组播节目时，主机会发送 MLD 成员关系报告报文，接入交换机收到该报文后对其携带的 IPv6 组播组和 IPv6 组播源地址进行策略检查，若该报文通过检查则允许该用户加入该 IPv6 组播组；否则，接入交换机将丢弃该报文。

- 当用户停止点播 IPv6 组播节目时，主机会发送 MLD 离开报文，接入交换机收到该报文后对其携带的 IPv6 组播组和 IPv6 组播源地址进行策略检查，若该报文通过检查则允许该用户离开该 IPv6 组播组；否则，接入交换机将丢弃该报文。

表1-33 配置 IPv6 组播用户控制策略

操作	命令	说明
进入系统视图	**system-view**	-
创建User Profile，并进入User-Profile视图	**user-profile** *profile-name*	-
配置IPv6组播用户控制策略	**mld-snooping access-policy** *acl6-number*	必选 缺省情况下，没有配置IPv6组播用户控制策略，即用户可以加入/离开任意合法的IPv6组播组
退回系统视图	**quit**	-
激活该User Profile	**user-profile** *profile-name* **enable**	必选 缺省情况下，User Profile处于未激活状态

说明

- 有关 **user-profile** 和 **user-profile enable** 命令的详细介绍，请参见“安全命令参考”中的“User Profile”。
- IPv6 组播用户控制策略与 IPv6 组播组过滤器在功能上类似，二者的区别在于：前者是基于用户权限的组播控制，需要与认证、授权功能结合使用，能够控制组播用户的加入与离开；后者是基于设备端口的组播控制，无需与认证、授权功能结合使用，只能控制组播用户的加入。

1.7.10 配置MLD Snooping主机跟踪功能

通过使能 MLD Snooping 主机跟踪功能，可以使交换机能够记录正在接收 IPv6 组播数据的成员主机信息（包括主机的 IPv6 地址、运行时间和超时时间等），以便于网络管理员对这些主机进行监控和管理。

1. 全局配置MLD Snooping主机跟踪功能

表1-34 全局配置 MLD Snooping 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
全局使能MLD Snooping主机跟踪功能	**host-tracking**	必选 缺省情况下，MLD Snooping主机跟踪功能处于关闭状态

2. 在VLAN内配置MLD Snooping主机跟踪功能

表1-35 在 VLAN 内配置 MLD Snooping 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
在 VLAN 内使能 MLD Snooping主机跟踪功能	**mld-snooping host-tracking**	必选 缺省情况下，MLD Snooping主机跟踪功能处于关闭状态

1.7.11 配置MLD Snooping协议中发送的MLD报文的DSCP优先级

在 IPv6 报文头中，包含一个 8bit 的 Traffic class 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 MLD Snooping 协议中发送的 MLD 报文的 DSCP 优先级进行配置。

表1-36 配置发送的 MLD 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD-Snooping视图	**mld-snooping**	-
配置发送的MLD报文的DSCP优先级	**dscp** *dscp-value*	必选 缺省情况下，发送的MLD报文的DSCP优先级为48

说明

此配置仅用于配置交换机自身生成的 MLD 报文的 DSCP 优先级，不会对交换机转发的 MLD 报文的 DSCP 优先级进行修改。

1.8 MLD Snooping显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 MLD Snooping 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 IPv6 组播组信息。

表1-37 MLD Snooping 显示和维护

配置	命令
查看MLD Snooping组的信息	**display mld-snooping group** [**vlan** *vlan-id*] [**slot** *slot-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

配置	命令
查看MLD Snooping跟踪的主机信息	**display mld-snooping host vlan** *vlan-id* **group** *ipv6-group-address* [**source** *ipv6-source-address*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6静态组播MAC地址表信息	**display mac-address** [*mac-address* [**vlan** *vlan-id*] | [**multicast**] [**vlan** *vlan-id*] [**count**]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看MLD Snooping监听到的MLD报文的统计信息	**display mld-snooping statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除MLD Snooping组的动态加入记录	**reset mld-snooping group** { *ipv6-group-address* | **all** } [**vlan** *vlan-id*]
清除MLD Snooping监听到的所有MLD报文的统计信息	**reset mld-snooping statistics**

说明

- 有关 **display mac-address multicast** 命令的详细介绍，请参见“IP 组播命令参考”中的“IGMP Snooping”。
- **reset mld-snooping group** 命令只对使能了 MLD Snooping 的 VLAN 有效，而对 VLAN 接口上使能了 MLD 的 VLAN 无效。
- **reset mld-snooping group** 只能清除动态加入记录，而无法清除静态加入记录。

1.9 MLD Snooping典型配置举例

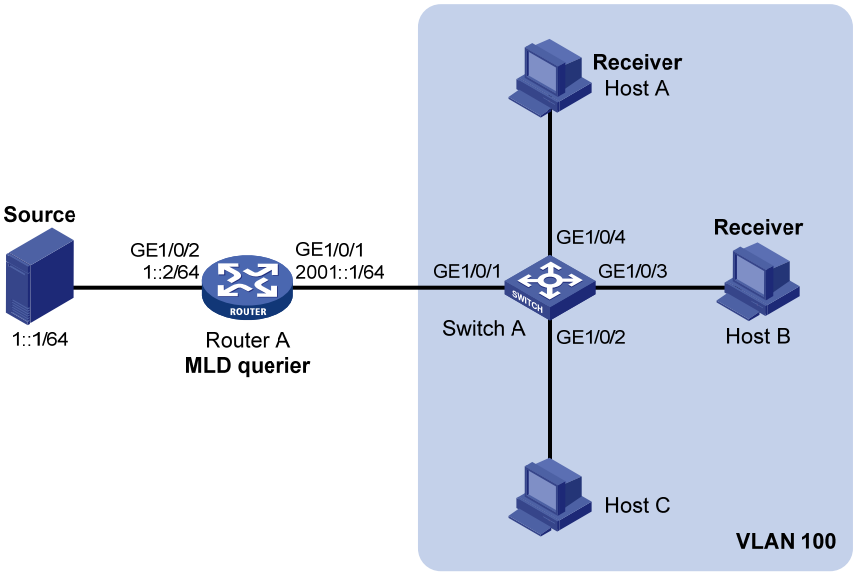
1.9.1 IPv6 组策略及模拟主机加入配置举例

1. 组网需求

- 如 图 1-4 所示，Router A通过GigabitEthernet1/0/2 端口连接IPv6 组播源（Source），通过GigabitEthernet1/0/1 端口连接Switch A；Router A上运行MLDv1，Switch A上运行版本 1 的MLD Snooping，并由Router A充当MLD查询器。
- 通过配置，使 Host A 和 Host B 能且只能接收发往 IPv6 组播组 FF1E::101 的 IPv6 组播数据，并且当 Host A 和 Host B 即使发生意外而临时中断接收 IPv6 组播数据时，发往 IPv6 组播组 FF1E::101 的 IPv6 组播数据也能不间断地通过 Switch A 的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/4 转发出去；同时，使 Switch A 将收到的未知 IPv6 组播数据直接丢弃，避免在其所属的 VLAN 内广播。

2. 组网图

图1-4 IPv6 组策略及模拟主机加入配置组网图



3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址

使能各设备的IPv6 转发功能，并按照 图 1-4 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 配置 Router A

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在端口 GigabitEthernet1/0/1 上使能 MLD。

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] mld enable
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm
[RouterA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch A

全局使能 MLD Snooping。

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中；在该 VLAN 内使能 MLD Snooping，并使能丢弃未知 IPv6 组播数据报文功能。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] mld-snooping drop-unknown
[SwitchA-vlan100] quit
```

配置 IPv6 组播组过滤器，以限定 VLAN 100 内的主机只能加入 IPv6 组播组 FF1E::101。

```
[SwitchA] acl ipv6 number 2001
[SwitchA-acl6-basic-2001] rule permit source ff1e::101 128
[SwitchA-acl6-basic-2001] quit
[SwitchA] mld-snooping
[SwitchA–mld-snooping] group-policy 2001 vlan 100
[SwitchA–mld-snooping] quit
```

在 GigabitEthernet1/0/3 和 GigabitEthernet1/0/4 上分别配置模拟主机加入 IPv6 组播组 FF1E::101。

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] mld-snooping host-join ff1e::101 vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
[SwitchA] interface gigabitethernet 1/0/4
[SwitchA-GigabitEthernet1/0/4] mld-snooping host-join ff1e::101 vlan 100
[SwitchA-GigabitEthernet1/0/4] quit
```

(4) 检验配置效果

查看 Switch A 上 VLAN 100 内 MLD Snooping 组的详细信息。

```
[SwitchA] display mld-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port(s).
            GE1/0/1                (D) ( 00:01:30 )
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:FF1E::101
        (::, FF1E::101):
          Attribute:    Host Port
          Host port(s):total 2 port(s).
            GE1/0/3                (D) ( 00:03:23 )
            GE1/0/4                (D) ( 00:04:10 )
    MAC group(s):
      MAC group address:3333-0000-0101
          Host port(s):total 2 port(s).
            GE1/0/3
            GE1/0/4
```

由此可见，Switch A 上的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/4 已经加入了 IPv6 组播组 FF1E::101。

1.9.2 静态端口配置举例

1. 组网需求

- 如 图 1-5 所示，Router A通过GigabitEthernet1/0/2 端口连接IPv6 组播源（Source），通过GigabitEthernet1/0/1 端口连接Switch A；Router A上运行MLDv1，Switch A、Switch B和Switch C上运行版本 1 的MLD Snooping，并由Router A充当MLD查询器。
- Host A 和 Host C 均为 IPv6 组播组 FF1E::101 的固定接收者（Receiver），通过将 Switch C 上的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/5 配置为 IPv6 组播组 FF1E::101 的静态成员端口，可以增强 IPv6 组播数据在传输过程中的可靠性。
- 假设由于受 STP 等链路层协议的影响，为了避免出现环路，Switch A—Switch C 的转发路径在正常情况下是阻断的，IPv6 组播数据只能通过 Switch A—Switch B—Switch C 的路径传递给连接在 Switch C 上的接收者；要求通过将 Switch A 的端口 GigabitEthernet1/0/3 配置为静态路由器端口，以保证当 Switch A—Switch B—Switch C 的路径出现阻断时，IPv6 组播数据可以几乎不间断地通过 Switch A—Switch C 的新路径传递给接收者。

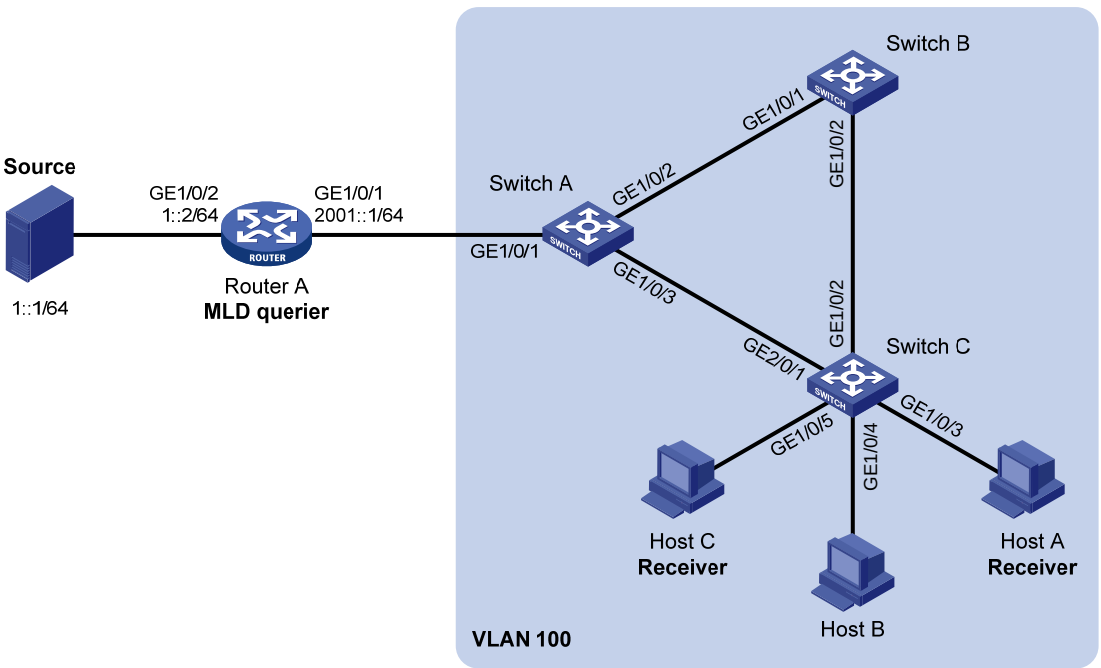
说明

如果没有配置静态路由器端口，那么当 Switch A—Switch B—Switch C 的路径出现阻断时，至少需要等待一个 MLD 查询和响应周期完成后，IPv6 组播数据才能通过 Switch A—Switch C 的新路径传递给接收者，IPv6 组播数据的传输在这个过程中将中断。

有关 STP（Spanning Tree Protocol，生成树协议）的详细介绍，请参见“二层技术-以太网交换配置指导”中的“生成树”。

2. 组网图

图1-5 静态端口配置组网图



3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址

使能各设备的IPv6 转发功能，并按照 图 1-5 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 配置 Router A

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在端口 GigabitEthernet1/0/1 上使能 MLD。

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] mld enable
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm
[RouterA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch A

全局使能 MLD Snooping。

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/3 添加到该 VLAN 中，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] quit
```

把 GigabitEthernet1/0/3 配置为静态路由器端口。

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] mld-snooping static-router-port vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
```

(4) 配置 Switch B

全局使能 MLD Snooping。

```
<SwitchB> system-view
[SwitchB] mld-snooping
[SwitchB-mld-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/2 添加到该 VLAN 中，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchB] vlan 100
[SwitchB-vlan100] port gigabitethernet 1/0/1 gigabitethernet 1/0/2
[SwitchB-vlan100] mld-snooping enable
[SwitchB-vlan100] quit
```

(5) 配置 Switch C

全局使能 MLD Snooping。

```
<SwitchC> system-view
[SwitchC] mld-snooping
```

```
[SwitchC-mld-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/5 添加到该 VLAN 中，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchC] vlan 100
[SwitchC-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/5
[SwitchC-vlan100] mld-snooping enable
[SwitchC-vlan100] quit
```

分别在端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/5 上配置静态加入 IPv6 组播组 FF1E::101。

```
[SwitchC] interface gigabitethernet 1/0/3
[SwitchC-GigabitEthernet1/0/3] mld-snooping static-group ff1e::101 vlan 100
[SwitchC-GigabitEthernet1/0/3] quit
[SwitchC] interface gigabitethernet 1/0/5
[SwitchC-GigabitEthernet1/0/5] mld-snooping static-group ff1e::101 vlan 100
[SwitchC-GigabitEthernet1/0/5] quit
```

(6) 检验配置效果

查看 Switch A 上 VLAN 100 内 MLD Snooping 组的详细信息。

```
[SwitchA] display mld-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 2 port(s).
            GE1/0/1                (D) ( 00:01:30 )
            GE1/0/3                (S)
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:FF1E::101
        (::, FF1E::101):
          Attribute:    Host Port
          Host port(s):total 1 port(s).
            GE1/0/2                (D) ( 00:03:23 )
    MAC group(s):
      MAC group address:3333-0000-0101
          Host port(s):total 1 port(s).
            GE1/0/2
```

由此可见，Switch A 上的端口 GigabitEthernet1/0/3 已经成为了静态路由器端口。

查看 Switch C 上 VLAN 100 内 MLD Snooping 组的详细信息。

```
[SwitchC] display mld-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
```

```
Port flags: D-Dynamic port, S-Static port, C-Copy port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
          GE1/0/2                (D) ( 00:01:23 )
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
      (::, FF1E::101):
        Attribute:    Host Port
        Host port(s):total 2 port(s).
          GE1/0/3                (S)
          GE1/0/5                (S)
  MAC group(s):
    MAC group address:3333-0000-0101
        Host port(s):total 2 port(s).
          GE1/0/3
          GE1/0/5
```

由此可见，Switch C 上的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/5 已经成为了 IPv6 组播组 FF1E::101 的静态成员端口。

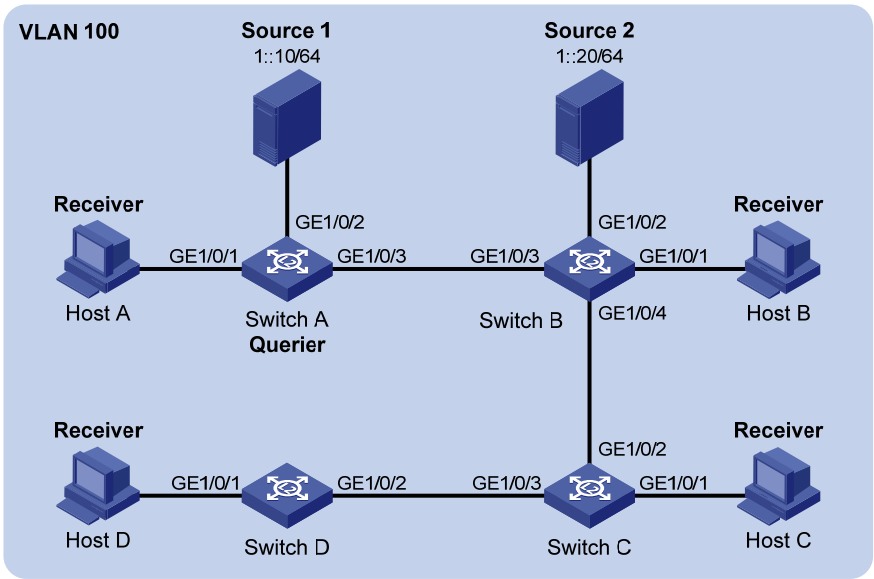
1.9.3 MLD Snooping查询器配置举例

1. 组网需求

- 如 图 1-6 所示，在一个没有三层设备的纯二层网络环境中，IPv6 组播源Source 1 和Source 2 分别向IPv6 组播组FF1E::101 和FF1E::102 发送IPv6 组播数据，Host A和Host C是IPv6 组播组FF1E::101 的接收者（Receiver），Host B和Host D则是IPv6 组播组FF1E::102 的接收者；所有接收者均使用MLDv1，所有交换机上都运行版本 1 的MLD Snooping，并选择距IPv6 组播源较近的Switch A来充当MLD Snooping查询器。
- 为防止交换机在没有二层 IPv6 组播转发表项时将 IPv6 组播数据在 VLAN 内广播，在所有交换机上都使能丢弃未知 IPv6 组播数据报文功能。

2. 组网图

图1-6 MLD Snooping 查询器配置组网图



3. 配置步骤

(1) 配置 Switch A

使能 IPv6 转发功能，并全局使能 MLD Snooping。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

创建 VLAN 100，并把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/3 添加到该 VLAN 中。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
```

在 VLAN 100 内使能 MLD Snooping，并使能丢弃未知 IPv6 组播数据报文功能。

```
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] mld-snooping drop-unknown
```

在 VLAN 100 内使能 MLD Snooping 查询器。

```
[SwitchA-vlan100] mld-snooping querier
[SwitchA-vlan100] quit
```

(2) 配置 Switch B

使能 IPv6 转发功能，并全局使能 MLD Snooping。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] mld-snooping
[SwitchB-mld-snooping] quit
```

创建 VLAN 100，并把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中。

```
[SwitchB] vlan 100
[SwitchB-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
```

在 VLAN 100 内使能 MLD Snooping，并使能丢弃未知 IPv6 组播数据报文功能。

```
[SwitchB-vlan100] mld-snooping enable
[SwitchB-vlan100] mld-snooping drop-unknown
[SwitchB-vlan100] quit
```

Switch C 和 Switch D 的配置与 Switch B 相似，配置过程略。

(3) 检验配置效果

当 MLD Snooping 查询器开始工作之后，除查询器以外的所有交换机都能收到 MLD 普遍组查询报文。通过使用 **display mld-snooping statistics** 命令可以查看 MLD 报文的统计信息，例如：

查看 Switch B 上收到的 MLD 报文的统计信息。

```
[SwitchB] display mld-snooping statistics
  Received MLD general queries:3.
  Received MLDv1 specific queries:0.
  Received MLDv1 reports:12.
  Received MLD dones:0.
  Sent     MLDv1 specific queries:0.
  Received MLDv2 reports:0.
  Received MLDv2 reports with right and wrong records:0.
  Received MLDv2 specific queries:0.
  Received MLDv2 specific sg queries:0.
  Sent     MLDv2 specific queries:0.
  Sent     MLDv2 specific sg queries:0.
  Received error MLD messages:0.
```

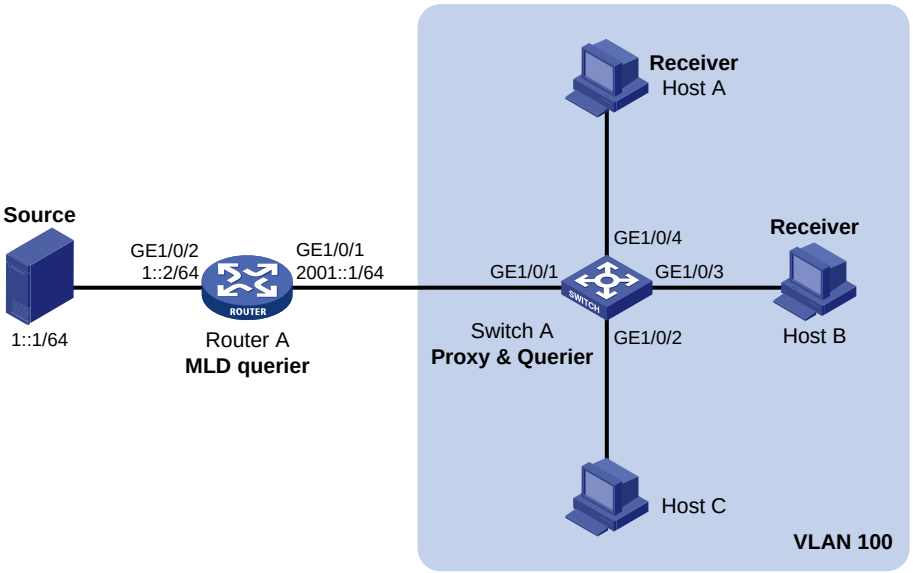
1.9.4 MLD Snooping Proxying配置举例

1. 组网需求

- 如 图 1-7 所示，Router A通过GigabitEthernet1/0/2 端口连接IPv6 组播源（Source），通过GigabitEthernet1/0/1 端口连接Switch A；Router A上运行MLDv1，Switch A上运行版本 1 的MLD Snooping，并由Router A充当MLD查询器。
- 通过配置，使 Switch A 能够代理下游主机向 Router A 发送的 MLD 报告报文和离开报文，以及响应 Router A 发来的 MLD 查询报文并向下游主机转发。

2. 组网图

图1-7 MLD Snooping Proxying 配置组网图



3. 配置步骤

(1) 配置 IPv6 地址

使能各设备的IPv6 转发功能，并按照 图 1-7 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 配置 Router A

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在端口 GigabitEthernet1/0/1 上使能 MLD。

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] mld enable
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm
[RouterA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch A

全局使能 MLD Snooping。

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中；在该 VLAN 内使能 MLD Snooping，并使能 MLD Snooping Proxying。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] mld-snooping proxying enable
[SwitchA-vlan100] quit
```

(4) 检验配置效果

当配置完成后，Host A 和 Host B 分别发送组地址为 FF1E::101 的 MLD 加入报文，Switch A 收到该报文后通过其路由器端口 GigabitEthernet1/0/1 向 Router A 也发送该组的加入报文。通过使用 **display mld-snooping group** 和 **display mld group** 命令可以分别查看 MLD Snooping 组和 MLD 组的信息，例如：

查看 Switch A 上 MLD Snooping 组的信息。

```
[SwitchA] display mld-snooping group
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port(s).
            GE1/0/1                (D)
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:FF1E::101
        (::, FF1E::101):
          Host port(s):total 2 port(s).
            GE1/0/3                (D)
            GE1/0/4                (D)
    MAC group(s):
      MAC group address:3333-0000-0101
          Host port(s):total 2 port(s).
            GE1/0/3
            GE1/0/4
```

查看 Router A 上 MLD 组的信息。

```
[RouterA] display mld group
Total 1 MLD Group(s).
Interface group report information
 GigabitEthernet1/0/1(2001::1):
  Total 1 MLD Group reported
   Group Address: FF1E::1
    Last Reporter: FE80::2FF:FFFF:FE00:1
    Uptime: 00:00:03
    Expires: 00:04:17
```

当 Host A 离开 IPv6 组播组 FF1E::101 时，向 Switch A 发送该组的 MLD 离开报文，但由于 Host B 仍未离开该组，因此 Switch A 并不会删除该组，也不会向 Router A 发送该组的离开报文，只是在该组对应转发表项的成员端口列表中将端口 GigabitEthernet1/0/4 删除。通过使用 **display mld-snooping group** 命令可以查看 MLD Snooping 组的信息，例如：

查看 Switch A 上 MLD Snooping 组的信息。

```
[SwitchA] display mld-snooping group
```

```
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
          GE1/0/1                 (D)
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
      (::, FF1E::101):
        Host port(s):total 1 port(s).
          GE1/0/3                 (D)
  MAC group(s):
    MAC group address:3333-0000-0101
        Host port(s):total 1 port(s).
          GE1/0/3
```

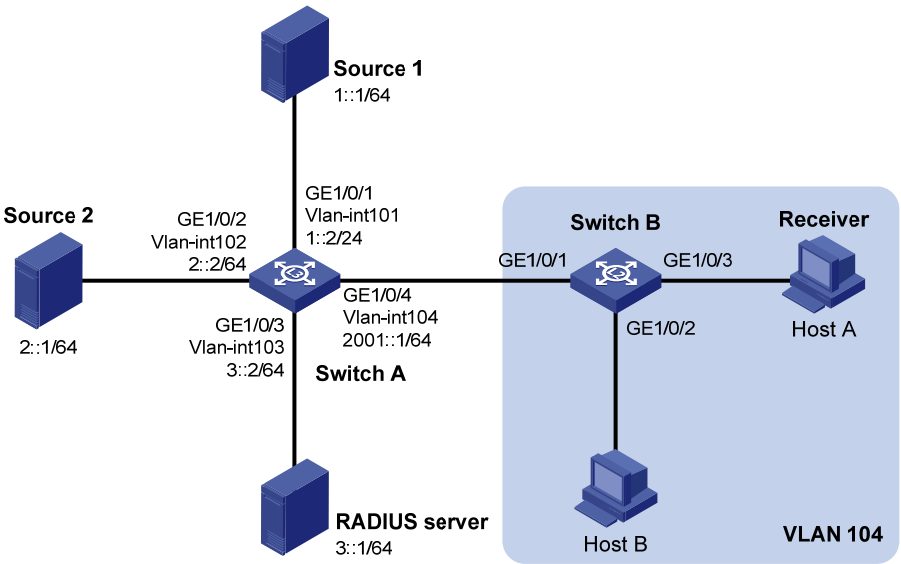
1.9.5 IPv6 组播源与组播用户控制策略配置举例

1. 组网需求

- 如 图 1-8 所示，三层交换机Switch A分别通过端口Vlan-interface101 和Vlan-interface102 连接IPv6 组播源Source 1 和Source 2，通过端口Vlan-interface103 连接RADIUS服务器，并通过接口Vlan-interface104 连接二层交换机Switch B；Switch A上运行MLDv1，Switch B上运行版本 1 的MLD Snooping，连接在Switch A上的IPV6 组播源和连接在Switch B上的主机都使用 802.1X进行认证。
- 通过在 Switch A 上配置 IPv6 组播源控制策略，禁止由 Source 2 发往 IPv6 组播组 FF1E::101 的 IPv6 组播数据进入网络；通过在 Switch B 上配置 IPv6 组播用户控制策略，使接收者（Receiver）Host A 只能加入/离开 IPv6 组播组 FF1E::101。

2. 组网图

图1-8 IPv6 组播源与组播用户控制策略配置组网图



3. 配置步骤

(1) 配置 IP 地址

使能各设备的IPv6 转发功能，并按照 图 1-8 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 配置 Switch A

分别创建VLAN 101～104，并将端口 GigabitEthernet1/0/1～GigabitEthernet1/0/4 依次加入VLAN 101～104 中。

```
<SwitchA> system-view
[SwitchA] vlan 101
[SwitchA-vlan101] port gigabitethernet 1/0/1
[SwitchA-vlan101] quit
[SwitchA] vlan 102
[SwitchA-vlan102] port gigabitethernet 1/0/2
[SwitchA-vlan102] quit
[SwitchA] vlan 103
[SwitchA-vlan103] port gigabitethernet 1/0/3
[SwitchA-vlan103] quit
[SwitchA] vlan 104
[SwitchA-vlan104] port gigabitethernet 1/0/4
[SwitchA-vlan104] quit
```

使能 IPv6 组播路由，在接口 Vlan-interface101、Vlan-interface102 和 Vlan-interface104 上使能 IPv6 PIM-DM，并在接口 Vlan-interface104 上使能 MLD。

```
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
```

```
[SwitchA-Vlan-interface102] pim ipv6 dm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim ipv6 dm
[SwitchA-Vlan-interface104] mld enable
[SwitchA-Vlan-interface104] quit
```

创建名为 policy1 的 QoS 策略，不允许来自 IPv6 组播源 2::1 的 IPv6 组播数据向 IPv6 组播组 FF1E::101 发送。

```
[SwitchA] acl ipv6 number 3001
[SwitchA-acl6-adv-3001] rule permit udp source 2::1 128 destination ff1e::101 128
[SwitchA-acl6-adv-3001] quit
[SwitchA] traffic classifier classifier1
[SwitchA-classifier-classifier1] if-match acl ipv6 3001
[SwitchA-classifier-classifier1] quit
[SwitchA] traffic behavior behavior1
[SwitchA-behavior-behavior1] filter deny
[SwitchA-behavior-behavior1] quit
[SwitchA] qos policy policy1
[SwitchA-qospolicy-policy1] classifier classifier1 behavior behavior1
[SwitchA-qospolicy-policy1] quit
```

创建名为 profile1 的 User Profile，在该 User Profile 下应用 QoS 策略 policy1，并将该 User Profile 激活。

```
[SwitchA] user-profile profile1
[SwitchA-user-profile-profile1] qos apply policy policy1 inbound
[SwitchA-user-profile-profile1] quit
[SwitchA] user-profile profile1 enable
```

创建名为 scheme1 的 RADIUS 认证方案，配置 RADIUS 服务器的类型为 extended 类型；指定主 RADIUS 认证/授权服务器和计费服务器的 IP 地址均为 3::1，共享密钥均为 123321；并配置发送给 RADIUS 服务器的用户名不需要携带域名。

```
[SwitchA] radius scheme scheme1
[SwitchA-radius-scheme1] server-type extended
[SwitchA-radius-scheme1] primary authentication 3::1
[SwitchA-radius-scheme1] key authentication 123321
[SwitchA-radius-scheme1] primary accounting 3::1
[SwitchA-radius-scheme1] key accounting 123321
[SwitchA-radius-scheme1] user-name-format without-domain
[SwitchA-radius-scheme1] quit
```

创建名为 domain1 的 ISP 域，指定 lan-access 用户的 RADIUS 认证、授权和计费方案均为 scheme1；并将该域指定为系统缺省的 ISP 域。

```
[SwitchA] domain domain1
[SwitchA-isp-domian1] authentication lan-access radius-scheme scheme1
[SwitchA-isp-domian1] authorization lan-access radius-scheme scheme1
[SwitchA-isp-domian1] accounting lan-access radius-scheme scheme1
[SwitchA-isp-domian1] quit
[SwitchA] domain default enable domain1
```

全局使能 802.1X，并在端口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/2 上分别使能 802.1X。

```
[SwitchA] dot1x
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] dot1x
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] dot1x
[SwitchA-GigabitEthernet1/0/2] quit
```

(3) 配置 Switch B

全局使能 MLD Snooping。

```
<SwitchB> system-view
[SwitchB] mld-snooping
[SwitchB-mld-snooping] quit
```

创建 VLAN 104 把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/3 添加到该 VLAN 中；在该 VLAN 内使能 MLD Snooping。

```
[SwitchB] vlan 104
[SwitchB-vlan104] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchB-vlan104] mld-snooping enable
[SwitchB-vlan104] quit
```

创建名为 profile2 的 User Profile，在该 User Profile 下配置只允许用户加入/离开 IPV6 组播组 FF1E::101，并将该 User Profile 激活。

```
[SwitchB] acl ipv6 number 2001
[SwitchB-acl6-basic-2001] rule permit source ff1e::101 128
[SwitchB-acl6-basic-2001] quit
[SwitchB] user-profile profile2
[SwitchB-user-profile-profile2] mld-snooping access-policy 2001
[SwitchB-user-profile-profile2] quit
[SwitchB] user-profile profile2 enable
```

创建名为 scheme2 的 RADIUS 认证方案，配置 RADIUS 服务器的类型为 extended 类型；指定主 RADIUS 认证/授权服务器和计费服务器的 IP 地址均为 3::1，共享密钥均为 321123；并配置发送给 RADIUS 服务器的用户名不需要携带域名。

```
[SwitchB] radius scheme scheme2
[SwitchB-radius-scheme2] server-type extended
[SwitchB-radius-scheme2] primBry authentication 3::1
[SwitchB-radius-scheme2] key authentication 321123
[SwitchB-radius-scheme2] primary accounting 3::1
[SwitchB-radius-scheme2] key accounting 321123
[SwitchB-radius-scheme2] user-name-format without-domain
[SwitchB-radius-scheme2] quit
```

创建名为 domain2 的 ISP 域，指定 lan-access 用户的 RADIUS 认证、授权和计费方案均为 scheme2；并将该域指定为系统缺省的 ISP 域。

```
[SwitchB] domain domain2
[SwitchB-isp-domian2] authentication lan-access radius-scheme scheme2
[SwitchB-isp-domian2] authorization lan-access radius-scheme scheme2
[SwitchB-isp-domian2] accounting lan-access radius-scheme scheme2
[SwitchB-isp-domian2] quit
[SwitchB] domain default enable domain2
```

全局使能 802.1X，并在端口 GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 上分别使能 802.1X。

```
[SwitchB] dot1x
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] dot1x
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] dot1x
[SwitchB-GigabitEthernet1/0/3] quit
```

(4) 配置 RADIUS 服务器

请在 RADIUS 服务器上进行与 Switch A 和 Switch B 相对应的配置，具体配置过程可参考 RADIUS 服务器的配置手册。

(5) 检验配置效果

配置完成后，所有的 IPv6 组播源和用户主机都向 RADIUS 服务器发起 802.1X 认证。当认证通过后，Source 1 向 IPv6 组播组 FF1E::101、Source 2 向 IPv6 组播组 FF1E::102 分别发送不同的 IPv6 组播数据，而 Host A 分别发起了对 IPv6 组播组 FF1E::101 和 FF1E::102 的加入。通过使用 **display mld-snooping group** 命令可以查看 MLD Snooping 组的信息：

查看 Switch B 上 VLAN 104 内 MLD Snooping 组的详细信息。

```
[SwitchB] display mld-snooping group vlan 104 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port(s).
            GE1/0/1                (D) ( 00:01:30 )
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:FF1E::101
        (::, FF1E::101):
          Attribute:    Host Port
          Host port(s):total 1 port(s).
            GE1/0/3                (D) ( 00:04:10 )
    MAC group(s):
      MAC group address:3333-0000-0101
          Host port(s):total 1 port(s).
            GE1/0/3
```

由此可见，Switch B 上的端口 GigabitEthernet1/0/3 只加入了 IPV6 组播组 FF1E::101，而未加入 IPV6 组播组 FF1E::102。

此后由于管理员的操作失误，Source 2 也开始向 IPv6 组播组 FF1E::101 发送 IPv6 组播数据。通过使用 **display multicast ipv6 forwarding-table** 命令可以查看 IPv6 组播转发表的信息：

查看 Switch A 上有关 IPv6 组播组 FF1E::101 的 IPv6 组播转发表信息。

```
[SwitchA] display multicast ipv6 forwarding-table ff1e::101
IPv6 Multicast Forwarding Table

Total 1 entry

Total 1 entry matched
00001. (1::1, FF1E::101)
     MID: 0, Flags: 0x0:0
     Uptime: 00:08:32, Timeout in: 00:03:26
     Incoming interface: Vlan-interface101
     List of 1 outgoing interfaces:
       1: Vlan-interface104
     Matched 19648 packets(20512512 bytes), Wrong If 0 packets
     Forwarded 19648 packets(20512512 bytes)
```

由此可见，Switch A 上只有 IPv6 组播源组（1::1，FF1E::101）的转发表项，而没有 IPv6 组播源组（2::1，FF1E::101）的转发表项，说明 Source 2 发往 IPv6 组播组 FF1E::101 的 IPv6 组播数据没能进入网络。

1.10 常见配置错误举例

1.10.1 交换机不能实现二层组播

1. 故障现象

交换机不能实现 MLD snooping 二层组播功能。

2. 分析

MLD Snooping 没有使能。

3. 处理过程

(1) 使用 **display current-configuration** 命令查看 MLD Snooping 的运行状态。

(2) 如果是没有使能 MLD Snooping，则需先在系统视图下使用 **mld-snooping** 命令全局使能 MLD Snooping，然后在 VLAN 视图下使用 **mld-snooping enable** 命令使能 VLAN 内的 MLD Snooping。

(3) 如果只是没有在相应 VLAN 下使能 MLD Snooping，则只需在 VLAN 视图下使用 **mld-snooping enable** 命令使能 VLAN 内的 MLD Snooping。

1.10.2 配置的IPv6 组播组策略不生效

1. 故障现象

配置了 IPv6 组播组策略，只允许主机加入某些特定的 IPv6 组播组，但主机仍然可以收到发往其它 IPv6 组播组的 IPv6 组播数据。

2. 分析

- IPv6 ACL 规则配置不正确；
- IPv6 组播组策略应用不正确；

- 没有使能丢弃未知 IPv6 组播数据报文的功能，使得属于过滤策略之外的 IPv6 组播数据报文（即未知 IPv6 组播数据报文）被广播。

3. 处理过程

(1) 使用 **display acl ipv6** 命令查看所配置的 IPv6 ACL 规则，检查其是否与所要实现的 IPv6 组播组过滤策略相符合。

(2) 在 MLD-Snooping 视图或相应的端口视图下使用 **display this** 命令查看是否应用了正确的 IPv6 组播组策略。如果没有，则使用 **group-policy** 或 **mld-snooping group-policy** 命令应用正确的 IPv6 组播组策略。

(3) 使用 **display current-configuration** 命令查看是否已使能丢弃未知 IPv6 组播数据报文的功能。如果没有使能，则使用 **mld-snooping drop-unknown** 命令使能丢弃未知 IPv6 组播数据报文功能。

1.11 附录

1.11.1 交换机对IPv6 组播协议报文的特殊处理规则

支持 MLD Snooping 功能的交换机使能了三层 IPv6 组播路由后，在不同的情况下对各种 IPv6 组播协议报文的处理有所差异。具体规则如下：

(1) 如果交换机上只使能了 MLD，或同时使能了 MLD 和 IPv6 PIM：

- 对于 MLD 报文，根据报文类型为其维护相应的动态成员端口或动态路由器端口；
- 对于 IPv6 PIM Hello 报文，为其维护相应的动态路由器端口。

(2) 如果交换机上只使能了 IPv6 PIM：

- 对于 MLD 报文，将其当作未知报文在 VLAN 内进行广播；
- 对于 IPv6 PIM Hello 报文，为其维护相应的动态路由器端口。

(3) 在交换机上关闭 MLD 时：

- 如果未使能 IPv6 PIM，则删除所有动态成员端口和动态路由器端口；
- 如果已使能 IPv6 PIM，则删除动态成员端口，只保留动态路由器端口。


说明

在使能了三层 IPv6 组播路由的交换机上，请使用命令 **display mld group port-info** 来查看二层端口信息。有关 **display mld group port-info** 命令的详细介绍，请参见“IP 组播命令参考”中的“MLD”。

(4) 在交换机上关闭 IPv6 PIM 时：

- 如果未使能 MLD，则删除动态路由器端口；
- 如果已使能 MLD，则保留所有动态成员端口和动态路由器端口。

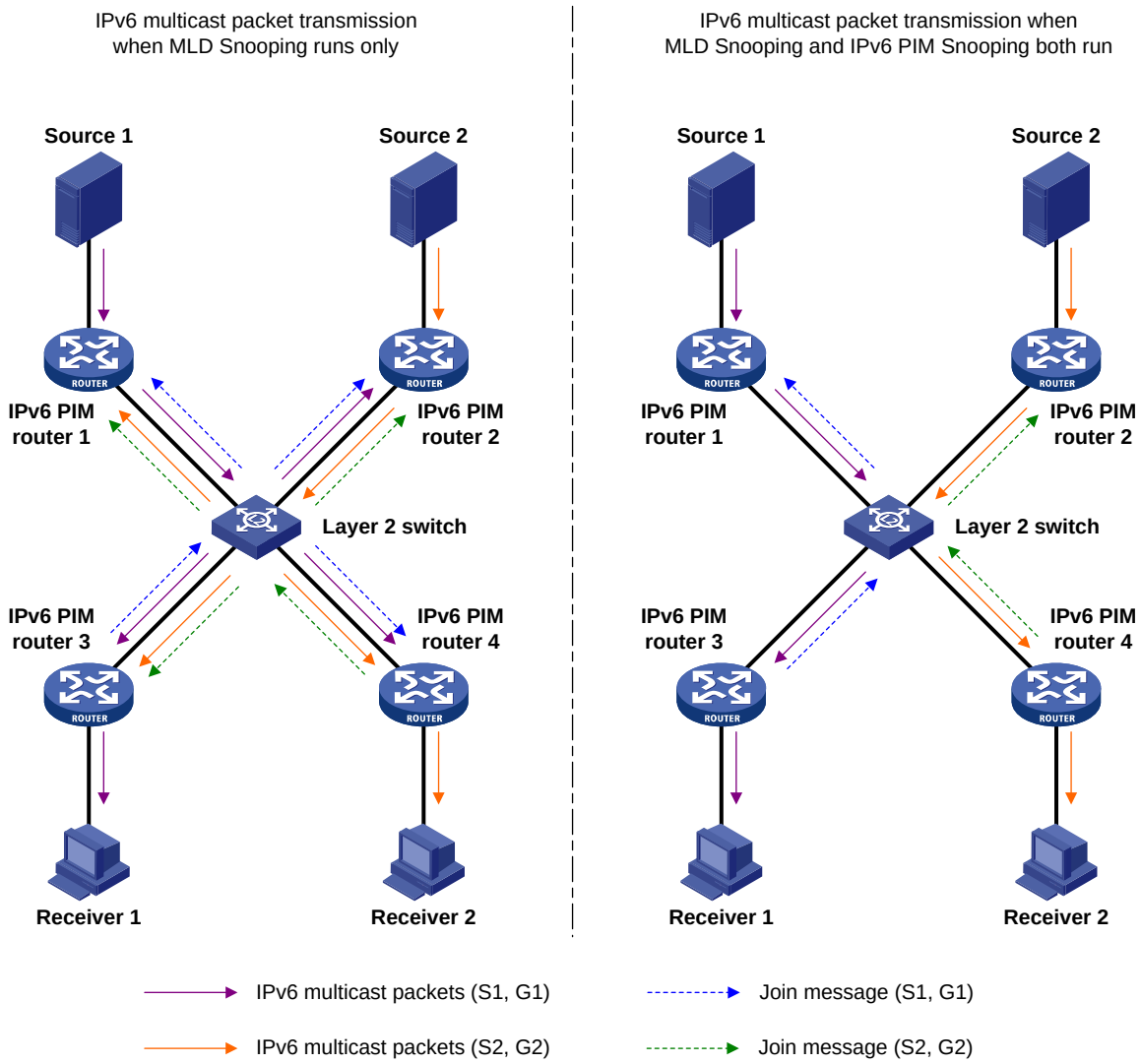
目 录

1 IPv6 PIM Snooping配置

1.1 IPv6 PIM Snooping简介

IPv6 PIM Snooping 是 IPv6 Protocol Independent Multicast Snooping（IPv6 协议无关组播窥探）的简称，运行 IPv6 PIM Snooping 的二层设备通过对收到的 IPv6 PIM 协议报文进行分析，将有接收需求的端口添加到相应的 IPv6 组播转发表项中，以实现 IPv6 组播报文的精确转发。

图1-1 二层设备运行 IPv6 PIM Snooping 前后的对比



如 图 1-1 所示，IPv6 组播源Source 1 和Source 2 分别向IPv6 组播组G1 和G2 发送IPv6 组播数据，而Receiver 1 和Receiver 2 则分别是G1 和G2 的接收者，二层设备上连接各IPv6 PIM路由器的端口都属于同一个VLAN：

- 当二层设备只运行 MLD Snooping 时，它通过监听 IPv6 PIM 路由器发出的 IPv6 PIM Hello 报文来维护路由器端口，并将所有 IPv6 组播数据报文向 VLAN 内的所有路由器端口转发；而对

于其它类型的 IPv6 PIM 协议报文，则在 VLAN 内广播。因此，无论 IPv6 PIM 路由器是否有接收需求，都会收到所有的 IPv6 PIM 协议报文和 IPv6 组播数据报文。

- 当二层设备同时运行了 MLD Snooping 和 IPv6 PIM Snooping 时，它通过监听 IPv6 PIM 路由器发出的 IPv6 PIM 协议报文来了解其接收需求，从而将有接收需求的 IPv6 PIM 路由器所在的端口加入到相应的 IPv6 组播转发表项中，使 IPv6 PIM 协议报文和 IPv6 组播数据报文能够被精确转发给有接收需求的 IPv6 PIM 路由器，从而节约了网络带宽。

说明

- 有关 MLD Snooping 和路由器端口的详细介绍，请参见“IP 组播配置指导”中的“MLD Snooping”。
- 有关 IPv6 PIM 的详细介绍，请参见“IP 组播配置指导”中的“IPv6 PIM”。

1.2 配置IPv6 PIM Snooping

在配置某 VLAN 内的 IPv6 PIM Snooping 时，需要在交换机上先全局使能 IPv6 转发功能和 MLD Snooping，然后在该 VLAN 内使能 MLD Snooping 和 IPv6 PIM Snooping。

表1-1 配置 IPv6 PIM Snooping

操作	命令	说明
进入系统视图	**system-view**	-
全局时能IPv6转发功能	**ipv6**	必选 缺省情况下，IPv6转发功能处于关闭状态
全局使能MLD Snooping，并进入MLD-Snooping视图	**mld-snooping**	必选 缺省情况下，MLD Snooping处于关闭状态
退回系统视图	**quit**	-
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能MLD Snooping	**mld-snooping enable**	必选 缺省情况下，VLAN内的MLD Snooping处于关闭状态
在VLAN内使能IPv6 PIM Snooping	**pim-snooping ipv6 enable**	必选 缺省情况下，VLAN内的IPv6 PIM Snooping处于关闭状态

说明

- 有关 **mld-snooping** 和 **mld-snooping enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“MLD Snooping”。
- 在 VLAN 内使能了 IPv6 PIM Snooping 之后，IPv6 PIM Snooping 功能只在属于该 VLAN 的端口上生效。
- 在 IPv6 组播 VLAN 的子 VLAN 内使能 IPv6 PIM Snooping 无效。有关 IPv6 组播 VLAN 的详细介绍，请参见“IP 组播配置指导”中的“IPv6 组播 VLAN”。
- 在部署了配置 IPv6 PIM Snooping 功能的交换机的网络中，请在接收者侧的边缘 IPv6 PIM 设备上配置加入/剪枝报文（**jp-pkt-size**）的最大长度不能大于 IPv6 路径 MTU，有关配置加入/剪枝报文的详细内容，请参见“IP 组播配置指导”中的“IPv6 PIM”。

1.3 IPv6 PIM Snooping显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IPv6 PIM Snooping 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 IPv6 PIM Snooping 的统计信息。

表1-2 IPv6 PIM Snooping 显示和维护

操作	命令
查看IPv6 PIM Snooping的邻居信息	**display pim-snooping ipv6 neighbor** [**vlan** *vlan-id*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6 PIM Snooping的路由信息	**display pim-snooping ipv6 routing-table** [**vlan** *vlan-id*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6 PIM Snooping监听到的PIM报文的统计信息	**display pim-snooping ipv6 statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IPv6 PIM Snooping监听到的PIM报文的统计信息	**reset pim-snooping ipv6 statistics**

1.4 IPv6 PIM Snooping典型配置举例

1. 组网需求

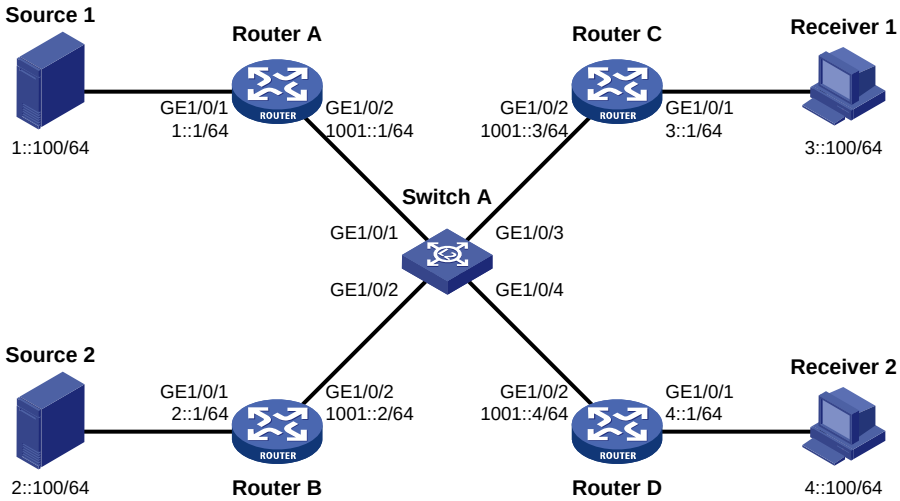
- 如 图 1-2 所示，Router A和Router B各自的GigabitEthernet1/0/1 接口分别连接IPv6 组播源 Source 1 和Source 2；Router C和Router D各自的GigabitEthernet1/0/1 接口分别连接接收者 Receiver 1 和 Receiver 2；Router A、Router B、Router C 和 Router D 各自的 GigabitEthernet1/0/2 接口都通过Switch A互连。
- Source 1 和 Source 2 分别通过 IPv6 组播组 FF1E::101 和 FF2E::101 发送 IPv6 组播数据，Receiver 1 和 Receiver 2 则分别接收来自 IPv6 组播组 FF1E::101 和 FF2E::101 的 IPv6 组播数据；Router C 和 Router D 各自的 GigabitEthernet1/0/1 接口上都运行 MLD，Router A、

Router B、Router C 和 Router D 上都运行 IPv6 PIM-SM，并由 Router A 的 GigabitEthernet1/0/2 接口充当 C-BSR 和 C-RP。

- 通过在 Switch A 上配置 MLD Snooping 和 IPv6 PIM Snooping，使 Switch A 将 IPv6 PIM 协议报文和 IPv6 组播数据报文只转发给有接收需求的路由器。

2. 组网图

图1-2 IPv6 PIM Snooping 典型配置组网图



3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址

使能各设备的IPv6 转发功能，并按照 图 1-2 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 配置 Router A

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并将接口 GigabitEthernet1/0/2 配置为 C-BSR 和 C-RP。

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] pim ipv6 sm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 sm
[RouterA-GigabitEthernet1/0/2] quit
[RouterA] pim ipv6
[RouterA-pim6] c-bsr 1001::1
[RouterA-pim6] c-rp 1001::1
```

(3) 配置 Router B

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM。

```
<RouterB> system-view
[RouterB] multicast ipv6 routing-enable
[RouterB] interface gigabitethernet 1/0/1
```

```
[RouterB-GigabitEthernet1/0/1] pim ipv6 sm
[RouterB-GigabitEthernet1/0/1] quit
[RouterB] interface gigabitethernet 1/0/2
[RouterB-GigabitEthernet1/0/2] pim ipv6 sm
```

(4) 配置 Router C

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并在接口 GigabitEthernet1/0/1 上使能 MLD。

```
<RouterC> system-view
[RouterC] multicast ipv6 routing-enable
[RouterC] interface gigabitethernet 1/0/1
[RouterC-GigabitEthernet1/0/1] pim ipv6 sm
[RouterC-GigabitEthernet1/0/1] mld enable
[RouterC-GigabitEthernet1/0/1] quit
[RouterC] interface gigabitethernet 1/0/2
[RouterC-GigabitEthernet1/0/2] pim ipv6 sm
```

(5) 配置 Router D

Router D 的配置与 Router C 相似，配置过程略。

(6) 配置 Switch A

全局使能 MLD Snooping。

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

创建 VLAN 100，把端口 GigabitEthernet1/0/1 到 GigabitEthernet1/0/4 添加到该 VLAN 中；在该 VLAN 内使能 MLD Snooping 和 IPv6 PIM Snooping。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] pim-snooping ipv6 enable
[SwitchA-vlan100] quit
```

(7) 检验配置效果

查看 Switch A 上 VLAN 100 内 IPv6 PIM Snooping 的邻居信息。

```
[SwitchA] display pim-snooping ipv6 neighbor vlan 100
  Total number of neighbors: 4

  VLAN ID: 100
    Total number of neighbors: 4
    Neighbor        Port                    Expires   Option Flags
    FE80::1         GE1/0/1                  02:02:23  LAN Prune Delay
    FE80::2         GE1/0/2                  03:00:05  LAN Prune Delay
    FE80::3         GE1/0/3                  02:22:13  LAN Prune Delay
    FE80::4         GE1/0/4                  03:07:22  LAN Prune Delay
```

由此可见，Router A、Router B、Router C 和 Router D 之间都建立起了 IPv6 PIM Snooping 邻居关系。

查看 Switch A 上 VLAN 100 内 IPv6 PIM Snooping 的路由信息。

```
[SwitchA] display pim-snooping ipv6 routing-table vlan 100 slot 1
  Total 2 entry(ies)
```

```
FSM Flag: NI-no info, J-join, PP-prune pending

VLAN ID: 100
  Total 2 entry(ies)
  (*, FF1E::101)
    Upstream neighbor: FE80::1
      Upstream port: GE1/0/1
      Total number of downstream ports: 1
        1: GE1/0/3
           Expires: 00:03:01, FSM: J
  (*, FF2E::101)
    Upstream neighbor: FE80::2
      Upstream port: GE1/0/2
      Total number of downstream ports: 1
        1: GE1/0/4
           Expires: 00:01:05, FSM: J
```

由此可见，Switch A 将向 Router C 转发 IPv6 组播组 FF1E::101 的 IPv6 组播数据，向 Router D 转发 IPv6 组播组 FF2E::101 的 IPv6 组播数据。

1.5 常见配置错误举例

1.5.1 交换机不能实现IPv6 PIM Snooping功能

1. 故障现象

交换机不能实现 IPv6 PIM Snooping 功能。

2. 分析

MLD Snooping 或 IPv6 PIM Snooping 没有使能。

3. 处理过程

(1) 使用 **display current-configuration** 命令查看 MLD Snooping 和 IPv6 PIM Snooping 的运行状态。

(2) 如果没有使能 MLD Snooping，请先在系统视图下使用 **mld-snooping** 命令全局使能 MLD Snooping，然后在 VLAN 视图下分别使用 **mld-snooping enable** 和 **pim-snooping ipv6 enable** 命令使能 VLAN 内的 MLD Snooping 和 IPv6 PIM Snooping。

(3) 如果没有使能 IPv6 PIM Snooping，请在 VLAN 视图下使用 **pim-snooping ipv6 enable** 命令使能 VLAN 内的 IPv6 PIM Snooping。

1.5.2 部分下游IPv6 PIM路由器无法收到IPv6 组播数据

1. 故障现象

在有分片加入/剪枝报文的网络中，部分下游 IPv6 PIM 路由器无法收到 IPv6 组播数据。

2. 分析

IPv6 PIM Snooping 不能对分片报文进行重组，因此无法维护分片加入/剪枝报文中携带的下游状态。为了保证系统功能正常，只能将分片加入/剪枝报文在 VLAN 内广播，因此需要在 VLAN 内连接 IPv6

PIM Snooping 交换机的所有 IPv6 PIM 路由器上都禁止加入报文抑制能力，以保证加入报文不被广播的分片加入/剪枝报文所抑制。假如存在未禁止该能力的 IPv6 PIM 路由器，被广播的分片加入/剪枝报文就会影响其它 IPv6 PIM 路由器的加入状态：如果某 IPv6 PIM 路由器有 IPv6 组播接收需求，但其发送的加入报文被抑制，那么该路由器将无法收到 IPv6 组播数据。

3. 处理过程

(1) 在 IPv6 PIM 路由器连接 IPv6 PIM Snooping 交换机的接口上使用 **pim ipv6 hello-option neighbor-tracking** 命令使能邻居跟踪功能。

(2) 如果存在不能够使能邻居跟踪功能的 IPv6 PIM 路由器，则需关闭 IPv6 PIM Snooping 交换机上的 IPv6 PIM Snooping 功能。

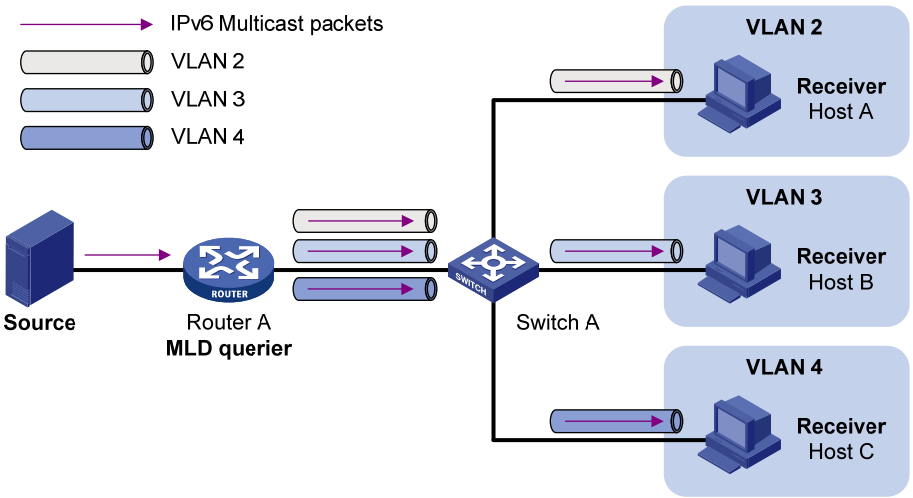
目 录

1 IPv6 组播VLAN配置

1.1 IPv6组播VLAN简介

如 图 1-1 所示，在传统的IPv6 组播点播方式下，当属于不同VLAN的主机Host A、Host B和Host C同时点播同一IPv6 组播组时，三层设备（Router A）需要把IPv6 组播数据在每个用户VLAN（即主机所属的VLAN）内都复制一份发送给二层设备（Switch A）。这样既造成了带宽的浪费，也给三层设备增加了额外的负担。

图1-1 未运行 IPv6 组播 VLAN 时的 IPv6 组播数据传输



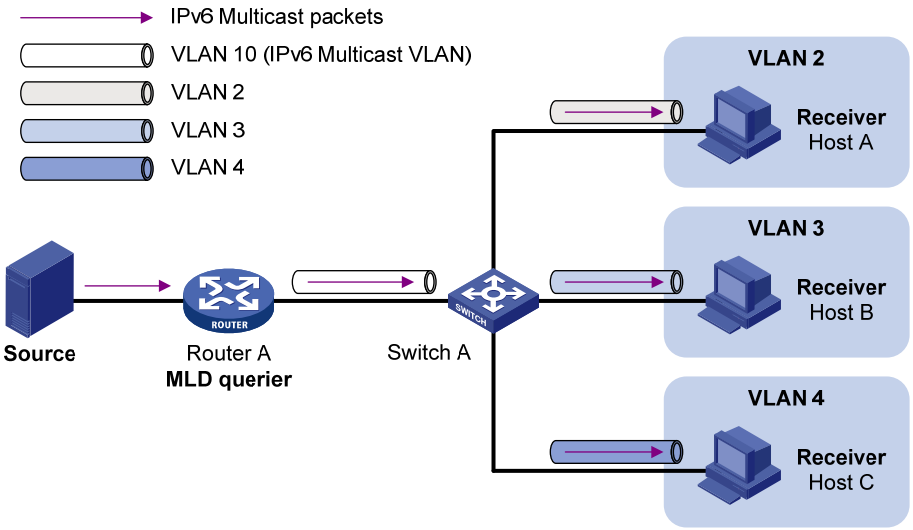
可以使用 IPv6 组播 VLAN 功能解决这个问题。在二层设备上配置了 IPv6 组播 VLAN 后，三层设备只需把 IPv6 组播数据在 IPv6 组播 VLAN 内复制一份发送给二层设备，而不必在每个用户 VLAN 内都复制一份，从而节省了网络带宽，也减轻了三层设备的负担。

IPv6 组播 VLAN 有以下两种实现和配置方式：

1. 基于子VLAN的IPv6 组播VLAN

如 图 1-2 所示，接收者主机Host A、Host B和Host C分属不同的用户VLAN。在Switch A上配置VLAN 10 为IPv6 组播VLAN，将所有的用户VLAN都配置为该IPv6 组播VLAN的子VLAN，并在IPv6 组播VLAN内使能MLD Snooping。

图1-2 基于子 VLAN 的 IPv6 组播 VLAN 示意图

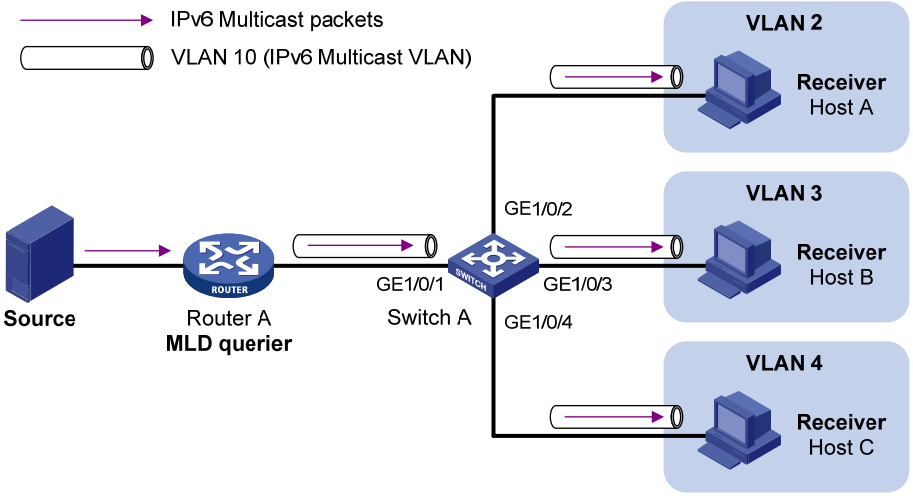


配置完成后，MLD Snooping 将在 IPv6 组播 VLAN 中对路由器端口进行维护，而在各子 VLAN 中对成员端口进行维护。这样，Router A 只需把 IPv6 组播数据在 IPv6 组播 VLAN 内复制一份发送给 Switch A 即可，Switch A 会将其分发给该 IPv6 组播 VLAN 内那些有接收者的子 VLAN。

2. 基于端口的IPv6 组播VLAN

如 图 1-3 所示，接收者主机Host A、Host B和Host C分属不同的用户VLAN，Switch A上的所有用户端口（即连接主机的端口）均为Hybrid类型。在Switch A上配置VLAN 10 为IPv6 组播VLAN，将所有用户端口都添加到该IPv6 组播VLAN内，并在IPv6 组播VLAN和所有用户VLAN内都使能MLD Snooping。

图1-3 基于端口的 IPv6 组播 VLAN 示意图



配置完成后，当 Switch A 上的用户端口收到来自主机的 MLD 报文时，会为其打上 IPv6 组播 VLAN 的 Tag 并上送给 MLD 查询器，于是 MLD Snooping 就可以在 IPv6 组播 VLAN 中对路由器端口和成员端口进行统一的维护。这样，Router A 只需把 IPv6 组播数据在 IPv6 组播 VLAN 内复制一份发送给 Switch A 即可，Switch A 会将其分发给该 IPv6 组播 VLAN 内的所有成员端口。

说明

- 有关 MLD Snooping 的相关配置，以及路由器端口和成员端口的详细介绍，请参见“IP 组播配置指导”中的“MLD Snooping”。
- 有关 VLAN Tag 的详细介绍，请参见“二层技术-以太网交换配置指导”中的“VLAN”。

1.2 IPv6组播VLAN配置任务简介

表1-1 IPv6 组播 VLAN 配置任务简介

配置任务		说明	详细配置
配置基于子VLAN的IPv6组播VLAN		二者必选其一	1.3
配置基于端口的IPv6组播VLAN	配置用户端口属性		1.4.2
	配置IPv6组播VLAN端口		1.4.3
配置IPv6组播VLAN内MLD Snooping转发表项的最大数量		可选	1.5

说明

若在设备上同时配置了基于子 VLAN 和基于端口的 IPv6 组播 VLAN，则基于端口的 IPv6 组播 VLAN 将优先生效。

1.3 配置基于子VLAN的IPv6组播VLAN

1.3.1 配置准备

在配置基于子 VLAN 的 IPv6 组播 VLAN 之前，需完成以下任务：

- 使能 IPv6 转发功能
- 创建相应的 VLAN
- 在欲配置为 IPv6 组播 VLAN 的 VLAN 内使能 MLD Snooping

1.3.2 配置基于子VLAN的IPv6 组播VLAN

首先需要把某个 VLAN 配置为 IPv6 组播 VLAN，再将用户 VLAN 添加到该 IPv6 组播 VLAN 内，使其成为 IPv6 组播 VLAN 的子 VLAN。

表1-2 配置基于子 VLAN 的 IPv6 组播 VLAN

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
配置指定VLAN为IPv6组播VLAN，并进入IPv6组播VLAN视图	**multicast-vlan ipv6** *vlan-id*	必选 缺省情况下，VLAN不是IPv6组播VLAN
向IPv6组播VLAN内添加子VLAN	**subvlan** *vlan-list*	必选 缺省情况下，IPv6组播VLAN内没有子VLAN

说明

- 在已使能了 IPv6 组播路由的设备上不允许再配置 IPv6 组播 VLAN。
- 要配置为 IPv6 组播 VLAN 的指定 VLAN 必须存在。
- 要添加到 IPv6 组播 VLAN 内的子 VLAN 必须存在，且不能是 IPv6 组播 VLAN 或其它 IPv6 组播 VLAN 的子 VLAN。
- IPv6 组播 VLAN 内子 VLAN 的总数不得超过系统限制。

1.4 配置基于端口的IPv6组播VLAN

在配置基于端口的 IPv6 组播 VLAN 时，需要先配置各用户端口的属性，然后再将配置好的用户端口添加到 IPv6 组播 VLAN 内。

说明

- 只允许将以太网端口或二层聚合接口类型的用户端口配置为 IPv6 组播 VLAN 的端口。
- 二层以太网端口视图下的配置只对当前端口有效；二层聚合接口视图下的配置对当前接口有效；端口组视图下的配置对当前端口组中的所有端口有效。

1.4.1 配置准备

在配置基于端口的 IPv6 组播 VLAN 之前，需完成以下任务：

- 使能 IPv6 转发功能
- 创建相应的 VLAN
- 在欲配置为 IPv6 组播 VLAN 的 VLAN 内使能 MLD Snooping
- 在所有的用户 VLAN 内都内使能 MLD Snooping

1.4.2 配置用户端口属性

首先,配置用户端口为Hybrid类型,允许用户VLAN的报文通过,缺省VLAN为其所属的用户VLAN。然后，配置用户端口允许 IPv6 组播 VLAN 的报文不带 Tag 通过。这样，当二层设备通过 IPv6 组播 VLAN 收到来自上游、打有 IPv6 组播 VLAN Tag 的 IPv6 组播数据报文时，会将其 Tag 去掉后再向下游转发。

表1-3 配置用户端口属性

操作		命令	说明
进入系统视图		**system-view**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
配置用户端口的链路类型为Hybrid类型		**port link-type hybrid**	必选 缺省情况下，端口的链路类型为Access类型
配置用户端口的缺省VLAN为其所属的用户VLAN		**port hybrid pvid vlan** *vlan-id*	必选 缺省情况下，Hybrid端口的缺省VLAN为VLAN 1
允许IPv6组播VLAN通过用户端口，且不携带Tag		**port hybrid vlan** *vlan-id-list* **untagged**	必选 缺省情况下，Hybrid端口只允许VLAN 1通过

说明

有关 **port link-type**、**port hybrid pvid vlan** 和 **port hybrid vlan** 命令的详细介绍，请参见“二层技术-以太网交换命令参考”中的“VLAN”。

1.4.3 配置IPv6 组播VLAN端口

首先需要把某个 VLAN 配置为 IPv6 组播 VLAN，再将用户端口添加到该 IPv6 组播 VLAN 内——既可以在 IPv6 组播 VLAN 内添加端口，也可以在端口上指定其所属的 IPv6 组播 VLAN——这两种配置方式是等效的。

1. 在IPv6 组播VLAN内配置IPv6 组播VLAN端口

表1-4 在 IPv6 组播 VLAN 内配置 IPv6 组播 VLAN 端口

操作	命令	说明
进入系统视图	**system-view**	-
配置指定VLAN为IPv6组播VLAN，并进入IPv6组播VLAN视图	**multicast-vlan ipv6** *vlan-id*	必选 缺省情况下，VLAN不是IPv6组播VLAN
向IPv6组播VLAN内添加端口	**port** *interface-list*	必选 缺省情况下，IPv6组播VLAN内没有端口

2. 在端口上配置IPv6 组播VLAN端口

表1-5 在端口上配置 IPv6 组播 VLAN 端口

操作		命令	说明
进入系统视图		**system-view**	-
配置指定VLAN为IPv6组播VLAN，并进入IPv6组播VLAN视图		**multicast-vlan ipv6** *vlan-id*	必选 缺省情况下，VLAN不是IPv6组播VLAN
退回系统视图		**quit**	-
进入相应视图	进入二层以太网或二层聚合接口视图	**interface** *interface-type interface-number*	二者必选其一
	进入端口组视图	**port-group manual** *port-group-name*	
指定端口所属的IPv6组播VLAN		**port multicast-vlan ipv6** *vlan-id*	必选 缺省情况下，端口不属于任何IPv6组播VLAN

说明

- 在已使能了 IPv6 组播路由的设备上不允许再配置 IPv6 组播 VLAN。
- 要配置为 IPv6 组播 VLAN 的指定 VLAN 必须存在。
- 一个端口只能属于一个 IPv6 组播 VLAN。

1.5 配置IPv6组播VLAN内MLD Snooping转发表项的最大数量

用户可以调整 IPv6 组播 VLAN 内 MLD Snooping 转发表项的最大数量，当所有 IPv6 组播 VLAN 内维护的表项总数达到最大数量后，将不再创建新的表项，直至有表项被老化或被手工删除。

表1-6 配置 IPv6 组播 VLAN 内 MLD Snooping 转发表项的最大数量

操作	命令	说明
进入系统视图	**system-view**	-
配置IPv6组播VLAN内MLD Snooping转发表项的最大数量	**multicast-vlan ipv6 entry-limit** *limit*	必选 缺省情况下，IPv6组播VLAN内MLD Snooping转发表项的最大数量为1000

说明

在对 IPv6 组播 VLAN 内 MLD Snooping 转发表项的最大数量进行配置时，如果所有 IPv6 组播 VLAN 内的表项总数已超过了配置值，系统将提示用户删除多余表项，但并不会自动删除任何已存在的表项，同时也不再继续创建新的表项。

1.6 IPv6组播VLAN显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IPv6 组播 VLAN 的运行情况，通过查看显示信息验证配置的效果。

表1-7 IPv6 组播 VLAN 显示和维护

操作	命令
查看IPv6组播VLAN的信息	**display multicast-vlan ipv6** [*vlan-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.7 IPv6组播VLAN典型配置举例

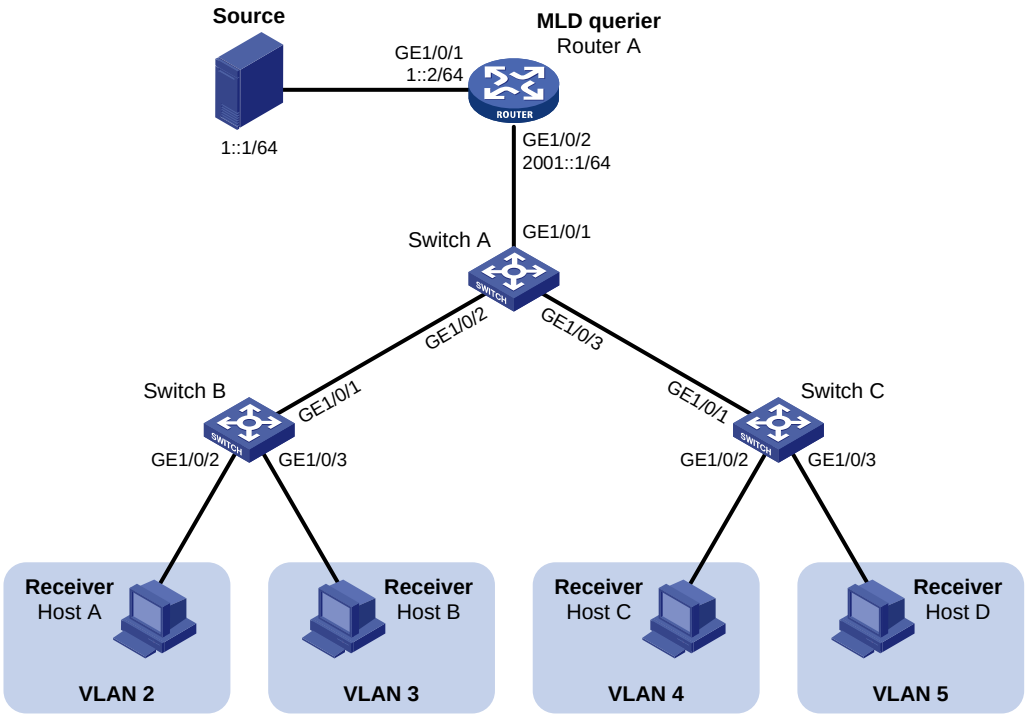
1.7.1 基于子VLAN的IPv6 组播VLAN配置举例

1. 组网需求

- 如 图 1-4 所示，Router A通过端口GigabitEthernet1/0/1 连接IPv6 组播源（Source），通过端口GigabitEthernet1/0/2 连接Switch A；Router A上运行MLDv1，Switch A～Switch C上都运行版本 1 的MLD Snooping，并由Router A充当MLD查询器。
- IPv6 组播源向 IPv6 组播组 FF1E::101 发送 IPv6 组播数据，Host A～Host D 都是该 IPv6 组播组的接收者（Receiver），且分别属于 VLAN 2～VLAN 5。
- 通过在 Switch A 上配置基于子 VLAN 的 IPv6 组播 VLAN，使 Router A 通过 IPv6 组播 VLAN 向 Switch A 下分属不同用户 VLAN 的主机分发 IPv6 组播数据。

2. 组网图

图1-4 基于子 VLAN 的 IPv6 组播 VLAN 配置组网图



3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址

使能各设备的IPv6 转发功能，并按照 图 1-4 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 配置 Router A

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在主机侧端口 GigabitEthernet1/0/2 上使能 MLD。

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm
[RouterA-GigabitEthernet1/0/2] mld enable
```

(3) 配置 Switch A

全局使能 MLD Snooping。

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

创建 VLAN 2～VLAN 5。

```
[SwitchA] vlan 2 to 5
```

配置端口 GigabitEthernet1/0/2 为 Trunk 端口，并允许 VLAN 2 和 VLAN 3 通过。

```
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type trunk
[SwitchA-GigabitEthernet1/0/2] port trunk permit vlan 2 3
[SwitchA-GigabitEthernet1/0/2] quit
```

配置端口 GigabitEthernet1/0/3 为 Trunk 端口，并允许 VLAN 4 和 VLAN 5 通过。

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-GigabitEthernet1/0/3] port trunk permit vlan 4 5
[SwitchA-GigabitEthernet1/0/3] quit
```

创建 VLAN 10，把端口 GigabitEthernet1/0/1 添加到该 VLAN 中，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] mld-snooping enable
[SwitchA-vlan10] quit
```

配置 VLAN 10 为 IPv6 组播 VLAN，并把 VLAN 2～VLAN 5 都配置为该 IPv6 组播 VLAN 的子 VLAN。

```
[SwitchA] multicast-vlan ipv6 10
[SwitchA-ipv6-mvlan-10] subvlan 2 to 5
[SwitchA-ipv6-mvlan-10] quit
```

(4) 配置 Switch B

全局使能 MLD Snooping。

```
<SwitchB> system-view
[SwitchB] mld-snooping
[SwitchB-mld-snooping] quit
```

创建 VLAN 2，把端口 GigabitEthernet1/0/2 添加到该 VLAN 中，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/2
[SwitchB-vlan2] mld-snooping enable
[SwitchB-vlan2] quit
```

创建 VLAN 3，把端口 GigabitEthernet1/0/3 添加到该 VLAN 中，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchB] vlan 3
[SwitchB-vlan3] port gigabitethernet 1/0/3
[SwitchB-vlan3] mld-snooping enable
[SwitchB-vlan3] quit
```

配置端口 GigabitEthernet1/0/1 为 Trunk 端口，并允许 VLAN 2 和 VLAN 3 通过。

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port link-type trunk
[SwitchB-GigabitEthernet1/0/1] port trunk permit vlan 2 3
```

(5) 配置 Switch C

Switch C 的配置与 Switch B 相似，配置过程略。

(6) 检验配置效果

查看 Switch A 上所有 IPv6 组播 VLAN 的信息。

```
[SwitchA] display multicast-vlan ipv6
 Total 1 IPv6 multicast-vlan(s)

 IPv6 Multicast vlan 10
   subvlan list:
    vlan 2-5
   port list:
    no port
```

查看 Switch A 上 MLD Snooping 组播组的信息。

```
[SwitchA] display mld-snooping group
  Total 5 IP Group(s).
  Total 5 IP Source(s).
  Total 5 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):2.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 0 port(s).
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:FF1E::101
        (::, FF1E::101):
          Host port(s):total 1 port(s).
            GE1/0/2                (D)
    MAC group(s):
      MAC group address:3333-0000-0101
          Host port(s):total 1 port(s).
            GE1/0/2

  Vlan(id):3.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 0 port(s).
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:FF1E::101
        (::, FF1E::101):
          Host port(s):total 1 port(s).
            GE1/0/2                (D)
    MAC group(s):
      MAC group address:3333-0000-0101
          Host port(s):total 1 port(s).
            GE1/0/2

  Vlan(id):4.
    Total 1 IP Group(s).
```

```
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 0 port(s).
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
      (::, FF1E::101):
        Host port(s):total 1 port(s).
          GE1/0/3               (D)
  MAC group(s):
    MAC group address:3333-0000-0101
        Host port(s):total 1 port(s).
          GE1/0/3

Vlan(id):5.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 0 port(s).
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
      (::, FF1E::101):
        Host port(s):total 1 port(s).
          GE1/0/3               (D)
  MAC group(s):
    MAC group address:3333-0000-0101
        Host port(s):total 1 port(s).
          GE1/0/3

Vlan(id):10.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
          GE1/0/1               (D)
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
      (::, FF1E::101):
        Host port(s):total 0 port(s).
  MAC group(s):
    MAC group address:3333-0000-0101
        Host port(s):total 0 port(s).
```

由此可见，MLD Snooping 在 IPv6 组播 VLAN（VLAN 10）中维护路由器端口，而在各子 VLAN（VLAN 2～VLAN 5）中维护各自的成员端口。

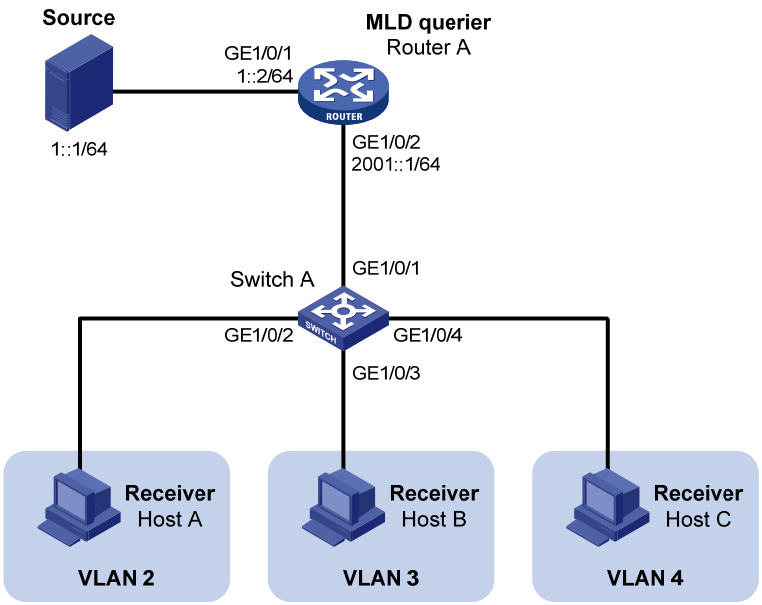
1.7.2 基于端口的IPv6 组播VLAN配置举例

1. 组网需求

- 如 图 1-5 所示，Router A通过端口GigabitEthernet1/0/1 连接IPv6 组播源（Source），通过端口GigabitEthernet1/0/2 连接Switch A；Router A上运行MLDv1，Switch A上运行版本 1 的MLD Snooping，并由Router A充当MLD查询器。
- IPv6 组播源向 IPv6 组播组 FF1E::101 发送 IPv6 组播数据，Host A～Host C 都是该 IPv6 组播组的接收者（Receiver），且分别属于 VLAN 2～VLAN 4。
- 通过在 Switch A 上配置基于端口的 IPv6 组播 VLAN，使 Router A 通过 IPv6 组播 VLAN 向 Switch A 下分属不同用户 VLAN 的主机分发 IPv6 组播数据。

2. 组网图

图1-5 基于端口的 IPv6 组播 VLAN 配置组网图



3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址

使能各设备的IPv6 转发功能，并按照 图 1-5 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 配置 Router A

使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在主机侧端口 GigabitEthernet1/0/2 上使能 MLD。

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] ipv6 pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] ipv6 pim dm
[RouterA-GigabitEthernet1/0/2] mld enable
```

(3) 配置 Switch A

全局使能 MLD Snooping。

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

创建 VLAN 10，把端口 GigabitEthernet1/0/1 添加到该 VLAN 中，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] mld-snooping enable
[SwitchA-vlan10] quit
```

创建 VLAN 2，并在该 VLAN 内使能 MLD Snooping。

```
[SwitchA] vlan 2
[SwitchA-vlan2] mld-snooping enable
[SwitchA-vlan2] quit
```

VLAN 3 与 VLAN 4 的配置与 VLAN 2 相似，配置过程略。

配置端口 GigabitEthernet1/0/2 为 Hybrid 类型，缺省 VLAN 为 VLAN 2；允许 VLAN 2 和 VLAN 10 的报文通过，且均不携带 Tag。

```
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type hybrid
[SwitchA-GigabitEthernet1/0/2] port hybrid pvid vlan 2
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 2 untagged
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 10 untagged
[SwitchA-GigabitEthernet1/0/2] quit
```

GigabitEthernet1/0/3 与 GigabitEthernet1/0/4 的配置与 GigabitEthernet1/0/2 相似，配置过程略。

配置 VLAN 10 为 IPv6 组播 VLAN。

```
[SwitchA] multicast-vlan ipv6 10
```

将端口 GigabitEthernet1/0/2 到 GigabitEthernet1/0/3 添加到 IPv6 组播 VLAN 10 内。

```
[SwitchA-ipv6-mvlan-10] port gigabitethernet 1/0/2 to gigabitethernet 1/0/3
[SwitchA-ipv6-mvlan-10] quit
```

配置端口 GigabitEthernet1/0/4 也属于 IPv6 组播 VLAN 10。

```
[SwitchA] interface gigabitethernet 1/0/4
[SwitchA-GigabitEthernet1/0/4] port multicast-vlan ipv6 10
[SwitchA-GigabitEthernet1/0/4] quit
```

(4) 检验配置效果

查看 Switch A 上所有 IPv6 组播 VLAN 的信息。

```
[SwitchA] display multicast-vlan ipv6
 Total 1 IPv6 multicast-vlan(s)

 IPv6 Multicast vlan 10
   subvlan list:
    no subvlan
   port list:
    GE1/0/2                 GE1/0/3                 GE1/0/4
```

查看 Switch A 上 MLD Snooping 组播组的信息。

```
[SwitchA] display mld-snooping group
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).

  Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
  Subvlan flags: R-Real VLAN, C-Copy VLAN
  Vlan(id):10.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
    Router port(s):total 1 port(s).
            GE1/0/1                (D)
    IP group(s):the following ip group(s) match to one mac group.
      IP group address:FF1E::101
        (::, FF1E::101):
          Host port(s):total 3 port(s).
            GE1/0/2                (D)
            GE1/0/3                (D)
            GE1/0/4                (D)
    MAC group(s):
      MAC group address:3333-0000-0101
          Host port(s):total 3 port(s).
            GE1/0/2
            GE1/0/3
            GE1/0/4
```

由此可见，MLD Snooping 统一在 IPv6 组播 VLAN（VLAN 10）中维护路由器端口和成员端口。

目 录

1 IPv6 组播路由与转发配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- IPv6 组播路由与转发功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

1.1 IPv6组播路由与转发简介

在 IPv6 组播实现中，组播路由和转发分为三种表：

- 每个 IPv6 组播路由协议都有一个协议自身的路由表，如 IPv6 PIM 路由表（IPv6 PIM Routing-Table）；
- 各 IPv6 组播路由协议的组播路由信息经过综合形成一个总的 IPv6 组播路由表（IPv6 Multicast Routing-Table）；
- IPv6 组播转发表（IPv6 Multicast Forwarding-Table）直接用于控制 IPv6 组播数据包的转发，是真正指导 IPv6 组播数据转发的转发表。

IPv6 组播路由表由一组（S，G）表项组成，其中（S，G）表示由源 S 向 IPv6 组播组 G 发送 IPv6 组播数据的路由信息。如果路由器支持多种 IPv6 组播路由协议，则其 IPv6 组播路由表中将包括由多种协议生成的组播路由。路由器根据组播路由和转发策略，从 IPv6 组播路由表中选出最优的组播路由，并下发到 IPv6 组播转发表中。

1.1.1 RPF检查机制

IPv6 组播路由协议依赖于现有的 IPv6 单播路由信息或 IPv6 MBGP 路由来创建 IPv6 组播路由表项。IPv6 组播路由协议在创建 IPv6 组播路由表项时，运用了 RPF（Reverse Path Forwarding，逆向路径转发）检查机制，以确保 IPv6 组播数据能够沿正确的路径传输，同时还能避免由于各种原因而造成的环路。

1. RPF检查过程

执行 RPF 检查的依据是 IPv6 单播路由或 IPv6 MBGP 路由：

- IPv6 单播路由表中汇集了到达各个目的网段的最短路径；
- IPv6 MBGP 路由表直接提供 IPv6 组播路由信息。

在执行 RPF 检查时，路由器同时查找 IPv6 单播路由表和 IPv6 MBGP 路由表，具体过程如下：

(1) 首先，分别从 IPv6 单播路由表和 IPv6 MBGP 路由表中各选出一条最优路由：

- 以“报文源”的 IPv6 地址为目的地址查找 IPv6 单播路由表，自动选取一条最优 IPv6 单播路由。对应表项中的出接口为 RPF 接口，下一跳为 RPF 邻居。路由器认为来自 RPF 邻居且由该 RPF 接口收到的 IPv6 组播报文所经历的路径是从源 S 到本地的最短路径。

- 以“报文源”的 IPv6 地址为目的地址查找 IPv6 MBGP 路由表，自动选取一条最优 IPv6 MBGP 路由。对应表项中的出接口为 RPF 接口，下一跳为 RPF 邻居。

(2) 然后，从这两条最优路由中选择一条作为 RPF 路由：

- 如果配置了按照最长匹配选择路由，则从这两条路由中选出最长匹配的那条路由；如果这两条路由的前缀长度一样，则选择其中优先级最高的那条路由；如果它们的优先级也相同，则按照 IPv6 MBGP 路由、IPv6 单播路由的顺序进行选择。
- 如果没有配置按照最长匹配选择路由，则从这两条路由中选出优先级最高的那条路由；如果它们的优先级相同，则按照 IPv6 MBGP 路由、IPv6 单播路由的顺序进行选择。

说明

根据 IPv6 组播报文传输的具体情况不同，“报文源”所代表的具体含义也不同：

- 如果当前报文沿从组播源到接收者或 RP（Rendezvous Point，汇集点）的 SPT（Shortest Path Tree，最短路径树）进行传输，则以组播源为“报文源”进行 RPF 检查；
- 如果当前报文沿从 RP 到接收者的 RPT（Rendezvous Point Tree，共享树）进行传输，或者沿从组播源到 RP 的组播源侧 RPT 进行传输，则都以 RP 为“报文源”进行 RPF 检查；
- 如果当前报文为 BSR（Bootstrap Router，自举路由器）报文，沿从 BSR 到各路由器的路径进行传输，则以 BSR 为“报文源”进行 RPF 检查。

有关 SPT、RPT、组播源侧 RPT、RP 和 BSR 的详细介绍，请参见“IP 组播配置指导”中的“IPv6 PIM”。

---

2. RPF检查在IPv6 组播转发中的应用

对每一个收到的 IPv6 组播数据报文都进行 RPF 检查会给路由器带来较大负担，而利用 IPv6 组播转发表可以解决这个问题。在建立 IPv6 组播路由和转发表时，会把 IPv6 组播数据报文（S，G）的 RPF 接口记录为（S，G）表项的入接口。当路由器收到 IPv6 组播数据报文（S，G）后，查找 IPv6 组播转发表：

(1) 如果 IPv6 组播转发表中不存在（S，G）表项，则对该报文执行 RPF 检查，将其 RPF 接口作为入接口，结合相关路由信息创建相应的表项，并下发到 IPv6 组播转发表中：

- 若该报文实际到达的接口正是其 RPF 接口，则 RPF 检查通过，向所有的出接口转发该报文；
- 若该报文实际到达的接口不是其 RPF 接口，则 RPF 检查失败，丢弃该报文。

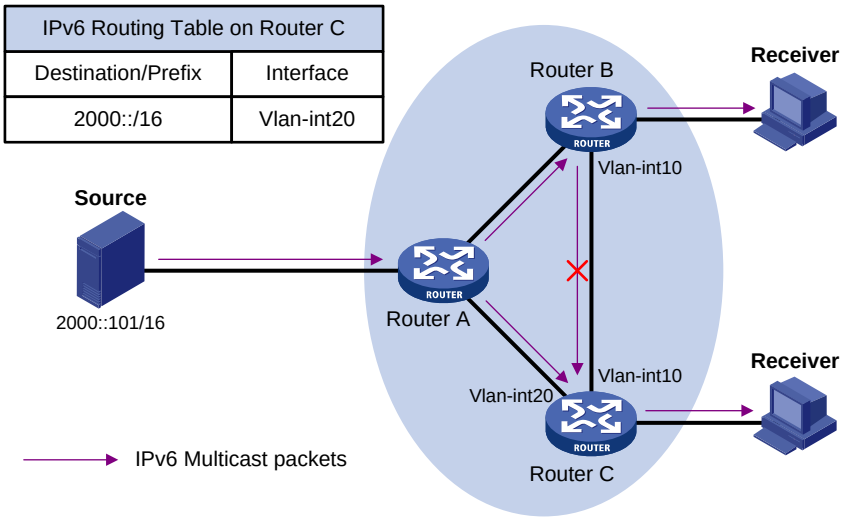
(2) 如果 IPv6 组播转发表中已存在（S，G）表项，且该报文实际到达的接口与入接口相匹配，则向所有的出接口转发该报文。

(3) 如果 IPv6 组播转发表中已存在（S，G）表项，但该报文实际到达的接口与入接口不匹配，则对此报文执行 RPF 检查：

- 若其 RPF 接口与入接口一致，则说明（S，G）表项正确，丢弃这个来自错误路径的报文；
- 若其 RPF 接口与入接口不符，则说明（S，G）表项已过时，于是把入接口更新为 RPF 接口。如果该报文实际到达的接口正是其 RPF 接口，则向所有的出接口转发该报文，否则将其丢弃。

如 图 1-1 所示，假设网络中IPv6 单播路由畅通，未配置IPv6 MBGP。IPv6 组播报文（S，G）沿从组播源（Source）到接收者（Receiver）的SPT进行传输。假定Router C上的IPv6 组播转发表中已存在（S，G）表项，其记录的入接口为Vlan-interface20。

图1-1 RPF 检查过程



- 如果该 IPv6 组播报文从接口 Vlan-interface20 到达 Router C，与（S，G）表项的入接口相匹配，则向所有的出接口转发该报文。
- 如果该 IPv6 组播报文从接口 Vlan-interface10 到达 Router C，与（S，G）表项的入接口不匹配，则对其执行 RPF 检查：通过查找 IPv6 单播路由表发现到达 Source 的出接口（即 RPF 接口）是 Vlan-interface20，与（S，G）表项的入接口一致。这说明（S，G）表项是正确的，该报文来自错误的路径，RPF 检查失败，于是丢弃该报文。

1.2 IPv6组播路由与转发配置任务简介

表1-1 IPv6 组播路由与转发配置任务简介

配置任务		说明	详细配置
使能IPv6组播路由		必选	1.3
配置IPv6组播路由与转发	配置IPv6组播路由策略	可选	1.4.2
	配置IPv6组播转发范围	可选	1.4.3
	配置IPv6组播转发表容量	可选	1.4.4

1.3 使能IPv6组播路由

在配置各项三层 IPv6 组播功能之前，必须首先使能 IPv6 组播路由。

表1-2 使能 IPv6 组播路由

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
使能IPv6组播路由	**multicast ipv6 routing-enable**	必选 缺省情况下，IPv6组播路由处于关闭状态

1.4 配置IPv6组播路由与转发

1.4.1 配置准备

在配置 IPv6 组播路由与转发之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通
- 配置 IPv6 PIM-DM（或 IPv6 PIM-SM）

在配置 IPv6 组播路由与转发之前，需准备以下数据：

- 单条 IPv6 组播转发表项的最大下行节点数目
- IPv6 组播转发表的最大表项数

1.4.2 配置IPv6 组播路由策略

可以配置组播路由器按照最长匹配原则来选择RPF路由，有关RPF路由选择的详细介绍，请参见“1.1.1 1. RPF检查过程”一节。此外，通过配置根据组播源或组播源组进行IPv6 组播流量的负载分担，还可以优化存在多条IPv6 组播数据流时的网络流量。

表1-3 配置 IPv6 组播路由策略

操作	命令	说明
进入系统视图	**system-view**	-
配置按照最长匹配选择RPF路由	**multicast ipv6 longest-match**	可选 缺省情况下，选择优先级最高的路由作为RPF路由
配置对IPv6组播流量进行负载分担	**multicast ipv6 load-splitting {source | source-group }**	可选 缺省情况下，不对IPv6组播流量进行负载分担

注意

multicast ipv6 load-splitting 命令对 IPv6 双向 PIM 不生效。

1.4.3 配置IPv6 组播转发范围

IPv6 组播信息在网络中的转发并不是漫无边际的，每个 IPv6 组播组对应的 IPv6 组播信息都必须在确定的范围内传递。目前可以在所有支持 IPv6 组播转发的接口上配置针对某个 IPv6 组播组或 Scope 值的转发边界。IPv6 组播转发边界为指定范围或 Scope 值的 IPv6 组播组划定了边界条件，如果 IPv6 组播报文的目的地址与边界条件匹配，就停止转发。当在一个接口上配置了 IPv6 组播转

发边界后，将不能从该接口转发 IPv6 组播报文（包括本机发出的 IPv6 组播报文），也不能从该接口接收 IPv6 组播报文。

表1-4 配置 IPv6 组播转发范围

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置IPv6组播转发边界	**multicast ipv6 boundary** { *ipv6-group-address prefix-length* | **scope** { *scope-id* | **admin-local** | **global** | **organization-local** | **site-local** } }	必选 缺省情况下，没有配置IPv6组播转发边界

1.4.4 配置IPv6 组播转发表容量

路由器为每个收到的 IPv6 组播数据报文都维护相应的转发表项。但是，IPv6 组播转发表项过多可能会耗尽路由器内存，从而导致路由器性能下降。用户可以根据实际组网情况和业务性能要求对 IPv6 组播转发表中的表项数量进行限制。如果 IPv6 组播转发表最大表项数的配置值小于当前值，则超出数目的表项并不会立刻被删除，而必须由 IPv6 组播路由协议来删除，同时也无法添加新的 IPv6 组播转发表项。

路由器为每个下行节点复制一份 IPv6 组播数据报文并发送出去，每个下行节点就形成 IPv6 组播分发树的一条分支。用户可以根据实际组网情况和业务性能要求对 IPv6 组播转发表中单条表项的下行节点数目（即出接口数目）进行限制，以缓解路由器的复制压力。如果单条 IPv6 组播转发表项的最大下行节点数目的配置值小于当前值，则超出数目的下行节点并不会被立刻删除，而必须由 IPv6 组播路由协议来删除，同时新增的下行节点将无法添加到该表项中。

表1-5 配置 IPv6 组播转发表容量

操作	命令	说明
进入系统视图	**system-view**	-
配置IPv6组播转发表的最大表项数	**multicast ipv6 forwarding-table route-limit** *limit*	可选 缺省情况下，IPv6组播转发表的最大表项数为1000
配置单条IPv6组播转发表项的最大下行节点数目	**multicast ipv6 forwarding-table downstream-limit** *limit*	可选 缺省情况下，单条IPv6组播转发表项的最大下行节点数目为128

1.5 IPv6组播路由与转发显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IPv6 组播路由与转发的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 IPv6 组播路由与转发的统计信息。

表1-6 IPv6 组播路由与转发显示和维护

操作	命令
查看IPv6组播边界信息	**display multicast ipv6 boundary** { **group** [*ipv6-group-address* [*prefix-length*]] | **scope** [*scope-id*] } [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6组播转发表信息	**display multicast ipv6 forwarding-table** [*ipv6-source-address* [*prefix-length*] | *ipv6-group-address* [*prefix-length*] | **incoming-interface** { *interface-type interface-number* | **register** } | **outgoing-interface** { **exclude** | **include** | **match** } { *interface-type interface-number* | **register** } | **statistics** | **slot** *slot-number*] * [**port-info**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6组播转发表的DF信息	**display multicast ipv6 forwarding-table df-info** [*rp-address*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6组播路由表信息	**display multicast ipv6 routing-table** [*ipv6-source-address* [*prefix-length*] | *ipv6-group-address* [*prefix-length*] | **incoming-interface** { *interface-type interface-number* | **register** } | **outgoing-interface** { { **exclude** | **include** | **match** } { *interface-type interface-number* | **register** } }] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6组播源的RPF信息	**display multicast ipv6 rpf-info** *ipv6-source-address* [*ipv6-group-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IPv6组播转发表中的转发项	**reset multicast ipv6 forwarding-table** { { *ipv6-source-address* [*prefix-length*] | *ipv6-group-address* [*prefix-length*] | **incoming-interface** { *interface-type interface-number* | **register** } } * | **all** }
清除IPv6组播路由表中的路由项	**reset multicast ipv6 routing-table** { { *ipv6-source-address* [*prefix-length*] | *ipv6-group-address* [*prefix-length*] | **incoming-interface** { *interface-type interface-number* | **register** } } * | **all** }

说明

有关 DF（Designated Forwarder，指定转发者）的详细介绍，请参见“IP 组播配置指导”中的“IPv6 PIM”。

注意

- 执行 **reset** 命令将清除 IPv6 组播路由表或 IPv6 组播转发表中的信息，可能导致 IPv6 组播信息无法正常传输；
- 清除 IPv6 组播路由表中的路由项后，IPv6 组播转发表中的相应表项也将随之删除；
- 清除 IPv6 组播转发表中的转发项后，IPv6 组播路由表中的相应表项也将随之删除。

1.6 常见配置错误举例

1.6.1 IPv6 组播数据异常终止

1. 故障现象

- 当某主机发送了加入 IPv6 组播组 G 的报文后，离该主机最近的路由器上却没有 IPv6 组播组 G 的组成员信息。中间路由器能成功接收 IPv6 组播数据，但数据不能到达末梢网络。
- 中间路由器的接口收到 IPv6 组播数据，但在 IPv6 PIM 路由表里没有相应的（S，G）表项。

2. 分析

- 命令 **multicast ipv6 boundary** 用来过滤接口上收到的 IPv6 组播报文。如果报文没有通过这个命令的 IPv6 ACL 匹配规则，则 IPv6 PIM 不能创建路由表项。
- 此外，IPv6 PIM 中的命令 **source-policy** 用来过滤接收的 IPv6 组播报文。如果报文没有通过这个命令的 IPv6 ACL 匹配规则，则 IPv6 PIM 也不能创建路由表项。

3. 处理过程

(1) 使用命令 **display current-configuration** 查看组播转发边界上配置的 IPv6 ACL 过滤规则。更改 **multicast ipv6 boundary** 命令的 IPv6 ACL 规则，使 IPv6 组播数据的源地址和 IPv6 组播组地址通过 IPv6 ACL 过滤。

(2) 检查组播过滤器配置。使用命令 **display current-configuration** 查看组播过滤器的配置，更改 **source-policy** 命令的 IPv6 ACL 规则，使 IPv6 组播数据的源地址和 IPv6 组播组地址通过 IPv6 ACL 过滤。

目 录

1 MLD配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- MLD 功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

1.1 MLD简介

MLD 是 Multicast Listener Discovery Protocol（组播侦听者发现协议）的简称，它用于 IPv6 路由器在其直连网段上发现组播侦听者。组播侦听者（Multicast Listener）是那些希望接收组播数据的主机节点。

路由器通过 MLD 协议，可以了解自己的直连网段上是否有 IPv6 组播组的侦听者，并在数据库里做相应记录。同时，路由器还维护与这些 IPv6 组播地址相关的定时器信息。

MLD 路由器使用 IPv6 单播链路本地地址作为源地址发送 MLD 报文。MLD 使用 ICMPv6（Internet Control Message Protocol for IPv6，IPv6 互联网控制消息协议）报文类型。所有的 MLD 报文被限制在本地链路上，跳数为 1。

1.1.1 MLD的版本

到目前为止，MLD 有两个版本：

- MLDv1（由 RFC 2710 定义），源自 IGMPv2
- MLDv2（由 RFC 3810 定义），源自 IGMPv3

所有版本的 MLD 协议都支持 ASM（Any-Source Multicast，任意信源组播）模型；MLDv2 可以直接应用于 SSM（Source-Specific Multicast，指定信源组播）模型，而 MLDv1 则需要在 MLD SSM Mapping 技术的支持下才能应用于 SSM 模型。

说明

有关 ASM 和 SSM 模型的介绍，请参见“IP 组播配置指导”中的“组播概述”。

1.1.2 MLDv1 原理简介

MLDv1 主要基于查询和响应机制完成对 IPv6 组播组成员的管理。

1. 查询器选举机制

当一个网段内有多台 IPv6 组播路由器时，由于它们都能从主机那里收到 MLD 成员关系报告报文（Multicast Listener Report Message），因此只需要其中一台路由器发送 MLD 查询报文（Query Message）即可，该路由器就称为 MLD 查询器（Querier）。这就需要有一个查询器的选举机制来确定由哪台路由器作为 MLD 查询器，其选举过程如下：

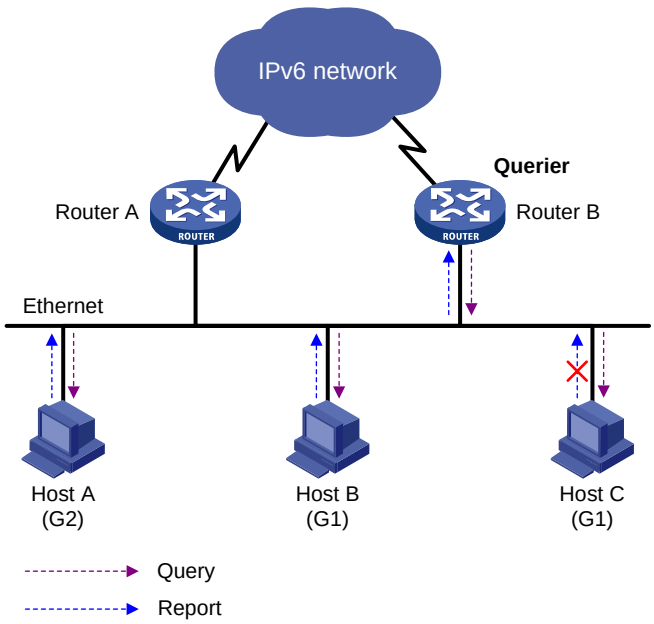
(1) 所有 MLD 路由器在初始时都认为自己是查询器，并向本地网段内的所有主机和路由器发送 MLD 普遍组查询（General Query）报文（目的地址为 FF02::1）；

(2) 本地网段中的其它 MLD 路由器在收到该报文后，将报文的源 IPv6 地址与自己的接口地址作比较。通过比较，IPv6 地址最小的路由器将成为查询器，其它路由器成为非查询器（Non-Querier）；

(3) 所有非查询器上都会启动一个定时器（即其它查询器存在时间定时器 Other Querier Present Timer）。在定时器超时前，如果收到了来自查询器的 MLD 查询报文，则重置该定时器；否则，就认为原查询器失效，并发起新的查询器选举过程。

2. 加入IPv6 组播组机制

图1-1 MLD 查询响应示意图



如 图 1-1 所示，假设Host B与Host C想要收到发往IPv6 组播组G1 的IPv6 组播数据，而Host A想要收到发往IPv6 组播组G2 的IPv6 组播数据，那么主机加入IPv6 组播组以及MLD查询器（Router B）维护IPv6 组播组成员关系的基本过程如下：

(1) 主机会主动向其要加入的 IPv6 组播组发送 MLD 成员关系报告报文以声明加入，而不必等待 MLD 查询器发来的 MLD 查询报文；

(2) MLD 查询器（Router B）周期性地以组播方式向本地网段内的所有主机和路由器发送普遍组查询报文（目的地址为 FF02::1）；

(3) 在收到该查询报文后，关注 G1 的 Host B 与 Host C 其中之一（这取决于谁的延迟定时器先超时）——譬如 Host B 会首先以组播方式向 G1 发送 MLD 成员关系报告报文，以宣告其属于 G1。由于本地网段中的所有主机都能收到 Host B 发往 G1 的报告报文，因此当 Host C 收到

该报告报文后，将不再发送同样针对 G1 的报告报文，因为 MLD 路由器（Router A 和 Router B）已知道本地网段中有对 G1 感兴趣的主机了。这个机制称为主机上的 MLD 成员关系报告抑制机制，该机制有助于减少本地网段的信息流量；

(4) 与此同时，由于 Host A 关注的是 G2，所以它仍将以组播方式向 G2 发送报告报文，以宣告其属于 G2；

(5) 经过以上的查询和响应过程，MLD 路由器了解到本地网段中有 G1 和 G2 的成员，于是由 IPv6 组播路由协议（如 IPv6 PIM）生成（*，G1）和（*，G2）组播转发项作为 IPv6 组播数据的转发依据，其中的“*”代表任意 IPv6 组播源；

(6) 当由 IPv6 组播源发往 G1 或 G2 的 IPv6 组播数据经过组播路由到达 MLD 路由器时，由于 MLD 路由器上存在（*，G1）和（*，G2）组播转发项，于是将该 IPv6 组播数据转发到本地网段，接收者主机便能收到该 IPv6 组播数据了。

3. 离开IPv6 组播组机制

当一个主机离开某 IPv6 组播组时：

(1) 该主机向本地网段内的所有 IPv6 组播路由器（目的地址为 FF02::2）发送离开组（Done）报文；

(2) 当查询器收到该报文后，向该主机所声明要离开的那个 IPv6 组播组发送特定组查询（Multicast-Address-Specific Query）报文（目的地址字段和组地址字段均填充为所要查询的 IPv6 组播组地址）；

(3) 如果该网段内还有该 IPv6 组播组的其它成员，则这些成员在收到特定组查询报文后，会在该报文中所设定的最大响应时间（Maximum Response Delay）内发送成员关系报告报文；

(4) 如果在最大响应时间内收到了该 IPv6 组播组其它成员发送的成员关系报告报文，查询器就会继续维护该IPv6组播组的成员关系；否则，查询器将认为该网段内已无该IPv6组播组的成员，于是不再维护这个 IPv6 组播组的成员关系。

1.1.3 MLDv2 原理简介

MLDv2 的原理与 MLDv1 基本相同，并新增了以下特性：

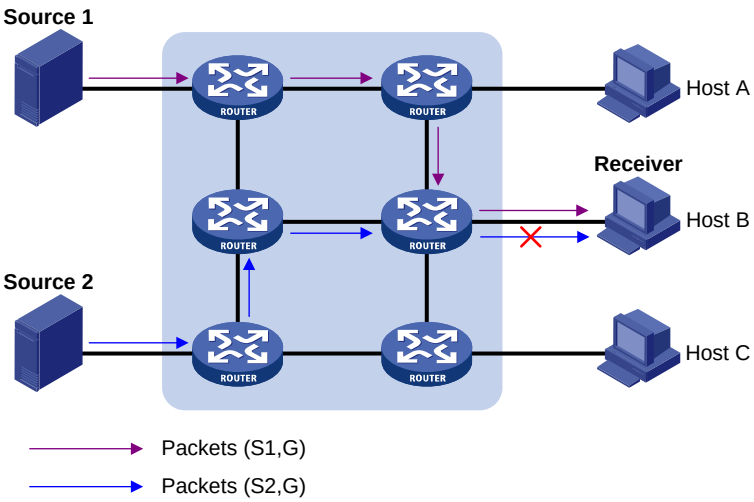
1. 对IPv6 组播源的过滤

MLDv2 增加了针对 IPv6 组播源的过滤模式（INCLUDE/EXCLUDE），使主机在加入某 IPv6 组播组 G 的同时，能够明确要求接收或拒绝来自某特定 IPv6 组播源 S 的 IPv6 组播信息。当主机加入 IPv6 组播组时：

- 若要求只接收来自指定 IPv6 组播源如 S1、S2、……发来的 IPv6 组播信息，则其报告报文中可以标记为 INCLUDE Sources（S1，S2，……）；
- 若拒绝接收来自指定 IPv6 组播源如 S1、S2、……发来的 IPv6 组播信息，则其报告报文中可以标记为 EXCLUDE Sources（S1，S2，……）。

如 图 1-2 所示，网络中存在Source 1（S1）和Source 2（S2）两个IPv6 组播源，均向IPv6 组播组G发送IPv6 组播报文。Host B仅对从Source 1 发往G的信息感兴趣，而对来自Source 2 的信息没有兴趣。

图1-2 指定源组的 IPv6 组播流路经



如果主机与路由器之间运行的是 MLDv1，Host B 加入 IPv6 组播组 G 时无法对 IPv6 组播源进行选择，因此无论 Host B 是否需要，来自 Source 1 和 Source 2 的 IPv6 组播信息都将传递给 Host B。当主机与路由器之间运行了 MLDv2 之后，Host B 就可以要求只接收来自 Source 1、发往 G 的 IPv6 组播信息（S1，G），或要求拒绝来自 Source 2、发往 G 的 IPv6 组播信息（S2，G），这样就只有来自 Source 1 的 IPv6 组播信息才能传递给 Host B 了。

2. MLD状态

运行 MLDv2 的组播路由器按每条直连链路上的组播地址（per multicast address per attached link）来保持 IPv6 组播组的状态。IPv6 组播组的状态包括：

- 过滤模式：保持对 INCLUDE 或 EXCLUDE 的状态跟踪。
- 源列表：保持对新增或删除 IPv6 组播源的跟踪。
- 定时器：表示 IPv6 组播地址超时后切换到 INCLUDE 模式的过滤定时器、关于源记录的源定时器等。

3. 接收者主机的状态侦听

运行 MLDv2 的组播路由器通过侦听接收者主机的状态，记录和维护网段上加入到源组的主机的信息。

1.1.4 MLD报文类型

下面以 MLDv2 为例对 MLD 的报文类型进行介绍：

1. MLD查询报文

MLD查询器通过发送MLD查询报文来了解相邻接口的组播侦听状态。MLD查询报文的格式如 图 1-3 所示，图中深蓝色部分为MLDv1 的报文格式，各字段的含义如 表 1-1 所示。

图1-3 MLDv2 查询报文格式

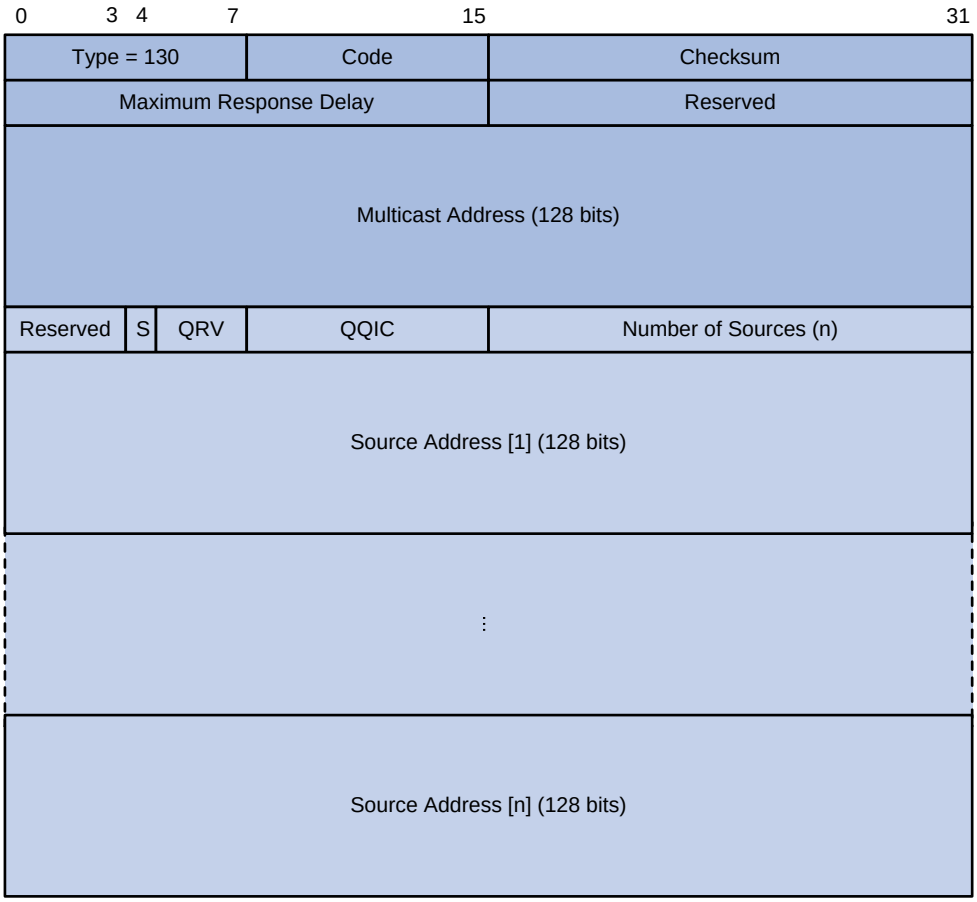


表1-1 MLDv2 查询报文各字段含义

字段	描述
Type = 130	报文类型，130代表查询报文
Code	初始化为0
Checksum	标准的IPv6校验和
Maximum Response Delay	主机发送报告报文前允许的最大响应时间
Reserved	保留字段，初始化为0
Multicast Address	• 普遍组查询中，此字段设置为 0 • 特定组或特定源组查询中，此字段设置为待查询的 IPv6 组播组地址
S	标识位，表示路由器接收到查询报文后是否对定时器更新进行抑制
QRV	查询器的健壮性变量（Querier’s Robustness Variable）
QQIC	查询器发送普遍组查询报文的查询间隔（Querier’s Query Interval Code）
Number of Sources	• 普遍组查询或特定组查询中，此字段设置为 0 • 特定源组查询中，此字段表示查询报文中包含的源地址个数

字段	描述
Source Address(i)	特定源组查询中的IPv6组播源地址（i＝1, 2, …, n，其中n表示源地址的个数）

2. MLD报告报文

主机通过发送MLD报告报文来汇报当前的组播侦听状态。MLD报告报文的格式如 图 1-4 所示，各字段的含义如 表 1-2 所示。

图1-4 MLDv2 报告报文格式

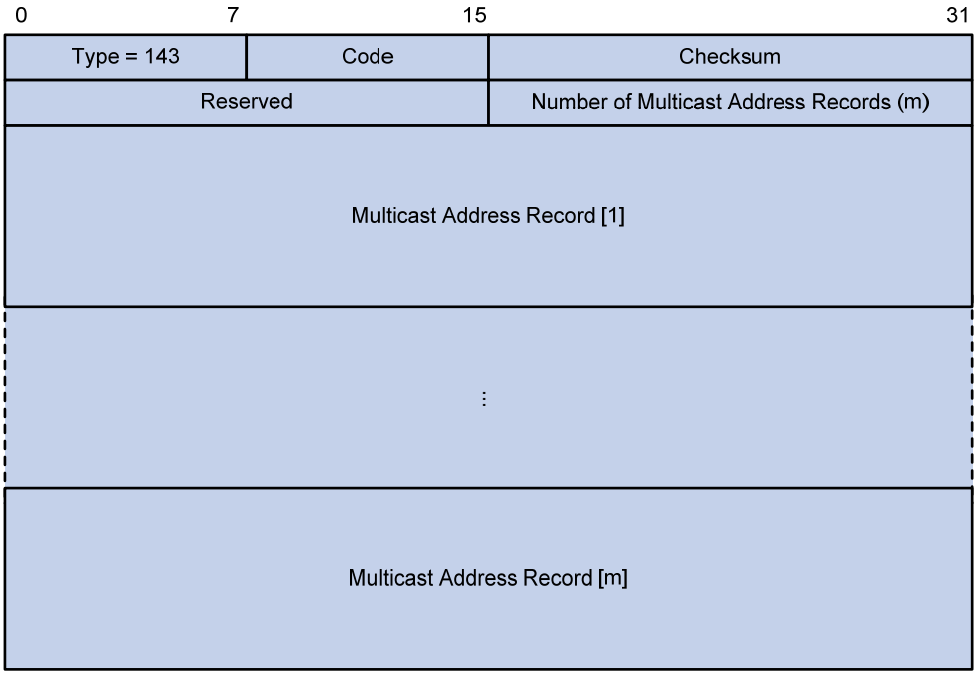


表1-2 MLDv2 报告报文各字段含义

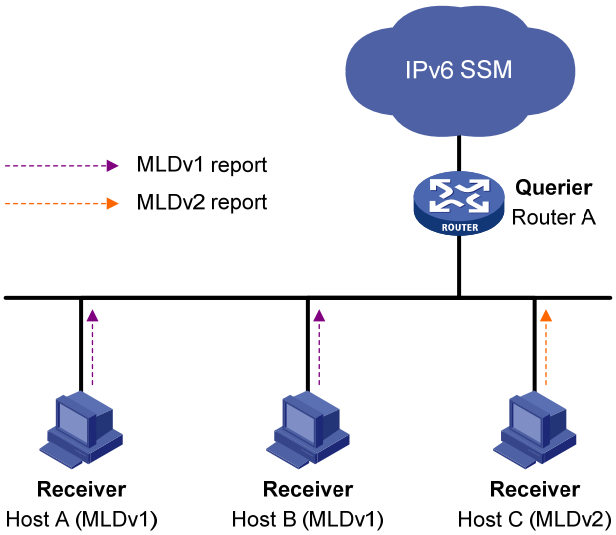
字段	描述
Type = 143	报文类型，143代表报告报文
Reserved	保留字段，发送时设置为0，接收时忽略此值
Checksum	标准的IPv6校验和
Number of Multicast Address Records	IPv6组播地址记录的个数
Multicast Address Record(i)	组播地址记录，表示主机在接口上侦听到的每个IPv6组播地址信息，包括记录类型、IPv6组播地址、IPv6源地址等（i＝1, 2, …, m，其中m表示IPv6组播地址记录的个数）

1.1.5 MLD SSM Mapping

MLD SSM Mapping 通过在路由器上配置 SSM 静态映射规则，从而为运行 MLDv1 的接收者主机提供对 SSM 模型的支持。

SSM 模型要求在接收者主机所在的网段，路由器能够了解主机加入 IPv6 组播组时所指定的 IPv6 组播源。如果接收者主机上运行的是 MLDv2，则可以在 MLDv2 的报告报文中直接指定 IPv6 组播源的地址；如果某些接收者主机只能运行 MLDv1，则在 MLDv1 的报告报文中无法指定 IPv6 组播源的地址。这种情况下需要通过在路由器上配置 MLD SSM Mapping 功能，将 MLDv1 报告报文中所包含的（*，G）信息映射为（G，INCLUDE，（S1，S2...））信息。

图1-5 MLD SSM Mapping 组网图



在如 图 1-5 所示的IPv6 SSM网络中，Host A、Host B和Host C上分别运行MLDv1 和MLDv2。在不允许将Host A和Host B升级为MLDv2 的情况下，若要为Host A和Host B也提供SSM组播服务，则需在Router A上配置MLD SSM Mapping功能。

配置完成后，当 Router A 收到来自主机的 MLDv1 报告报文时，首先检查该报文中所携带的 IPv6 组播组地址 G，然后根据检查结果的不同分别进行处理：

(1) 如果 G 不在 IPv6 SSM 组地址范围内，则提供 ASM 组播服务。

(2) 如果 G 在 IPv6 SSM 组地址范围内：

- 若 Router A 上没有 G 对应的 MLD SSM Mapping 规则，则无法提供 SSM 组播服务，丢弃该报文；
- 若 Router A 上有 G 对应的 MLD SSM Mapping 规则，则依据规则将报告报文中所包含的(*，G）信息映射为（G，INCLUDE，（S1，S2...））信息，可以提供 SSM 组播服务。

说明

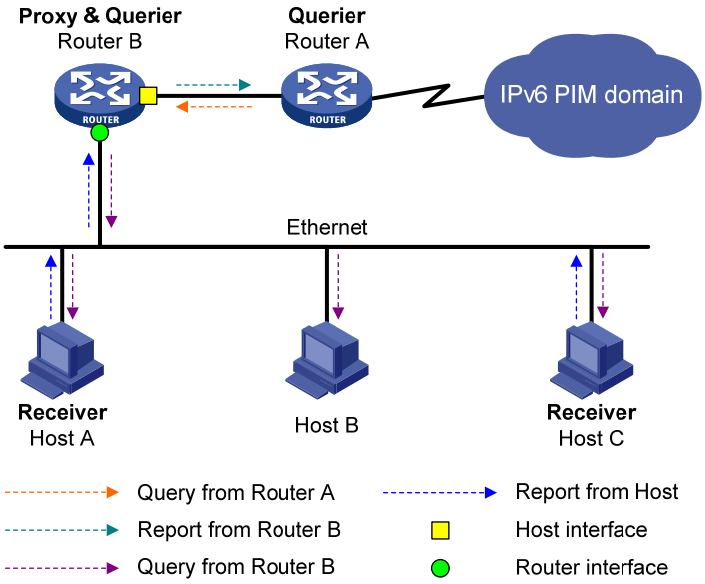
- MLD SSM Mapping 不对 MLDv2 的报告报文进行处理。
- 有关 IPv6 SSM 组地址范围的介绍，请参见“IP 组播配置指导”中的“IPv6 PIM”。

1.1.6 MLD Proxying

在一些简单的树型网络拓扑中，边缘设备上并不需要运行复杂的IPv6组播路由协议(如IPv6 PIM)，可以通过在这些设备上配置 MLD Proxying（MLD 代理）功能，使其代理下游主机来发送 MLD 报

文及维护组成员关系，并基于该关系进行 IPv6 组播转发。在上游设备看来，配置了 MLD Proxying 功能的设备（称为 MLD 代理设备）不再是一个 IPv6 PIM 邻居，而只是一台主机。

图1-6 MLD Proxying 组网图



如 图 1-6 所示，MLD Proxying模型中定义了以下两种接口类型：

- 上行接口：又称代理接口，指 MLD 代理设备上运行 MLD Proxying 功能的接口，即朝向组播分发树树根方向的接口。由于该接口执行 MLD 协议的主机行为，因此也称为主机接口（Host Interface）。
- 下行接口：指 MLD 代理设备上除上行接口外其它运行 MLD 协议的接口，即背向组播分发树树根方向的接口。由于该接口执行 MLD 协议的路由器行为，因此也称为路由器接口（Router Interface）。

MLD 代理设备上维护着一个组成员关系数据库（Membership Database），将所有下行接口维护的组成员关系记录都存到这个数据库中。组成员关系记录的结构如下：（Multicast-address，Filter-mode，Source-list），每条记录都是各下行接口上具有相同组地址的成员关系记录的合集。

上行接口正是依据这个数据库来执行主机行为——当收到查询报文时根据当前数据库状态响应报告报文，或者当数据库变化时主动发送报告或离开报文；而下行接口则执行路由器行为——参与查询器的选举、发送查询报文并根据报告报文维护组成员关系等。

1.1.7 协议规范

与 MLD 相关的协议规范有：

- RFC 2710：Multicast Listener Discovery (MLD) for IPv6
- RFC 3810：Multicast Listener Discovery Version 2 (MLDv2) for IPv6
- RFC 4605：Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")

1.2 MLD配置任务简介

表1-3 MLD 配置任务简介

配置任务		说明	详细配置
配置MLD基本功能	使能MLD	必选	1.3.2
	配置MLD版本	可选	1.3.3
	配置静态加入	可选	1.3.4
	配置IPv6组播组过滤器	可选	1.3.5
	配置接口加入的IPv6组播组最大数量	可选	1.3.6
调整MLD性能	配置MLD报文选项	可选	1.4.2
	配置MLD查询和响应	可选	1.4.3
	配置MLD快速离开	可选	1.4.4
	配置MLD主机跟踪功能	可选	1.4.5
	配置MLD报文的DSCP优先级	可选	1.4.6
配置MLD SSM Mapping	使能MLD SSM Mapping	可选	1.5.2
	配置MLD SSM Mapping规则	可选	1.5.3
配置MLD Proxying	使能MLD Proxying	可选	1.6.2
	配置下行接口的IPv6组播转发能力	可选	1.6.3

说明

- MLD 视图下的配置具有全局性，接口视图下的配置仅对所在接口有效。
- 若没有在接口视图下进行配置，则该接口继承 MLD 视图下的全局配置；若两个视图都进行了配置，则优先采用该接口视图下所作的配置。

1.3 配置MLD基本功能

1.3.1 配置准备

在配置 MLD 基本功能之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通
- 配置 IPv6 PIM-DM（或 IPv6 PIM-SM）

在配置 MLD 基本功能之前，需准备以下数据：

- MLD 的版本
- 以静态方式加入的 IPv6 组播组和 IPv6 组播源的地址
- IPv6 组播组过滤的 ACL 规则

- 接口加入的 IPv6 组播组最大数量

1.3.2 使能MLD

在需要建立和维护 IPv6 组播组成员关系的接口上使能 MLD。

表1-4 使能 MLD

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6组播路由	**multicast ipv6 routing-enable**	必选 缺省情况下，IPv6组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能MLD	**mld enable**	必选 缺省情况下，没有使能MLD

有关 **multicast ipv6 routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。

1.3.3 配置MLD版本

由于不同版本 MLD 协议的报文结构与种类不同，因此需要为同一网段上的所有路由器配置相同版本的 MLD，否则 MLD 将不能正常运行。

1. 全局配置MLD版本

表1-5 全局配置 MLD 版本

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD视图	**mld**	-
配置MLD的版本	**version** *version-number*	必选 缺省情况下，MLD的版本为MLDv1

2. 在接口上配置MLD版本

表1-6 在接口上配置 MLD 版本

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
配置MLD的版本	**mld version** *version-number*	必选 缺省情况下，MLD的版本为MLDv1

1.3.4 配置静态加入

在配置了静态加入 IPv6 组播组或组播源组后，接口将作为该 IPv6 组播组的虚拟组成员存在，从而可以接收发往该组的 IPv6 组播数据，以测试 IPv6 组播数据的转发。

表1-7 配置静态加入

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置静态加入IPv6组播组或组播源组	**mld static-group** *ipv6-group-address* [**source** *ipv6-source-address*]	必选 缺省情况下，接口没有以静态方式加入任何IPv6组播组或组播源组

说明

- 在运行 IPv6 PIM-SM 的设备上配置静态加入时，如果待配接口上使能了 IPv6 PIM-SM，则该接口必须为 IPv6 PIM-SM 的 DR，否则该接口将不能加入 IPv6 组播组或组播源组；如果待配接口上未使能 IPv6 PIM-SM 但使能了 MLD，则该接口必须为 MLD 查询器，否则该接口也不能加入 IPv6 组播组或组播源组。有关 IPv6 PIM-SM 和 DR 的介绍，请参见“IP 组播配置指导”中的“IPv6 PIM”。
- 在配置了静态加入后，接口并不会对 MLD 查询器发出的查询报文进行响应；当配置静态加入或取消静态加入的配置时，接口也不会主动发送 MLD 成员关系报告报文或 MLD 离开组报文。也就是说，该接口并没有真正成为该 IPv6 组播组的成员。

1.3.5 配置IPv6 组播组过滤器

如果不希望接口所在网段上的主机加入某些 IPv6 组播组，可在该接口上配置 IPv6 ACL 规则作为过滤器，接口将按照该规则对收到的 MLD 成员关系报告报文进行过滤，只为该规则所允许的 IPv6 组播组维护组成员关系。

表1-8 配置 IPv6 组播组过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
配置IPv6组播组过滤器	**mld group-policy** *acl6-number* [*version-number*]	必选 缺省情况下，接口上没有配置IPv6组播组过滤器，即该接口下的主机可以加入任意合法的IPv6组播组

1.3.6 配置接口加入的IPv6 组播组最大数量

通过配置接口加入的 IPv6 组播组最大数量，可以方便用户灵活控制接口所能加入的 IPv6 组播组数量。

表1-9 配置接口加入的 IPv6 组播组最大数量

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置接口加入的IPv6组播组最大数量	**mld group-limit** *limit*	必选 缺省情况下，接口加入的IPv6组播组最大数量为1000

说明

本配置只限制接口动态加入的 IPv6 组播组数量，而不限制其静态加入的 IPv6 组播组数量。

1.4 调整MLD性能

说明

对于本节中的配置任务来说:

- 在 MLD 视图下所作的配置对所有接口生效，在接口视图下所作的配置仅对当前接口生效;
- 如果在这两个视图下进行了相同功能或参数的配置，则不论配置先后，接口视图下的配置将被优先采用。

1.4.1 配置准备

在调整 MLD 性能之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通
- 配置 MLD 基本功能

在调整 MLD 性能之前，需准备以下数据：

- MLD 查询器的启动查询间隔
- MLD 查询器的启动查询次数
- 发送 MLD 普遍组查询报文的时间间隔
- MLD 查询器的健壮系数
- MLD 普遍组查询的最大响应时间
- 最后组成员查询间隔
- MLD 其它查询器的存在时间
- MLD 报文的 DSCP 优先级

1.4.2 配置MLD报文选项

由于 MLD 特定组查询报文和 MLD 特定源组查询报文的存在，而 IPv6 组播组是千变万化的，设备不可能加入所有的组，因此 MLD 需要借助 Router-Alert（路由器报警）选项来将送达本地、但没有加入的IPv6组播报文送往上层协议进行处理。有关Router-Alert选项的详细介绍，请参考RFC 2113。

对于 IPv6 头中是否携带 Router-Alert 选项的 MLD 报文，设备会做出不同的处理：

- 出于兼容性考虑，缺省情况下设备不对 Router-Alert 选项进行检查，即处理所有收到的 MLD 报文。此时，MLD 报文中无论是否携带有 Router-Alert 选项，设备都会将其送给上层协议进行处理。
- 为了提高设备性能、减少不必要的开支，同时出于协议安全性的考虑，可以配置设备丢弃未携带 Router-Alert 选项的 MLD 报文，此时，当设备收到 MLD 报文时，会检查该报文的 Router-Alert 选项，如果没有携带该选项，就丢弃该报文。

1. 全局配置MLD报文选项

表1-10 全局配置 MLD 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD视图	**mld**	-
配置丢弃未携带Router-Alert选项的MLD报文	**require-router-alert**	必选 缺省情况下，设备不对Router-Alert选项进行检查
配置在发送的MLD报文中携带Router-Alert选项	**send-router-alert**	必选 缺省情况下，在发送的MLD报文中携带Router-Alert选项

2. 在接口上配置MLD报文选项

表1-11 在接口上配置 MLD 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
配置丢弃未携带Router-Alert选项的MLD报文	**mld require-router-alert**	必选 缺省情况下，设备不对Router-Alert选项进行检查
配置在发送的MLD报文中携带Router-Alert选项	**mld send-router-alert**	必选 缺省情况下，在发送的MLD报文中携带Router-Alert选项

1.4.3 配置MLD查询和响应

MLD 查询器的健壮系数是为了弥补可能发生的网络丢包而设置的报文重传次数，健壮系数越大，MLD 查询器就越“健壮”，但是 IPv6 组播组超时所需的时间也就越长。

当 MLDv1/v2 查询器启动时，会以“MLD 查询器启动查询间隔”为时间间隔发送“MLD 查询器启动查询次数”次 MLD 普遍组查询报文。

MLDv1/v2 查询器会周期性地发送 MLD 普遍组查询报文，以判断网络上是否有 IPv6 组播组成员，发送间隔即为“发送 MLD 普遍组查询报文的时间间隔”。可以根据网络的实际情况来修改周期性发送 MLD 普遍组查询报文的时间间隔。

当 MLDv1 查询器收到 MLD 离开组报文后，会以“最后组成员查询间隔”为时间间隔发送“最后组成员查询次数”次 MLD 特定组查询报文；当 MLDv2 查询器收到改变 IPv6 组播组与 IPv6 组播源列表关系的 MLD 报告报文后，也会以“最后组成员查询间隔”为时间间隔发送“最后组成员查询次数”次 MLD 特定源组查询报文。最后组成员查询次数在数值上等于 MLD 查询器的健壮系数。

在收到 MLD 查询报文（包括普遍组查询、特定组查询和特定源组查询）后，主机会为其所加入的每个 IPv6 组播组都启动一个延迟定时器，其值在 0 到最大响应时间（该时间值从 MLD 查询报文的最大响应时间字段获得）中随机选定，当定时器的值减为 0 时，主机就会向该定时器对应的 IPv6 组播组发送 MLD 成员关系报告报文。

合理配置 MLD 查询的最大响应时间，既可以使主机对 MLD 查询报文做出快速响应，又可以减少由于定时器同时超时，造成大量主机同时发送报告报文而引起的网络拥塞：

- 对于 MLD 普遍组查询报文来说，通过配置 MLD 普遍组查询的最大响应时间来填充其最大响应时间字段；
- 对于 MLD 特定组查询报文和 MLD 特定源组查询报文来说，所配置的最后组成员查询间隔将被填充到其最大响应时间字段。也就是说，MLD 特定组查询的最大响应时间在数值上等于最后组成员查询间隔。

当同一网段上有多台组播路由器时，由 MLD 查询器负责发送 MLD 查询报文。如果非查询器在“其它查询器存在时间”超时前没有收到来自查询器的 MLD 查询报文，就会认为原有查询器失效，从而触发新的查询器选举过程；否则，非查询器将重置“其它查询器存在时间定时器”。

1. 全局配置MLD查询和响应

表1-12 全局配置 MLD 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD视图	**mld**	-

操作	命令	说明
配置MLD查询器的健壮系数	**robust-count** *robust-value*	必选 缺省情况下，MLD查询器的健壮系数为2
配置MLD查询器的启动查询间隔	**startup-query-interval** *interval*	必选 缺省情况下，MLD查询器的启动查询间隔为发送MLD普遍组查询报文时间间隔的1/4
配置MLD查询器的启动查询次数	**startup-query-count** *value*	必选 缺省情况下，MLD查询器的启动查询次数等于MLD查询器的健壮系数
配置发送MLD普遍组查询报文的时间间隔	**timer query** *interval*	必选 缺省情况下，发送MLD普遍组查询报文的时间间隔为125秒
配置MLD普遍组查询的最大响应时间	**max-response-time** *interval*	必选 缺省情况下，MLD普遍组查询的最大响应时间为10秒
配置最后组成员查询间隔	**last-listener-query-interval** *interval*	必选 缺省情况下，最后组成员查询间隔为1秒
配置MLD其它查询器的存在时间	**timer other-querier-present** *interval*	必选 缺省情况下，MLD其它查询器的存在时间＝发送MLD普遍组查询报文的时间间隔×MLD查询器的健壮系数＋MLD普遍组查询的最大响应时间÷2

2. 在接口上配置MLD查询和响应

表1-13 在接口上配置 MLD 查询和响应

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置MLD查询器的健壮系数	**mld robust-count** *robust-value*	必选 缺省情况下，MLD查询器的健壮系数为2
配置MLD查询器的启动查询间隔	**mld startup-query-interval** *interval*	必选 缺省情况下，MLD查询器的启动查询间隔为发送MLD普遍组查询报文时间间隔的1/4
配置MLD查询器的启动查询次数	**mld startup-query-count** *value*	必选 缺省情况下，MLD查询器的启动查询次数等于MLD查询器的健壮系数
配置发送MLD普遍组查询报文的时间间隔	**mld timer query** *interval*	必选 缺省情况下，发送MLD普遍组查询报文的时间间隔为125秒

操作	命令	说明
配置MLD普遍组查询的最大响应时间	**mld max-response-time** *interval*	必选 缺省情况下，MLD普遍组查询的最大响应时间为10秒
配置最后组成员查询间隔	**mld last-listener-query-interval** *interval*	必选 缺省情况下，最后组成员查询间隔为1秒
配置MLD其它查询器的存在时间	**mld timer other-querier-present** *interval*	必选 缺省情况下，MLD其它查询器的存在时间＝发送MLD普遍组查询报文的时间间隔×MLD查询器的健壮系数＋MLD普遍组查询的最大响应时间÷2

注意

- 应确保 MLD 其它查询器的存在时间大于发送 MLD 普遍组查询报文的时间间隔，否则有可能导致网络内的 MLD 查询器反复变化。
- 应确保发送 MLD 普遍组查询报文的时间间隔大于 MLD 普遍组查询的最大响应时间，否则有可能造成对 IPv6 组播组成员的误删。

1.4.4 配置MLD快速离开

在某些应用（如 ADSL 拨号上网）中，MLD 查询器的一个端口唯一对应着一台接收者主机，当主机在多个 IPv6 组播组间频繁切换（如进行电视选台）时，为了快速响应主机的离开组报文，可以在 MLD 查询器上开启 MLD 快速离开功能。

在开启了 MLD 快速离开功能之后，当 MLD 查询器收到来自主机的离开组报文时，不再发送 MLD 特定组查询报文或 MLD 特定源组查询报文，而是直接向上游发送离开通告，这样一方面减小了响应延迟，另一方面也节省了网络带宽。

1. 全局配置MLD快速离开

表1-14 全局配置 MLD 快速离开

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD视图	**mld**	-
配置IPv6组播组成员快速离开	**fast-leave** [**group-policy** *acl6-number*]	必选 缺省情况下，IPv6组播组成员快速离开功能是关闭的

2. 在接口上配置MLD快速离开

表1-15 在接口上配置 MLD 快速离开

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置IPv6组播组成员快速离开	**mld fast-leave** [**group-policy** *acl6-number*]	必选 缺省情况下，IPv6组播组成员快速离开功能是关闭的

注意

MLD 快速离开配置只对除 VLAN 接口以外的其它三层接口生效，如三层以太网端口、三层聚合接口及 Tunnel 接口。

1.4.5 配置MLD主机跟踪功能

通过使能 MLD 主机跟踪功能，可以使设备能够记录正在接收 IPv6 组播数据的成员主机信息（包括主机的 IPv6 地址、运行时间和超时时间等），以便于网络管理员对这些主机进行监控和管理。

1. 全局配置MLD主机跟踪功能

表1-16 全局配置 MLD 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD视图	**mld**	-
全局使能MLD主机跟踪功能	**host-tracking**	必选 缺省情况下，MLD主机跟踪功能处于关闭状态

2. 在接口上配置MLD主机跟踪功能

表1-17 在接口上配置 MLD 主机跟踪功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
在接口上使能MLD主机跟踪功能	**mld host-tracking**	必选 缺省情况下，MLD主机跟踪功能处于关闭状态

1.4.6 配置MLD报文的DSCP优先级

在 IPv6 报文头中，包含一个 8bit 的 Traffic class 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 MLD 报文的 DSCP 优先级进行配置。

表1-18 配置 MLD 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD视图	**mld**	-
配置MLD报文的DSCP优先级	**dscp** *dscp-value*	可选 缺省情况下，MLD报文的DSCP优先级为48

1.5 配置MLD SSM Mapping

在 IPv6 SSM 网络中，由于各种可能的限制，某些接收者主机只能运行 MLDv1。为了向这些仅支持 MLDv1 的接收者主机提供 SSM 服务，可以在路由器上配置 MLD SSM Mapping 功能。

1.5.1 配置准备

在配置 MLD SSM Mapping 之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通
- 配置 MLD 基本功能

1.5.2 使能MLD SSM Mapping

表1-19 使能 MLD SSM Mapping

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能MLD SSM Mapping功能	**mld ssm-mapping enable**	必选 缺省情况下，接口上的MLD SSM Mapping功能处于关闭状态

说明

为保证本网段内运行任意版本 MLD 的接收者主机都能得到 SSM 服务，建议在处于该网段的接口上运行 MLDv2。

1.5.3 配置MLD SSM Mapping规则

通过多次配置可以实现同一 IPv6 组播组到多个 IPv6 组播源的映射。

表1-20 配置 MLD SSM Mapping 规则

操作	命令	说明
进入系统视图	**system-view**	-
进入MLD视图	**mld**	-
配置MLD SSM Mapping规则	**ssm-mapping** *ipv6-group-address prefix-length ipv6-source-address*	必选 缺省情况下，未配置MLD SSM Mapping规则

注意

若 VLAN 接口上运行的是 MLDv2，则在使用 **mld-snooping host-join** 命令配置模拟主机加入时即便不指定 IPv6 组播源，模拟主机也仍将发送 MLDv2 的报告报文，这种情况下不会依据 MLD SSM Mapping 规则创建相应的 IPv6 组播组。有关 **mld-snooping host-join** 命令的详细介绍，请参见“IP 组播命令参考”中的“MLD Snooping”。

1.6 配置MLD Proxying

1.6.1 配置准备

在配置 MLD Proxying 之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通
- 使能 IPv6 组播路由

1.6.2 使能MLD Proxying

在设备朝向组播分发树树根方向的接口上使能了 MLD 代理功能之后，该设备就成为了 MLD 代理设备。

表1-21 使能 MLD Proxying

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能MLD代理功能	**mld proxying enable**	必选 缺省情况下，接口上的MLD代理功能处于关闭状态

说明

- 一台设备上只允许有一个接口使能 MLD 代理功能。
- 在进行 MLD 的相关配置时，在已使能 MLD 代理功能的接口上不允许再使能 MLD，且只有 **mld require-router-alert**、**mld send-router-alert** 和 **mld version** 命令可在该接口上生效。
- 在已使能 MLD 代理功能的设备上不允许再使能其它 IPv6 组播路由协议（如不能在接口上使能 IPv6 PIM-DM 或 IPv6 PIM-SM，但在 IPv6 PIM 视图下配置的 **source-lifetime**、**source-policy** 和 **ssm-policy** 命令仍会生效），反之亦然。
- 若某 VLAN 内已使能了 MLD Snooping，则该 VLAN 对应的 VLAN 接口上不能再使能 MLD 代理功能，反之亦然。

1.6.3 配置下行接口的IPv6 组播转发能力

通常，IPv6 组播数据只能被查询器转发，非查询器不具备 IPv6 组播转发能力，这样可以避免 IPv6 组播数据被重复转发。对于 MLD 代理设备也同样如此，只有当其下行接口是查询器时，IPv6 组播数据才能通过该接口转发给下游主机。

但在某些情况下，MLD 代理设备的下行接口不能在查询器竞选中获胜，此时应使能该下行接口在非查询器状态下的 IPv6 组播转发能力。

表1-22 配置下行接口的 IPv6 组播转发能力

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能下行接口在非查询器状态下的IPv6组播转发能力	**mld proxying forwarding**	必选 缺省情况下，当MLD代理设备的下行接口处于非查询器状态时，不转发IPv6组播数据

注意

在共享网段内存在多台 MLD 代理设备的情况下，当其中一台 MLD 代理设备的下行接口已竞选为查询器时，不应再在其它 MLD 代理设备的下行接口上进行此配置，否则该网段内将收到多份重复的 IPv6 组播数据流。

1.7 MLD显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 MLD 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 MLD 的统计信息。

表1-23 MLD 显示和维护

操作	命令
查看MLD组的信息	**display mld group** [*ipv6-group-address* | **interface** *interface-type interface-number*] [**static** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看MLD组的二层端口信息	**display mld group port-info** [**vlan** *vlan-id*] [**slot** *slot-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上MLD跟踪的主机信息	**display mld host interface** *interface-type interface-number* **group** *ipv6-group-address* [**source** *ipv6-source-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看二层端口上MLD跟踪的主机信息	**display mld host port-info vlan** *vlan-id* **group** *ipv6-group-address* [**source** *ipv6-source-address*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上MLD配置和运行信息	**display mld interface** [*interface-type interface-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看MLD代理组的信息	**display mld proxying group** [*group-address*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看MLD路由表的信息	**display mld routing-table** [*ipv6-source-address* [*prefix-length*] | *ipv6-group-address* [*prefix-length*] | **flags** { **act** | **suc** }] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看MLD SSM Mapping规则	**display mld ssm-mapping** *ipv6-group-address* [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看依据MLD SSM Mapping规则创建的IPv6组播组信息	**display mld ssm-mapping group** [*ipv6-group-address* | **interface** *interface-type interface-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上依据MLD SSM Mapping规则加入的主机信息	**display mld ssm-mapping host interface** *interface-type interface-number* **group** *ipv6-group-address* **source** *ipv6-source-address* [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除MLD组的动态加入记录	**reset mld group** { **all** | **interface** interface-type interface-number { **all** | ipv6-group-address [prefix-length] [ipv6-source-address [prefix-length]] } }
清除MLD组二层端口的动态加入记录	**reset mld group port-info** { **all** | *ipv6-group-address* } [**vlan** *vlan-id*]
清除依据MLD SSM Mapping规则创建的IPv6组播组信息	**reset mld ssm-mapping group** { **all** | **interface** interface-type interface-number { **all** | ipv6-group-address [prefix-length] [ipv6-source-address [prefix-length]] } }

说明

- **reset mld group** 和 **reset mld group port-info** 命令只能清除动态加入记录，而无法清除静态加入记录。
- **display mld host interface** 只能查看除 Vlan 接口外的三层接口上 MLD 跟踪的主机信息。
- **display mld ssm-mapping host interface** 只能查看除 Vlan 接口外的三层接口上依据 MLD SSM Mapping 规则加入的主机信息。

注意

执行 **reset mld group** 命令可能导致接收者中断组播信息的接收。

1.8 MLD典型配置举例

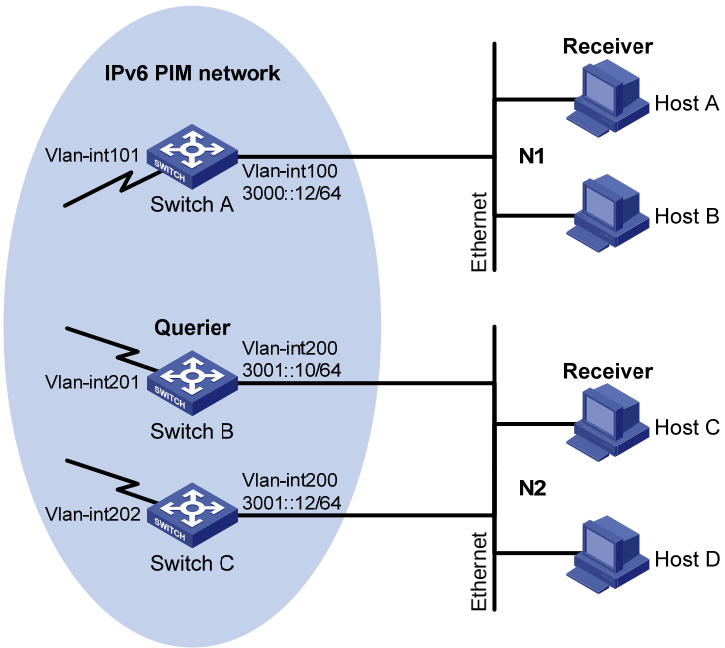
1.8.1 MLD基本功能配置举例

1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者组成末梢网络 N1 和 N2，Host A 与 Host C 分别为 N1 和 N2 中的组播信息接收者。
- IPv6 PIM 网络中的 Switch A 连接 N1，Switch B 与 Switch C 共同连接 N2。
- Switch A 通过 Vlan-interface100 连接 N1，通过 Vlan-interface101 连接 IPv6 PIM 网络中的其它设备。
- Switch B 与 Switch C 分别通过各自的 Vlan-interface200 连接 N2，并分别通过 Vlan-interface201 和 Vlan-interface202 连接 IPv6 PIM 网络中的其它设备。
- Switch A 与 N1 之间运行 MLDv1；Switch B、Switch C 与 N2 之间也分别运行 MLDv1，且由于 Switch B 的接口 IPv6 地址较小，因此在 N2 中通常由其来充当 MLD 查询器。
- 通过配置，使 N1 中的主机只能加入 IPv6 组播组 FF1E::101，而对 N2 中的主机则无任何限制。

2. 组网图

图1-7 MLD 基本功能配置组网图



3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址和 IPv6 单播路由协议

使能各交换机的IPv6 转发功能，并按照 图 1-7 配置各接口的IP地址和前缀长度，具体配置过程略。

配置 IPv6 PIM 网络内的各交换机之间采用 OSPFv3 协议进行互连，确保 IPv6 PIM 网络内部在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IPv6 组播路由，并使能 IPv6 PIM-DM 和 MLD

在 Switch A 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在主机侧接口 Vlan-interface100 上使能 MLD。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 dm
[SwitchA-Vlan-interface101] quit
```

在 Switch B 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在主机侧接口 Vlan-interface200 上使能 MLD。

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] mld enable
[SwitchB-Vlan-interface200] pim ipv6 dm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 201
[SwitchB-Vlan-interface201] pim ipv6 dm
[SwitchB-Vlan-interface201] quit
```

在 Switch C 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在主机侧接口 Vlan-interface200 上使能 MLD。

```
<SwitchC> system-view
[SwitchC] multicast ipv6 routing-enable
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] mld enable
[SwitchC-Vlan-interface200] pim ipv6 dm
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 202
[SwitchC-Vlan-interface202] pim ipv6 dm
[SwitchC-Vlan-interface202] quit
```

(3) 配置 IPv6 组播组过滤器

在 Switch A 上限定接口 Vlan-interface100 下的主机只能加入 IPv6 组播组 FF1E::101。

```
[SwitchA] acl ipv6 number 2001
[SwitchA-acl6-basic-2001] rule permit source ff1e::101 128
[SwitchA-acl6-basic-2001] quit
[SwitchA] interface vlan-interface 100
```

```
[SwitchA-Vlan-interface100] mld group-policy 2001
[SwitchA-Vlan-interface100] quit
```

(4) 检查配置效果

通过使用 **display mld interface** 命令可以查看各交换机接口上 MLD 的配置和运行情况。例如：

查看 Switch B 在 Vlan-interface200 上的 MLD 信息。

```
[SwitchB] display mld interface vlan-interface 200
 Vlan-interface200(FE80::200:5EFF:FE66:5100):
   MLD is enabled
   Current MLD version is 1
   Value of query interval for MLD(in seconds): 125
   Value of other querier present interval for MLD(in seconds): 255
   Value of maximum query response time for MLD(in seconds): 10
   Querier for MLD: FE80::200:5EFF:FE66:5100 (this router)
  Total 1 MLD Group reported
```

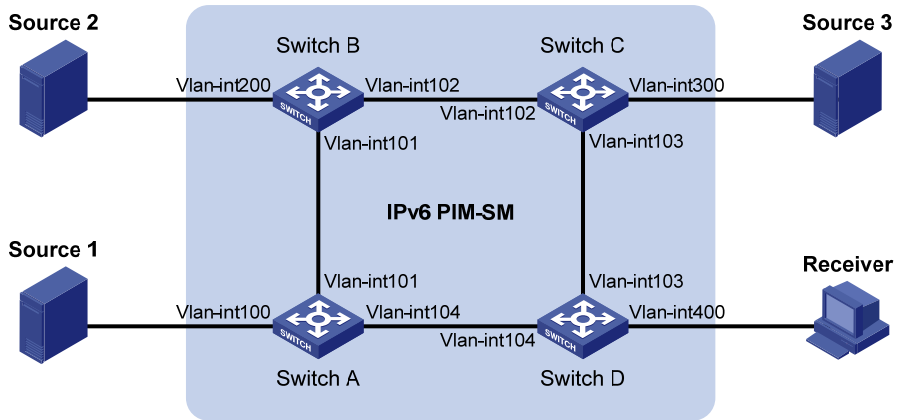
1.8.2 MLD SSM Mapping功能配置举例

1. 组网需求

- IPv6 PIM-SM 网络中同时采用 ASM 和 SSM 方式提供 IPv6 组播服务，将 Switch D 的接口 Vlan-interface104 配置为 C-BSR 和 C-RP，IPv6 SSM 组播组的范围为 FF3E::/64。
- Switch D 的接口 Vlan-interface400 上运行 MLDv2，接收者主机上运行 MLDv1，且不能升级至 MLDv2，因此该主机在加入 IPv6 组播组时无法指定 IPv6 组播源。
- Source 1、Source 2 和 Source 3 都向 IPv6 SSM 组范围内的组播组发送 IPv6 组播数据，要求通过在 Switch D 上配置 MLD SSM Mapping 功能，使接收者主机只能接收来自 Source 1 和 Source 3 的 IPv6 组播数据。

2. 组网图

图1-8 MLD SSM Mapping 功能配置组网图



设备	接口	IPv6地址	设备	接口	IPv6地址
Source 1	-	1001::1/64	Source 3	-	3001::1/64
Source 2	-	2001::1/64	Receiver	-	4001::1/64
Switch A	Vlan-int100	1001::2/64	Switch C	Vlan-int300	3001::2/64
	Vlan-int101	1002::1/64		Vlan-int103	3002::1/64
	Vlan-int104	1003::1/64		Vlan-int102	2002::2/64
Switch B	Vlan-int200	2001::2/64	Switch D	Vlan-int400	4001::2/64
	Vlan-int101	1002::2/64		Vlan-int103	3002::2/64

	Vlan-int102	2002::1/64		Vlan-int104	1003::2/64

3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址和 IPv6 单播路由协议

使能各交换机的IPv6 转发功能，并按照 图 1-8 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

配置 IPv6 PIM-SM 域内的各交换机之间采用 OSPFv3 协议进行互连，确保 IPv6 PIM-SM 域内部在网络层互通，并且各交换机之间能够借助 IPv6 单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IPv6 组播路由，并使能 IPv6 PIM-SM、MLD 和 MLD SSM Mapping 功能

在 Switch D 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并在主机侧接口 Vlan-interface400 上使能 MLD 和 MLD SSM Mapping 功能，配置 MLD 版本为 2。

```
<SwitchD> system-view
[SwitchD] multicast ipv6 routing-enable
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] mld enable
[SwitchD-Vlan-interface400] mld version 2
[SwitchD-Vlan-interface400] mld ssm-mapping enable
[SwitchD-Vlan-interface400] pim ipv6 sm
[SwitchD-Vlan-interface400] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim ipv6 sm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 104
[SwitchD-Vlan-interface104] pim ipv6 sm
[SwitchD-Vlan-interface104] quit
```

在 Switch A 上使能 IPv6 组播路由，并在各接口上使能 IPv6 PIM-SM。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim ipv6 sm
[SwitchA-Vlan-interface104] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP

在 Switch D 上配置 C-BSR 和 C-RP 的位置。

```
[SwitchD] pim ipv6
[SwitchD-pim6] c-bsr 1003::2
[SwitchD-pim6] c-rp 1003::2
[SwitchD-pim6] quit
```

(4) 配置 IPv6 SSM 组播组的地址范围

在 Switch D 上配置 IPv6 SSM 组播组的地址范围为 FF3E::/64。

```
[SwitchD] acl ipv6 number 2000
[SwitchD-acl6-basic-2000] rule permit source ff3e:: 64
[SwitchD-acl6-basic-2000] quit
[SwitchD] pim ipv6
[SwitchD-pim6] ssm-policy 2000
[SwitchD-pim6] quit
```

Switch A、Switch B 和 Switch C 的配置与 Switch D 相似，配置过程略。

(5) 配置 MLD SSM Mapping 规则

在 Switch D 上配置 MLD SSM Mapping 规则。

```
[SwitchD] mld
[SwitchD-mld] ssm-mapping ff3e:: 64 1001::1
[SwitchD-mld] ssm-mapping ff3e:: 64 3001::1
[SwitchD-mld] quit
```

(6) 检验配置效果

通过使用 **display mld ssm-mapping** 命令可以查看交换机上的 MLD SSM Mapping 规则。例如：

查看 Switch D 上 IPv6 组播组 FF3E::101 的 MLD SSM Mapping 规则。

```
[SwitchD] display mld ssm-mapping ff3e::101
Group: FF3E::101
Source list:
       1001::1
       3001::1
```

通过使用 **display mld ssm-mapping group** 命令可以查看交换机上依据 MLD SSM Mapping 规则创建的 IPv6 组播组信息。例如：

查看 Switch D 上依据 MLD SSM Mapping 规则创建的 IPv6 组播组信息。

```
[SwitchD] display mld ssm-mapping group
Total 1 MLD SSM-mapping Group(s).
Interface group report information
 Vlan-interface400 (4001::2):
  Total 1 MLD SSM-mapping Group reported
   Group Address: FF3E::101
    Last Reporter: 4001::1
    Uptime: 00:02:04
    Expires: off
```

通过使用 **display pim ipv6 routing-table** 命令可以查看交换机的 IPv6 PIM 路由表信息。例如：

查看 Switch D 上 IPv6 PIM 路由表信息。

```
[SwitchD] display pim ipv6 routing-table
Total 0 (*, G) entry; 2 (S, G) entry

 (1001::1, FF3E::101)
     Protocol: pim-ssm, Flag:
     UpTime: 00:13:25
     Upstream interface: Vlan-interface104
         Upstream neighbor: 1003::1
         RPF prime neighbor: 1003::1
```

```
    Downstream interface(s) information:
      Total number of downstreams: 1
        1: Vlan-interface400
                Protocol: mld, UpTime: 00:13:25, Expires: -

(3001::1, FF3E::101)
    Protocol: pim-ssm, Flag:
    UpTime: 00:13:25
    Upstream interface: Vlan-interface103
        Upstream neighbor: 3002::1
        RPF prime neighbor: 3002::1
    Downstream interface(s) information:
      Total number of downstreams: 1
        1: Vlan-interface400
                Protocol: mld, UpTime: 00:13:25, Expires: -
```

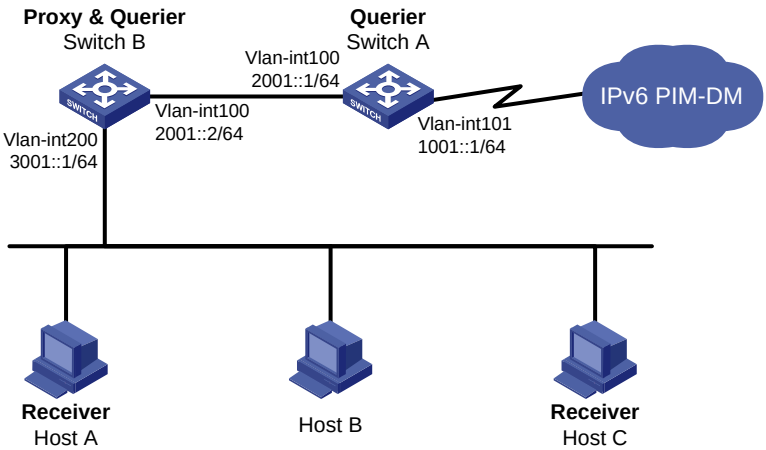
1.8.3 MLD Proxying功能配置举例

1. 组网需求

- 核心网络中运行 IPv6 PIM-DM，末梢网络中的接收者 Host A 和 Host C 通过 IPv6 组播组 FF3E::101 点播视频节目。
- 要求通过在 Switch B 上配置 MLD Proxying 功能，使其在不运行 IPv6 PIM-DM 的情况下实现组成员关系的维护和 IPv6 组播数据的正常转发。

2. 组网图

图1-9 MLD Proxying 功能配置组网图



3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址

使能各交换机的IPv6 转发功能，并按照 图 1-9 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

(2) 使能 IPv6 组播路由，并使能 IPv6 PIM-DM、MLD 和 MLD Proxying

在 Switch A 上使能 IPv6 组播路由，在接口 Vlan-interface101 上使能 IPv6 PIM-DM，并在接口 Vlan-interface100 上使能 MLD。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 dm
[SwitchA-Vlan-interface100] quit
```

在 Switch B 上使能 IPv6 组播路由，在接口 Vlan-interface100 上使能 MLD Proxying，并在接口 Vlan-interface200 上使能 MLD。

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] mld proxying enable
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] mld enable
[SwitchB-Vlan-interface200] quit
```

(3) 检验配置效果

通过使用 **display mld interface** 命令可以查看各交换机接口上 MLD 的配置和运行情况。例如：

查看 Switch B 在 Vlan-interface100 上 MLD 配置和运行的详细信息。

```
[SwitchB] display mld interface vlan-interface 100 verbose
 Vlan-interface100(2001::2):
   MLD proxy is enabled
   Current MLD version is 1
   Multicast routing on this interface: enabled
   Require-router-alert: disabled
```

通过使用 **display mld group** 命令可以查看 MLD 组的信息。例如：

查看 Switch A 上 MLD 组的信息。

```
[SwitchA] display mld group
Total 1 MLD Group(s).
Interface group report information
 Vlan-interface100(2001::1):
  Total 1 MLD Groups reported
   Group Address   Last Reporter   Uptime      Expires
   ff3e::101       2001::2         00:02:04    00:01:15
```

由此可见，主机的 MLD 报告通过 Switch B 的代理接口 Vlan-interface100 发给 Switch A。

1.9 常见配置错误举例

1.9.1 接收者侧路由器上无组成员信息

1. 故障现象

当某主机发送了加入 IPv6 组播组 G 的报文后，离该主机最近的路由器上却没有 IPv6 组播组 G 的组成员信息。

2. 分析

- 组网、接口连线的正确与否以及接口的协议层是否 up 将直接影响 IPv6 组播组成员信息的生成；
- 在路由器上必须使能 IPv6 组播路由，在连接主机的接口上必须使能 MLD；
- 如果路由器接口上运行的 MLD 版本比主机的低，那么路由器将无法识别主机发来的较高版本的 MLD 报告报文；
- 如果在接口上使用命令 **mld group-policy** 对加入 IPv6 组播组 G 进行了限制后，该接口将不再接收未通过过滤的那些要求加入 IPv6 组播组 G 的报文。

3. 处理过程

(1) 检查组网是否正确，接口间的连线是否正确，以及接口状态是否正常，是否配置了正确的 IPv6 地址。通过命令 **display mld interface** 查看接口信息。若无接口信息输出，说明接口状态异常，原因通常是接口上配置了 **shutdown** 命令，或者接口连线不正确，或者接口上没有配置正确的 IPv6 地址。

(2) 检查是否使能了 IPv6 组播路由。通过命令 **display current-configuration** 查看是否配置了命令 **multicast ipv6 routing-enable**。若缺少该配置，则需要在系统视图下执行命令 **multicast ipv6 routing-enable** 使能 IPv6 组播路由，同时也需要在相应接口上使能 MLD。

(3) 检查接口上运行的 MLD 版本。通过命令 **display mld interface** 来检查接口上运行的 MLD 版本是否低于主机所使用的版本。

(4) 检查接口上是否配置了 IPv6 ACL 规则来限制主机加入 IPv6 组播组 G。通过命令 **display current-configuration interface** 观察是否配置了 **mld group-policy** 命令。如果配置的 IPv6 ACL 规则对加入 IPv6 组播组 G 进行了限制，则需要修改该 IPv6 ACL 规则，允许接受 IPv6 组播组 G 的报告报文。

1.9.2 同一网段各路由器上组成员关系不一致

1. 故障现象

在同一网段的不同 MLD 路由器上，各自维护的组成员关系不一致。

2. 分析

- 运行 MLD 的路由器为每个接口维护多个参数，各参数之间相互影响，非常复杂。如果同一网段路由器的 MLD 接口参数配置不一致，必然导致组成员关系的混乱。
- 另外，MLD 目前有 2 个版本，版本不同的 MLD 路由器与主机之间虽然可以兼容，但是连接在同一网段的所有路由器必须运行相同版本的 MLD。如果同一网段路由器的 MLD 版本不一致，也将导致 MLD 组成员关系的混乱。

3. 处理过程

(1) 检查 MLD 配置。通过命令 **display current-configuration** 观察接口上 MLD 的配置信息。

(2) 在同一网段的所有路由器上执行命令 **display mld interface** 来检查 MLD 相关定时器的参数，确保配置一致。

(3) 通过命令 **display mld interface** 来检查各路由器上运行的 MLD 版本是否一致。

目 录

1 IPv6 PIM配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- IPv6 PIM 功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

1.1 IPv6 PIM简介

IPv6 PIM 是 Protocol Independent Multicast for IPv6（IPv6 协议无关组播）的简称，表示可以利用 IPv6 单播静态路由或者任意 IPv6 单播路由协议（包括 RIPng、OSPFv3、IS-ISv6、BGP4+等）所生成的 IPv6 单播路由表为 IPv6 组播提供路由。IPv6 组播路由与所采用的 IPv6 单播路由协议无关，只要能够通过 IPv6 单播路由协议产生相应的 IPv6 组播路由表项即可。IPv6 PIM 借助 RPF（Reverse Path Forwarding，逆向路径转发）机制实现对 IPv6 组播报文的转发。当 IPv6 组播报文到达本地设备时，首先对其进行 RPF 检查：若 RPF 检查通过，则创建相应的 IPv6 组播路由表项，从而进行 IPv6 组播报文的转发；若 RPF 检查失败，则丢弃该报文。有关 RPF 的详细介绍，请参见“IP 组播配置指导”中的“IPv6 组播路由与转发”。

根据实现机制的不同，IPv6 PIM 分为以下几种类型：

- IPv6 PIM-DM（Protocol Independent Multicast-Dense Mode for IPv6，IPv6 协议无关组播—密集模式）
- IPv6 PIM-SM（Protocol Independent Multicast-Sparse Mode for IPv6，IPv6 协议无关组播—稀疏模式）
- IPv6 BIDIR-PIM（Bidirectional Protocol Independent Multicast for IPv6，IPv6 双向协议无关组播，简称 IPv6 双向 PIM）
- IPv6 PIM-SSM（Protocol Independent Multicast Source-Specific Multicast for IPv6，IPv6 协议无关组播－指定源组播）

说明

为了描述的方便，本文中把由支持 IPv6 PIM 协议的组播路由器所组成的网络简称为“IPv6 PIM 域”。

1.1.1 IPv6 PIM-DM简介

IPv6 PIM-DM 属于密集模式的 IPv6 组播路由协议，使用“推（Push）模式”传送 IPv6 组播数据，通常适用于 IPv6 组播组成员相对比较密集的小型网络，其基本原理如下：

- IPv6 PIM-DM 假设网络中的每个子网都存在至少一个 IPv6 组播组成员，因此 IPv6 组播数据将被扩散（Flooding）到网络中的所有节点。然后，IPv6 PIM-DM 对没有 IPv6 组播数据转发

的分支进行剪枝（Prune），只保留包含接收者的分支。这种“扩散—剪枝”现象周期性地发生，被剪枝的分支也可以周期性地恢复成转发状态。

- 当被剪枝分支的节点上出现了 IPv6 组播组的成员时，为了减少该节点恢复成转发状态所需的时间，IPv6 PIM-DM 使用嫁接（Graft）机制主动恢复其对 IPv6 组播数据的转发。

一般说来，密集模式下数据包的转发路径是有源树（Source Tree，即以 IPv6 组播源为“根”、IPv6 组播组成员为“枝叶”的一棵转发树）。由于有源树使用的是从 IPv6 组播源到接收者的最短路径，因此也称为最短路径树（Shortest Path Tree，SPT）。

IPv6 PIM-DM 的工作机制可以概括如下：

- 邻居发现
- 构建 SPT
- 嫁接
- 断言

1. 邻居发现

在 IPv6 PIM 域中，路由器通过周期性地的向本网段的所有 IPv6 PIM 路由器以组播方式发送 IPv6 PIM Hello 报文（以下简称 Hello 报文），以发现 IPv6 PIM 邻居，维护各路由器之间的 IPv6 PIM 邻居关系，从而构建和维护 SPT。

说明

路由器每个运行了 IPv6 PIM 协议的接口都会周期性地发送 Hello 报文，从而了解与该接口相关的 IPv6 PIM 邻居信息。

---

2. 构建SPT

构建 SPT 的过程也就是“扩散—剪枝”的过程：

(1) 在 IPv6 PIM-DM 域中，IPv6 组播源 S 向 IPv6 组播组 G 发送 IPv6 组播报文时，首先对 IPv6 组播报文进行扩散：路由器对该报文的 RPF 检查通过后，便创建一个（S，G）表项，并将该报文向网络中的所有下游节点转发。经过扩散，IPv6 PIM-DM 域内的每个路由器上都会创建（S，G）表项。

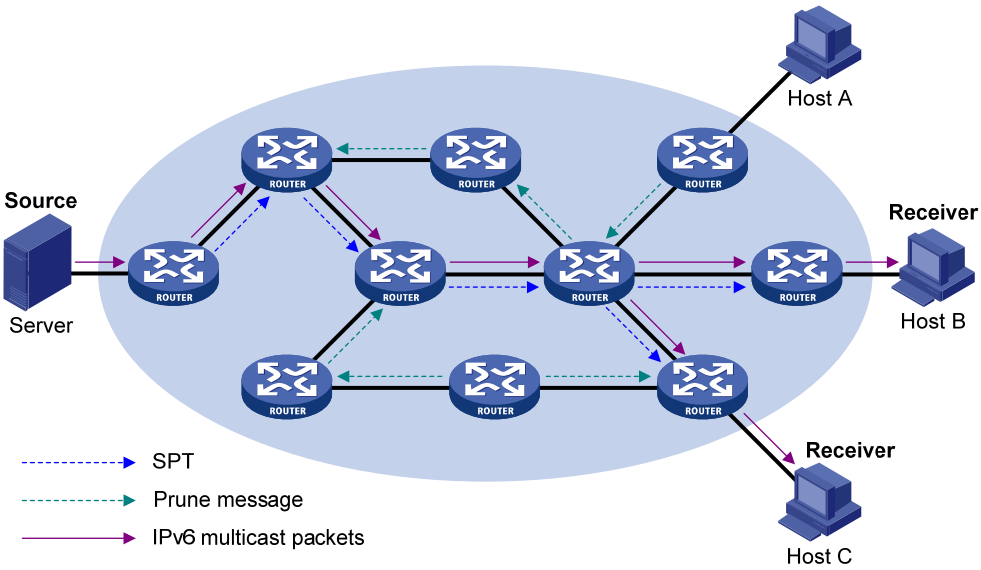
(2) 然后对那些下游没有接收者的节点进行剪枝：由没有接收者的下游节点向上游节点发剪枝报文（Prune Message），以通知上游节点将相应的接口从其组播转发表项（S，G）所对应的出接口列表中删除，并不再转发该 IPv6 组播组的报文至该节点。

说明

- （S，G）表项包括 IPv6 组播源的地址 S、IPv6 组播组的地址 G、出接口列表和入接口等。
- 路由器上收到 IPv6 组播数据的接口称为“上游”，转发 IPv6 组播数据的接口称为“下游”。

剪枝过程最先由叶子路由器发起，如 图 1-1 所示，没有接收者（Receiver）的路由器（如与Host A直连的路由器）主动发起剪枝，并一直持续到IPv6 PIM-DM域中只剩下必要的分支，这些分支共同构成了SPT。

图1-1 IPv6 PIM-DM 中构建 SPT 示意图



“扩散—剪枝”的过程是周期性发生的。各个被剪枝的节点提供超时机制，当剪枝超时后便重新开始这一过程。

3. 嫁接

当被剪枝的节点上出现了 IPv6 组播组的成员时，为了减少该节点恢复成转发状态所需的时间，IPv6 PIM-DM 使用嫁接机制主动恢复其对 IPv6 组播数据的转发，过程如下：

(1) 需要恢复接收 IPv6 组播数据的节点向其上游节点发送嫁接报文（Graft Message）以申请重新加入到 SPT 中；

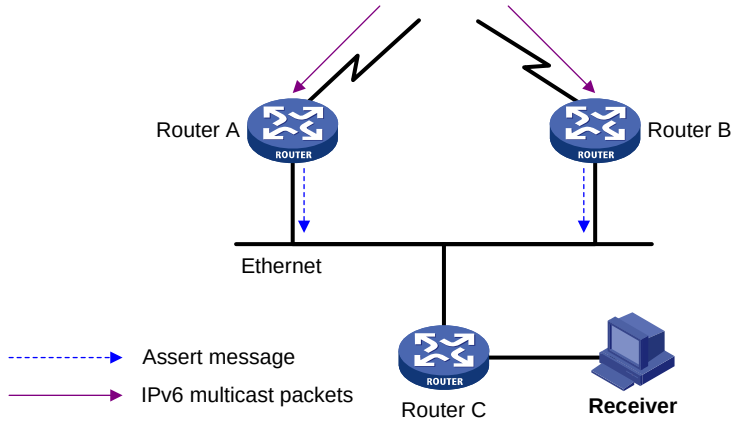
(2) 当上游节点收到该报文后恢复该下游节点的转发状态，并向其回应一个嫁接应答报文（Graft-Ack Message）以进行确认；

(3) 如果发送嫁接报文的下游节点没有收到来自其上游节点的嫁接应答报文，将重新发送嫁接报文直到被确认为止。

4. 断言

在一个网段内如果存在多台组播路由器，则相同的 IPv6 组播报文可能会被重复发送到该网段。为了避免出现这种情况，就需要通过断言（Assert）机制来选定唯一的 IPv6 组播数据转发者。

图1-2 Assert 机制示意图



如 图 1-2 所示，当Router A和Router B从上游节点收到（S，G）的IPv6 组播报文后，都会向本地网段转发该报文，于是处于下游的节点Router C就会收到两份相同的IPv6 组播报文，Router A和Router B也会从各自的下游接口收到对方转发来的该IPv6 组播报文。此时，Router A和Router B会通过下游接口向本网段的所有IPv6 PIM路由器以组播方式发送断言报文（Assert Message），该报文中携带有以下信息：IPv6 组播源地址S、IPv6 组播地址G、到IPv6 组播源的IPv6 单播路由/IPv6 MBGP路由的优先级和度量值。通过一定的规则对这些参数进行比较后，Router A和Router B中的获胜者将成为（S，G）IPv6 组播报文在本网段的转发者，比较规则如下：

(1) 到 IPv6 组播源的优先级较高者获胜；

(2) 如果到 IPv6 组播源的优先级相等，那么到 IPv6 组播源的度量值较小者获胜；

(3) 如果到 IPv6 组播源的度量值也相等，则下游接口 IPv6 链路本地地址较大者获胜。

1.1.2 IPv6 PIM-SM简介

IPv6 PIM-DM 使用以“扩散—剪枝”方式构建的 SPT 来传送 IPv6 组播数据。尽管 SPT 的路径最短，但是其建立的过程效率较低，并不适合大中型网络。

IPv6 PIM-SM 属于稀疏模式的 IPv6 组播路由协议，使用“拉（Pull）模式”传送 IPv6 组播数据，通常适用于 IPv6 组播组成员分布相对分散、范围较广的大中型网络，其基本原理如下：

- IPv6 PIM-SM 假设所有主机都不需要接收 IPv6 组播数据，只向明确提出需要 IPv6 组播数据的主机转发。IPv6 PIM-SM 实现组播转发的核心任务就是构造并维护 RPT（Rendezvous Point Tree，共享树），RPT 选择 IPv6 PIM 域中某台路由器作为公用的根节点 RP（Rendezvous Point，汇集点），IPv6 组播数据通过 RP 沿着 RPT 转发给接收者；
- 连接接收者的路由器向某 IPv6 组播组对应的 RP 发送加入报文，该报文被逐跳送达 RP，所经过的路径就形成了 RPT 的分支；
- IPv6 组播源如果要向某 IPv6 组播组发送 IPv6 组播数据，首先由 IPv6 组播源侧 DR（Designated Router，指定路由器）负责向 RP 进行注册，把注册报文（Register Message）通过单播方式发送给 RP，该报文到达 RP 后触发建立 SPT。之后 IPv6 组播源把 IPv6 组播数据沿着 SPT 发向 RP，当 IPv6 组播数据到达 RP 后，被复制并沿着 RPT 发送给接收者。

说明

复制仅发生在分发树的分支处，这个过程能够自动重复直到数据包最终到达接收者。

IPv6 PIM-SM 的工作机制可以概括如下：

- 邻居发现
- DR 选举
- RP 发现
- 嵌入式 RP
- 构建 RPT
- IPv6 组播源注册
- SPT 切换
- 断言

1. 邻居发现

IPv6 PIM-SM使用与IPv6 PIM-DM类似的邻居发现机制，具体请参见“1.1.1 1. 邻居发现”一节。

2. DR选举

借助 Hello 报文还可以为共享网络（如 Ethernet）选举 DR，DR 将作为该共享网络中 IPv6 组播数据的唯一转发者。

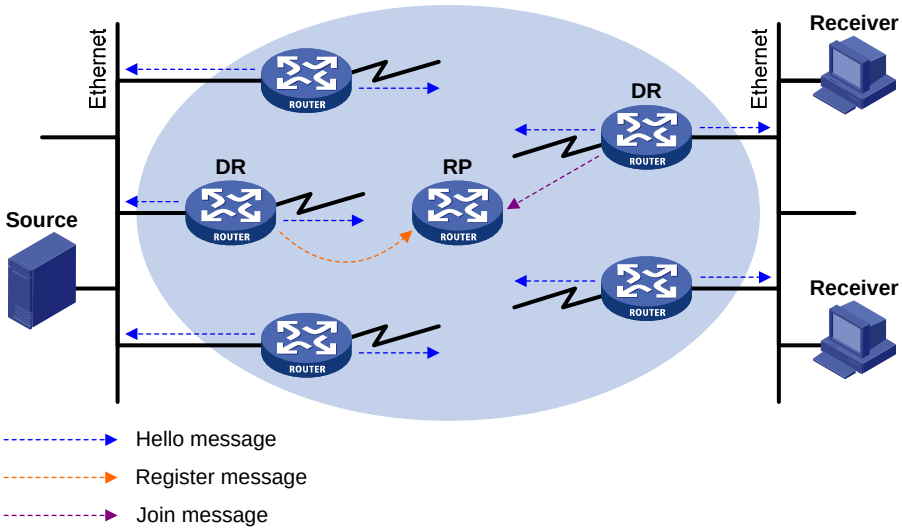
无论是与 IPv6 组播源相连的网络，还是与接收者相连的网络，都需要选举 DR。接收者侧的 DR 负责向 RP 发送加入报文；IPv6 组播源侧的 DR 负责向 RP 发送注册报文。

说明

- 各路由器之间通过比较 Hello 报文中所携带的优先级和 IPv6 链路本地地址，可以为多路由器网段选举 DR。
- 在充当接收者侧 DR 的设备上必须使能 MLD，否则连接在该 DR 上的接收者将不能通过该 DR 加入 IPv6 组播组。

有关 MLD 的介绍，请参见“IP 组播配置指导”中的“MLD”。

图1-3 DR 选举示意图



如 图 1-3 所示，DR的选举过程如下：

(1) 共享网络上的各路由器相互之间发送 Hello 报文（携带有竞选 DR 优先级的参数），拥有最高优先级的路由器将成为 DR；

(2) 如果优先级相同，或者网络中至少有一台路由器不支持在 Hello 报文中携带竞选 DR 优先级的参数，则根据各路由器的 IPv6 链路本地地址大小来竞选 DR，IPv6 链路本地地址最大的路由器将成为 DR。

当 DR 出现故障时，其余路由器在超时后仍没有收到来自 DR 的 Hello 报文，则会触发新的 DR 选举过程。

3. RP发现

RP 是 IPv6 PIM-SM 域中的核心设备。在结构简单的小型网络中，IPv6 组播信息量少，整个网络仅依靠一个 RP 进行 IPv6 组播信息的转发即可，此时可以在 IPv6 PIM-SM 域中的各路由器上静态指定 RP 的位置；但是在更多的情况下，IPv6 PIM-SM 域的规模都很大，通过 RP 转发的 IPv6 组播信息量巨大。为了缓解 RP 的负担并优化 RPT 的拓扑结构，可以在 IPv6 PIM-SM 域中配置多个 C-RP（Candidate-RP，候选 RP），通过自举机制来动态选举 RP，使不同的 RP 服务于不同的组播组，此时需要配置 BSR（Bootstrap Router，自举路由器）。BSR 是 IPv6 PIM-SM 域中的管理核心，一个 IPv6 PIM-SM 域内只能有一个 BSR，但可以配置多个 C-BSR（Candidate-BSR，候选 BSR）。这样，一旦 BSR 发生故障，其余 C-BSR 能够通过自动选举产生新的 BSR，从而确保业务免受中断。

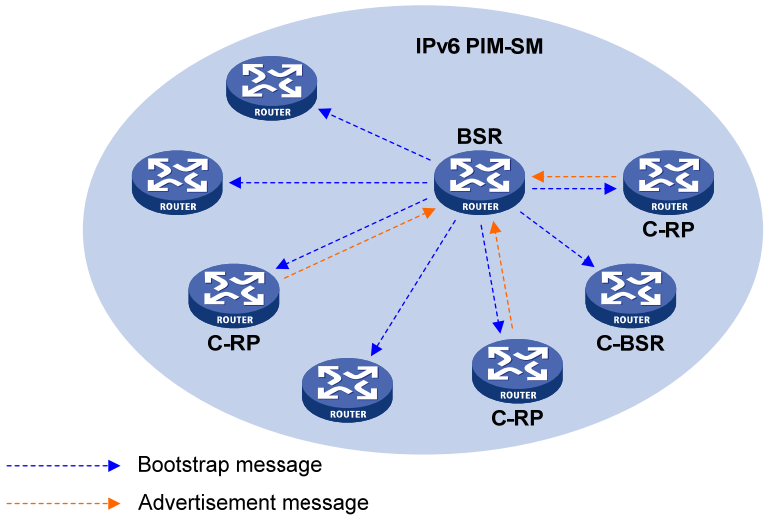
说明

- 一个 RP 可以同时服务于多个 IPv6 组播组，但一个 IPv6 组播组只能唯一对应一个 RP。
- 一台设备可以同时充当 C-RP 和 C-BSR。

如 图 1-4 所示，BSR负责收集网络中由C-RP发来的宣告报文（Advertisement Message），该报文中携带有C-RP的地址和优先级以及其服务的IPv6 组范围，BSR将这些信息汇总为RP-Set（RP集，

即IPv6 组播组与RP的映射关系数据库），封装在自举报文（Bootstrap Message，BSM）中并发布到整个IPv6 PIM-SM域。

图1-4 RP 与 BSR 信息交互示意图



网络中的各路由器将依据 RP-Set 提供的信息，使用相同的规则从众多 C-RP 中为特定 IPv6 组播组选择其对应的 RP，具体规则如下：

(1) 首先比较 C-RP 所服务的 IPv6 组范围，所服务的 IPv6 组范围较小者获胜。

(2) 若服务的 IPv6 组范围相同，再比较 C-RP 的优先级，优先级较高者获胜。

(3) 若优先级也相同，再使用哈希（Hash）函数计算哈希值，哈希值较大者获胜。

(4) 若哈希值也相同，则 C-RP 的 IPv6 地址较大者获胜。

哈希函数的表达式为：Value (G, M, C_i) = (1103515245 * ((1103515245 * (G & M) + 12345) XOR C_i) + 12345) mod 2^{31}，其中各符号的含义如 表 1-1 所示。

表1-1 哈希函数中各符号含义

符号	含义
Value	哈希值
G	将IPv6组播组地址的每32比特作为一节，各节之间通过异或运算而得的数值。假设IPv6组播组地址为FF0E:C20:1A3:63::101，则G = 0xFF0E0C20 XOR 0x01A30063 XOR 0x00000000 XOR 0x00000101
M	哈希掩码长度（Hash Mask Length）
C_i	将C-RP IPv6地址的每32比特作为一节，各节之间通过异或运算而得的数值。假设C-RP的IPv6地址为3FFE:B00:C18:1::10，则C_i = 0x3FFE0B00 XOR 0x0C180001 XOR 0x00000000 XOR 0x00000010
&	逻辑运算符，表示与运算
XOR	逻辑运算符，表示异或运算
mod	算术运算符，表示整除取余

4. 嵌入式RP

通过嵌入式 RP（Embedded RP）机制可以从 IPv6 组播地址中解析出 RP 的地址，从而实现 IPv6 组播组到 RP 的映射，以取代静态配置的 RP 或由 BSR 机制动态计算出来的 RP，DR 不再需要预先知道 RP 的信息，只需对组播报文进行分析即可知道 RP 的地址。其工作原理如下：

- 接收者侧：

(1) 接收者主机发送 MLD 报告报文声明加入某 IPv6 组播组；

(2) 接收者侧的 DR 提取内嵌在 IPv6 组播地址中的 RP 地址，并向该 RP 发送加入报文（Join Message）。

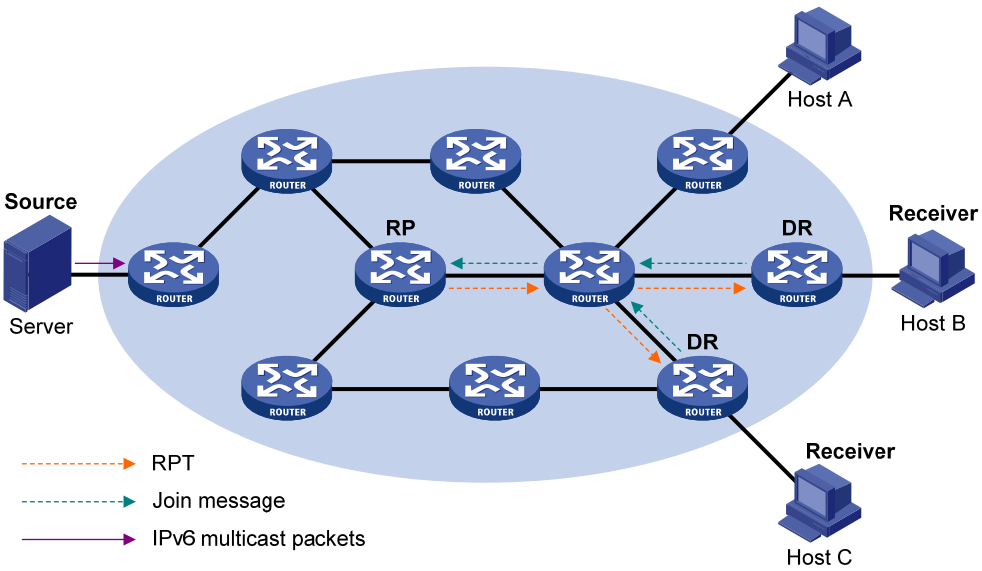
- IPv6 组播源侧：

(1) IPv6 组播源要向某 IPv6 组播组发送 IPv6 组播数据；

(2) IPv6 组播源侧的 DR 提取内嵌在 IPv6 组播地址中的 RP 地址，并向该 RP 发送注册报文。

5. 构建RPT

图1-5 IPv6 PIM-SM 中构建 RPT 示意图



如 图 1-5 所示，RPT的构建过程如下：

(1) 当接收者加入一个 IPv6 组播组 G 时，先通过 MLD 报文通知与其直连的 DR；

(2) DR 掌握了 IPv6 组播组 G 的接收者的信息后，向该组所对应的 RP 方向逐跳发送加入报文；

(3) 从 DR 到 RP 所经过的路由器就形成了 RPT 的分支，这些路由器都在其转发表中生成了（*，G）表项，这里的“*”表示来自任意 IPv6 组播源。RPT 以 RP 为根，以 DR 为叶子。

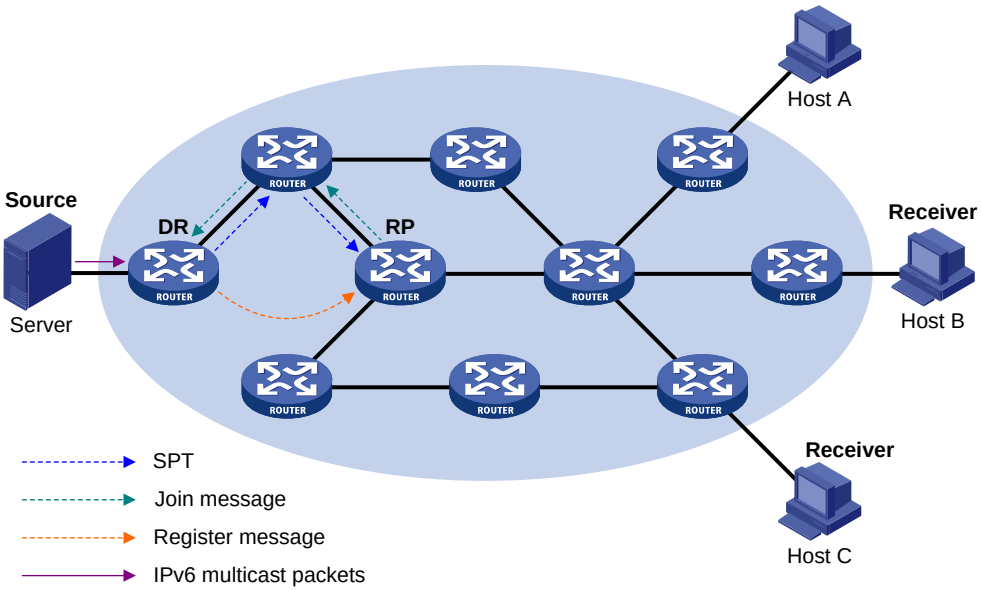
当发往 IPv6 组播组 G 的 IPv6 组播数据流经 RP 时，数据就会沿着已建立好的 RPT 到达 DR，进而到达接收者。

当某接收者对 IPv6 组播组 G 的信息不再感兴趣时，与其直连的 DR 会逆着 RPT 向该组的 RP 方向逐跳发送剪枝报文；上游节点收到该报文后在其出接口列表中删除与下游节点相连的接口，并检查自己是否拥有该 IPv6 组播组的接收者，如果没有则继续向其上游转发该剪枝报文。

6. IPv6 组播源注册

IPv6 组播源注册的目的是向 RP 通知 IPv6 组播源的存在。

图1-6 IPv6 组播源注册示意图



如 图 1-6 所示，IPv6 组播源向RP注册的过程如下：

(1) 当 IPv6 组播源 S 向 IPv6 组播组 G 发送了一个 IPv6 组播报文时，与 IPv6 组播源直连的 DR 在收到该报文后，就将其封装成注册报文，并通过单播方式发送给相应的 RP；

(2) 当 RP 收到该报文后，一方面解封装注册报文并将封装在其中的 IPv6 组播报文沿着 RPT 转发给接收者，另一方面向 IPv6 组播源方向逐跳发送（S，G）加入报文。这样，从 RP 到 IPv6 组播源所经过的路由器就形成了 SPT 的分支，这些路由器都在其转发表中生成了（S，G）表项。SPT 以 IPv6 组播源侧的 DR 为根，以 RP 为叶子。

(3) IPv6 组播源发出的 IPv6 组播数据沿着已建立好的 SPT 到达 RP，然后由 RP 把 IPv6 组播数据沿着 RPT 向接收者进行转发。当 RP 收到沿着 SPT 转发来的 IPv6 组播数据后，通过单播方式向与 IPv6 组播源直连的 DR 发送注册停止报文（Register-Stop Message），IPv6 组播源注册过程结束。

说明

上述描述中假定允许 RP 发起 SPT 切换，否则 IPv6 组播源侧 DR 将一直用注册报文封装 IPv6 组播报文，注册过程不会结束，除非 RP 上（S，G）表项的出接口变为空。

7. SPT切换

在 IPv6 PIM-SM 域中，一个 IPv6 组播组唯一对应一个 RP 和一棵 RPT。在 SPT 切换前，所有发往该组的 IPv6 组播报文都必须先由 IPv6 组播源侧 DR 封装在注册报文中发往 RP，由 RP 解封装后再沿 RPT 分发给接收者侧的 DR，RP 是所有 IPv6 组播数据必经的中转站。这个过程存在以下三个问题：

- IPv6 组播源侧的 DR 和 RP 必须对 IPv6 组播数据进行繁琐的封装/解封装处理。
- IPv6 组播数据的转发路径不一定是从 IPv6 组播源到接收者的最短路径。
- 当 IPv6 组播流量变大时，RP 负担增大，容易引发故障。

为了解决上述问题，IPv6 PIM-SM 允许由 RP 或接收者侧的 DR 发起 SPT 切换：

(1) RP 发起的 SPT 切换

RP 收到第一个 IPv6 组播数据包后，立即向 IPv6 组播源方向发送（S，G）加入报文，在 IPv6 组播源侧 DR 与 RP 之间建立起 SPT 分支，后续的 IPv6 组播报文都直接沿该分支到达 RP。

说明

由RP发起的SPT切换的详细过程，请参见“1.1.2 6. IPv6 组播源注册”一节。

(2) 接收者侧 DR 发起的 SPT 切换

接收者侧 DR 收到第一个 IPv6 组播数据包后，立即发起 SPT 切换，过程如下：

- 首先，接收者侧 DR 向 IPv6 组播源方向逐跳发送（S，G）加入报文，并最终送达 IPv6 组播源侧 DR，沿途经过的所有路由器在其转发表中都生成了（S，G）表项，从而建立了 SPT 分支；
- 随后，当 IPv6 组播数据沿 SPT 到达 RPT 与 SPT 分叉的路由器时，该路由器开始丢弃沿 RPT 到达的 IPv6 组播数据，同时向 RP 逐跳发送含 RP 位的剪枝报文，RP 收到该报文后继续向 IPv6 组播源方向发送剪枝报文（假设此时只有这一个接收者），从而完成了 SPT 切换；
- 最终，IPv6 组播数据将沿 SPT 从 IPv6 组播源到达到接收者。

通过 SPT 切换，IPv6 PIM-SM 能够以比 IPv6 PIM-DM 更经济的方式建立 SPT。

8. 断言

IPv6 PIM-SM使用与IPv6 PIM-DM类似的断言机制，具体请参见“1.1.1 4. 断言”一节。

1.1.3 IPv6 双向PIM简介

在某些组网应用（譬如多方电视电话会议）中，同时存在多个接收者和多个 IPv6 组播源，在这种情况下，如果使用传统的 IPv6 PIM-DM 或 IPv6 PIM-SM 按 SPT 转发 IPv6 组播数据，需在每台路由器上针对每个 IPv6 组播源都创建（S，G）表项，这将占用大量的系统资源。为了解决这个问题，提出了 IPv6 双向 PIM 的概念。IPv6 双向 PIM 由 IPv6 PIM-SM 发展而来，它通过建立以 RP 为中心、分别连接 IPv6 组播源和接收者的双向 RPT，使 IPv6 组播数据沿着双向 RPT 从 IPv6 组播源经由 RP 转发到接收者。这样，在每台路由器上只需维护（*，G）表项即可，从而节约了系统资源。

IPv6 双向 PIM 主要适用于 IPv6 组播源和接收者都比较密集的网络，其工作机制可以概括如下：

- 邻居发现
- RP 发现
- DF 选举
- 构建双向 RPT

1. 邻居发现

IPv6 双向PIM使用与IPv6 PIM-SM完全相同的邻居发现机制，具体请参见“1.1.1 1. 邻居发现”一节。

2. RP发现

IPv6 双向PIM使用与IPv6 PIM-SM完全相同的RP发现机制，具体请参见“1.1.2 3. RP发现”一节。

IPv6 PIM-SM 的 RP 必须指定为一个实际存在的 IPv6 地址，而 IPv6 双向 PIM 的 RP 则可以指定为一个实际不存在的 IPv6 地址，简称 RPA（Rendezvous Point Address，汇集点地址）。RPA 所属网段对应的链路就称为 RPL（Rendezvous Point Link，汇集点链路），连接到 RPL 上的所有接口都可以充当 RP，且互为备份。

说明

IPv6 双向 PIM 中的 RPF 接口是指向 RP 的接口，RPF 邻居自然是到达 RP 的下一跳地址。

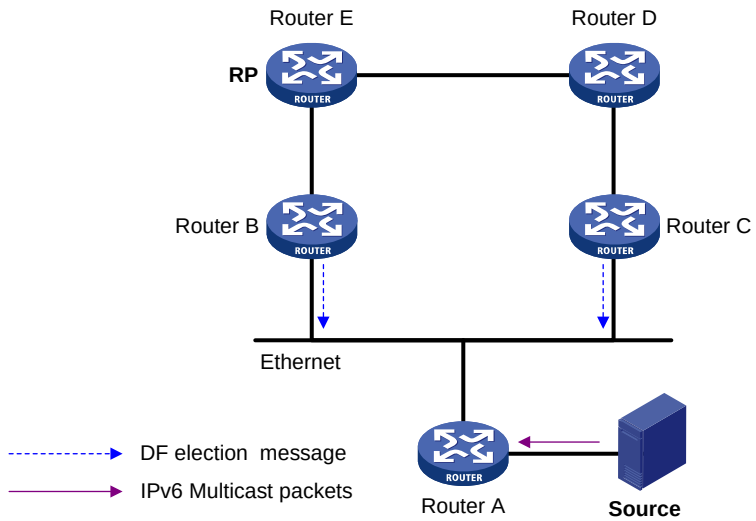
3. DF选举

DF（Designated Forwarder，指定转发者）是 IPv6 双向 PIM 中的重要角色，IPv6 组播数据由 IPv6 组播源向 RP 转发的动力来自于 DF，也就是说只有 DF 才有能力将 IPv6 组播数据向 RP 方向转发。因此，每个 RP 在每个网段都需要有其对应的 DF，以负责将该网段的 IPv6 组播数据向该 RP 转发；此外，在有多台 IPv6 组播路由器的网段，DF 的唯一性也可以避免相同的 IPv6 组播报文被重复发往 RP。

说明

在 RPL 上不需要选举 DF。

图1-7 DF 选举示意图



如 图 1-7 所示，Router B和Router C都可以从Router A收到由IPv6 组播源向IPv6 组播组G发送的IPv6 组播报文，如果它们都向下游节点转发该报文，RP最终将收到两份相同的IPv6 组播报文。因此，Router B和Router C一旦获得RP的信息，就会为该RP发起DF的选举：Router B和Router C将分别向本网段的所有IPv6 PIM路由器以组播方式发送DF选举报文（DF Election Message），该报文携带有以下信息：RP的地址、到RP的IPv6 单播路由/IPv6 MBGP路由的优先级和度量值。通过一定规则对这些参数进行比较后，Router B和Router C中的获胜者将成为DF，具体的比较规则如下：

(1) 到 RP 的优先级较高者获胜；

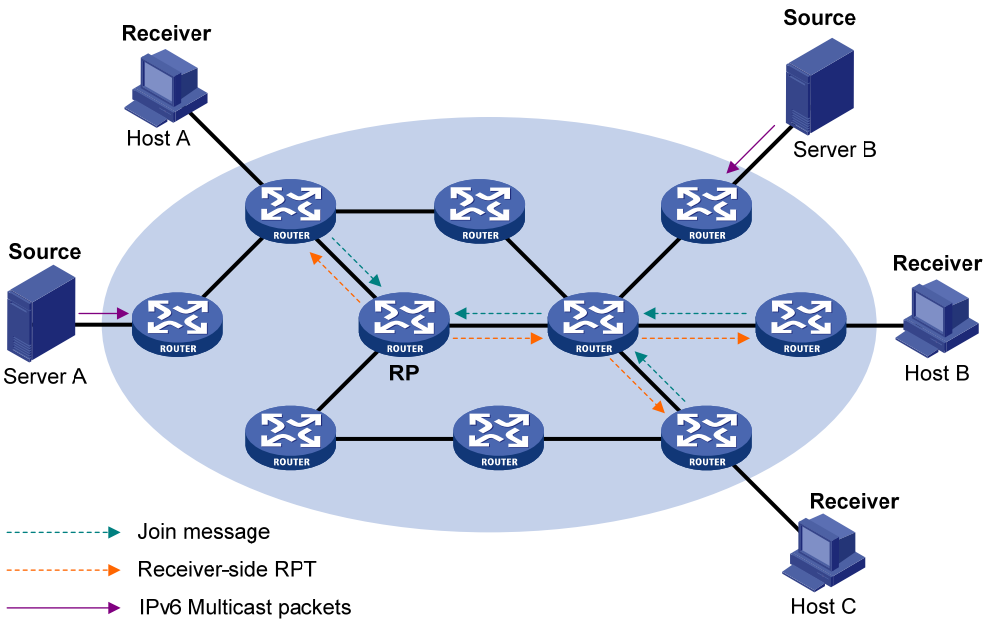
(2) 如果到 RP 的优先级相等，那么到 RP 的度量值较小者获胜；

(3) 如果到 RP 的度量值也相等，则接口的 IPv6 链路本地地址较大者获胜。

4. 构建双向RPT

双向 RPT 由两部分构成：一部分是以 RP 为根、以直连接收者的路由器为叶子的 RPT，简称接收者侧 RPT；而另一部分则是以 RP 为根、以直连 IPv6 组播源的路由器为叶子的 RPT，简称组播源侧 RPT。这两部分 RPT 的构建过程不同，下面分别加以介绍。

图1-8 接收者侧 RPT 构建示意图



接收者侧RPT的构建过程与IPv6 PIM-SM中RPT的构建过程类似，如 图 1-8 所示，其构建过程如下：

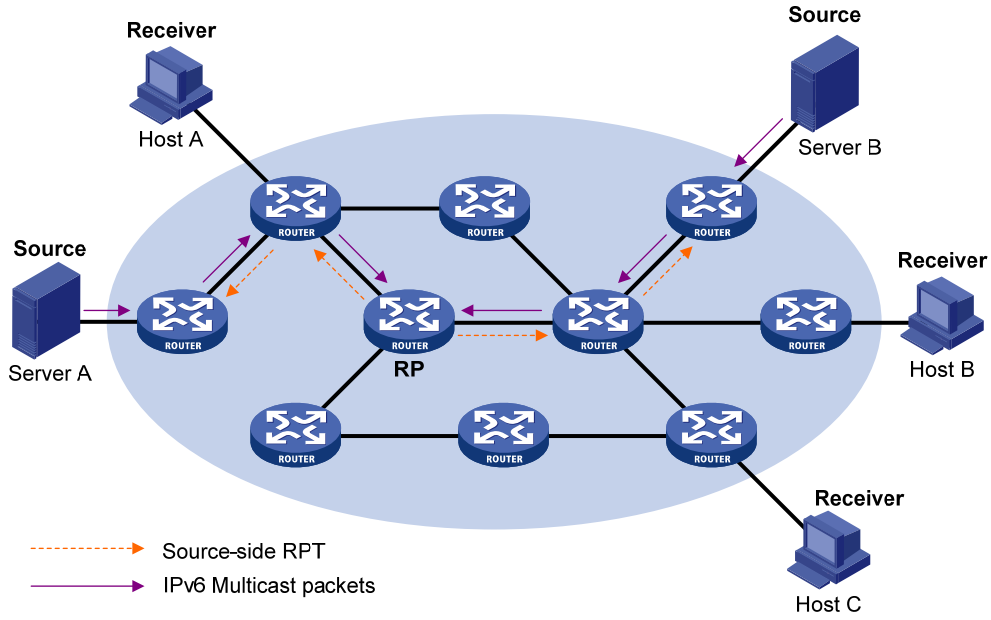
(1) 当接收者加入一个 IPv6 组播组 G 时，先通过 MLD 报文通知与其直连的路由器；

(2) 该路由器掌握了 IPv6 组播组 G 的接收者的信息后，向该组所对应的 RP 方向逐跳发送加入报文；

(3) 从直连接收者的路由器到 RP 所经过的路由器就形成了接收者侧 RPT 的分支，这些路由器都在其转发表中生成了（*，G）表项。

当某接收者对 IPv6 组播组 G 的信息不再感兴趣时，与其直连的路由器会逆着接收者侧 RPT 向该组的 RP 方向逐跳发送剪枝报文；上游节点收到该报文后在其出接口列表中删除与下游节点相连的接口，并检查自己是否拥有该 IPv6 组播组的接收者，如果没有则继续向其上游转发该剪枝报文。

图1-9 组播源侧 RPT 构建示意图



组播源侧RPT的构建过程则相对简单，如 图 1-9 所示，其构建过程如下：

(1) IPv6 组播源发向 IPv6 组播组 G 的 IPv6 组播数据在途径的每个网段，都被该网段的 DF 无条件地向 RP 转发；

(2) 从直连组播源的路由器到 RP 所经过的路由器就形成了 IPv6 组播源侧 RPT 的分支，这些路由器都在其转发表中生成了（*，G）表项。

当双向 RPT 构建完成之后，由 IPv6 组播源发出的 IPv6 组播数据将依次沿着 IPv6 组播源侧 RPT 和接收者侧 RPT，经由 RP 转发至接收者。

说明

当接收者和 IPv6 组播源位于 RP 同一侧时，组播源侧 RPT 与接收者侧 RPT 有可能在到达 RP 之前就已汇合。在这种情况下，由该 IPv6 组播源发往该接收者的 IPv6 组播数据将在此汇合点直接被转发给该接收者，而不必经由 RP。

1.1.4 IPv6 管理域机制简介

1. 两种域机制的划分

一般情况下，在一个 IPv6 PIM-SM/IPv6 双向 PIM 域内只能有一个 BSR，并由该 BSR 负责在整个 IPv6 PIM-SM/IPv6 双向 PIM 域内宣告 RP-Set 信息，所有 IPv6 组播组的信息都在此 BSR 管理的网络范围内进行转发，我们称之为 IPv6 非管理域机制。

考虑到管理的精细化，可以将整个 IPv6 PIM-SM/IPv6 双向 PIM 域划分为一个 IPv6 Global 域（IPv6 Global-scope Zone）和多个 IPv6 管理域（IPv6 Admin-scope Zone），一方面可以有效分担单一 BSR 的管理压力，另一方面可以使用私有组地址为特定区域提供专门的服务。相应地，我们称之为 IPv6 管理域机制。

IPv6 管理域与特定 Scope 值的 IPv6 组播组相对应，针对不同的 Scope 值划分相应的 IPv6 管理域。IPv6 管理域的边界由 ZBR（Zone Border Router，区域边界路由器）构成，每个 IPv6 管理域各维护一个 BSR，为特定 Scope 值的 IPv6 组播组服务，属于此范围的 IPv6 组播协议报文（如断言报文、BSR 自举报文等）无法通过 IPv6 管理域边界。不同 IPv6 管理域所服务的 IPv6 组播组范围可以重叠，该范围内的 IPv6 组播组只在本 IPv6 管理域内有效，相当于私有组地址。而 IPv6 Global 域则可视为一种特殊的 IPv6 管理域，其维护的 BSR 为 Scope 值为 14 的 IPv6 组播组提供服务。

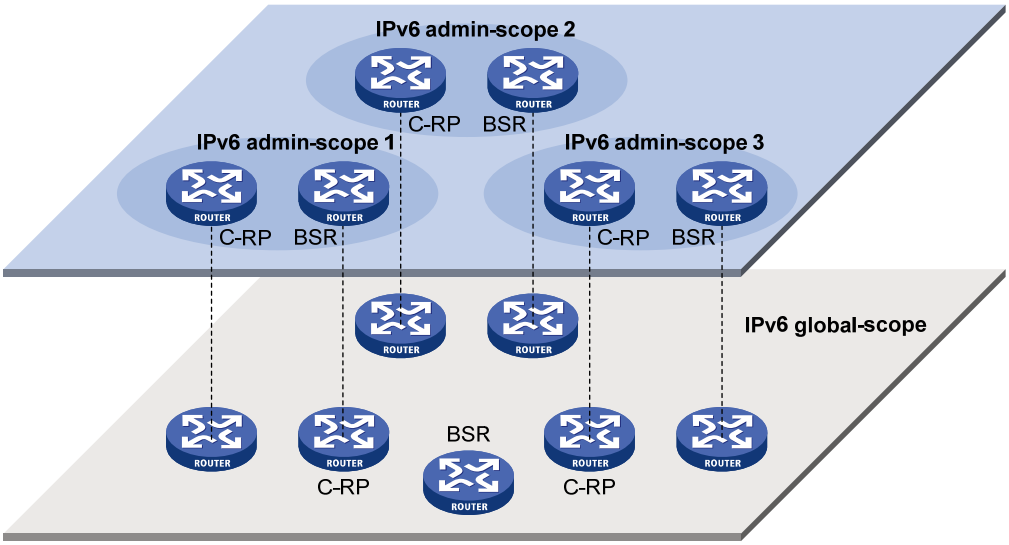
2. 管理域与Global域的关系

每个 IPv6 管理域以及 IPv6 Global 域都有独立的 C-RP 和 BSR 设备，这些设备仅在其所属的域有效，也就是说 BSR 机制与 RP 选举在各 IPv6 管理域之间是隔离的；每个 IPv6 管理域都有自己的边界，各 IPv6 管理域所服务 IPv6 组播组范围内的 IPv6 组播信息不能进、出该边界。为了更清晰地理解 IPv6 管理域和 IPv6 Global 域之间的关系，可以从以下两个角度进行考虑：

(1)　地域空间角度

IPv6 管理域是针对特定 Scope 值的逻辑管理区域，属于此范围的 IPv6 组播报文只能在本 IPv6 管理域的域内或域外传播，无法跨过 IPv6 管理域的边界。

图1-10 地域空间上 IPv6 管理域与 IPv6 Global 域的关系

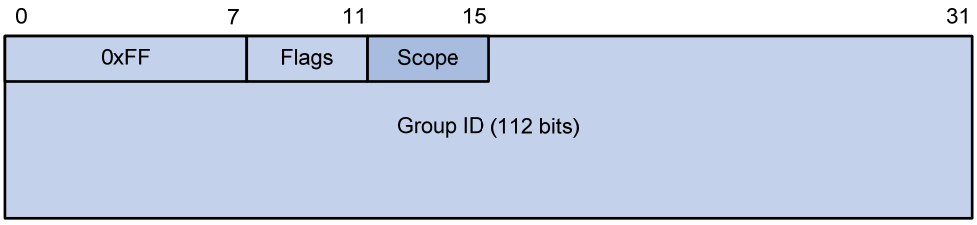


如 图 1-10 所示，对于同一Scope值的IPv6 组播组而言，各IPv6 管理域在地域上必须相互独立、相互隔离。而IPv6 Global域则包含了IPv6 PIM-SM/IPv6 双向PIM域内的所有路由器，不属于任何IPv6 管理域服务范围的IPv6 组播报文，可以在整个IPv6 PIM-SM/IPv6 双向PIM域范围内传播。

(2)　Scope 值角度

如 图 1-11 所示，IPv6 组播通过其地址结构中的Scope字段来表明该IPv6 组播组属于哪个域。

图1-11 IPv6 组播地址结构



Scope值较大的域包含Scope值较小的域，Scope值为E所对应的域（即IPv6 Global域）最大。Scope字段可能的取值及其含义如 表 1-2 所示。

表1-2 Scope 字段的取值及其含义

取值	含义	所属域
0、F	保留（Reserved）	-
1	接口本地范围（Interface-Local Scope）	-
2	链路本地范围（Link-Local Scope）	-
3	子网本地范围（Subnet-Local Scope）	IPv6管理域
4	管理本地范围（Admin-Local Scope）	IPv6管理域
5	站点本地范围（Site-Local Scope）	IPv6管理域
6、7、9～D	未分配（Unassigned）	IPv6管理域
8	机构本地范围（Organization-Local Scope）	IPv6管理域
E	全球范围（Global Scope）	IPv6 Global域

1.1.5 IPv6 PIM-SSM简介

SSM（Source-Specific Multicast，指定信源组播）模型和 ASM（Any-Source Multicast，任意信源组播）模型是两个完全对等的模型。目前，ASM 模型包括 IPv6 PIM-DM 和 IPv6 PIM-SM 两种模式，SSM 模型能够借助 IPv6 PIM-SM 的部分技术来实现，也称为 IPv6 PIM-SSM。

SSM 模型为指定源组播提供了解决方案，通过 MLDv2 来维护主机与路由器之间的关系。在实际应用中，通常采用 MLDv2 以及 IPv6 PIM-SM 的一部分技术来实现 SSM 模型。由于接收者已经通过其它渠道（如广告咨询等）知道了 IPv6 组播源的具体位置，因此在 SSM 模型中无需 RP，无需构建 RPT，也无需 IPv6 组播源注册过程来发现其它 IPv6 PIM 域内的 IPv6 组播源。

IPv6 PIM-SSM 的工作机制可以概括如下：

- 邻居发现
- DR 选举
- 构建 SPT

1. 邻居发现

IPv6 PIM-SSM使用与IPv6 PIM-SM完全相同的邻居发现机制，具体请参见“1.1.1 1. 邻居发现”一节。

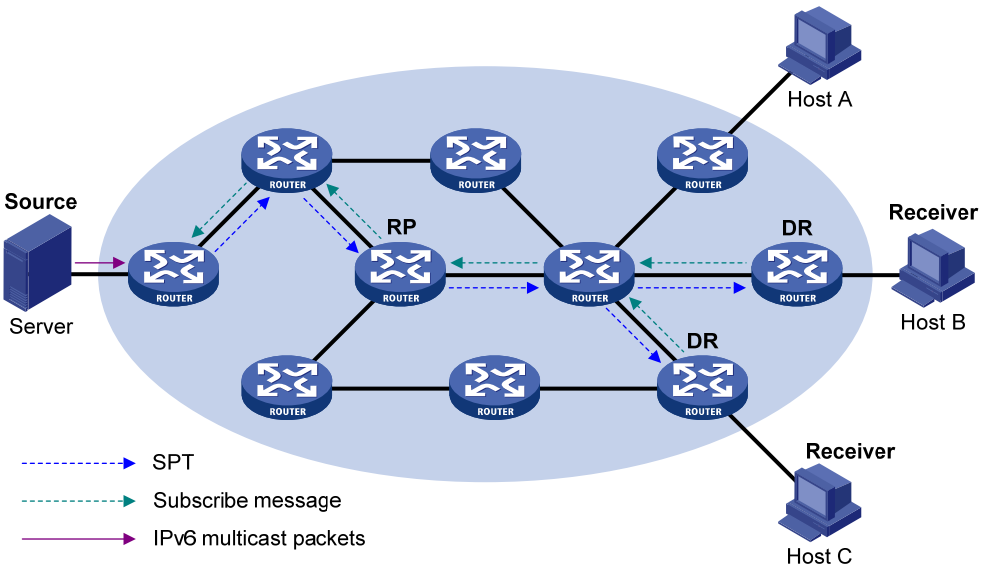
2. DR选举

IPv6 PIM-SSM使用与IPv6 PIM-SM完全相同的DR选举机制，具体请参见 “1.1.2 2. DR选举”一节。

3. 构建SPT

构建为 IPv6 PIM-SM 服务的 RPT，还是构建为 IPv6 PIM-SSM 服务的 SPT，关键在于接收者准备加入的IPv6组播组是否属于IPv6 SSM组地址范围（IANA 保留的IPv6 SSM组地址范围为 FF3x::/32，其中 x 表示任意合法的 scope）。

图1-12 IPv6 PIM-SSM 中构建 SPT 示意图



如 图 1-12 所示，Host B和Host C为IPv6 组播信息的接收者（Receiver），由其借助MLDv2 的报告报文向DR报告自己对来自IPv6 组播源S、发往IPv6 组播组G的信息感兴趣。收到该报告报文的DR先判断该报文中的IPv6 组地址是否在IPv6 SSM组地址范围内：

- 如果在 IPv6 SSM 组地址范围内，则构建 IPv6 PIM-SSM，并向 IPv6 组播源 S 逐跳发送通道（Channel）的订阅报文（Subscribe Message）。沿途所有路由器上都创建（S，G）表项，从而在网络内构建了一棵以 IPv6 组播源 S 为根、以接收者为叶子的 SPT，该 SPT 就是 IPv6 PIM-SSM 中的传输通道；
- 如果不在 IPv6 SSM 组地址范围内，则仍旧按照 IPv6 PIM-SM 的流程进行后续处理，此时接收者侧 DR 需要向 RP 发送（*，G）加入报文，同时 IPv6 组播源侧 DR 需要进行 IPv6 组播源的注册。

说明

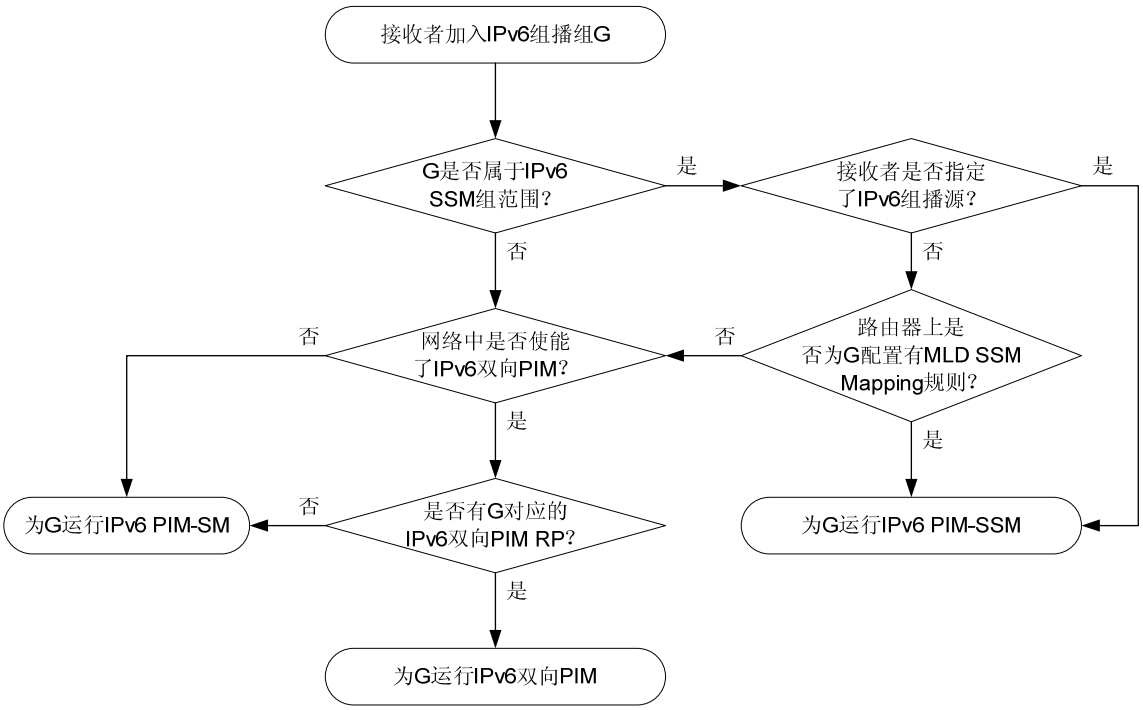
在 IPv6 PIM-SSM 中，借助“通道”的概念表示 IPv6 组播组，借助“订阅报文”的概念表示加入报文。

1.1.6 各IPv6 PIM协议运行关系

在一个 IPv6 PIM 网络中，不允许 IPv6 PIM-DM 与其它类型的 IPv6 PIM 协议（IPv6 PIM-SM、IPv6 双向 PIM 和 IPv6 PIM-SSM）同时运行，但允许同时运行 IPv6 PIM-SM、IPv6 双向 PIM 和 IPv6 PIM-SSM。

当网络中同时运行IPv6 PIM-SM、IPv6 双向PIM和IPv6 PIM-SSM时，针对具体的组加入行为运行哪种类型的IPv6 PIM协议，其判断过程如 图 1-13 所示。

图1-13 各 IPv6 PIM 协议运行关系示意图



说明

有关 MLD SSM Mapping 的详细介绍，请参见“IP 组播配置指导”中的“MLD”。

1.1.7 协议规范

与 IPv6 PIM 相关的协议规范有：

- RFC 3973：Protocol Independent Multicast-Dense Mode (PIM-DM): Protocol Specification (Revised)

- RFC 4601：Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification (Revised)
- RFC 3956：Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address
- RFC 5015：Bidirectional Protocol Independent Multicast (BIDIR-PIM)
- RFC 5059：Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)
- RFC 4607：Source-Specific Multicast for IP
- draft-ietf-ssm-overview-05：An Overview of Source-Specific Multicast (SSM)

1.2 配置IPv6 PIM-DM

1.2.1 IPv6 PIM-DM配置任务简介

表1-3 IPv6 PIM-DM 配置任务简介

配置任务	说明	详细配置
使能IPv6 PIM-DM	必选	1.2.3
使能状态刷新能力	可选	1.2.4
配置状态刷新参数	可选	1.2.5
配置IPv6 PIM-DM定时器	可选	1.2.6
配置IPv6 PIM公共特性	可选	1.6

1.2.2 配置准备

在配置 IPv6 PIM-DM 之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通

在配置 IPv6 PIM-DM 之前，需准备以下数据：

- 发送状态刷新报文的时间间隔
- 等待接收新状态刷新报文的最小时间
- 状态刷新报文的 Hop Limit 值
- 嫁接报文的重传时间

1.2.3 使能IPv6 PIM-DM

在接口上使能了 IPv6 PIM-DM 后，路由器会定期发送 Hello 报文以发现 IPv6 PIM 邻居，并对收到的来自 IPv6 PIM 邻居的报文进行处理。在部署 IPv6 PIM-DM 域时，建议在其所有非边界接口上均使能 IPv6 PIM-DM。

表1-4 使能 IPv6 PIM-DM

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
使能IPv6组播路由	**multicast ipv6 routing-enable**	必选 缺省情况下，IPv6组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能IPv6 PIM-DM	**pim ipv6 dm**	必选 缺省情况下，IPv6 PIM-DM处于关闭状态

注意

- 同一台设备所有接口上所启用的 IPv6 PIM 模式必须相同。
- IPv6 PIM-DM 不能与处于 IPv6 SSM 组地址范围内的 IPv6 组播组同时使用。

说明

有关 **multicast ipv6 routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。

1.2.4 使能状态刷新能力

为了避免各路由器上被剪枝的接口因为超时而恢复转发，与 IPv6 组播源直连的路由器会周期性地发送（S，G）状态刷新报文，该报文沿着 IPv6 PIM-DM 域最初的扩散路径逐跳进行转发，从而刷新沿途所有路由器上的剪枝定时器的状态。只有当一个共享网段内的所有 IPv6 PIM 路由器上都使能了状态刷新能力时，该共享网段才具备状态刷新能力。

请在 IPv6 PIM-DM 域内的所有路由器上进行如下配置。

表1-5 使能状态刷新能力

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能状态刷新能力	**pim ipv6 state-refresh-capable**	可选 缺省情况下，状态刷新能力处于使能状态

1.2.5 配置状态刷新参数

在与 IPv6 组播源直连的路由器上，会以一定的时间间隔周期性地发送（S，G）状态刷新报文，可以通过配置来改变这个时间间隔。

路由器可能在短时间内收到多个状态刷新报文，而其中有些报文可能是重复的。为了避免接收这些重复的报文，可以配置接收新状态刷新报文的等待时间：路由器将丢弃在该时间内收到的状态刷新

报文；当该时间超时后，路由器将正常接收新的状态刷新报文，并更新自己的 IPv6 PIM-DM 状态，同时重置该等待时间。

在收到状态刷新报文时，路由器会将该报文的 Hop Limit 值减 1 后转发给其下游，直至该报文的 Hop Limit 值减为 0，当网络规模很小时，状态刷新报文将在网络中循环传递。因此，为了有效控制刷新报文的传递范围，需要根据网络规模大小配置合适的 Hop Limit 值。

请在 IPv6 PIM-DM 域内的所有路由器上进行如下配置。

表1-6 配置状态刷新参数

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置发送状态刷新报文的时间间隔	**state-refresh-interval** *interval*	可选 缺省情况下，发送状态刷新报文的时间间隔为60秒
配置接收新状态刷新报文的等待时间	**state-refresh-rate-limit** *interval*	可选 缺省情况下，接收新状态刷新报文的等待时间为30秒
配置状态刷新报文的Hop Limit值	**state-refresh-hoplimit** *hoplimit-value*	可选 缺省情况下，状态刷新报文的Hop Limit值为255

1.2.6 配置IPv6 PIM-DM定时器

嫁接报文是 IPv6 PIM-DM 中唯一使用确认机制的报文。在 IPv6 PIM-DM 域中，下游路由器发出嫁接报文后，如果在指定时间内没有收到来自其上游路由器的嫁接应答报文，则会重发嫁接报文，直到被确认。

表1-7 配置 IPv6 PIM-DM 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置嫁接报文的重传时间	**pim ipv6 timer graft-retry** *interval*	可选 缺省情况下，嫁接报文的重传时间为3秒

说明

有关IPv6 PIM-DM其它定时器的相关配置，请参见“1.6.7 配置IPv6 PIM公共定时器”。

1.3 配置IPv6 PIM-SM

1.3.1 IPv6 PIM-SM配置任务简介

表1-8 IPv6 PIM-SM 配置任务简介

配置任务		说明	详细配置
使能IPv6 PIM-SM		必选	1.3.3
配置RP	配置静态RP	三者必选其一	1.3.4 1.
	配置C-RP		1.3.4 2.
	使能嵌入式RP		1.3.4 3.
	全局配置C-RP定时器	可选	1.3.4 4.
配置BSR	配置C-BSR	必选	1.3.5 1.
	配置BSR服务边界	可选	1.3.5 2.
	全局配置C-BSR参数	可选	1.3.5 3.
	配置C-BSR定时器	可选	1.3.5 4.
	关闭自举报文语义分片功能	可选	1.3.5 5.
配置IPv6管理域	使能IPv6管理域机制	可选	1.3.6 1.
	配置IPv6管理域边界	可选	1.3.6 2.
	配置IPv6管理域的C-BSR	可选	1.3.6 3.
配置IPv6组播源注册		可选	1.3.7
配置禁止SPT切换		可选	1.3.8
配置IPv6 PIM公共特性		可选	1.6

1.3.2 配置准备

在配置 IPv6 PIM-SM 之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通

在配置 IPv6 PIM-SM 之前，需准备以下数据：

- 静态 RP 的 IPv6 地址及表示其所服务 IPv6 组范围的 IPv6 ACL 规则
- C-RP 的优先级及表示其所服务 IPv6 组范围的 IPv6 ACL 规则
- 合法 C-RP 的地址范围及表示其所服务 IPv6 组范围的 IPv6 ACL 规则
- 发送宣告报文的时间间隔
- C-RP 超时时间
- C-BSR 的优先级
- 哈希掩码长度
- 表示合法 BSR 地址范围的 IPv6 ACL 规则

- 自举时间间隔
- 自举超时时间
- 表示注册报文过滤规则的 IPv6 ACL 规则
- 注册抑制时间
- 注册探测时间
- 禁止发起 SPT 切换的 IPv6 ACL 规则和排序规则

1.3.3 使能IPv6 PIM-SM

在接口上使能了 IPv6 PIM-SM 后，路由器会定期发送 Hello 报文以发现 IPv6 PIM 邻居，并对收到的来自 IPv6 PIM 邻居的报文进行处理。在部署 IPv6 PIM-SM 域时，建议在其所有非边界接口上均使能 IPv6 PIM-SM。

表1-9 使能 IPv6 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6组播路由	**multicast ipv6 routing-enable**	必选 缺省情况下，IPv6组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能IPv6 PIM-SM	**pim ipv6 sm**	必选 缺省情况下，IPv6 PIM-SM处于关闭状态

注意

同一台设备所有接口上所启用的 IPv6 PIM 模式必须相同。

说明

有关 **multicast ipv6 routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。

1.3.4 配置RP

RP 可以通过手工方式静态配置，也可以通过 BSR 机制动态选举。对于大型 IPv6 PIM 网络，配置静态 RP 将会非常繁琐。所以，通常静态 RP 是作为动态选举 RP 机制的备份手段，以提高网络的健壮性，增强组播网络的运营管理能力。

注意

当 IPv6 PIM 网络中同时运行 IPv6 PIM-SM 和 IPv6 双向 PIM 时，请勿使一个 RP 同时为 IPv6 PIM-SM 和 IPv6 双向 PIM 工作，否则可能引起 IPv6 PIM 路由表出错。

1. 配置静态RP

当网络内仅有一个动态 RP 时，通过手工配置静态 RP 可以避免因单一节点故障而引起的通信中断，同时也可以避免 C-RP 与 BSR 之间频繁的信息交互而占用带宽。

请在 IPv6 PIM-SM 域内的所有路由器上进行如下配置。

表1-10 配置静态 RP

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置服务于IPv6 PIM-SM的静态RP	**static-rp** *ipv6-rp-address* [*acl6-number*] [**preferred**]	必选 缺省情况下，没有配置静态RP

注意

为了让静态 RP 功能正常发挥作用，必须在所有路由器上指定相同的静态 RP 地址。

2. 配置C-RP

在 PIM-SM 域中，可以把有意成为 RP 的路由器配置为 C-RP。BSR 通过接收来自 C-RP 的 C-RP 信息，或者接收来自其它路由器的自动 RP 宣告，收集 C-RP 信息并将其汇总为 RP-Set 信息，然后在全网内扩散。之后，网络内的其它路由器根据 RP-Set 信息计算出特定组播组范围所对应的 RP。建议在骨干网路由器上配置 C-RP。

为了防止 C-RP 欺骗，需要在 BSR 上配置合法的 C-RP 地址范围及其所服务的组播组范围。同时由于每个 C-BSR 都可能成为 BSR，因此必须在 IPv6 PIM-SM 域内的所有 C-BSR 上都配置相同的过滤策略。

表1-11 配置 C-RP

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置某接口为服务于IPv6 PIM-SM的C-RP	**c-rp** *ipv6-address* [{ **group-policy** *acl6-number* | **scope** *scope-id* } | **priority** *priority* | **holdtime** *hold-interval* | **advertisement-interval** *adv-interval*] *	必选 缺省情况下，没有配置C-RP

操作	命令	说明
合法C-RP的地址范围及其所服务的IPv6组播组范围	**crp-policy** *acl6-number*	可选 缺省情况下，C-RP的地址范围及其所服务的IPv6组播组范围不受任何限制

说明

- 在配置 C-RP 时，应在 C-RP 与 IPv6 PIM-SM 域中的其它设备之间保留较大的通信带宽。
- 一个 RP 可以为多个 IPv6 组播组服务，也可以为所有 IPv6 组播组服务。每个 IPv6 组播组在任意时刻，只能由唯一的一个 RP 为其转发数据，而不能由多个 RP 转发数据。

3. 使能嵌入式RP

嵌入式 RP 机制可以从 IPv6 组播地址中直接解析出 RP，从而取代静态配置的 RP 或由 BSR 机制动态计算出来的 RP，DR 不再需要预先知道 RP 的信息，只需对组播报文进行分析即可知道 RP 的地址。

请在 IPv6 PIM-SM 域内的所有路由器上进行如下配置。

表1-12 使能嵌入式 RP

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
使能嵌入式RP功能	**embedded-rp** [*acl6-number*]	可选 缺省情况下，默认嵌入式RP地址范围内的IPv6组播组均可以使用嵌入式RP功能

说明

默认的嵌入式 RP 地址范围为 FF7x::/12 和 FFFx::/12，其中 x 表示任意合法的 scope。有关 scope 字段的详细介绍，请参见“IP 组播配置指导”中的“组播概述”。

4. 全局配置C-RP定时器

为了使 BSR 能够在 IPv6 PIM-SM 域内分发 RP-Set 信息，C-RP 必须周期性地向 BSR 发送宣告报文，BSR 从该报文中学习 RP-Set 信息，并将该信息与自己的 IPv6 地址一起封装在自举报文中向域中的所有 IPv6 PIM 路由器进行宣告。

C-RP 在其宣告报文中封装一个保持时间，BSR 在收到该报文后，从中获得该时间值并启动 C-RP 超时定时器，如果超时后 BSR 仍没有收到来自 C-RP 后续的宣告报文，则认为目前网络中的 C-RP 失效或不可达。

请在已配置为 C-RP 的路由器上进行如下配置。

表1-13 全局配置 C-RP 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置发送宣告报文的时间间隔	**c-rp advertisement-interval** *interval*	可选 缺省情况下，发送宣告报文的时间间隔为60秒
配置C-RP超时时间	**c-rp holdtime** *interval*	可选 缺省情况下，C-RP的超时时间为150秒

说明

有关IPv6 PIM-SM其它定时器的相关配置，请参见“1.6.7 配置IPv6 PIM公共定时器”。

1.3.5 配置BSR

在一个 IPv6 PIM-SM 域中只能有一个 BSR，但需要配置至少一个 C-BSR。任意一台路由器都可以被配置为 C-BSR。在 C-BSR 之间通过自动选举产生 BSR，BSR 负责在 IPv6 PIM-SM 域中收集并发布 RP 信息。

1. 配置C-BSR

C-BSR 应配置在骨干网的路由器上，在将路由器配置为 C-BSR 时，必须同时指定一个使能了 IPv6 PIM-SM 的接口的 IPv6 地址。C-BSR 间的自动选举机制简单描述如下：

- 最初，每个 C-BSR 都认为自己是本 IPv6 PIM-SM 域的 BSR，并使用接口的 IPv6 地址作为 BSR 地址，发送自举报文；
- 当某 C-BSR 收到其它 C-BSR 发来的自举报文时，首先比较自己与后者的优先级，优先级较高者获胜；在优先级相同的情况下，再比较自己与后者的 BSR 地址，拥有较大 IPv6 地址者获胜。如果后者获胜，则用后者的 BSR 地址替换自己的 BSR 地址，并不再认为自己是 BSR；否则，保留自己的 BSR 地址，并继续认为自己是 BSR。

通过在路由器上配置合法 BSR 的地址范围，可以对收到的自举报文按照地址范围进行过滤，从而防止某些恶意主机非法伪装成 BSR，以避免合法的 BSR 被恶意取代。必须在 IPv6 PIM-SM 域内的所有路由器上进行相同的配置。通常针对以下两类情况实施预防措施：

- 某些恶意主机通过伪造自举报文以欺骗路由器，试图更改 RP 映射关系。这种攻击通常发生在边缘路由器上，由于 BSR 处于网络内部，主机在网络外部，因此边缘路由器通过对收到的自举报文进行邻居检查和 RPF 检查，丢弃不符合要求的报文，就可以避免外部网络用户对内部网络 BSR 的攻击；
- 网络中某台路由器被攻击者控制，或者有非法接入的路由器时，攻击者可以将这样的路由器配置为 C-BSR，并使其在竞争中获胜，从而控制网络中 RP 信息的发布权。由于在被配置为 C-BSR 后，路由器会自动向整个网络扩散自举报文，而自举报文是 Hop Limit 值为 1 的 IPv6 组播报文，所以只要其邻居路由器不接收该自举报文，就不会影响整个网络。因此，通过在

整个网络的所有路由器上都配置合法 BSR 的地址范围，从而丢弃合法范围之外的自举报文，就可以防止此类攻击。

以上两种预防措施可以部分地保护网络中 BSR 的安全。但是如果某台合法的 BSR 路由器被攻击者控制，还是可能导致问题。

表1-14 配置 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置某接口为C-BSR	**c-bsr** *ipv6-address* [*hash-length* [*priority*]]	必选 缺省情况下，没有配置C-BSR
配置合法的BSR地址范围	**bsr-policy** *acl6-number*	可选 缺省情况下，BSR的地址范围不受任何限制

说明

由于 BSR 与 IPv6 PIM-SM 域中的其它设备需要交换大量信息，因此应在 C-BSR 与 IPv6 PIM-SM 域中的其它设备之间保留较大的通信带宽。

2. 配置BSR服务边界

BSR 作为 IPv6 PIM-SM 域中的管理核心，负责将收集到的 RP-Set 信息以自举报文的形式发向 IPv6 PIM-SM 域中的所有路由器。

BSR 的服务边界，即 IPv6 PIM-SM 域的边界。BSR 是针对特定的服务范围而言的，众多的 BSR 服务边界接口将网络划分成不同的 IPv6 PIM-SM 域，自举报文无法通过 IPv6 PIM-SM 域的边界，BSR 服务边界之外的路由器也不能参与本 IPv6 PIM-SM 域内的组播转发。

请在欲配置为 BSR 服务边界的路由器上进行如下配置。

表1-15 配置 BSR 服务边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置BSR的服务边界	**pim ipv6 bsr-boundary**	必选 缺省情况下，没有配置BSR的服务边界

3. 全局配置C-BSR参数

在一个 IPv6 PIM-SM 域中，从众多 C-BSR 中选举出唯一的 BSR。IPv6 PIM-SM 域内的 C-RP 向 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向本 IPv6 PIM-SM 域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定 IPv6 组播组所对应 RP 的地址。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-16 全局配置 C-BSR 参数

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置哈希掩码长度	**c-bsr hash-length** *hash-length*	可选 缺省情况下，哈希掩码长度为126
配置C-BSR的优先级	**c-bsr priority** *priority*	可选 缺省情况下，C-BSR的优先级为64

4. 配置C-BSR定时器

当某 C-BSR 竞选成为 BSR 后，它会通过自举报文向其所服务的区域以组播方式发送自己的 IPv6 地址和 RP-Set 信息。BSR 以自举时间间隔周期性地向网络发送自举报文，收到该报文的 C-BSR 会在自举超时时间内将其保持，此时 BSR 选举过程暂停；当自举超时时间超时后，C-BSR 之间会触发新一轮的 BSR 选举过程。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-17 配置 C-BSR 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置自举时间间隔	**c-bsr interval** *interval*	可选 缺省取值请参见表后的说明
配置自举超时时间	**c-bsr holdtime** *interval*	可选 缺省取值请参见表后的说明

说明

关于自举时间间隔的取值:

- 如果没有进行手工配置，则其取值由如下公式决定：自举时间间隔 =（自举超时时间 − 10）÷ 2。缺省情况下，自举超时时间为 130 秒，则自举时间间隔的缺省值 =（130 − 10）÷ 2 = 60（秒）。
- 如果进行了手工配置，则取配置值。

关于自举超时时间的取值:

- 如果没有进行手工配置，则其取值由如下公式决定：自举超时时间 = 自举时间间隔 × 2 + 10。缺省情况下，自举时间间隔为 60 秒，则自举超时时间的缺省值 = 60 × 2 + 10 = 130（秒）。
- 如果进行了手工配置，则取配置值。

注意

配置时，应保证自举时间间隔小于自举超时时间。

5. 关闭自举报文语义分片功能

BSR 周期性地向所在 IPv6 PIM-SM 域发送自举报文以通告 RP-Set 信息。当 RP-Set 信息较少时，自举报文被封装在一个 IPv6 报文中发送出去；而当 RP-Set 信息较多时，自举报文的大小可能超过接口的 MTU（Maximum Transmission Unit，最大传输单元）值，从而触发其在 IP 层的分片。在这种情况下，一个 IP 分片的丢失就会导致整个自举报文都被丢弃。

自举报文语义分片功能可以解决上述问题：当自举报文大于接口 MTU 时，会被分解为多个自举报文分片（Bootstrap Message Fragment，BSMF）。非 BSR 收到自举报文分片后，若发现某组范围对应的 RP 信息都在这一个分片中，便立即更新该组范围对应的 RP-Set；若发现某组范围映射的 RP 信息被分在了多个分片中，则待收齐了这些分片后再更新该组范围对应的 RP-Set。这样，由于不同分片所含组范围对应的 RP 信息不同，因此个别分片的丢失只影响该分片所含组范围对应的 RP 信息，而不会导致整个自举报文都被丢弃。

自举报文语义分片功能是缺省使能的，但由于不支持该功能的设备会将自举报文分片当作完整的自举报文处理，从而导致其学到的 RP-Set 信息不完整，因此当 IPv6 PIM-SM 域中存在此类设备时，请在已配置为 C-BSR 的路由器上关闭本功能。

表1-18 关闭自举报文语义分片功能

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
关闭自举报文语义分片功能	**undo bsm-fragment enable**	必选 缺省情况下，自举报文语义分片功能处于使能状态

说明

通常，BSR 根据其 BSR 接口的 MTU 值对自举报文进行语义分片；而对由于新学到 IPv6 PIM 邻居而触发的自举报文发送，则根据发送接口的 MTU 值进行语义分片。

1.3.6 配置IPv6 管理域

在 IPv6 非管理域机制下，一个 IPv6 PIM-SM 域中只有唯一的 BSR，整个网络都在该 BSR 的管理范围之内。为了更有针对性地管理，可以将整个 IPv6 PIM-SM 域划分为多个 IPv6 管理域：每个 IPv6 管理域中各维护一个 BSR，服务于特定 Scope 值的 IPv6 组播组；IPv6 Global 域则可视为一种特殊的 IPv6 管理域，其维护的 BSR 为 Scope 值为 14 的 IPv6 组播组提供服务。

1. 使能IPv6 管理域机制

在配置 IPv6 管理域各项功能之前，必须先使能 IPv6 管理域机制。

请在 IPv6 PIM-SM 域内的所有路由器上进行如下配置。

表1-19 使能 IPv6 管理域机制

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
使能IPv6管理域机制	**c-bsr admin-scope**	必选 缺省情况下，IPv6管理域机制处于关闭状态

2. 配置IPv6 管理域边界

在 IPv6 管理域机制中，各 IPv6 管理域的边界由 ZBR 构成，每个 IPv6 管理域各维护一个 BSR，服务于特定 Scope 值的 IPv6 组播组，属于此范围的 IPv6 组播协议报文（如断言报文、BSR 自举报文等）无法通过 IPv6 管理域边界。

请在欲配置为 ZBR 的路由器上进行如下配置。

表1-20 配置 IPv6 管理域边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置IPv6组播转发边界	**multicast ipv6 boundary** { *ipv6-group-address prefix-length* | **scope** { *scope-id* | **admin-local** | **global** | **organization-local** | **site-local** } }	必选 缺省情况下，没有配置IPv6组播转发边界

说明

有关 **multicast ipv6 boundary** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。

3. 配置IPv6 管理域的C-BSR

在应用了 IPv6 管理域机制的网络中，从众多 C-BSR 中选举出针对不同 Scope 值的 BSR。网络内的 C-RP 只向对应的 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向其所服务 IPv6 管理域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定 IPv6 组播组所对应 RP 的地址。

请在欲配置为 IPv6 管理域 C-BSR 的路由器上进行如下配置。

表1-21 配置 IPv6 管理域的 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-

操作	命令	说明
配置IPv6管理域的C-BSR	**c-bsr scope** { *scope-id* | **admin-local** | **global** | **organization-local** | **site-local** } [**hash-length** *hash-length* | **priority** *priority*] *	必选 缺省情况下，没有配置IPv6管理域的C-BSR

说明

哈希掩码长度和 C-BSR 的优先级可以在全局和 IPv6 管理域这两种范围内进行配置：

- 如果在 IPv6 管理域下进行了配置，则取其配置值；
- 如果未在 IPv6 管理域下进行配置，则取相应的全局值。

有关全局C-BSR参数的相关配置，请参见“1.3.5 3. 全局配置C-BSR参数”。

1.3.7 配置IPv6 组播源注册

在 IPv6 PIM-SM 域内，IPv6 组播源侧 DR 向 RP 发送注册报文，而这些注册报文拥有不同的 IPv6 组播源或 IPv6 组播地址。为了让 RP 服务于特定的 IPv6 组播组，可以对注册报文进行过滤。如果某个（S，G）表项被过滤规则拒绝，或者过滤规则中没有定义对它的操作，RP 都会向 DR 发送注册停止报文，以停止该 IPv6 组播数据的注册过程。

出于对注册报文在传递过程中完整性的考虑，可以配置根据整个报文来计算校验和。但为了减少往注册报文中封装数据报文的工作量并考虑到互通性，一般情况下不建议配置根据注册报文的全部内容来计算校验和的方式。

当接收者不再通过 RP 接收发往某 IPv6 组播组的数据（即 RP 不再服务于该 IPv6 组播组），或 RP 开始接收 IPv6 组播源沿着 SPT 发来的 IPv6 组播数据时，RP 将向 IPv6 组播源侧 DR 发送注册停止报文，DR 收到该报文后将停止发送封装有 IPv6 组播数据的注册报文并启动注册停止定时器（Register-Stop Timer）。在注册停止定时器超时之前，DR 会向 RP 发送一个空注册报文（Null-Register Message，即不封装 IPv6 组播数据的注册报文）：如果 DR 在注册探测时间（Register_Probe_Time）内收到了来自 RP 的注册停止报文，DR 将刷新其注册停止定时器；否则，DR 将重新开始发送封装有 IPv6 组播数据的注册报文。

注册停止定时器的超时时间是一个随机值，由其它两个时间值决定：注册抑制时间（Register_Suppression_Time）和注册探测时间。其具体取值范围如下：（0.5×注册抑制时间，1.5×注册抑制时间）－注册探测时间。

请在所有已配置为 C-RP 的路由器上配置注册报文的过滤规则和根据注册报文的全部内容来计算校验和；请在所有可能成为 IPv6 组播源侧 DR 的路由器上配置注册抑制时间和注册探测时间。

表1-22 配置 IPv6 组播源注册

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-

操作	命令	说明
配置注册报文的过滤规则	**register-policy** *acl6-number*	可选 缺省情况下，没有配置注册报文的过滤规则
配置根据注册报文的全部内容来计算校验和	**register-whole-checksum**	可选 缺省情况下，仅根据注册报文头来计算校验和
配置注册抑制时间	**register-suppression-timeout** *interval*	可选 缺省情况下，注册抑制时间为60秒
配置注册探测时间	**probe-interval** *interval*	可选 缺省情况下，注册探测时间为5秒

1.3.8 配置禁止SPT切换

缺省情况下，当本系列交换机作为 RP 或接收者侧 DR 时，设备会在收到第一个 IPv6 组播数据包后立即发起 SPT 切换，可以通过下面的配置关闭从 RPT 到 SPT 的切换。

表1-23 配置禁止 SPT 切换

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置禁止发起SPT切换的条件	**spt-switch-threshold infinity** [**group-policy** *acl6-number* [**order** *order-value*]]	可选 缺省情况下，设备收到第一个IPv6组播数据包后便立即向SPT切换

注意

由于某些设备无法将 IPv6 组播报文封装在注册报文中发给 RP，因此在可能成为 RP 的设备上不建议配置永不发起 SPT 切换，以免导致 IPv6 组播报文转发失败。

1.4 配置IPv6双向PIM

1.4.1 IPv6 双向PIM配置任务简介

表1-24 IPv6 双向 PIM 配置任务简介

配置任务		说明	详细配置
使能IPv6 PIM-SM		必选	1.4.3
使能IPv6双向PIM		必选	1.4.4
配置RP	配置静态RP	三者必选其一	1.4.5 1.

配置任务		说明	详细配置
	配置C-RP		1.4.5 2.
	使能嵌入式RP		1.4.5 3.
	全局配置C-RP定时器	可选	1.4.5 4.
配置BSR	配置C-BSR	必选	1.4.6 1.
	配置BSR服务边界	可选	1.4.6 2.
	全局配置C-BSR参数	可选	1.4.6 3.
	配置C-BSR定时器	可选	1.4.6 4.
	关闭自举报文语义分片功能	可选	1.4.6 5.
配置IPv6管理域	使能IPv6管理域机制	可选	1.4.7 1.
	配置IPv6管理域边界	可选	1.4.7 2.
	配置IPv6管理域的C-BSR	可选	1.4.7 3.
配置IPv6 PIM公共特性		可选	1.6

1.4.2 配置准备

在配置 IPv6 双向 PIM 之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通

在配置 IPv6 双向 PIM 之前，需准备以下数据：

- 静态 RP 的 IPv6 地址及表示其所服务 IPv6 组范围的 IPv6 ACL 规则
- C-RP 的优先级及表示其所服务 IPv6 组范围的 IPv6 ACL 规则
- 合法 C-RP 的地址范围及表示其所服务 IPv6 组范围的 IPv6 ACL 规则
- 发送宣告报文的时间间隔
- C-RP 超时时间
- C-BSR 的优先级
- 哈希掩码长度
- 表示合法 BSR 地址范围的 IPv6 ACL 规则
- 自举时间间隔
- 自举超时时间

1.4.3 使能IPv6 PIM-SM

由于 IPv6 双向 PIM 是在 IPv6 PIM-SM 的基础上实现的，因此在使能 IPv6 双向 PIM 之前须先使能 IPv6 PIM-SM。在部署 IPv6 双向 PIM 域时，建议在其所有非边界接口上均使能 IPv6 PIM-SM。

表1-25 使能 IPv6 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6组播路由	**multicast ipv6 routing-enable**	必选 缺省情况下，IPv6组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能IPv6 PIM-SM	**pim ipv6 sm**	必选 缺省情况下，IPv6 PIM-SM处于关闭状态

注意

同一台设备所有接口上所启用的 IPv6 PIM 模式必须相同。

说明

有关 **multicast ipv6 routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。

1.4.4 使能IPv6 双向PIM

请在 IPv6 双向 PIM 域内的所有路由器上进行如下配置。

表1-26 使能 IPv6 双向 PIM

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
使能IPv6双向PIM	**bidir-pim enable**	必选 缺省情况下，IPv6双向PIM处于关闭状态

1.4.5 配置RP

RP 可以通过手工方式静态配置，也可以通过 BSR 机制动态选举。对于大型 IPv6 PIM 网络，配置静态 RP 将会非常繁琐。所以，通常静态 RP 是作为动态选举 RP 机制的备份手段，以提高网络的健壮性，增强组播网络的运营管理能力。

注意

当 IPv6 PIM 网络中同时运行 IPv6 PIM-SM 和 IPv6 双向 PIM 时，请勿使一个 RP 同时为 IPv6 PIM-SM 和 IPv6 双向 PIM 工作，否则可能引起 IPv6 PIM 路由表出错。

1. 配置静态RP

当网络内仅有一个动态 RP 时，通过手工配置静态 RP 可以避免因单一节点故障而引起的通信中断，同时也可以避免 C-RP 与 BSR 之间频繁的信息交互而占用带宽。

请在 IPv6 双向 PIM 域内的所有路由器上进行如下配置。

表1-27 配置静态 RP

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置服务于IPv6双向PIM的静态RP	**static-rp** *ipv6-rp-address* [*acl6-number*] [**preferred**] **bidir**	必选 缺省情况下，没有配置静态RP

注意

为了让静态 RP 功能正常发挥作用，必须在所有路由器上指定相同的静态 RP 地址。

说明

IPv6 双向 PIM 允许将静态 RP 的 IPv6 地址指定为一个实际不存在的 IPv6 地址。譬如，一条链路两端接口的 IPv6 地址分别为 1001::1/64 和 1001::2/64，可以将静态 RP 的 IPv6 地址指定为同网段但实际不存在的一个地址，如 1001::100/64，该链路就成为了 RPL。

2. 配置C-RP

在 IPv6 双向 PIM 域中，可以把有意成为 RP 的路由器配置为 C-RP。BSR 通过接收来自 C-RP 的 C-RP 信息，或者接收来自其它路由器的自动 RP 宣告，收集 C-RP 信息并将其汇总为 RP-Set 信息，然后在全网内扩散。之后，网络内的其它路由器根据 RP-Set 信息计算出特定组播组范围所对应的 RP。建议在骨干网路由器上配置 C-RP。

为了防止 C-RP 欺骗，需要在 BSR 上配置合法的 C-RP 地址范围及其所服务的组播组范围。同时由于每个 C-BSR 都可能成为 BSR，因此必须在 IPv6 双向 PIM 域内的所有 C-BSR 上都配置相同的过滤策略。

表1-28 配置 C-RP

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入IPv6 PIM视图	**pim ipv6**	-
配置某接口为服务于IPv6双向PIM的C-RP	**c-rp** *ipv6-address* [{ **group-policy** *acl6-number* | **scope** *scope-id* } | **priority** *priority* | **holdtime** *hold-interval* | **advertisement-interval** *adv-interval*] * **bidir**	必选 缺省情况下，没有配置C-RP

说明

- 在配置 C-RP 时，应在 C-RP 与 IPv6 双向 PIM 域中的其它设备之间保留较大的通信带宽。
- 一个 RP 可以为多个 IPv6 组播组服务，也可以为所有 IPv6 组播组服务。每个 IPv6 组播组在任意时刻，只能由唯一的一个 RP 为其转发数据，而不能由多个 RP 转发数据。

3. 使能嵌入式RP

嵌入式 RP 机制可以从 IPv6 组播地址中直接解析出 RP，从而取代静态配置的 RP 或由 BSR 机制动态计算出来的 RP，DR 不再需要预先知道 RP 的信息，只需对组播报文进行分析即可知道 RP 的地址。

请在 IPv6 双向 PIM 域内的所有路由器上进行如下配置。

表1-29 使能嵌入式 RP

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
使能嵌入式RP功能	**embedded-rp** [*acl6-number*]	可选 缺省情况下，默认嵌入式RP地址范围内的IPv6组播组均可以使用嵌入式RP功能

说明

默认的嵌入式 RP 地址范围为 FF7x::/12 和 FFFx::/12，其中 x 表示任意合法的 scope。有关 scope 字段的详细介绍，请参见“IP 组播配置指导”中的“组播概述”。

4. 全局配置C-RP定时器

为了使 BSR 能够在 IPv6 双向 PIM 域内分发 RP-Set 信息，C-RP 必须周期性地向 BSR 发送宣告报文，BSR 从该报文中学习 RP-Set 信息，并将该信息与自己的 IPv6 地址一起封装在自举报文中向域中的所有 IPv6 PIM 路由器进行宣告。

C-RP 在其宣告报文中封装一个保持时间，BSR 在收到该报文后，从中获得该时间值并启动 C-RP 超时定时器，如果超时后 BSR 仍没有收到来自 C-RP 后续的宣告报文，则认为目前网络中的 C-RP 失效或不可达。

请在已配置为 C-RP 的路由器上进行如下配置。

表1-30 全局配置 C-RP 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置发送宣告报文的时间间隔	**c-rp advertisement-interval** *interval*	可选 缺省情况下，发送宣告报文的时间间隔为60秒
配置C-RP超时时间	**c-rp holdtime** *interval*	可选 缺省情况下，C-RP的超时时间为150秒

说明

有关IPv6 双向PIM其它定时器的相关配置，请参见“1.6.7 配置IPv6 PIM公共定时器”。

1.4.6 配置BSR

在一个 IPv6 双向 PIM 域中只能有一个 BSR，但需要配置至少一个 C-BSR。任意一台路由器都可以被配置为 C-BSR。在 C-BSR 之间通过自动选举产生 BSR，BSR 负责在 IPv6 双向 PIM 域中收集并发布 RP 信息。

1. 配置C-BSR

C-BSR 应配置在骨干网的路由器上，在将路由器配置为 C-BSR 时，必须同时指定一个使能了 IPv6 PIM-SM 的接口的 IPv6 地址。C-BSR 间的自动选举机制简单描述如下：

- 最初，每个 C-BSR 都认为自己是本 IPv6 双向 PIM 域的 BSR，并使用接口的 IPv6 地址作为 BSR 地址，发送自举报文；
- 当某 C-BSR 收到其它 C-BSR 发来的自举报文时，首先比较自己与后者的优先级，优先级较高者获胜；在优先级相同的情况下，再比较自己与后者的 BSR 地址，拥有较大 IPv6 地址者获胜。如果后者获胜，则用后者的 BSR 地址替换自己的 BSR 地址，并不再认为自己是 BSR；否则，保留自己的 BSR 地址，并继续认为自己是 BSR。

通过在路由器上配置合法 BSR 的地址范围，可以对收到的自举报文按照地址范围进行过滤，从而防止某些恶意主机非法伪装成 BSR，以避免合法的 BSR 被恶意取代。必须在 IPv6 双向 PIM 域内的所有路由器上进行相同的配置。通常针对以下两类情况实施预防措施：

- 某些恶意主机通过伪造自举报文以欺骗路由器，试图更改 RP 映射关系。这种攻击通常发生在边缘路由器上，由于 BSR 处于网络内部，主机在网络外部，因此边缘路由器通过对收到的自举报文进行邻居检查和 RPF 检查，丢弃不符合要求的报文，就可以避免外部网络用户对内部网络 BSR 的攻击；
- 网络中某台路由器被攻击者控制，或者有非法接入的路由器时，攻击者可以将这样的路由器配置为 C-BSR，并使其在竞争中获胜，从而控制网络中 RP 信息的发布权。由于在被配置为 C-BSR 后，路由器会自动向整个网络扩散自举报文，而自举报文是 Hop Limit 值为 1 的 IPv6

组播报文，所以只要其邻居路由器不接收该自举报文，就不会影响整个网络。因此，通过在整个网络的所有路由器上都配置合法 BSR 的地址范围，从而丢弃合法范围之外的自举报文，就可以防止此类攻击。

以上两种预防措施可以部分地保护网络中 BSR 的安全。但是如果某台合法的 BSR 路由器被攻击者控制，还是可能导致问题。

表1-31 配置 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置某接口为C-BSR	**c-bsr** *ipv6-address* [*hash-length* [*priority*]]	必选 缺省情况下，没有配置C-BSR
配置合法的BSR地址范围	**bsr-policy** *acl6-number*	可选 缺省情况下，BSR的地址范围不受任何限制

说明

由于 BSR 与 IPv6 双向 PIM 域中的其它设备需要交换大量信息，因此应在 C-BSR 与 IPv6 双向 PIM 域中的其它设备之间保留较大的通信带宽。

2. 配置BSR服务边界

BSR 作为 IPv6 双向 PIM 域中的管理核心，负责将收集到的 RP-Set 信息以自举报文的形式发向 IPv6 双向 PIM 域中的所有路由器。

BSR 的服务边界，即 IPv6 双向 PIM 域的边界。BSR 是针对特定的服务范围而言的，众多的 BSR 服务边界接口将网络划分成不同的 IPv6 双向 PIM 域，自举报文无法通过 IPv6 双向 PIM 域的边界，BSR 服务边界之外的路由器也不能参与本 IPv6 双向 PIM 域内的组播转发。

请在欲配置为 BSR 服务边界的路由器上进行如下配置。

表1-32 配置 BSR 服务边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置BSR的服务边界	**pim ipv6 bsr-boundary**	必选 缺省情况下，没有配置BSR的服务边界

3. 全局配置C-BSR参数

在一个 IPv6 双向 PIM 域中，从众多 C-BSR 中选举出唯一的 BSR。IPv6 双向 PIM 域内的 C-RP 向 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向本 IPv6 双向 PIM 域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定 IPv6 组播组所对应 RP 的地址。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-33 全局配置 C-BSR 参数

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置哈希掩码长度	**c-bsr hash-length** *hash-length*	可选 缺省情况下，哈希掩码长度为126
配置C-BSR的优先级	**c-bsr priority** *priority*	可选 缺省情况下，C-BSR的优先级为64

4. 配置C-BSR定时器

当某 C-BSR 竞选成为 BSR 后，它会通过自举报文向其所服务的区域以组播方式发送自己的 IPv6 地址和 RP-Set 信息。BSR 以自举时间间隔周期性地向网络发送自举报文，收到该报文的 C-BSR 会在自举超时时间内将其保持，此时 BSR 选举过程暂停；当自举超时时间超时后，C-BSR 之间会触发新一轮的 BSR 选举过程。

请在已配置为 C-BSR 的路由器上进行如下配置。

表1-34 配置 C-BSR 定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置自举时间间隔	**c-bsr interval** *interval*	可选 缺省取值请参见表后的说明
配置自举超时时间	**c-bsr holdtime** *interval*	可选 缺省取值请参见表后的说明

说明

关于自举时间间隔的取值：

- 如果没有进行手工配置，则其取值由如下公式决定：自举时间间隔 =（自举超时时间 − 10）÷ 2。缺省情况下，自举超时时间为 130 秒，则自举时间间隔的缺省值 =（130 − 10）÷ 2 = 60（秒）。
- 如果进行了手工配置，则取配置值。

关于自举超时时间的取值：

- 如果没有进行手工配置，则其取值由如下公式决定：自举超时时间 = 自举时间间隔 × 2 + 10。缺省情况下，自举时间间隔为 60 秒，则自举超时时间的缺省值 = 60 × 2 + 10 = 130（秒）。
- 如果进行了手工配置，则取配置值。

注意

配置时，应保证自举时间间隔小于自举超时时间。

5. 关闭自举报文语义分片功能

BSR 周期性地向所在 IPv6 双向 PIM 域发送自举报文以通告 RP-Set 信息。当 RP-Set 信息较少时，自举报文被封装在一个 IPv6 报文中发送出去；而当 RP-Set 信息较多时，自举报文的大小可能超过接口的 MTU（Maximum Transmission Unit，最大传输单元）值，从而触发其在 IP 层的分片。在这种情况下，一个 IP 分片的丢失就会导致整个自举报文都被丢弃。

自举报文语义分片功能可以解决上述问题：当自举报文大于接口 MTU 时，会被分解为多个自举报文分片（Bootstrap Message Fragment，BSMF）。非 BSR 收到自举报文分片后，若发现某组范围对应的 RP 信息都在这一个分片中，便立即更新该组范围对应的 RP-Set；若发现某组范围映射的 RP 信息被分在了多个分片中，则待收齐了这些分片后再更新该组范围对应的 RP-Set。这样，由于不同分片所含组范围对应的 RP 信息不同，因此个别分片的丢失只影响该分片所含组范围对应的 RP 信息，而不会导致整个自举报文都被丢弃。

自举报文语义分片功能是缺省使能的，但由于不支持该功能的设备会将自举报文分片当作完整的自举报文处理，从而导致其学到的 RP-Set 信息不完整，因此当 IPv6 双向 PIM 域中存在此类设备时，请在已配置为 C-BSR 的路由器上关闭本功能。

表1-35 关闭自举报文语义分片功能

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
关闭自举报文语义分片功能	**undo bsm-fragment enable**	必选 缺省情况下，自举报文语义分片功能处于使能状态

说明

通常，BSR 根据其 BSR 接口的 MTU 值对自举报文进行语义分片；而对由于新学到 IPv6 PIM 邻居而触发的自举报文发送，则根据发送接口的 MTU 值进行语义分片。

1.4.7 配置IPv6 管理域

在 IPv6 非管理域机制下，一个 IPv6 双向 PIM 域中只有唯一的 BSR，整个网络都在该 BSR 的管理范围之内。为了更有针对性地管理，可以将整个 IPv6 双向 PIM 域划分为多个 IPv6 管理域：每个 IPv6 管理域中各维护一个 BSR，服务于特定 Scope 值的 IPv6 组播组；IPv6 Global 域则可视为一种特殊的 IPv6 管理域，其维护的 BSR 为 Scope 值为 14 的 IPv6 组播组提供服务。

1. 使能IPv6 管理域机制

在配置 IPv6 管理域各项功能之前，必须先使能 IPv6 管理域机制。

请在 IPv6 双向 PIM 域内的所有路由器上进行如下配置。

表1-36 使能 IPv6 管理域机制

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
使能IPv6管理域机制	**c-bsr admin-scope**	必选 缺省情况下，IPv6管理域机制处于关闭状态

2. 配置IPv6 管理域边界

在 IPv6 管理域机制中，各 IPv6 管理域的边界由 ZBR 构成，每个 IPv6 管理域各维护一个 BSR，服务于特定 Scope 值的 IPv6 组播组，属于此范围的 IPv6 组播协议报文（如断言报文、BSR 自举报文等）无法通过 IPv6 管理域边界。

请在欲配置为 ZBR 的路由器上进行如下配置。

表1-37 配置 IPv6 管理域边界

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置IPv6组播转发边界	**multicast ipv6 boundary** { *ipv6-group-address prefix-length* | **scope** { *scope-id* | **admin-local** | **global** | **organization-local** | **site-local** } }	必选 缺省情况下，没有配置IPv6组播转发边界

说明

有关 **multicast ipv6 boundary** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。

3. 配置IPv6 管理域的C-BSR

在应用了 IPv6 管理域机制的网络中，从众多 C-BSR 中选举出针对不同 Scope 值的 BSR。网络内的 C-RP 只向对应的 BSR 发送宣告报文，由 BSR 汇总为 RP-Set，并向其所服务 IPv6 管理域内的所有路由器进行宣告。所有路由器都使用统一的哈希算法，得到特定 IPv6 组播组所对应 RP 的地址。

请在欲配置为 IPv6 管理域 C-BSR 的路由器上进行如下配置。

表1-38 配置 IPv6 管理域的 C-BSR

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-

操作	命令	说明
配置IPv6管理域的C-BSR	**c-bsr scope** { *scope-id* | **admin-local** | **global** | **organization-local** | **site-local** } [**hash-length** *hash-length* | **priority** *priority*] *	必选 缺省情况下，没有配置IPv6管理域的C-BSR


说明

哈希掩码长度和 C-BSR 的优先级可以在全局和 IPv6 管理域这两种范围内进行配置：

- 如果在 IPv6 管理域下进行了配置，则取其配置值；
- 如果未在 IPv6 管理域下进行配置，则取相应的全局值。

有关全局C-BSR参数的相关配置，请参见“1.4.6 3. 全局配置C-BSR参数”。

1.5 配置IPv6 PIM-SSM


说明

IPv6 PIM-SSM 模型需要 MLDv2 的支持，因此应确保连接有接收者的 IPv6 PIM 路由器上使能了 MLDv2。

1.5.1 IPv6 PIM-SSM配置任务简介

表1-39 IPv6 PIM-SSM 配置任务简介

配置任务	说明	详细配置
使能IPv6 PIM-SM	必选	1.5.3
配置IPv6 SSM组播组范围	可选	1.5.4
配置IPv6 PIM公共特性	可选	1.6

1.5.2 配置准备

在配置 IPv6 PIM-SSM 之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通

在配置 IPv6 PIM-SSM 之前，需准备以下数据：

- IPv6 SSM 组播组范围

1.5.3 使能IPv6 PIM-SM

由于 IPv6 PIM-SSM 是通过 IPv6 PIM-SM 的部分子集功能实现的，因此在配置 IPv6 PIM-SSM 之前须先使能 IPv6 PIM-SM。在部署 IPv6 PIM-SSM 域时，建议在其所有非边界接口上均使能 IPv6 PIM-SM。

表1-40 使能 IPv6 PIM-SM

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6组播路由	**multicast ipv6 routing-enable**	必选 缺省情况下，IPv6组播路由处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	-
使能IPv6 PIM-SM	**pim ipv6 sm**	必选 缺省情况下，IPv6 PIM-SM处于关闭状态

注意

同一台设备所有接口上所启用的 IPv6 PIM 模式必须相同。

说明

有关 **multicast ipv6 routing-enable** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。

1.5.4 配置IPv6 SSM组播组范围

在把来自 IPv6 组播源的信息传递给接收者的过程中，是采用 IPv6 PIM-SSM 模型还是 IPv6 PIM-SM 模型，这取决于接收者订阅通道（S，G）中的 IPv6 组播组是否在 IPv6 SSM 组播组范围之内，所有使能了 IPv6 PIM-SM 的接口将会认为属于该范围内的 IPv6 组播组采用了 IPv6 SSM 模型。

请在 IPv6 PIM-SSM 域内的所有路由器上进行如下配置。

表1-41 配置 IPv6 SSM 组播组范围

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置IPv6 SSM组播组的范围	**ssm-policy** *acl6-number*	可选 缺省情况下，IPv6 SSM组播组的范围为FF3x::/32，其中x表示任意合法的scope

注意

- 应确保域内所有路由器上配置的 IPv6 SSM 组播组地址范围都一致，否则组播信息将无法通过 IPv6 SSM 模型进行传输。
- 如果某 IPv6 组播组属于 IPv6 SSM 组播组范围，但该组成员使用 MLDv1 发送加入报文，则设备不会触发（*，G）加入报文。

1.6 配置IPv6 PIM公共特性

说明

对于本节中那些既可在 IPv6 PIM 视图、又可在接口视图下进行配置的功能或参数来说:

- 在 IPv6 PIM 视图下所作的配置对所有接口生效，在接口视图下所作的配置仅对当前接口生效;
- 如果在这两个视图下对某功能或参数均进行了配置，则不论配置先后，接口视图下的配置将被优先采用。

1.6.1 IPv6 PIM公共特性配置任务简介

表1-42 IPv6 PIM 公共特性配置任务简介

配置任务	说明	详细配置
配置IPv6组播数据过滤器	可选	1.6.3
配置Hello报文过滤器	可选	1.6.4
配置Hello报文选项	可选	1.6.5
配置剪枝延迟时间	可选	1.6.6
配置IPv6 PIM公共定时器	可选	1.6.7
配置加入/剪枝报文规格	可选	1.6.8
配置IPv6 PIM与BFD联动	可选	1.6.9
配置IPv6 PIM报文的DSCP优先级	可选	1.6.10

1.6.2 配置准备

在配置 IPv6 PIM 公共特性之前，需完成以下任务：

- 使能 IPv6 转发功能并配置任一 IPv6 单播路由协议，实现域内网络层互通
- 配置 IPv6 PIM-DM（或 IPv6 PIM-SM、IPv6 PIM-SSM）

在配置 IPv6 PIM 公共特性之前，需准备以下数据：

- 表示 IPv6 组播数据过滤规则的 IPv6 ACL 规则
- 表示合法 Hello 报文的源地址范围的 IPv6 ACL 规则
- 竞选 DR 的优先级（全局值/接口值）
- 保持 IPv6 PIM 邻居可达状态的时间（全局值/接口值）
- 发送剪枝报文的延迟时间（全局值/接口值）
- 剪枝否决时间（全局值/接口值）
- 剪枝延迟时间
- 发送 Hello 报文的时间间隔（全局值/接口值）
- 发送 Hello 报文的最大延迟时间（接口值）
- 保持断言状态的时间（全局值/接口值）
- 发送加入/剪枝报文的时间间隔（全局值/接口值）
- 保持加入/剪枝状态的时间（全局值/接口值）
- IPv6 组播源生存时间
- 加入/剪枝报文的最大长度
- 加入/剪枝报文中（S，G）表项的最大数量
- IPv6 PIM 报文的 DSCP 优先级

1.6.3 配置IPv6 组播数据过滤器

无论在 IPv6 PIM-DM 还是 IPv6 PIM-SM 域内，各路由器都可以对流经自己的 IPv6 组播数据进行检查，通过比较是否符合过滤规则，从而决定是否继续转发 IPv6 组播数据。也就是说 IPv6 PIM 域内的路由器能够成为 IPv6 组播数据的过滤器。过滤器的存在一方面有助于实现信息流量控制，另一方面可以在安全性方面限定下游接收者能够获得的信息。

表1-43 配置 IPv6 组播数据过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置IPv6组播数据过滤器	**source-policy** *acl6-number*	必选 缺省情况下，没有配置IPv6组播数据过滤器

说明

- 通常，过滤器的位置离 IPv6 组播源距离越近，过滤影响越明显。
- 过滤器不仅过滤独立的 IPv6 组播数据，还过滤封装在注册报文中的 IPv6 组播数据。

1.6.4 配置Hello报文过滤器

随着 IPv6 PIM 协议的推广和应用，对其安全性的要求也越来越高。建立正确的 IPv6 PIM 邻居是 IPv6 PIM 协议安全应用的前提。如果在接口上指定了合法 Hello 报文的源地址范围，便能够保证 IPv6 PIM 邻居的正确建立，从而有效防止各种 IPv6 PIM 协议报文攻击，提高设备对 IPv6 PIM 协议报文处理的安全性。

表1-44 配置 Hello 报文过滤器

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置合法Hello报文的源地址范围	**pim ipv6 neighbor-policy** *acl6-number*	必选 缺省情况下，Hello报文的源地址范围不受任何限制

说明

当 Hello 报文过滤器的配置生效后，对于之前已建立的 IPv6 PIM 邻居，若由于其 Hello 报文被过滤而导致无法收到后续的 Hello 报文，将会在老化超时后被自动删除

1.6.5 配置Hello报文选项

无论在 IPv6 PIM-DM 域还是在 IPv6 PIM-SM 域内，各路由器之间发送的 Hello 报文都包含很多可供配置的选项，对各选项的介绍如下：

- DR_Priority（仅用于 IPv6 PIM-SM）：表示竞选 DR 的优先级，优先级高的设备被选举为 DR。可以在与 IPv6 组播源或接收者直连的共享网络中的所有路由器上都配置此参数。
- Holdtime：表示保持 IPv6 PIM 邻居可达状态的时间，若超时后仍没有收到 Hello 报文，则认为 IPv6 PIM 邻居失效或不可达。
- LAN_Prune_Delay：表示在共享网络上传递剪枝报文的延迟时间，该选项由三部分组成：LAN-delay（发送剪枝报文的延迟时间）、Override-interval（剪枝否决时间）和禁止加入报文抑制能力。当共享网段内各 IPv6 PIM 路由器的 LAN-delay 或 Override-interval 不同时，取其中最大的值；当要禁止加入报文抑制能力时，须在共享网段内的所有 IPv6 PIM 路由器上都禁止该能力。

LAN-delay 表示路由器从收到下游路由器发来的剪枝报文到继续向上游路由器发送剪枝报文的延迟时间，Override-interval 则表示允许下游路由器否决剪枝动作的时间。路由器在收到下游路由器发来的剪枝报文后并不立即执行剪枝动作，而是仍将当前的转发状态保持 LAN-delay＋Override-interval 时间。如果下游路由器需要继续接收 IPv6 组播数据，则必须在 Override-interval 时间内向上游路由器发送加入报文以否决这个剪枝动作，这就称为剪枝否决；如果 Override-interval 时间超时后未收到任何加入报文，上游路由器就会在 LAN-delay＋Override-interval 时间超时后执行剪枝动作。

IPv6 PIM 路由器发送 Hello 报文时，会生成一个随机数 Generation ID 并携带在该报文中。一台 IPv6 PIM 路由器的 Generation ID 一般不会变化，除非该路由器的状态发生了改变（如接口刚使能了 IPv6 PIM 或设备进行了重启），此时，当其开始或重新开始发送 Hello 报文时，会生成一个新的 Generation ID。这样，如果 IPv6 PIM 路由器发现其上游邻居发来的 Hello 报文中 Generation ID 发生了改变，就会认为该邻居的状态已经丢失或其上游邻居已经改变，从而触发发送加入报文以进行状态刷新。

在禁止加入报文抑制能力（即使能邻居跟踪）时，应在共享网段的所有 IPv6 PIM 路由器上都禁止该能力，否则上游路由器无法跟踪每台下游路由器的加入报文。

1. 全局配置Hello报文选项

表1-45 全局配置 Hello 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置竞选DR的优先级	**hello-option dr-priority** *priority*	可选 缺省情况下，竞选DR的优先级为1
配置保持IPv6 PIM邻居可达状态的时间	**hello-option holdtime** *interval*	可选 缺省情况下，保持IPv6 PIM邻居可达状态的时间为105秒
配置发送剪枝报文的延迟时间	**hello-option lan-delay** *interval*	可选 缺省情况下，发送剪枝报文的延迟时间为500毫秒
配置剪枝否决时间	**hello-option override-interval** *interval*	可选 缺省情况下，剪枝否决时间为2500毫秒
禁止加入报文抑制能力	**hello-option neighbor-tracking**	必选 缺省情况下，加入报文抑制能力处于使能状态

2. 在接口上配置Hello报文选项

表1-46 在接口上配置 Hello 报文选项

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置竞选DR的优先级	**pim ipv6 hello-option dr-priority** *priority*	可选 缺省情况下，竞选DR的优先级为1
配置保持IPv6 PIM邻居可达状态的时间	**pim ipv6 hello-option holdtime** *interval*	可选 缺省情况下，保持IPv6 PIM邻居可达状态的时间为105秒
配置发送剪枝报文的延迟时间	**pim ipv6 hello-option lan-delay** *interval*	可选 缺省情况下，发送剪枝报文的延迟时间为500毫秒

操作	命令	说明
配置剪枝否决时间	**pim ipv6 hello-option override-interval** *interval*	可选 缺省情况下，剪枝否决时间为2500毫秒
禁止加入报文抑制能力	**pim ipv6 hello-option neighbor-tracking**	必选 缺省情况下，加入报文抑制能力处于使能状态
配置不接受无Generation ID的Hello报文	**pim ipv6 require-genid**	必选 缺省情况下，接受无Generation ID的Hello报文

1.6.6 配置剪枝延迟时间

剪枝延迟时间（Prune Delay）的作用是使共享网段中的上游路由器在收到下游路由器发来的剪枝报文后不会立即执行剪枝动作，而将当前的转发状态继续保持 Prune Delay 时间：如果在该时间内收到了下游路由器发来的加入报文，则否决这个剪枝动作；否则，便执行这个执行剪枝动作。

表1-47 配置剪枝延迟时间

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置剪枝延迟时间	**prune delay** *interval*	可选 缺省情况下，未配置剪枝延迟时间

1.6.7 配置IPv6 PIM公共定时器

IPv6 PIM 路由器通过周期性地发送 Hello 报文，以发现 IPv6 PIM 邻居，并维护各路由器之间的 IPv6 PIM 邻居关系。

为了避免多个 IPv6 PIM 路由器同时发送 Hello 报文而导致冲突，当 IPv6 PIM 路由器在收到 Hello 报文时，将延迟一段时间后再发送 Hello 报文，该时间值为小于“触发 Hello 报文的最大延迟时间”的一个随机值。

IPv6 PIM 路由器通过周期性地向其上游路由器发送加入/剪枝报文以更新状态，在该报文中携带有保持时间，上游路由器为被剪枝的下游接口设置保持加入/剪枝状态定时器。

在断言中落选的路由器将会剪掉其下游的转发接口，并把这种断言状态保持一段时间。超时后，落选的路由器会重新恢复转发 IPv6 组播数据。

当路由器没有收到来自 IPv6 组播源 S 的后续 IPv6 组播数据时，不会立即删除（S，G）表项，而是将其维持一段时间后再删除，这段时间就称为 IPv6 组播源的生存时间。

1. 全局配置IPv6 PIM公共定时器

表1-48 全局配置 IPv6 PIM 公共定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置发送Hello报文的时间间隔	**timer hello** *interval*	可选 缺省情况下，发送Hello报文的时间间隔为30秒
配置发送加入/剪枝报文的时间间隔	**timer join-prune** *interval*	可选 缺省情况下，发送加入/剪枝报文的时间间隔为60秒
配置保持加入/剪枝状态的时间	**holdtime join-prune** *interval*	可选 缺省情况下，保持加入/剪枝状态的时间为210秒
配置保持断言状态的时间	**holdtime assert** *interval*	可选 缺省情况下，保持断言状态的时间为180秒
配置IPv6组播源生存时间	**source-lifetime** *interval*	可选 缺省情况下，IPv6组播源的生存时间为210秒

2. 在接口上配置IPv6 PIM公共定时器

表1-49 在接口上配置 IPv6 PIM 公共定时器

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置发送Hello报文的时间间隔	**pim ipv6 timer hello** *interval*	可选 缺省情况下，发送Hello报文的时间间隔为30秒
配置触发Hello报文的最大延迟时间	**pim ipv6 triggered-hello-delay** *interval*	可选 缺省情况下，触发Hello报文的最大延迟时间为5秒
配置发送加入/剪枝报文的时间间隔	**pim ipv6 timer join-prune** *interval*	可选 缺省情况下，发送加入/剪枝报文的时间间隔为60秒
配置保持加入/剪枝状态的时间	**pim ipv6 holdtime join-prune** *interval*	可选 缺省情况下，保持加入/剪枝状态的时间为210秒
配置保持断言状态的时间	**pim ipv6 holdtime assert** *interval*	可选 缺省情况下，保持断言状态的时间为180秒

说明

如果网络没有特殊要求，建议采用缺省值。

1.6.8 配置加入/剪枝报文规格

如果加入/剪枝报文的尺寸较大，则丢失一个报文将导致较多信息的遗失；如果加入/剪枝报文的尺寸较小，则单个报文的丢失所产生的影响也将降低。

通过控制加入/剪枝报文中（S，G）表项的数目，可以有效减少单位时间内发送的（S，G）表项数量。

表1-50 配置加入/剪枝报文规格

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置加入/剪枝报文的最大长度	**jp-pkt-size** *packet-size*	可选 缺省情况下，加入/剪枝报文的最大长度为8100字节
配置加入/剪枝报文中（S，G）表项的最大数量	**jp-queue-size** *queue-size*	可选 缺省情况下，加入/剪枝报文中（S，G）表项的最大数量为1020个

注意

如果 IPv6 PIM 网络中部署了配置 IPv6 PIM Snooping 功能的交换机，请在接收者侧的边缘 IPv6 PIM 设备上配置加入/剪枝报文的最大长度不能大于 IPv6 路径 MTU。

1.6.9 配置IPv6 PIM与BFD联动

IPv6 PIM 借助 Hello 报文在共享网段中选举出 DR，使其成为该网段中 IPv6 组播数据的唯一转发者。当 DR 出现故障时，只有待其老化后才会触发新的 DR 选举过程，这个过程通常比较长。为了实现 DR 的快速切换，可以在共享网段的 IPv6 PIM 邻居之间引入 BFD 机制进行链路状态的快速检测。通过在共享网段内的所有 IPv6 PIM 路由器上都使能 IPv6 PIM 与 BFD 联动功能，可以使这些 IPv6 PIM 邻居快速感知 DR 故障并重新选举 DR。

表1-51 配置 IPv6 PIM 与 BFD 联动

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
使能IPv6 PIM BFD功能	**pim ipv6 bfd enable**	必选 缺省情况下，IPv6 PIM BFD功能处于关闭状态

说明

- 只有在接口上先使能了 IPv6 PIM-DM 或 IPv6 PIM-SM，本配置才能生效。
- 有关 BFD 的详细介绍，请参见“可靠性配置指导”中的“BFD”。

1.6.10 配置IPv6 PIM报文的DSCP优先级

在 IPv6 报文头中，包含一个 8bit 的 Traffic class 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 IPv6 PIM 报文的 DSCP 优先级进行配置。

表1-52 配置 IPv6 PIM 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入IPv6 PIM视图	**pim ipv6**	-
配置IPv6 PIM报文的DSCP优先级	**dscp** *dscp-value*	可选 缺省情况下，IPv6 PIM报文的DSCP优先级为48

1.7 IPv6 PIM显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IPv6 PIM 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 IPv6 PIM 统计信息。

表1-53 IPv6 PIM 显示和维护

操作	命令
查看IPv6 PIM-SM域中的BSR信息，以及本地配置并生效的C-RP信息	**display pim ipv6 bsr-info** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6 PIM所使用的IPv6单播路由信息	**display pim ipv6 claimed-route** [*ipv6-source-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
查看IPv6 PIM控制报文的数量	**display pim ipv6 control-message counters** [**message-type** { **probe** | **register** | **register-stop** } | [**interface** *interface-type interface-number* | **message-type** { **assert** | **bsr** | **crp** | **graft** | **graft-ack** | **hello** | **join-prune** | **state-refresh** }] *] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6双向PIM的DF信息	**display pim ipv6 df-info** [*rp-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看尚未确认的IPv6 PIM-DM嫁接信息	**display pim ipv6 grafts** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看接口上的IPv6 PIM信息	**display pim ipv6 interface** [*interface-type interface-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看待发送的加入/剪枝报文信息	**display pim ipv6 join-prune mode** { **sm** [**flags** *flag-value*] | **ssm** } [**interface** *interface-type interface-number* | **neighbor** *ipv6-neighbor-address*] * [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6 PIM邻居信息	**display pim ipv6 neighbor** [**interface** *interface-type interface-number* | *ipv6-neighbor-address* | **verbose**] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看IPv6 PIM路由表的内容	**display pim ipv6 routing-table** [*ipv6-group-address* [*prefix-length*] | *ipv6-source-address* [*prefix-length*] | **incoming-interface** [*interface-type interface-number* | **register**] | **outgoing-interface** { **include** | **exclude** | **match** } { *interface-type interface-number* | **register** } | **mode** *mode-type* | **flags** *flag-value* | **fsm**] * [**|** { **begin** | **exclude** | **include** } *regular-expression*]
查看RP的信息	**display pim ipv6 rp-info** [*ipv6-group-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
重置IPv6 PIM控制报文计数器	**reset pim ipv6 control-message counters** [**interface** *interface-type interface-number*]

1.8 IPv6 PIM典型配置举例

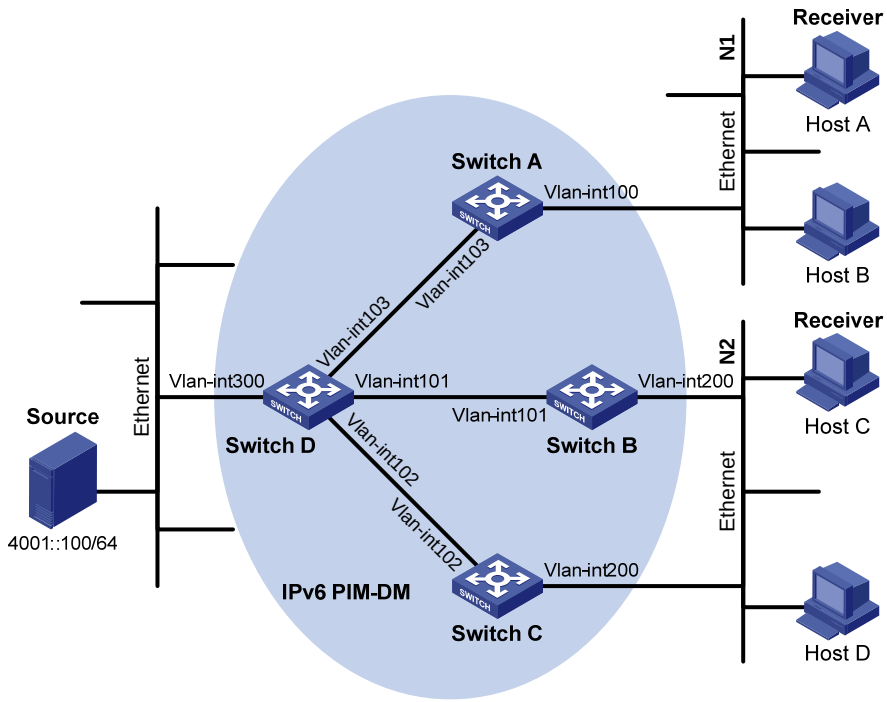
1.8.1 IPv6 PIM-DM典型配置举例

1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者群体组成末梢网络，每个末梢网络中都存在至少一个接收者，整个 IPv6 PIM 域采用 DM 方式。
- Host A 和 Host C 为两个末梢网络中的 IPv6 组播信息接收者；Switch D 通过 Vlan-interface300 接口与 IPv6 组播源（Source）所在的网络连接；Switch A 通过 Vlan-interface100 接口连接末梢网络 N1，通过 Vlan-interface103 接口连接 Switch D；Switch B 和 Switch C 通过各自的 Vlan-interface200 接口连接末梢网络 N2，分别通过 Vlan-interface101 和 Vlan-interface102 接口连接 Switch D。
- Switch A 与末梢网络 N1 之间运行 MLDv1；Switch B 和 Switch C 与末梢网络 N2 之间也运行 MLDv1。

2. 组网图

图1-14 IPv6 PIM-DM 典型配置组网图



设备	接口	IPv6地址	设备	接口	IPv6地址
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int103	1002::1/64		Vlan-int103	1002::2/64
Switch B	Vlan-int200	2001::1/64		Vlan-int101	2002::2/64
	Vlan-int101	2002::1/64		Vlan-int102	3001::2/64
Switch C	Vlan-int200	2001::2/64			
	Vlan-int102	3001::1/64			

3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址和 IPv6 单播路由协议

使能各交换机的IPv6 转发功能，并按照 图 1-14 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

配置 IPv6 PIM-DM 域内的各交换机之间采用 OSPFv3 协议进行互连，确保 IPv6 PIM-DM 域内部在网络层互通，并且各交换机之间能够借助 IPv6 单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IPv6 组播路由，并使能 IPv6 PIM-DM 和 MLD

在 Switch A 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-DM，并在其连接末梢网络的接口 Vlan-interface100 上使能 MLD。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim ipv6 dm
```

```
[SwitchA-Vlan-interface103] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，配置过程略。

在 Switch D 上使能 IPv6 组播路由，并在各接口上使能 IPv6 PIM-DM。

```
<SwitchD> system-view
[SwitchD] multicast ipv6 routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] pim ipv6 dm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim ipv6 dm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 101
[SwitchD-Vlan-interface101] pim ipv6 dm
[SwitchD-Vlan-interface101] quit
[SwitchD] interface vlan-interface 102
[SwitchD-Vlan-interface102] pim ipv6 dm
[SwitchD-Vlan-interface102] quit
```

(3) 检验配置效果

通过使用 **display pim ipv6 interface** 命令可以查看交换机接口上 IPv6 PIM 的配置和运行情况。例如：

查看 Switch D 上 IPv6 PIM 的配置信息。

```
[SwitchD] display pim ipv6 interface
 Interface          NbrCnt HelloInt   DR-Pri     DR-Address
 Vlan300            0      30         1          FE80::A01:201:1
                                                 (local)
 Vlan103            0      30         1          FE80::A01:201:2
                                                 (local)
 Vlan101            1      30         1          FE80::A01:201:3
                                                 (local)
 Vlan102            1      30         1          FE80::A01:201:4
                                                 (local)
```

通过使用 **display pim ipv6 neighbor** 命令可以查看交换机之间的 IPv6 PIM 邻居关系。例如：

查看 Switch D 上 IPv6 PIM 的邻居关系信息。

```
[SwitchD] display pim ipv6 neighbor
 Total Number of Neighbors = 3

 Neighbor        Interface           Uptime   Expires  Dr-Priority
 FE80::A01:101:1 Vlan103             00:04:00 00:01:29 1
 FE80::B01:102:2 Vlan101             00:04:16 00:01:29 3
 FE80::C01:103:3 Vlan102             00:03:54 00:01:17 5
```

假如 Host A 需要接收 IPv6 组播组 G（FF0E::101）的信息，当 IPv6 组播源 S（4001::100/64）向 IPv6 组播组 G 发送 IPv6 组播数据时，通过扩散生成 SPT，SPT 路径中各交换机（Switch A 和 Switch D）上都存在（S，G）表项，Host A 向 Switch A 发送 MLD 报告以加入 IPv6 组播组 G，在 Switch A 上生成（*，G）表项。通过使用 **display pim ipv6 routing-table** 命令可以查看交换机的 IPv6 PIM 路由表信息。例如：

查看 Switch A 上的 IPv6 PIM 路由表信息。

```
[SwitchA] display pim ipv6 routing-table
 Total 1 (*, G) entry; 1 (S, G) entry

 (*, FF0E::101)
     Protocol: pim-dm, Flag: WC
     UpTime: 00:01:24
     Upstream interface: NULL
         Upstream neighbor: NULL
         RPF prime neighbor: NULL
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: mld, UpTime: 00:01:20, Expires: never

 (4001::100, FF0E::101)
     Protocol: pim-dm, Flag: ACT
     UpTime: 00:01:20
     Upstream interface: Vlan-interface103
         Upstream neighbor: 1002::2
         RPF prime neighbor: 1002::2
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: pim-dm, UpTime: 00:01:20, Expires: never
```

查看 Switch D 上的 IPv6 PIM 路由表信息。

```
[SwitchD] display pim ipv6 routing-table
 Total 0 (*, G) entry; 1 (S, G) entry
 (4001::100, FF0E::101)
     Protocol: pim-dm, Flag: LOC ACT
     UpTime: 00:02:19
     Upstream interface: Vlan-interface300
         Upstream neighbor: NULL
         RPF prime neighbor: NULL
     Downstream interface(s) information:
     Total number of downstreams: 2
         1: Vlan-interface103
             Protocol: pim-dm, UpTime: 00:02:19, Expires: never
         2: Vlan-interface102
             Protocol: pim-dm, UpTime: 00:02:19, Expires: never
```

1.8.2 IPv6 PIM-SM非管理域典型配置举例

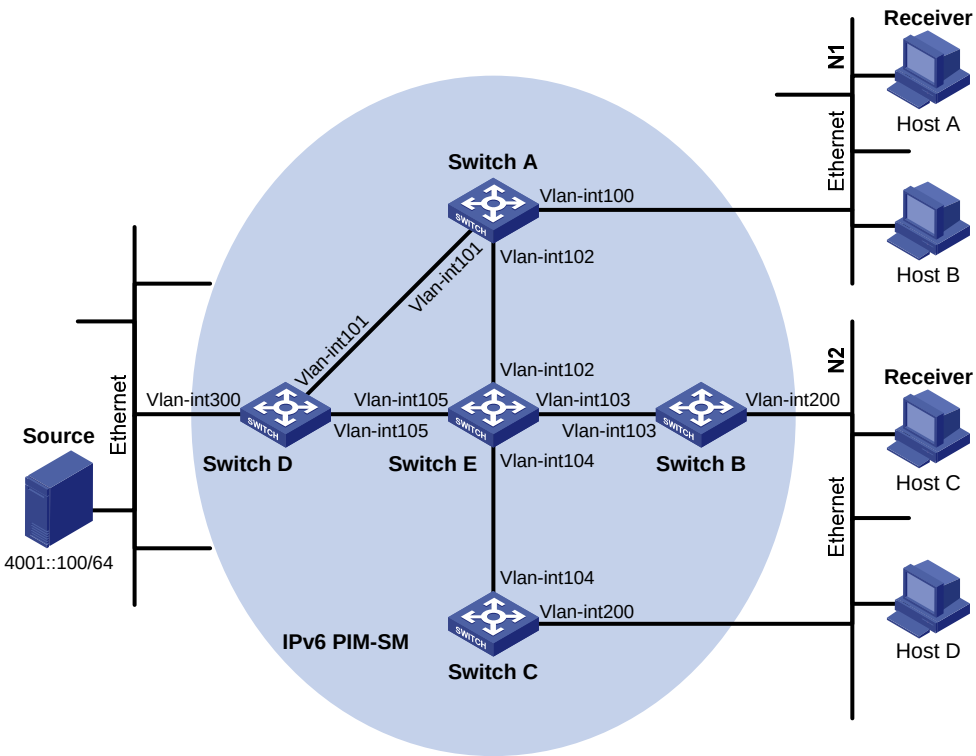
1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者群体组成末梢网络，每个末梢网络中都存在至少一个接收者，整个 IPv6 PIM 域采用 SM 方式。

- Host A和Host C为两个末梢网络中的IPv6组播信息接收者；Switch D通过Vlan-interface300接口与IPv6组播源（Source）所在网络连接；Switch A通过Vlan-interface100接口连接末梢网络N1，通过Vlan-interface101接口和Vlan-interface102接口分别连接Switch D和Switch E；Switch B和Switch C通过各自的Vlan-interface200接口连接末梢网络N2，分别通过Vlan-interface103和Vlan-interface104接口连接Switch E。
- 将Switch D的Vlan-interface105接口和Switch E的Vlan-interface102都配置为C-BSR和C-RP，其中Switch E上C-BSR的优先级较高。C-RP所服务的IPv6组播组范围为FF0E::101/64，通过改变哈希掩码长度使此范围内的IPv6组地址间隔映射到这两个C-RP上。
- Switch A与末梢网络N1之间运行MLDv1；Switch B和Switch C与末梢网络N2之间也运行MLDv1。

2. 组网图

图1-15 IPv6 PIM-SM非管理域典型配置组网图



设备	接口	IPv6地址	设备	接口	IPv6地址
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int101	1002::1/64		Vlan-int101	1002::2/64
	Vlan-int102	1003::1/64		Vlan-int105	4002::1/64
Switch B	Vlan-int200	2001::1/64	Switch E	Vlan-int104	3001::2/64
	Vlan-int103	2002::1/64		Vlan-int103	2002::2/64
Switch C	Vlan-int200	2001::2/64		Vlan-int102	1003::2/64
	Vlan-int104	3001::1/64		Vlan-int105	4002::2/64

3. 配置步骤

(1) 使能IPv6转发功能，并配置IPv6地址和IPv6单播路由协议

使能各交换机的IPv6转发功能，并按照 图 1-15 配置各接口的IPv6地址和前缀长度，具体配置过程略。

配置 IPv6 PIM-SM 域内的各交换机之间采用 OSPFv3 协议进行互连，确保 IPv6 PIM-SM 域内部在网络层互通，并且各交换机之间能够借助 IPv6 单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IPv6 组播路由，并使能 IPv6 PIM-SM 和 MLD

在 Switch A 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并在其连接末梢网络的接口 Vlan-interface300 上使能 MLD。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim ipv6 sm
[SwitchA-Vlan-interface102] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，Switch D 和 Switch E 除了不需要在相应接口上使能 MLD 外，其它的配置也与 Switch A 相似，配置过程略。

(3) 配置 C-BSR 和 C-RP

在 Switch D 上配置 RP 通告的服务范围，以及 C-BSR 和 C-RP 的位置，并指定哈希掩码长度为 128，C-BSR 的优先级为 10。

```
<SwitchD> system-view
[SwitchD] acl ipv6 number 2005
[SwitchD-acl6-basic-2005] rule permit source ff0e::101 64
[SwitchD-acl6-basic-2005] quit
[SwitchD] pim ipv6
[SwitchD-pim6] c-bsr 4002::1 128 10
[SwitchD-pim6] c-rp 4002::1 group-policy 2005
[SwitchD-pim6] quit
```

在 Switch E 上配置 RP 通告的服务范围，以及 C-BSR 和 C-RP 的位置，并指定哈希掩码长度为 128，C-BSR 的优先级为 20。

```
<SwitchE> system-view
[SwitchE] acl ipv6 number 2005
[SwitchE-acl6-basic-2005] rule permit source ff0e::101 64
[SwitchE-acl6-basic-2005] quit
[SwitchE] pim ipv6
[SwitchE-pim6] c-bsr 1003::2 128 20
[SwitchE-pim6] c-rp 1003::2 group-policy 2005
[SwitchE-pim6] quit
```

(4) 检验配置效果

通过使用 **display pim ipv6 interface** 命令可以查看交换机接口上 IPv6 PIM 的配置和运行情况。例如：

查看 Switch A 上 IPv6 PIM 的配置信息。

```
[SwitchA] display pim ipv6 interface
```

```
 Interface            NbrCnt HelloInt   DR-Pri    DR-Address
 Vlan100              0      30         1         FE80::A01:201:1
                                                  (local)
 Vlan101              1      30         1         FE80::A01:201:2
 Vlan102              1      30         1         FE80::A01:201:3
```

通过使用 **display pim ipv6 bsr-info** 命令可以查看交换机上 BSR 选举的信息，以及本地配置并生效的 C-RP 信息。例如：

查看 Switch A 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchA] display pim ipv6 bsr-info
 Elected BSR Address: 1003::2
     Priority: 20
     Hash mask length: 128
     State: Accept Preferred
     Uptime: 00:04:22
     Expires: 00:01:46
```

查看 Switch D 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchD] display pim ipv6 bsr-info
 Elected BSR Address: 1003::2
     Priority: 20
     Hash mask length: 128
     State: Elected
     Uptime: 00:05:26
     Expires: 00:01:45
 Candidate BSR Address: 4002::1
     Priority: 10
     Hash mask length: 128
     State: Candidate

 Candidate RP: 4002::1(Vlan-interface105)
     Priority: 192
     HoldTime: 130
     Advertisement Interval: 60
     Next advertisement scheduled at: 00:00:48
```

查看 Switch E 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchE] display pim ipv6 bsr-info
 Elected BSR Address: 1003::2
     Priority: 20
     Hash mask length: 128
     State: Elected
     Uptime: 00:01:10
     Next BSR message scheduled at: 00:01:48
 Candidate BSR Address: 1003::2
     Priority: 20
     Hash mask length: 128
     State: Elected

 Candidate RP: 1003::2(Vlan-interface102)
```

```
    Priority: 192
    HoldTime: 130
    Advertisement Interval: 60
    Next advertisement scheduled at: 00:00:48
```

通过使用 **display pim ipv6 rp-info** 命令可以查看交换机上获取的 RP 信息。例如：

查看 Switch A 上的 RP 信息。

```
[SwitchA] display pim ipv6 rp-info
 PIM-SM BSR RP information:
 prefix/prefix length: FF0E::101/64
     RP: 4002::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:05:19
     Expires: 00:02:11

     RP: 1003::2
     Priority: 192
     HoldTime: 130
     Uptime: 00:05:19
     Expires: 00:02:11
```

假如 Host A 需要接收 IPv6 组播组 G（FF0E::100）的信息，由于根据哈希算法得出 G 对应的 RP 为 Router E，因此 Switch A 和 Switch E 之间会生成 RPT。当 IPv6 组播源 S（4001::100/64）向 RP 发起注册后，Switch D 和 Switch E 之间会生成 SPT。当 Switch A 收到 IPv6 组播数据后立即执行从 RPT 到 SPT 的切换。RPT 路径中的交换机（Switch A 和 Switch E）上存在（*，G）表项，而 SPT 路径中的交换机（Switch A 和 Switch D）上存在（S，G）表项，通过使用 **display pim ipv6 routing-table** 命令可以查看交换机的 IPv6 PIM 路由表信息。例如：

查看 Switch A 上的 IPv6 PIM 路由表信息。

```
[SwitchA] display pim ipv6 routing-table
 Total 1 (*, G) entry; 1 (S, G) entry

 (*, FF0E::100)
     RP: 1003::2
     Protocol: pim-sm, Flag: WC
     UpTime: 00:03:45
     Upstream interface: Vlan-interface102
         Upstream neighbor: 1003::2
         RPF prime neighbor: 1003::2
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface100
             Protocol: mld, UpTime: 00:02:15, Expires: 00:03:06

 (4001::100, FF0E::100)
     RP: 1003::2
     Protocol: pim-sm, Flag: SPT ACT
     UpTime: 00:02:15
     Upstream interface: Vlan-interface101
```

```
        Upstream neighbor: 1002::2
        RPF prime neighbor: 1002::2
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface100
            Protocol: pim-sm, UpTime: 00:02:15, Expires: 00:03:06
```

查看 Switch D 上的 IPv6 PIM 路由表信息。

```
[SwitchD] display pim ipv6 routing-table
 Total 0 (*, G) entry; 1 (S, G) entry

 (4001::100, FF0E::100)
    RP: 1003::2
    Protocol: pim-sm, Flag: SPT LOC ACT
    UpTime: 00:14:44
    Upstream interface: Vlan-interface300
        Upstream neighbor: NULL
        RPF prime neighbor: NULL
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface105
            Protocol: mld, UpTime: 00:14:44, Expires: 00:02:26
```

查看 Switch E 上的 IPv6 PIM 路由表信息。

```
[SwitchE] display pim ipv6 routing-table
 Total 1 (*, G) entry; 0 (S, G) entry

 (*, FF0E::100)
    RP: 1003::2 (local)
    Protocol: pim-sm, Flag: WC
    UpTime: 00:16:56
    Upstream interface: Register
        Upstream neighbor: 4002::1
        RPF prime neighbor: 4002::1
    Downstream interface(s) information:
    Total number of downstreams: 1
        1: Vlan-interface102
            Protocol: pim-sm, UpTime: 00:16:56, Expires: 00:02:34
```

1.8.3 IPv6 PIM-SM管理域典型配置举例

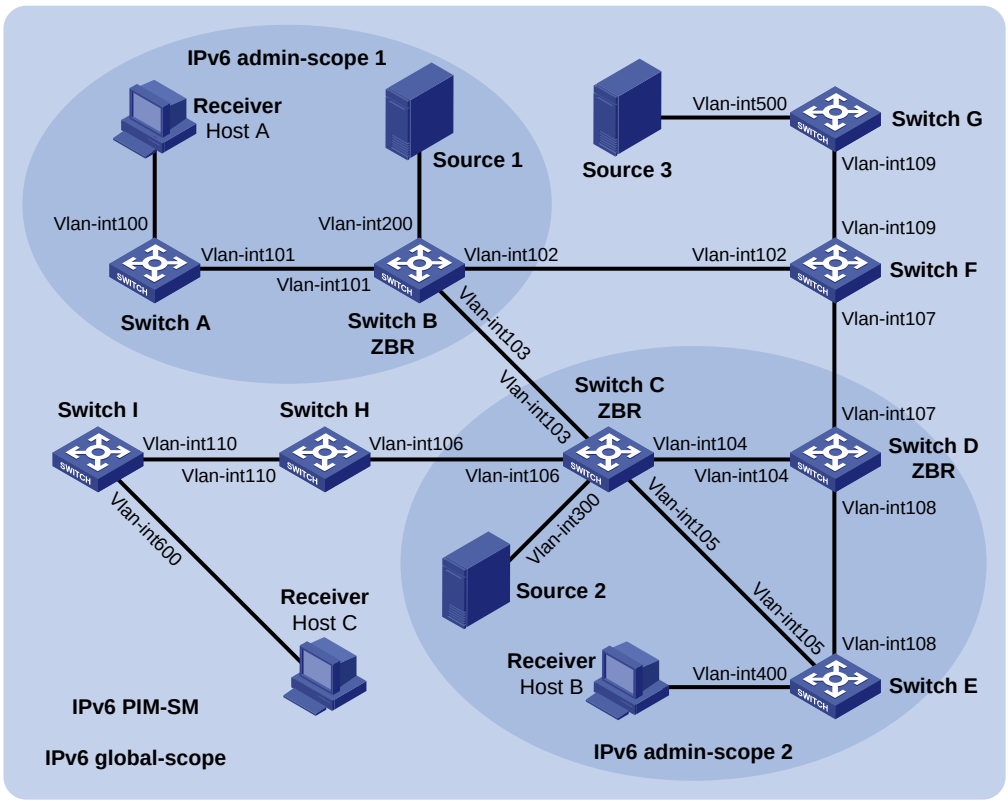
1. 组网需求

- 接收者通过组播方式接收视频点播信息，整个 IPv6 PIM 域采用 SM 管理域方式，划分为 IPv6 管理域 1（Scope 值为 4）、IPv6 管理域 2（Scope 值为 4）和 IPv6 Global 域，Switch B、Switch C 和 Switch D 为各 IPv6 管理域的 ZBR。
- Source 1 和 Source 2 分别向 IPv6 组播组 FF14::101 发送内容不同的 IPv6 组播信息，Host A 和 Host B 则分别只接收来自 Source 1 和 Source 2 的 IPv6 组播信息；Source 3 向 IPv6 组播组 FF1E::202 发送 IPv6 组播信息，Host C 为其接收者。

- Switch B 的 Vlan-interface101 接口为 IPv6 管理域 1 的 C-BSR 和 C-RP，服务于 Scope 值为 4 的 IPv6 组播组；Switch D 的 Vlan-interface104 接口为 IPv6 管理域 2 的 C-BSR 和 C-RP，服务于 Scope 值为 4 的 IPv6 组播组；Switch F 的 Vlan-interface109 接口为 IPv6 Global 域的 C-BSR 和 C-RP，服务于 Scope 值为 14 的 IPv6 组播组。
- Switch A、Switch E 和 Switch I 分别与各自所连接的接收者之间运行 MLDv1。

2. 组网图

图1-16 IPv6 PIM-SM 管理域配置组网图



设备	接口	IPv6地址	设备	接口	IPv6地址
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int104	3002::2/64
	Vlan-int101	1002::1/64		Vlan-int108	6001::1/64
Switch B	Vlan-int200	2001::1/64		Vlan-int107	6002::1/64
	Vlan-int101	1002::2/64	Switch E	Vlan-int400	7001::1/64
	Vlan-int103	2002::1/64		Vlan-int105	3003::2/64
	Vlan-int102	2003::1/64		Vlan-int108	6001::2/64
Switch C	Vlan-int300	3001::1/64	Switch F	Vlan-int109	8001::1/64
	Vlan-int104	3002::1/64		Vlan-int107	6002::2/64
	Vlan-int105	3003::1/64		Vlan-int102	2003::2/64
	Vlan-int103	2002::2/64	Switch G	Vlan-int500	9001::1/64
	Vlan-int106	3004::1/64		Vlan-int109	8001::2/64
Switch H	Vlan-int110	4001::1/64	Source 1	-	2001::100/64
	Vlan-int106	3004::2/64	Source 2	-	3001::100/64
Switch I	Vlan-int600	5001::1/64	Source 3	-	9001::100/64
	Vlan-int110	4001::2/64			

3. 配置步骤

(1) 配置 IPv6 地址和 IPv6 单播路由协议

使能各交换机的IPv6 转发功能，请按照 图 1-16 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

配置 IPv6 PIM-SM 域内的各交换机之间采用 OSPFv3 协议进行互连，确保 IPv6 PIM-SM 域内部在网络层互通，并且各交换机之间能够借助 IPv6 单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IPv6 组播路由和 IPv6 管理域机制，并使能 IPv6 PIM-SM 和 MLD

在 Switch A 上使能 IPv6 组播路由和 IPv6 管理域机制，在各接口上使能 IPv6 PIM-SM，并在其连接有接收者的接口 Vlan-interface100 上使能 MLD。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] pim ipv6
[SwitchA-pim6] c-bsr admin-scope
[SwitchA-pim6] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
```

Switch E 和 Switch I 的配置与 Switch A 相似，配置过程略。

在 Switch B 上使能 IPv6 组播路由和 IPv6 管理域机制，并在各接口上使能 IPv6 PIM-SM。

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] pim ipv6
[SwitchB-pim6] c-bsr admin-scope
[SwitchB-pim6] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] pim ipv6 sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim ipv6 sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim ipv6 sm
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim ipv6 sm
[SwitchB-Vlan-interface103] quit
```

Switch C、Switch D、Switch F、Switch G 和 Switch H 的配置与 Switch B 相似，配置过程略。

(3) 配置 IPv6 管理域边界

在 Switch B 上将接口 Vlan-interface102 和 Vlan-interface103 配置为 IPv6 管理域 1 的边界。

```
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] multicast ipv6 boundary scope 4
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] multicast ipv6 boundary scope 4
```

```
[SwitchB-Vlan-interface103] quit
```

在 Switch C 上将接口 Vlan-interface103 和 Vlan-interface106 配置为 IPv6 管理域 2 的边界。

```
<SwitchC> system-view
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] multicast ipv6 boundary scope 4
[SwitchC-Vlan-interface103] quit
[SwitchC] interface vlan-interface 106
[SwitchC-Vlan-interface106] multicast ipv6 boundary scope 4
[SwitchC-Vlan-interface106] quit
```

在 Switch D 上将接口 Vlan-interface107 配置为 IPv6 管理域 2 的边界。

```
<SwitchD> system-view
[SwitchD] interface vlan-interface 107
[SwitchD-Vlan-interface107] multicast ipv6 boundary scope 4
[SwitchD-Vlan-interface107] quit
```

(4) 配置 C-BSR 和 C-RP

在 Switch B 上配置 RP 通告的服务范围，并将接口 Vlan-interface101 配置为 IPv6 管理域 1 的 C-BSR 和 C-RP。

```
[SwitchB] pim ipv6
[SwitchB-pim6] c-bsr scope 4
[SwitchB-pim6] c-bsr 1002::2
[SwitchB-pim6] c-rp 1002::2 scope 4
[SwitchB-pim6] quit
```

在 Switch D 上配置 RP 通告的服务范围，并将接口 Vlan-interface104 配置为 IPv6 管理域 2 的 C-BSR 和 C-RP。

```
[SwitchD] pim ipv6
[SwitchD-pim6] c-bsr scope 4
[SwitchD-pim6] c-bsr 3002::2
[SwitchD-pim6] c-rp 3002::2 scope 4
[SwitchD-pim6] quit
```

在 Switch F 上将接口 Vlan-interface109 配置为 IPv6 Global 域的 C-BSR 和 C-RP。

```
<SwitchF> system-view
[SwitchF] pim ipv6
[SwitchF-pim6] c-bsr scope global
[SwitchF-pim6] c-bsr 8001::1
[SwitchF-pim6] c-rp 8001::1
[SwitchF-pim6] quit
```

(5) 检验配置效果

通过使用 **display pim ipv6 bsr-info** 命令可以查看交换机上 BSR 选举的信息，以及本地配置并生效的 C-RP 信息。例如：

查看 Switch B 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchB] display pim ipv6 bsr-info
 Elected BSR Address: 8001::1
     Priority: 64
     Hash mask length: 126
     State: Accept Preferred
     Scope: 14
```

```
    Uptime: 00:01:45
    Expires: 00:01:25
 Elected BSR Address: 1002::2
    Priority: 64
    Hash mask length: 126
    State: Elected
    Scope: 4
    Uptime: 00:04:54
    Next BSR message scheduled at: 00:00:06
 Candidate BSR Address: 1002::2
    Priority: 64
    Hash mask length: 126
    State: Elected
    Scope: 4

 Candidate RP: 1002::2(Vlan-interface101)
    Priority: 192
    HoldTime: 130
    Advertisement Interval: 60
    Next advertisement scheduled at: 00:00:15
```

查看 Switch D 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchD] display pim ipv6 bsr-info
 Elected BSR Address: 8001::1
    Priority: 64
    Hash mask length: 126
    State: Accept Preferred
    Scope: 14
    Uptime: 00:01:45
    Expires: 00:01:25
 Elected BSR Address: 3002::2
    Priority: 64
    Hash mask length: 126
    State: Elected
    Scope: 4
    Uptime: 00:03:48
    Next BSR message scheduled at: 00:01:12
 Candidate BSR Address: 3002::2
    Priority: 64
    Hash mask length: 126
    State: Elected
    Scope: 4

 Candidate RP: 3002::2(Vlan-interface104)
    Priority: 192
    HoldTime: 130
    Advertisement Interval: 60
    Next advertisement scheduled at: 00:00:10
```

查看 Switch F 上的 BSR 信息，以及本地配置并生效的 C-RP 信息。

```
[SwitchF] display pim ipv6 bsr-info
 Elected BSR Address: 8001::1
     Priority: 64
     Hash mask length: 126
     State: Elected
     Scope: 14
     Uptime: 00:01:11
     Next BSR message scheduled at: 00:00:49
 Candidate BSR Address: 8001::1
     Priority: 64
     Hash mask length: 126
     State: Elected
     Scope: 14

 Candidate RP: 8001::1(Vlan-interface109)
     Priority: 192
     HoldTime: 130
     Advertisement Interval: 60
     Next advertisement scheduled at: 00:00:55
```

通过使用 **display pim ipv6 rp-info** 命令可以查看交换机上获取的 RP 信息。例如：

查看 Switch B 上的 RP 信息。

```
[SwitchB] display pim ipv6 rp-info
 PIM-SM BSR RP information:
 prefix/prefix length: FF0E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF1E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF2E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF3E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
```

```
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF4E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF5E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF6E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF7E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF8E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF9E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFAE::/16
    RP: 8001::1
    Priority: 192
```

```
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFBE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFCE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFDE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFEE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFFE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF04::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF14::/16
    RP: 1002::2
```

```
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF24::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF34::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF44::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF54::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF64::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF74::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF84::/16
```

```
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF94::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFA4::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFB4::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFC4::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFD4::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFE4::/16
    RP: 1002::2
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51
```

```
 prefix/prefix length: FFF4::/16
     RP: 1002::2
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51
```

查看 Switch F 上的 RP 信息。

```
[SwitchF] display pim rp-info
 PIM-SM BSR RP information:
 prefix/prefix length: FF0E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF1E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF2E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF3E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF4E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
     Uptime: 00:03:39
     Expires: 00:01:51

 prefix/prefix length: FF5E::/16
     RP: 8001::1
     Priority: 192
     HoldTime: 130
```

```
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF6E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF7E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF8E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FF9E::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFAE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFBE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFCE::/16
    RP: 8001::1
    Priority: 192
```

```
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFDE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFEE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51

prefix/prefix length: FFFE::/16
    RP: 8001::1
    Priority: 192
    HoldTime: 130
    Uptime: 00:03:39
    Expires: 00:01:51
```

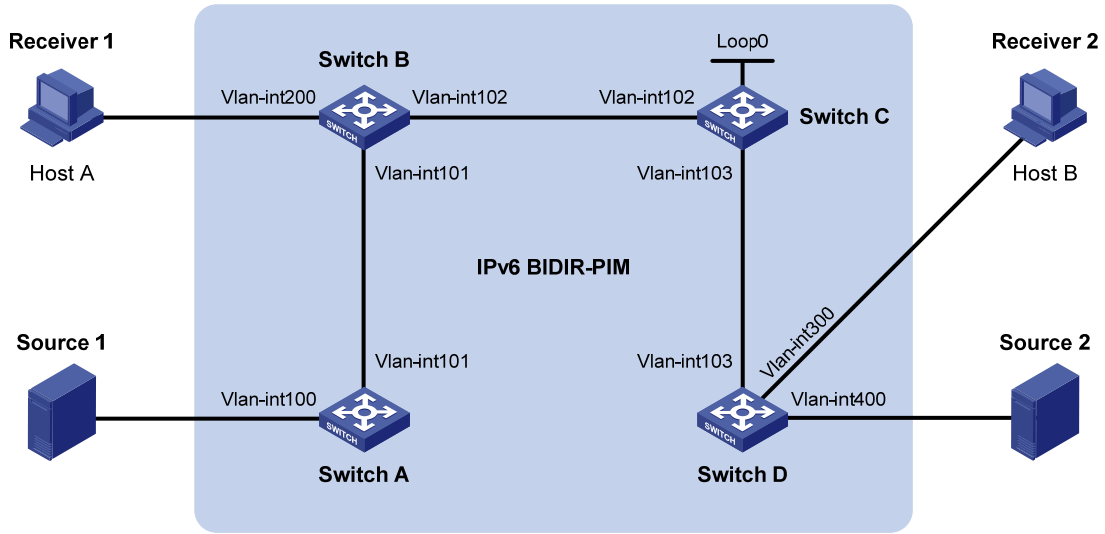
1.8.4 IPv6 双向PIM典型配置举例

1. 组网需求

- 整个 IPv6 PIM 域采用 BIDIR 方式，Source 1 和 Source 2 都向 IPv6 组播组 FF14::101 发送 IPv6 组播信息，Host A 和 Host B 为 IPv6 组播信息的接收者。
- 将 Switch C 的 Vlan-interface102 接口配置为 C-BSR，Loopback0 接口配置为服务于 IPv6 双向 PIM 的 C-RP。
- Switch B 和 Switch D 分别与各自所连接的接收者之间运行 MLDv1。

2. 组网图

图1-17 IPv6 双向 PIM 典型配置组网图



设备	接口	IPv6地址	设备	接口	IPv6地址
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int101	1002::1/64		Vlan-int400	5001::1/64
Switch B	Vlan-int200	2001::1/64		Vlan-int103	3001::2/64
	Vlan-int101	1002::2/64	Source 1	-	1001::2/64
	Vlan-int102	2002::1/64	Source 2	-	5001::2/64
Switch C	Vlan-int102	2002::2/64	Receiver 1	-	2001::2/64
	Vlan-int103	3001::1/64	Receiver 2	-	4001::2/64
	Loop0	6001::1/128			

3. 配置步骤

(1) 配置 IPv6 转发功能，并配置 IPv6 地址和 IPv6 单播路由协议

使能各路由器的IPv6 转发功能，并按照 图 1-17 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

配置 IPv6 双向 PIM 域内的各交换机之间采用 OSPFv3 协议进行互连，确保 IPv6 双向 PIM 域内部在网络层互通，并且各交换机之间能够借助 IPv6 单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IPv6 组播路由，并使能 IPv6PIM-SM、IPv6 双向 PIM 和 MLD

在 Switch A 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并使能 IPv6 双向 PIM。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] pim ipv6
[SwitchA-pim6] bidir-pim enable
[SwitchA-pim6] quit
```

在 Switch B 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，在其连接有接收者的接口 Vlan-interface200 上使能 MLD，并使能 IPv6 双向 PIM。

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] mld enable
[SwitchB-Vlan-interface200] pim ipv6 sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim ipv6 sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim ipv6 sm
[SwitchB-Vlan-interface102] quit
[SwitchB] pim ipv6
[SwitchB-pim6] bidir-pim enable
[SwitchB-pim6] quit
```

在 Switch C 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并使能 IPv6 双向 PIM。

```
<SwitchC> system-view
[SwitchC] multicast ipv6 routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim ipv6 sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] pim ipv6 sm
[SwitchC-Vlan-interface103] quit
[SwitchC] interface loopback 0
[SwitchC-LoopBack0] pim ipv6 sm
[SwitchC-LoopBack0] quit
[SwitchC] pim ipv6
[SwitchC-pim6] bidir-pim enable
```

在 Switch D 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，在其连接有接收者的接口 Vlan-interface300 上使能 MLD，并使能 IPv6 双向 PIM。

```
<SwitchD> system-view
[SwitchD] multicast ipv6 routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] mld enable
[SwitchD-Vlan-interface300] pim ipv6 sm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] pim ipv6 sm
[SwitchD-Vlan-interface400] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim ipv6 sm
[SwitchD-Vlan-interface103] quit
[SwitchD] pim ipv6
[SwitchD-pim6] bidir-pim enable
```

```
[SwitchD-pim6] quit
```

(3) 配置 C-BSR 和 C-RP

在 Switch C 上将接口 Vlan-interface102 配置为 C-BSR，并将接口 Loopback0 配置为服务于 IPv6 双向 PIM 的 C-RP。

```
[SwitchC-pim6] c-bsr 2002::2
[SwitchC-pim6] c-rp 6001::1 bidir
[SwitchC-pim6] quit
```

(4) 检验配置效果

通过使用 **display pim ipv6 df-info** 命令可以查看交换机上 IPv6 双向 PIM 的 DF 信息。例如：

查看 Switch A 上 IPv6 双向 PIM 的 DF 信息。

```
[SwitchA] display pim ipv6 df-info
 RP Address: 6001::1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Vlan100             Win     100        2          01:08:50  FE80::200:5EFF:
                                                              FE71:2800 (local)
  Vlan101             Lose    100        1          01:07:49  FE80::20F:E2FF:
                                                              FE38:4E01
```

查看 Switch B 上 IPv6 双向 PIM 的 DF 信息。

```
[SwitchB] display pim ipv6 df-info
 RP Address: 6001::1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Vlan200             Win     100        1          01:24:09  FE80::200:5EFF:
                                                              FE71:2801 (local)
  Vlan101             Win     100        1          01:24:09  FE80::20F:E2FF:
                                                              FE38:4E01 (local)
  Vlan102             Lose    0          0          01:23:12  FE80::20F:E2FF:
                                                              FE15:5601
```

查看 Switch C 上 IPv6 双向 PIM 的 DF 信息。

```
[SwitchC] display pim ipv6 df-info
 RP Address: 6001::1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Loop0               -       -          -          -         -
  Vlan102             Win     0          0          01:06:07  FE80::20F:E2FF:
                                                              FE15:5601 (local)
  Vlan103             Win     0          0          01:06:07  FE80::20F:E2FF:
                                                              FE15:5602 (local)
```

查看 Switch D 上 IPv6 双向 PIM 的 DF 信息。

```
[SwitchD] display pim ipv6 df-info
 RP Address: 6001::1
  Interface           State   DF-Pref    DF-Metric  DF-Uptime DF-Address
  Vlan300             Win     100        1          01:19:53  FE80::200:5EFF:
                                                              FE71:2803 (local)
  Vlan400             Win     100        1          00:39:34  FE80::200:5EFF:
                                                              FE71:2802 (local)
  Vlan103             Lose    0          0          01:21:40  FE80::20F:E2FF:
                                                              FE15:5602
```

通过使用 **display multicast ipv6 forwarding-table df-info** 命令可以查看交换机上 IPv6 组播转发表的 DF 信息，有关 **display multicast ipv6 forwarding-table df-info** 命令的详细介绍，请参见“IP 组播命令参考”中的“IPv6 组播路由与转发”。例如：

查看 Switch A 上 IPv6 组播转发表的 DF 信息。

```
[SwitchA] display multicast ipv6 forwarding-table df-info
Multicast DF information
Total 1 RP

Total 1 RP matched

00001. RP Address: 6001::1
     MID: 0, Flags: 0x2100000:0
     Uptime: 00:08:32
     RPF interface: Vlan-interface101
     List of 1 DF interfaces:
       1: Vlan-interface100
```

查看 Switch B 上 IPv6 组播转发表的 DF 信息。

```
[SwitchB] display multicast ipv6 forwarding-table df-info
Multicast DF information
Total 1 RP

Total 1 RP matched

00001. RP Address: 6001::1
     MID: 0, Flags: 0x2100000:0
     Uptime: 00:06:24
     RPF interface: Vlan-interface102
     List of 2 DF interfaces:
       1: Vlan-interface101
       2: Vlan-interface200
```

查看 Switch C 上 IPv6 组播转发表的 DF 信息。

```
[SwitchC] display multicast ipv6 forwarding-table df-info
Multicast DF information
Total 1 RP

Total 1 RP matched

00001. RP Address: 6001::1
     MID: 0, Flags: 0x2100000:0
     Uptime: 00:07:21
     RPF interface: LoopBack0
     List of 2 DF interfaces:
       1: Vlan-interface102
       2: Vlan-interface103
```

查看 Switch D 上 IPv6 组播转发表的 DF 信息。

```
[SwitchD] display multicast ipv6 forwarding-table df-info
Multicast DF information
```

```
Total 1 RP

Total 1 RP matched

00001. RP Address: 6001::1
     MID: 0, Flags: 0x2100000:0
     Uptime: 00:05:12
     RPF interface: Vlan-interface103
     List of 2 DF interfaces:
       1: Vlan-interface300
       2: Vlan-interface400
```

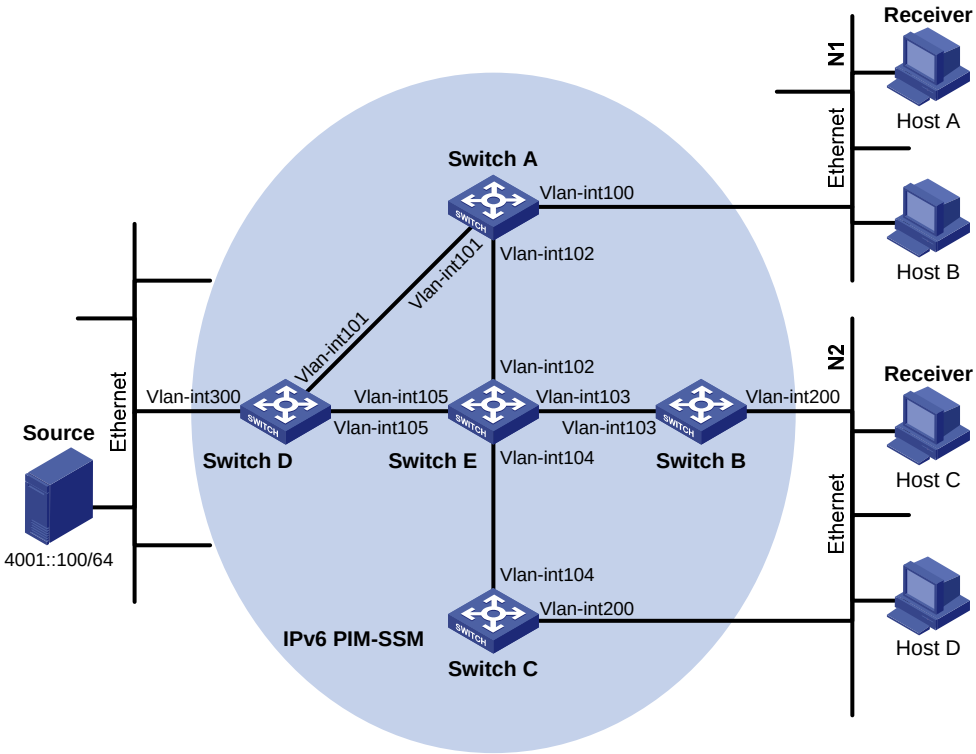
1.8.5 IPv6 PIM-SSM典型配置举例

1. 组网需求

- 接收者通过组播方式接收视频点播信息，不同组织的接收者群体组成末梢网络，每个末梢网络中都存在至少一个接收者，整个 IPv6 PIM 域采用 SSM 方式。
- Host A 和 Host C 为两个末梢网络中的 IPv6 组播信息接收者；Switch D 通过 Vlan-interface300 接口与 IPv6 组播源（Source）所在网络连接；Switch A 通过 Vlan-interface100 接口连接末梢网络 N1，通过 Vlan-interface101 接口和 Vlan-interface102 接口分别连接 Switch D 和 Switch E；Switch B 和 Switch C 通过各自的 Vlan-interface200 接口连接末梢网络 N2，分别通过 Vlan-interface103 和 Vlan-interface104 接口连接 Switch E；Switch E 连接 Switch A、Switch B、Switch C 和 Switch D。
- SSM 组播组的范围是 FF3E::/64。
- Switch A 与末梢网络 N1 之间运行 MLDv2；Switch B 和 Switch C 与末梢网络 N2 之间也运行 MLDv2。

2. 组网图

图1-18 IPv6 PIM-SSM 典型配置组网图



设备	接口	IPv6地址	设备	接口	IPv6地址
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int101	1002::1/64		Vlan-int101	1002::2/64
	Vlan-int102	1003::1/64		Vlan-int105	4002::1/64
Switch B	Vlan-int200	2001::1/64	Switch E	Vlan-int104	3001::2/64
	Vlan-int103	2002::1/64		Vlan-int103	2002::2/64
Switch C	Vlan-int200	2001::2/64		Vlan-int102	1003::2/64
	Vlan-int104	3001::1/64		Vlan-int105	4002::2/64

3. 配置步骤

(1) 使能 IPv6 转发功能，并配置 IPv6 地址和 IPv6 单播路由协议

使能各交换机的IPv6 转发功能，并按照 图 1-18 配置各接口的IPv6 地址和前缀长度，具体配置过程略。

配置 IPv6 PIM-SSM 域内的各交换机之间采用 OSPFv3 协议进行互连，确保 IPv6 PIM-SSM 域内部在网络层互通，并且各交换机之间能够借助 IPv6 单播路由协议实现动态路由更新，具体配置过程略。

(2) 使能 IPv6 组播路由，并使能 IPv6 PIM-SM 和 MLD

在 Switch A 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并在其连接末梢网络的接口 Vlan-interface100 上使能 MLD，且配置其版本为 2。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] mld version 2
```

```
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim ipv6 sm
[SwitchA-Vlan-interface102] quit
```

Switch B 和 Switch C 的配置与 Switch A 相似，Switch D 和 Switch E 除了不需要在相应接口上使能 MLD 外，其它的配置也与 Switch A 相似，配置过程略。

(3) 配置 IPv6 SSM 组播组的地址范围

在 Switch A 上配置 IPv6 SSM 组播组的地址范围为 FF3E::/64。

```
[SwitchA] acl ipv6 number 2000
[SwitchA-acl6-basic-2000] rule permit source ff3e:: 64
[SwitchA-acl6-basic-2000] quit
[SwitchA] pim ipv6
[SwitchA-pim6] ssm-policy 2000
[SwitchA-pim6] quit
```

Switch B、Switch C、Switch D 和 Switch E 的配置与 Switch A 相似，配置过程略。

(4) 检验配置效果

通过使用 **display pim ipv6 interface** 命令可以查看交换机接口上 IPv6 PIM 的配置和运行情况。例如：

查看 Switch A 上 IPv6 PIM 的配置信息。

```
[SwitchA] display pim ipv6 interface
 Interface             NbrCnt HelloInt   DR-Pri   DR-Address
 Vlan100               0      30         1        FE80::A01:201:1
                                                  (local)
 Vlan101               1      30         1        FE80::A01:201:2
 Vlan102               1      30         1        FE80::A01:201:3
```

假如 Host A 需要接收指定 IPv6 组播源 S（4001::100/64）发往 IPv6 组播组 G（FF3E::101）的信息，Switch A 会向 IPv6 组播源方向构造 SPT，SPT 路径中的交换机（Switch A 和 Switch D）上生成（S，G）表项，而 SPT 路径之外的交换机（Switch E）上没有 IPv6 组播路由项，通过使用 **display pim ipv6 routing-table** 命令可以查看交换机的 IPv6 PIM 路由表信息。例如：

查看 Switch A 上的 IPv6 PIM 路由表信息。

```
[SwitchA] display pim ipv6 routing-table
 Total 0 (*, G) entry; 1 (S, G) entry

 (4001::100, FF3E::101)
     Protocol: pim-ssm, Flag:
     UpTime: 00:00:11
     Upstream interface: Vlan-interface101
         Upstream neighbor: 1002::2
         RPF prime neighbor: 1002::2
     Downstream interface(s) information:
     Total number of downstreams: 1
```

```
        1: Vlan-interface100
            Protocol: mld, UpTime: 00:00:11, Expires: 00:03:25
```

查看 Switch D 上的 IPv6 PIM 路由表信息。

```
[SwitchD] display pim ipv6 routing-table
 Total 0 (*, G) entry; 1 (S, G) entry

 (4001::100, FF3E::101)
     Protocol: pim-ssm, Flag: LOC
     UpTime: 00:08:02
     Upstream interface: Vlan-interface300
         Upstream neighbor: NULL
         RPF prime neighbor: NULL
     Downstream interface(s) information:
     Total number of downstreams: 1
         1: Vlan-interface105
             Protocol: pim-ssm, UpTime: 00:08:02, Expires: 00:03:25
```

1.9 常见配置错误举例

1.9.1 无法正确建立组播分发树

1. 故障现象

网络中各路由器（包括直连 IPv6 组播源或接收者的路由器）上都没有 IPv6 组播转发项，也就是说无法正确建立组播分发树，客户端无法接收 IPv6 组播数据。

2. 分析

- 无论运行哪个 IPv6 PIM 模式，IPv6 单播路由是创建 IPv6 PIM 路由项的基础。只有单播运行顺畅，组播才能运行成功。
- IPv6 PIM 协议需要 RPF 接口支持 IPv6 PIM 协议。RPF 邻居也必须是 IPv6 PIM 邻居。如果 RPF 接口或 RPF 邻居没有使能 IPv6 PIM，组播分发树就不能正确建立，导致组播数据转发异常。
- IPv6 PIM 协议需要整个网络运行相同的 IPv6 PIM 模式，即 DM 或 SM。否则，组播分发树不能正确建立，导致组播数据转发异常。

3. 处理过程

(1) 检查 IPv6 单播路由。使用命令 **display ipv6 routing-table** 命令检查是否有到达 IPv6 组播源或 RP 的 IPv6 单播路由项。

(2) 检查 RPF 接口上是否使能 IPv6 PIM。通过命令 **display pim ipv6 interface** 命令查看接口的 IPv6 PIM 信息。若接口上没有使能 IPv6 PIM，请使用 **pim ipv6 dm** 或 **pim ipv6 sm** 命令使能 IPv6 PIM。

(3) 检查 RPF 邻居是否是 IPv6 PIM 邻居。通过命令 **display pim ipv6 neighbor** 查看 IPv6 PIM 邻居信息。

(4) 检查直连 IPv6 组播源、直连接收者的路由器接口上是否使能了 IPv6 PIM 和 MLD。

(5) 检查 IPv6 PIM 模式是否一致。通过命令 **display pim ipv6 interface verbose** 检查 RPF 接口和 RPF 邻居所在路由器的对应接口上是否使能了相同模式的 IPv6 PIM 协议。

(6) 检查整个网络中各路由器上的 IPv6 PIM 模式是否一致。通过命令 **display current-configuration** 查看接口上的 IPv6 PIM 模式信息，确保所有路由器配置相同模式的 IPv6 PIM 协议，即要么全部配置为 **pim ipv6 sm**，要么全部配置为 **pim ipv6 dm**。

1.9.2 IPv6 组播数据异常终止在中间路由器

1. 故障现象

IPv6 组播数据可以到达中间路由器，但无法到达最后一跳路由器。中间路由器某接口上收到 IPv6 组播数据，但 IPv6 PIM 路由表中没有创建相应的（S，G）表项。

2. 分析

- 命令 **multicast ipv6 boundary** 用于在接口上设置 IPv6 组播转发边界，如果 IPv6 组播数据无法通过该边界，IPv6 PIM 是无法创建路由项的。
- 此外，**source-policy** 命令用于过滤接收到的 IPv6 组播数据报文。如果 IPv6 组播数据报文无法通过该命令的 IPv6 ACL 规则，IPv6 PIM 也是无法创建路由项的。

3. 处理过程

(1) 检查 IPv6 组播转发边界的配置。通过命令 **display current-configuration** 查看 IPv6 组播转发边界上的设置，使用 **multicast ipv6 boundary** 命令更改 IPv6 组播转发边界的设置，使 IPv6 组播数据能够通过该边界。

(2) 检查 IPv6 组播过滤器配置。通过命令 **display current-configuration** 查看 IPv6 组播过滤器的配置，更改 **source-policy** 命令的 IPv6 ACL 规则，使 IPv6 组播数据的源/组地址通过 IPv6 ACL 过滤。

1.9.3 IPv6 PIM-SM中RP无法加入SPT

1. 故障现象

共享树不能正确建立，或者 RP 不能加入到达 IPv6 组播源的源树。

2. 分析

- RP 是 IPv6 PIM-SM 网络的核心，为指定 IPv6 组播组服务。多个 RP 可以在网络中共存。请确保所有路由器上的 RP 信息是一致的；同一个指定组是被映射到相同的 RP。否则，IPv6 组播异常。
- 对于静态 RP 机制，整个网络中的所有路由器包括静态 RP 本身必须用静态 RP 命令配置相同的 RP。否则，IPv6 组播异常。

3. 处理过程

(1) 检查是否有到达 RP 的路由。通过命令 **display ipv6 routing-table** 查看各路由器上是否有到达 RP 的路由。

(2) 检查动态 RP 信息。通过命令 **display pim ipv6 rp-info** 查看各路由器上的 RP 信息是否一致。如果不一致，请配置成一致。

(3) 检查静态 RP 的配置。通过命令 **display pim ipv6 rp-info** 查看全网所有路由器上是否配置了相同的静态 RP。

1.9.4 IPv6 PIM-SM中无法建立RPT或无法进行源注册

1. 故障现象

C-RP 无法向 BSR 单播通告报文，BSR 没有发布包含 C-RP 的自举报文，BSR 上没有到达各 C-RP 的单播路由，共享树无法正确建立，或者 DR 无法向 RP 进行源注册。

2. 分析

- C-RP 周期性用单播模式向 BSR 通告报文。如果 C-RP 没有到达 BSR 的路由，BSR 就不能接收 C-RP 发来的通告报文，也就不会通告包含该 C-RP 的自举消息。
- RP 是 IPv6 PIM-SM 网络的核心。请确保所有路由器上的 RP 信息是一致的，指定 IPv6 组播组 G 被映射到相同的 RP，到 RP 的单播路由是可达的。

3. 处理过程

(1) 检查到各 C-RP、RP 和 BSR 的单播路由是否可达。通过命令 **display ipv6 routing-table** 查看各路由器上是否有到达 RP 和 BSR 的路由，及 C-RP 和 BSR 之间的路由。请确保各 C-RP 上存在到达 BSR 的单播路由，BSR 上存在到达各 C-RP 的单播路由，整个网络中所有路由器上存在到达 RP 的单播路由。

(2) 检查RP和BSR信息。IPv6 PIM-SM协议需要RP和BSR的支持，首先使用命令**display pim ipv6 bsr-info** 查看各路由器上是否有 BSR 信息，使用 **display pim ipv6 rp-info** 命令查看各路由器上 RP 信息是否正确。

(3) 检查 IPv6 PIM 邻居关系。通过命令 **display pim ipv6 neighbor** 查看各路由器是否建立了正常的邻居关系。

目　录

1 IPv6 MBGP配置

说明

在以下路由协议的介绍中所指的路由器及路由器图标，代表了一般意义下的路由器以及运行了路由协议的以太网交换机。

1.1 IPv6 MBGP简介

为了提供对多种网络层协议的支持，IETF（Internet Engineering Task Force，互联网工程任务组）对 BGP-4 进行了扩展，形成 MP-BGP（Multiprotocol Border Gateway Protocol，多协议边界网关协议），使 BGP 能够为多种路由应用提供路由信息。

由于 IPv6 组播拓扑和 IPv6 单播拓扑有可能不同，需要通过 MP-BGP 扩展使得 BGP 能够将 IPv6 单播 NLRI（Network Layer Reachability Information，网络层可达性信息）和 IPv6 组播 NLRI 分开运载，其中组播 NLRI 专用于执行 RPF（Reverse Path Forwarding，逆向路径转发）功能。这使得在 IPv6 单播路由表和组播路由表中，对于相同的前缀有不同的路径选择，维护了一致的单播转发并使域间组播正常工作。

目前的 MP-BGP 标准是 RFC 2858（Multiprotocol Extensions for BGP-4，BGP-4 的多协议扩展）。MP-BGP 在 IPv6 组播上的应用简称为 IPv6 MBGP（IPv6 Multicast BGP，IPv6 组播 BGP）。

说明

- 本章主要介绍 MP-BGP 应用于 IPv6 组播的配置事项，即 IPv6 MBGP 配置事项，有关 BGP 的详细内容以及配置事项，请参见“三层技术-IP 路由配置指导”中的“BGP”。
- 有关 RPF 检查的详细内容以及配置事项，请参见“IP 组播配置参考”中的“组播路由与转发”。

1.2 IPv6 MBGP配置任务简介

表1-1 IPv6 MBGP 配置任务简介

配置任务		说明	详细配置
配置IPv6 MBGP基本功能	配置IPv6 MBGP对等体	必选	1.3.2
	配置路由首选值	可选	1.3.3
控制路由信息的发布与接收	配置IPv6 MBGP发布本地IPv6路由	可选	1.4.2
	配置IPv6 MBGP引入其他路由	可选	1.4.3
	配置IPv6 MBGP路由聚合	可选	1.4.4

配置任务		说明	详细配置
	配置向MBGP对等体/对等体组发送缺省路由	可选	1.4.5
	配置IPv6 MBGP路由信息的发布策略	可选	1.4.6
	配置IPv6 MBGP路由信息的接收策略	可选	1.4.7
	配置IPv6 MBGP路由衰减	可选	1.4.8
配置IPv6 MBGP的路由属性	配置IPv6 MBGP路由管理优先级	可选	1.5.2
	配置本地优先级缺省值	可选	1.5.3
	配置MED属性	可选	1.5.4
	配置发布路由时将自身地址作为下一跳	可选	1.5.5
	配置AS_PATH属性	可选	1.5.6
调整和优化IPv6 MBGP网络	配置IPv6 MBGP软复位	可选	1.6.2
	配置IPv6 MBGP ORF能力	可选	1.6.3
	配置最大等价路由的条数	可选	1.6.4
组建大型IPv6 MBGP网络	配置IPv6 MBGP对等体组	可选	1.7.2
	配置IPv6 MBGP团体	可选	1.7.3
	配置IPv6 MBGP路由反射器	可选	1.7.4

1.3 配置IPv6 MBGP的基本功能

1.3.1 配置准备

由于 IPv6 MBGP 是 MP-BGP 的一种应用，因此在配置 IPv6 MBGP 之前，需完成以下任务：

- 使能 IPv6 功能
- 配置接口的网络层地址
- 配置 BGP 基本功能

1.3.2 配置IPv6 MBGP对等体

表1-2 配置 IPv6 MBGP 对等体

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
启动BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统没有运行BGP
进入IPv6地址族视图	**ipv6-family**	-
配置对等体的AS号	**peer** *ipv6-address* **as-number** *as-number*	必选 缺省情况下，对等体无AS号
创建并进入IPv6 MBGP地址族视图	**ipv6-family multicast**	必选
激活指定对等体	**peer** *ipv6-address* **enable**	必选 缺省情况下，对等体在IPv6 MBGP地址族视图下处于非激活状态

1.3.3 配置路由首选值

表1-3 配置路由首选值

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
为从IPv6 MBGP对等体/对等体组接收的路由分配首选值	**peer** { *ipv6-group-name* | *ipv6-address* } **preferred-value** *value*	可选 缺省情况下，从IPv6 MBGP对等体/对等体组接收的路由的首选值为0

注意

- 如果同时通过路由策略和 **peer preferred-value** 命令为从 IPv6 MBGP 对等体/对等体组接收的 IPv6 MBGP 路由设置首选值，通过路由策略过滤的 IPv6 MBGP 路由的首选值将取路由策略中设置的首选值而不是 **peer preferred-value** 命令设置的首选值;
- 只有当路由策略里设置的首选值为 0 时，IPv6 MBGP 路由的首选值才会取 **peer preferred-value** 命令设置的值;
- 没有通过路由策略过滤的 IPv6 组播路由的首选值仍取 **peer preferred-value** 命令设置的首选值。
- 通过路由策略配置 BGP IPv6 组播路由信息首选值的相关配置可参考命令 **peer route-policy** 和“三层技术-IP 路由命令参考/路由策略”中的 **apply preferred-value**。

1.4 控制路由信息的发布与接收

控制 IPv6 MBGP 的路由信息的发布和接收，包括对路由信息进行过滤、应用路由策略和路由衰减。

1.4.1 配置准备

在控制路由信息的发布与接收之前，需完成以下任务：配置 IPv6 MBGP 基本功能。

1.4.2 配置IPv6 MBGP发布本地IPv6 路由

表1-4 配置 IPv6 MBGP 发布本地 IPv6 路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
将路由通告到IPv6 MBGP路由表中	**network** *ipv6-address prefix-length* [**route-policy** *route-policy-name* | **short-cut**]	必选 缺省情况下，IPv6 MBGP不通告任何路由

1.4.3 配置IPv6 MBGP引入其他路由

表1-5 配置 IPv6 MBGP 引入其他路由

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
允许将缺省路由引入到IPv6 MBGP路由表中	**default-route imported**	可选 缺省情况下，IPv6 MBGP不允许将缺省路由引入到IPv6 MBGP路由表中
配置引入其它协议路由信息并通告	**import-route** *protocol* [*process-id* [**med** *med-value* | **route-policy** *route-policy-name*] *]	必选 缺省情况下，IPv6 MBGP不引入且不通告其它协议的路由

说明

通过 **import-route** 命令引入 IGP 路由时，不能引入 IGP 的缺省路由，必须通过配置 **default-route imported** 命令将缺省路由引入到 IPv6 MBGP 路由表中。

1.4.4 配置IPv6 MBGP路由聚合

在中型或大型 BGP 网络中，在向对等体发布 IPv6 组播路由信息时，可以配置路由聚合，减小对等体路由表中的路由数量。BGP 仅支持 IPv6 组播路由的手动聚合。

表1-6 配置路由聚合

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置手动路由聚合	**aggregate** *ipv6-address prefix-length* [**as-set** | **attribute-policy** *route-policy-name* | **detail-suppressed** | **origin-policy** *route-policy-name* | **suppress-policy** *route-policy-name*] *	必选 缺省情况下，不进行路由聚合

1.4.5 配置向IPv6 MBGP对等体/对等体组发送缺省路由

表1-7 配置向 IPv6 MBGP 对等体/对等体组发送缺省路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置向IPv6 MBGP对等体/对等体组发送缺省路由	**peer** { *ipv6-group-name* | *ipv6-address* } **default-route-advertise** [**route-policy** *route-policy-name*]	必选 缺省情况下，不向IPv6 MBGP对等体/对等体组发布缺省路由

说明

执行 **peer default-route-advertise** 命令后，不论本地路由表中是否存在缺省路由，都将向指定IPv6 MBGP 对等体/对等体组发布一条下一跳地址为本地地址的缺省路由。

1.4.6 配置IPv6 MBGP路由信息的发布策略

表1-8 配置 IPv6 MBGP 路由信息的发布策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
对发布的路由信息进行过滤	**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* } **export** [*protocol process-id*]	任选其一 缺省情况下，不对发布的路由信息进行过滤 可以根据需求选择过滤策略，同时配置几种过滤策略时，按照如下顺序执行： • **filter-policy export** • **peer filter-policy export**
配置基于IPv6 ACL的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **filter-policy** *acl6-number* **export**	
配置基于AS路径过滤列表的IPv6 BGP路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **as-path-acl** *as-path-acl-number* **export**	

操作	命令	说明
配置基于IPv6地址前缀列表的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **ipv6-prefix** *ipv6-prefix-name* **export**	• **peer as-path-acl export** • **peer ipv6-prefix export** • **peer route-policy export** 只有通过前面的过滤策略，才能继续执行后面的过滤策略；只有通过所有配置的过滤策略后，路由信息才能被发布
配置出方向路由策略	**peer** { *ipv6-group-name* | *ipv6-address* } **route-policy** *route-policy-name* **export**	

说明

- IPv6 MBGP 对等体组的成员必须与所在的组使用相同的出方向路由更新策略，即对外发布路由时，一个对等体组遵循的策略是相同的。
- IPv6 MBGP 对引入的路由信息进行过滤后，会将符合条件的路由信息发布给 IPv6 MBGP 对等体。

1.4.7 配置IPv6 MBGP路由信息的接收策略

对 IPv6 MBGP 接收的路由进行过滤，只有满足某些条件的路由才能被 IPv6 MBGP 接收，并加到路由表中。

表1-9 配置 IPv6 MBGP 路由信息的接收策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
对接收的路由信息进行过滤	**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* } **import**	任选其一 缺省情况下，不对接收的路由信息进行过滤 可以根据需求选择过滤策略，同时配置几种过滤策略时，按照如下顺序执行： • **filter-policy import**
对接收的路由信息应用路由策略	**peer** { *ipv6-group-name* | *ipv6-address* } **route-policy** *route-policy-name* **import**	

操作	命令	说明
配置基于IPv6 ACL的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **filter-policy** *acl6-number* **import**	• **peer filter-policy import** • **peer as-path-acl import** • **peer ip-prefix import** **peer route-policy import**只有通过前面的过滤策略，才能继续执行后面的过滤策略；只有通过所有配置的过滤策略后，路由信息才能被接收
配置基于AS路径过滤列表的IPv6 BGP路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **as-path-acl** *as-path-acl-number* **import**	
配置基于IPv6地址前缀列表的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **ipv6-prefix** *ipv6-prefix-name* **import**	
配置允许从对等体/对等体组接收的最大IPv6地址前缀数	**peer** { *ipv6-group-name* | *ipv6-address* } **route-limit** *limit* [*percentage*]	可选 缺省情况下，允许从组播对等体/对等体组接收的最大路由数无限制

说明

IPv6 MBGP 对等体组的成员可以与所在的组使用不同的入方向路由策略，即接收路由时，各对等体可以选择自己的策略。

1.4.8 配置IPv6 MBGP路由衰减

表1-10 配置 IPv6 MBGP 路由衰减

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置IPv6 MBGP路由衰减	**dampening** [*half-life-reachable half-life-unreachable reuse suppress ceiling* | **route-policy** *route-policy-name*] *	可选 缺省情况下，没有配置IPv6 MBGP路由衰减

1.5 配置IPv6 MBGP的路由属性

本节主要介绍使用各类路由属性来改变 IPv6 MBGP 的选路策略。包括如下属性：

- IPv6 MBGP 路由管理优先级
- 缺省 LOCAL_PREF 属性值
- MED 属性
- NEXT_HOP 属性
- AS_PATH 属性

1.5.1 配置准备

在配置 IPv6 MBGP 的路由属性之前，需完成以下任务：配置 IPv6 MBGP 基本功能。

1.5.2 配置IPv6 MBGP路由管理优先级

表1-11 配置 IPv6 MBGP 路由管理优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置IPv6 MBGP路由管理优先级	**preference** { *external-preference internal-preference local-preference* | **route-policy** *route-policy-name* }	可选 缺省情况下，EBGP路由的管理优先级为255，IBGP路由的管理优先级为255，本地产生的IPv6 BGP路由的管理优先级为130

1.5.3 配置本地优先级缺省值

表1-12 配置本地优先级缺省值

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置本地优先级缺省值	**default local-preference** *value*	可选 缺省情况下，本地优先级缺省值为100

1.5.4 配置MED属性

表1-13 配置 MED 属性

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置系统的缺省MED值	**default med** *med-value*	可选 缺省情况下，系统缺省的MED值为0
配置允许比较来自不同AS邻居的路由的MED属性值	**compare-different-as-med**	可选 缺省情况下，不允许比较来自不同AS邻居的路由的MED属性值
配置根据路由来自的AS进行分组对MED排序优选	**bestroute compare-med**	可选 缺省情况下，不根据路由来自的AS进行分组对MED排序优选
配置允许比较联盟对等体的路由按MED值进行优选	**bestroute med-confederation**	可选 缺省情况下，比较联盟对等体的路由时不考虑MED值

1.5.5 配置发布路由时将自身地址作为下一跳

在某些组网环境中，为保证 IPv6 组播 IBGP 邻居能够找到正确的下一跳，可以配置在向 IPv6 组播 IBGP 对等体/对等体组发布路由时，改变下一跳地址为自身地址。如果配置了 IPv6 MBGP 负载分担，则不论是否配置了 **peer next-hop-local** 命令，本地路由器向 IPv6 组播 IBGP 对等体/对等体组发布路由时都先将下一跳地址改变为自身地址。

在第三方下一跳（即两个 IPv6 MBGP 连接在同一网段的广播网）这种特殊的组网环境中，缺省情况下，向 IPv6 组播 EBGP 对等体/对等体组发布路由时，不将自身地址作为下一跳。

表1-14 配置发布路由时将自身地址做为下一跳

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-

操作	命令	说明
配置发布路由时将自身地址作为下一跳	**peer** { *ipv6-group-name* | *ipv6-address* } **next-hop-local**	可选 缺省情况下，向EBGP对等体/对等体组发布路由时，将自身地址作为下一跳；向IBGP对等体/对等体发布路由时，不将自身地址作为下一跳

1.5.6 配置AS_PATH属性

表1-15 配置 AS_PATH 属性

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置允许本地AS号在所接收的路由的AS_PATH属性中出现，并可同时配置允许重复的次数	**peer** { *ipv6-group-name* | *ipv6-address* } **allow-as-loop** [*number*]	可选 缺省情况下，不允许本地AS号重复
禁止路由器将AS_PATH当作选路算法中的一个因素	**bestroute as-path-neglect**	可选 缺省情况下，路由器可以将AS_PATH当作选路算法中的一个因素
配置发送IPv6 MBGP更新报文时不携带私有AS编号	**peer** { *ipv6-group-name* | *ipv6-address* } **public-as-only**	可选 缺省情况下，发送IPv6 MBGP更新报文时携带私有自治系统号

1.6 调整和优化IPv6 MBGP网络

1.6.1 配置准备

在调整和优化 IPv6 MBGP 网络之前，需完成以下任务：

- 使能 IPv6 能力
- 配置 IPv6 MBGP 基本功能

1.6.2 配置IPv6 MBGP软复位

IPv6 MBGP 的选路策略改变后，为了使新的策略生效，必须复位 IPv6 MBGP 连接，但这样会造成短暂的 IPv6 MBGP 连接中断。

通过使能 Route-refresh 功能，当策略改变后，系统可以在不中断 IPv6 MBGP 连接的情况下，自动对 IPv6 MBGP 路由表进行动态刷新。

如果对等体不支持 Route-Refresh 功能，则可以将从对等体接收的所有路由更新保存在本地，当选路策略发生改变后，在不中断连接的情况下重新刷新 IPv6 MBGP 路由表，并应用新的策略。

1. 通过Route-Refresh实现IPv6 MBGP软复位

在对等体支持并使能 Route-Refresh 功能的情况下，如果 IPv6 MBGP 的路由策略发生了变化，本地路由器会向 IPv6 MBGP 对等体发布 Route-Refresh 消息，收到此消息的对等体会将其路由信息重新发给本地路由器。这样，在不中断 IPv6 MBGP 连接的情况下，就可以对 IPv6 MBGP 路由表进行动态更新，并应用新的选路策略。

表1-16 通过 Route-Refresh 实现 IPv6 MBGP 软复位

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6地址族视图	**ipv6-family**	-
使能IPv6 BGP路由刷新功能	**peer** { *ipv6-group-name* | *ipv6-address* } **capability-advertise route-refresh**	可选 缺省情况下，路由刷新功能处于使能状态

2. 通过将所有路由更新保存在本地实现IPv6 MBGP软复位

当对等体不支持 Route-Refresh 功能时，可通过配置 **peer keep-all-routes** 命令实现软复位功能。用户也可以通过执行 **refresh bgp ipv6 multicast** 命令对保存在本地的所有路由重新过一遍策略。

表1-17 通过将所有路由更新保存在本地实现 IPv6 MBGP 软复位

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6地址族视图	**ipv6-family**	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-

操作	命令	说明
保存所有来自对等体/对等体组的原始路由信息，即使这些路由没有通过已配置的入口策略	**peer** { *ipv6-group-name* | *ipv6-address* } **keep-all-routes**	必选 缺省情况下，不保存对等体/对等体组的原始路由信息
手工对IPv6 MBGP连接进行软复位	**refresh bgp ipv6 multicast** { **all** | *ipv6-address* | **group** *ipv6-group-name* | **external** | **internal** } { **export** | **import** }	可选

1.6.3 使能IPv6 MBGP ORF能力功能

BGP ORF 特性是将本地入口策略通过 Route-refresh 报文发送给邻居，当邻居需要向 BGP 对等体发送 Update 更新报文时，通过本地的路由策略后还需要进行 ORF 策略的过滤，只有通过 ORF 策略的路由信息才会发给 BGP 对等体，以达到减少 BGP 邻居间 Update 更新报文的交互，节省网络资源的目的。

使能 BGP ORF 能力后，本地和 BGP 对等体会通过 Open 报文协商 ORF 能力（即收发的报文里是否允许携带 ORF 信息，如果允许携带，是否可以携带非标准的 ORF 信息），当协商完毕并成功建立邻居关系后，可以通过特殊的 Route-refresh 报文交互 ORF 信息。

ORF能力协商成功需要两端的配置来保证，关于两端参数的选择请参见 表 1-19。

表1-18 配置 IPv6 MBGP ORF 能力

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
使能BGP路由刷新功能	**peer** { *group-name* | *ipv6-address* } **capability-advertise route-refresh**	可选 缺省情况下，BGP路由刷新功能处于使能状态 如果该功能当前处于未使能状态，则必须配置该命令
使能BGP ORF非标准功能	**peer** { *group-name* | *ipv6-address* } **capability-advertise orf non-standard**	可选 缺省情况下，BGP ORF能力支持RFC5291、RFC5292的标准能力 如果该功能当前处于未使能状态，则必须配置该命令
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-

操作	命令	说明
使能BGP ORF地址前缀能力协商功能	**peer** { *group-name* | *ipv6-address* } **capability-advertise orf ip-prefix** { **both** | **receive** | **send** }	必选 缺省情况下，BGP不支持ORF 地址前缀的能力协商

表1-19 both、send、receive 参数选择以及配置效果描述表

本地选择参数	对端选择参数	协商成功后
send	**receive**	本端的ORF发送能力，对端的ORF接收能力
	both	
receive	**send**	本端的ORF接收能力，对端的ORF发送能力
	both	
both	**both**	双向的ORF发送和接收能力

1.6.4 配置最大等价路由的条数

表1-20 配置最大等价路由的条数

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
配置进行IPv6 MBGP负载分担的路由条数	**balance** *number*	必选 缺省情况下，不进行负载分担

1.7 配置大型IPv6 MBGP网络

1.7.1 配置准备

在配置 IPv6 MBGP 对等体组之前，需完成以下任务：配置 IPv6 MBGP 基本功能。

1.7.2 配置IPv6 MBGP对等体组

为方便管理，减少重复配置操作，管理员将一些享有相同更新策略的 IPv6 BGP 对等体划分到一个逻辑组织，这些组织就称为对等体组。只须对一个对等体组配置策略，该策略对组内所有成员都有效。

表1-21 配置 IPv6 MBGP 对等体组

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	-
进入IPv6地址族视图	**ipv6-family**	-
创建BGP对等体组	**group** *ipv6-group-name* [**external** | **internal**]	必选
将对等体加入已存在的对等体组	**peer** *ipv6-address* **group** *ipv6-group-name* [**as-number** *as-number*]	必选 缺省情况下，对等体不属于任何对等体组
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-
使能IPv6单播对等体组	**peer** *ipv6-group-name* **enable**	必选
在已使能的对等体组中加入使能的IPv6组播对等体	**peer** *ipv6-address* **group** *ipv6-group-name*	必选 缺省情况下，对等体不属于任何对等体组

注意

配置 IPv6 MBGP 对等体组:

- 需要先在 IPv6 地址族视图下将对等体加入对等体组;
- 然后在 BGP IPv6 组播地址族视图下使能 IPv6 MBGP 对等体组，并把 IPv6 MBGP 对等体加入 IPv6 MBGP 对等体组。

1.7.3 配置IPv6 MBGP团体

对等体组可以使一组对等体共享相同的策略，而利用团体可以使多个 AS 中的一组 IPv6 MBGP 路由器共享相同的策略。团体是一个路由属性，在 IPv6 MBGP 对等体之间传播，它并不受到 AS 范围的限制。

BGP 路由器在将带有团体属性的路由发布给其它对等体之前，可以改变此路由原有的团体属性。除了使用公认的团体属性外，用户还可以使用团体属性列表自定义扩展团体属性，以便更为灵活地控制路由策略。

表1-22 配置 IPv6 MBGP 团体

操作		命令	说明
进入系统视图		**system-view**	-
启动BGP，进入BGP视图		**bgp** *as-number*	-
进入IPv6 MBGP地址族视图		**ipv6-family multicast**	-
配置向IPv6 MBGP对等体/IPv6 MBGP对等体组发布团体属性	配置发布团体属性	**peer** { *ipv6-group-name* | *ipv6-address* } **advertise-community**	必选 缺省情况下，不将团体属性发布给任何对等体/对等体组
	配置发布扩展团体属性	**peer** { *ipv6-group-name* | *ipv6-address* } **advertise-ext-community**	必选 缺省情况下，不将扩展团体属性发布给任何对等体/对等体组
对发布给MBGP对等体/MBGP对等体组的路由指定路由策略		**peer** { *ipv6-group-name* | *ipv6-address* } **route-policy** *route-policy-name* **export**	必选 缺省情况下，不指定对等体/对等体组的路由策略

说明

- 配置 IPv6 MBGP 团体时，必须使用路由策略来定义具体的团体属性，然后在发布路由信息时应用此路由策略。
- 关于路由策略的配置，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

1.7.4 配置IPv6 MBGP路由反射器

为保证 IBGP 对等体之间的连通性，需要在 IPv6 组播 IBGP 对等体之间建立全连接关系。当 IPv6 组播 IBGP 对等体数目很多时，建立全连接网的开销很大。路由反射器可以解决这个问题。

表1-23 配置 IPv6 MBGP 路由反射器

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6 MBGP地址族视图	**ipv6-family multicast**	-

操作	命令	说明
配置将本机作为路由反射器，并将IPv6 MBGP对等体/对等体组作为路由反射器的客户	**peer** { *ipv6-group-name* | *ipv6-address* } **reflect-client**	必选 缺省情况下，没有配置路由反射器及其客户
使能客户机之间的路由反射	**reflect between-clients**	可选 缺省情况下，允许客户到客户的路由反射
配置路由反射器的集群ID	**reflector cluster-id** *cluster-id*	可选 缺省情况下，每个路由反射器是使用自己的Router ID作为集群ID

说明

- 通常情况下，路由反射器的客户之间不要求是全连接的，路由通过反射器从一个客户反射到其它客户；如果客户之间是全连接的，可以禁止客户间的反射，以便减少开销。
- 当一个集群里有多个路由反射器时，需要给所有位于同一个集群内的路由反射器配置相同的*cluster-id*，以避免路由循环。

1.8 IPv6 MBGP显示和维护

1.8.1 IPv6 MBGP显示

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IPv6 MBGP 的运行情况，通过查看显示信息验证配置的效果。

表1-24 IPv6 MBGP 显示

操作	命令
显示对等体组信息	**display bgp ipv6 multicasst group** [*ipv6-group-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 MBGP发布的IPv6路由信息	**display bgp ipv6 multicast network** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示AS路径信息	**display bgp ipv6 multicast paths** [*as-regular-expression* | **|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 MBGP对等体/对等体组信息	**display bgp ipv6 multicast peer** [[*ipv6-address*] **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示收到的邻居ORF信息中的前缀信息	**display bgp ipv6 multicast peer** *ipv6-address* **received ipv6-prefix** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 MBGP路由信息	**display bgp ipv6 multicast routing-table** [*ipv6-address prefix-length*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定AS路径过滤列表的路由	**display bgp ipv6 multicast routing-table as-path-acl** *as-path-acl-number* [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定IPv6 MBGP团体的路由信息	**display bgp ipv6 multicast routing-table community** [*aa:nn*<1-13>] [**no-advertise** | **no-export** | **no-export-subconfed**]* [**whole-match**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定IPv6 MBGP团体列表的路由信息	**display bgp ipv6 multicast routing-table community-list** { { *basic-community-list-number* | *comm-list-name* } [**whole-match**] | *adv-community-list-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 MBGP衰减的路由	**display bgp ipv6 multicast routing-table dampened** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 MBGP路由衰减参数	**display bgp ipv6 multicast routing-table dampening parameter** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示来自不同自治系统的IPv6 MBGP路由	**display bgp ipv6 multicast routing-table different-origin-as** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 MBGP路由振荡统计信息	**display bgp ipv6 multicast routing-table flap-info** [**regular-expression** *as-regular-expression* | [**as-path-acl** *as-path-acl-number* | *ipv6-address prefix-length* [**longer-match**]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]]
显示向指定的IPv6 MBGP对等体发送或者从指定的IPv6 BGP对等体收到的路由信息	**display bgp ipv6 multicast routing-table peer** *ipv6-address* { **advertised-routes** | **received-routes** } [*network-address prefix-length* | **statistic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配AS正则表达式的IPv6组播路由信息	**display bgp ipv6 multicast routing-table regular-expression** *as-regular-expression*
显示IPv6 MBGP的路由统计信息	**display bgp ipv6 multicast routing-table statistic** [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示IPv6 MBGP路由表中的路由信息	**display ipv6 multicast routing-table** [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定目的地址的组播路由信息	**display ipv6 multicast routing-table** *ipv6-address prefix-length* [**longer-match**] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.8.2 复位IPv6 MBGP连接

当 MBGP 路由策略或协议发生变化后，如果需要通过复位 IPv6 MBGP 连接使新的配置生效，请在用户视图下进行下列配置。

表1-25 复位 IPv6 MBGP 连接

操作	命令
复位指定的IPv6 MBGP连接	**reset bgp ipv6 multicast** { *as-number* | *ipv6-address* [**flap-info**] | **all** | **group** *ipv6-group-name* | **external** | **internal** }

1.8.3 清除IPv6 MBGP信息

在用户视图下执行 **reset** 命令可清除 IPv6 MBGP 路由的统计信息。

表1-26 清除 IPv6 MBGP 信息

操作	命令
清除IPv6 MBGP路由的衰减信息并释放被抑制的路由	**reset bgp ipv6 multicast dampening** [*ipv6-address prefix-length*]
清除IPv6 MBGP路由的振荡统计信息	**reset bgp ipv6 multicast flap-info** [*ipv6-address/prefix-length* | **regexp** *as-path-regexp* | **as-path-acl** *as-path-acl-number*]

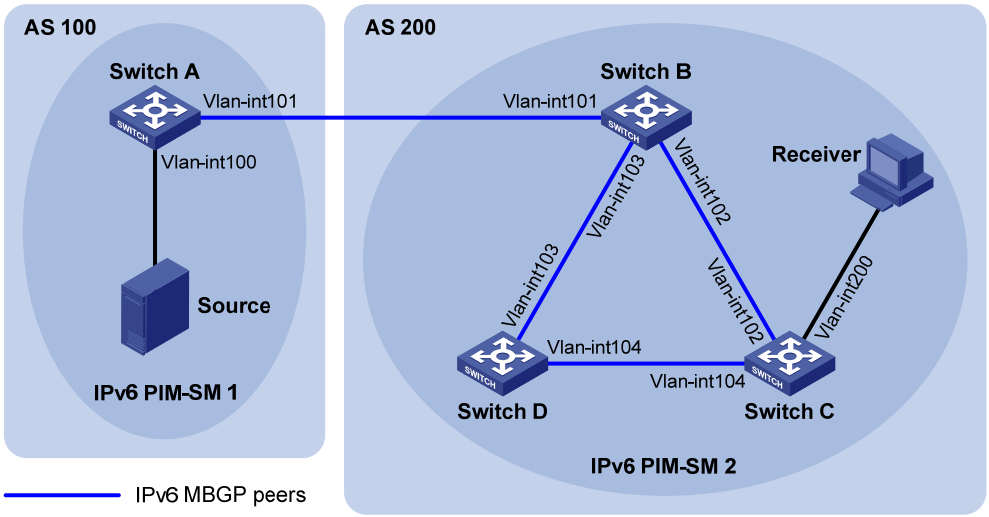
1.9 IPv6 MBGP典型配置举例

1. 组网需求

- 网络中存在两个自治系统：IPv6 PIM-SM 1 属于 AS 100，IPv6 PIM-SM 2 属于 AS 200。各 AS 内部采用 OSPFv3 进行互联，AS 之间采用 IPv6 MBGP 交换 IPv6 组播路由信息；
- IPv6 组播源属于 AS 100 内的 IPv6 PIM-SM 1，接收者则属于 AS 200 内的 IPv6 PIM-SM 2；
- 在 IPv6 PIM 域中的所有交换机上都使能嵌入式 RP 功能。

2. 组网图

图1-1 IPv6 MBGP 典型配置组网图



设备	接口	IP地址	设备	接口	IP地址
Source	-	1002::100/64	Switch C	Vlan-int200	3002::1/64
Switch A	Vlan-int100	1002::1/64		Vlan-int102	2001::2/64
	Vlan-int101	1001::1/64		Vlan-int104	3001::1/64
Switch B	Vlan-int101	1001::2/64	Switch D	Vlan-int103	2002::2/64
	Vlan-int102	2001::1/64		Vlan-int104	3001::2/64
	Vlan-int103	2002::1/64			

3. 配置步骤

(1) 配置 IPv6 转发功能，并配置 IPv6 地址和 IPv6 单播路由协议

- 使能各交换机的IPv6 转发功能，并按照 图 1-1 配置各接口的IPv6 地址和前缀长度，具体配置过程略。
- 配置 AS200 内的各交换机之间采用 OSPFv3 路由协议交换路由信息（AS 内各交换机创建的 OSPF 进程号为 1），确保各 AS 内部在网络层互通，具体配置过程略。

(2) 使能 IPv6 组播路由，使能 IPv6 PIM-SM 和 MLD，并配置 BSR 的服务边界

在 Switch A 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM。

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
```

Switch B 和 Switch D 上的配置与 Switch A 相似，配置过程略。

在 Switch C 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并在主机侧接口 Vlan-interface200 上使能 MLD。

```
<SwitchC> system-view
[SwitchC] multicast ipv6 routing-enable
```

```
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim ipv6 sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim ipv6 sm
[SwitchC-Vlan-interface104] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] pim ipv6 sm
[SwitchC-Vlan-interface200] mld enable
[SwitchC-Vlan-interface200] quit
```

在 Switch A 上配置 BSR 的服务边界。

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 bsr-boundary
[SwitchA-Vlan-interface101] quit
```

在 Switch B 上配置 BSR 的服务边界。

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim ipv6 bsr-boundary
[SwitchB-Vlan-interface101] quit
```

(3) 使能嵌入式 RP 功能

在 Switch A 上使能嵌入式 RP 功能。

```
[SwitchA] pim ipv6
[SwitchA-pim6] embedded-rp
[SwitchA-pim6] quit
```

Switch B、Switch C 和 Switch D 上的配置与 Switch A 相似，配置过程略。

(4) 配置 BGP 协议，并配置 IPv6 MBGP 对等体

在 Switch A 上配置 EBGP 邻接关系、并配置 IPv6 MBGP 对等体。

```
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 1001::2 as-number 200
[SwitchA-bgp-af-ipv6] import-route direct
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] ipv6-family multicast
[SwitchA-bgp-af-ipv6-mul] peer 1001::2 enable
[SwitchA-bgp-af-ipv6-mul] import-route direct
[SwitchA-bgp-af-ipv6-mul] quit
[SwitchA-bgp] quit
```

在 Switch B 上配置 EBGP 邻接关系、配置 IPv6 MBGP 对等体，并引入 OSPFv3 路由。

```
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 1001::1 as-number 100
[SwitchB-bgp-af-ipv6] import-route ospfv3 1
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] ipv6-family multicast
[SwitchB-bgp-af-ipv6-mul] peer 1001::1 enable
```

```
[SwitchB-bgp-af-ipv6-mul] import-route ospfv3 1
[SwitchB-bgp-af-ipv6-mul] quit
[SwitchB-bgp] quit
```

(5) 检验配置效果

通过使用 **display bgp ipv6 multicast peer** 命令可以查看交换机之间 IPv6 MBGP 对等体的关系。例如：

查看 Switch B 上 IPv6 MBGP 对等体关系的信息。

```
[SwitchB] display bgp ipv6 multicast peer

 BGP local router ID : 2.2.2.2
 Local AS number : 200
 Total number of peers : 3              Peers in established state : 3

  Peer       AS  MsgRcvd  MsgSent  OutQ PrefRcv Up/Down  State

  1001::1    100      56      56     0       0 00:40:54 Established
```