

组播技术白皮书

目 录

1 概述	1
1.1 产生背景.....	1
1.2 技术优点.....	1
2 组播技术实现	1
2.1 组播地址机制.....	1
2.1.1 IP 组播地址.....	1
2.1.2 组播 MAC 地址.....	2
2.2 组播模型分类.....	3
2.3 组成员关系管理.....	3
2.4 组播报文转发.....	5
2.4.1 组播转发树.....	5
2.4.2 组播报文转发机制.....	5
2.5 组播路由协议.....	5
2.5.1 域内组播路由协议.....	6
2.5.2 域间组播路由协议.....	7
2.6 二层组播协议.....	8
2.6.1 IGMP Snooping.....	9
2.6.2 PIM Snooping.....	9
2.6.3 组播 VLAN.....	9
3 典型组网应用	10
3.1 单域组播组网应用.....	10
3.2 跨域组播组网应用.....	10
3.2.1 PIM-SM&MBGP&MSDP 组合方案.....	10
3.2.2 PIM-SM 与隧道（MBGP&MSDP）组合方案.....	11
3.3 组播穿越防火墙组网应用.....	12

1 概述

1.1 产生背景

传统的 IP 通信有两种方式：一种是在源主机与目的主机之间点对点的通信，即单播；另一种是在源主机与同一网段中所有其它主机之间点对多点的通信，即广播。如果要将信息发送给多个主机而非所有主机，若采用广播方式实现，不仅会将信息发送给不必要的主机而浪费带宽，也不能实现跨网段发送；若采用单播方式实现，重复的 IP 包不仅会占用大量带宽，也会增加源主机的负载。所以，传统的单播和广播通信方式不能有效地解决单点发送、多点接收的问题。

组播是指在 IP 网络中将数据包以尽力传送的形式发送到某个确定的节点集合（即组播组），其基本思想是：源主机（即组播源）只发送一份数据，其目的地址为组播组地址；组播组中的所有接收者都可收到同样的数据拷贝，并且只有组播组内的主机可以接收该数据，其它主机无法接收。

1.2 技术优点

组播技术有效地解决了单点发送、多点接收的问题，实现了 IP 网络中点到多点的高效数据传送，能够大量节约网络带宽、降低网络负载。作为一种与单播和广播并列的通信方式，组播的意义不仅在于此。更重要的是，利用网络的组播特性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、网络电台和实时视频会议等互联网的信息服务领域。

2 组播技术实现

组播技术的实现需要解决以下几方面问题：

- 组播源向一组确定的接收者发送信息，而如何来标识这组确定的接收者？——这需要用到[组播地址机制](#)；
- 接收者通过加入组播组来实现对组播信息的接收，而接收者是如何动态地加入或离开组播组的？——即如何进行[组成员关系管理](#)；
- 组播报文在网络中是如何被转发并最终到达接收者的？——即[组播报文转发](#)的过程；
- 组播报文的转发路径（即组播转发树）是如何构建的？——这是由各[组播路由协议](#)来完成的。

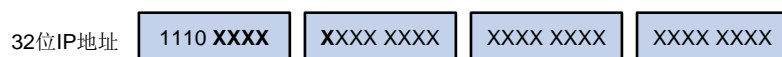
2.1 组播地址机制

2.1.1 IP 组播地址

IP 组播地址用于标识一个 IP 组播组。IANA（Internet Assigned Numbers Authority，互联网编号分配委员会）将 D 类地址空间分配给组播使用，范围从 224.0.0.0 到 239.255.255.255。

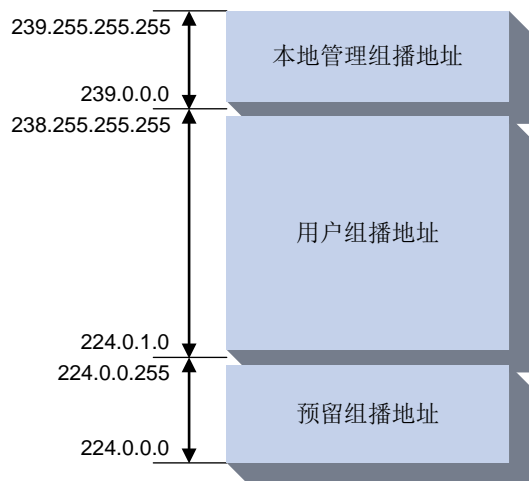
如[图 1](#)所示，IP 组播地址前四位均为“1110”。

图1 IP 组播地址格式



整个 IP 组播地址空间的划分如图 2 所示。

图2 IP 组播地址划分



- 224.0.0.0~224.0.0.255: 预留组播地址, 除 224.0.0.0 保留不做分配, 其它地址供路由协议、拓扑查找和维护协议使用。该范围内的地址属于局部范畴, 对于以该范围内组播地址为目的地址的数据包来说, 不论 TTL (Time to Live, 生存时间) 为多少, 都不会被转发出本地网段。
- 224.0.1.0~238.255.255.255: 用户组播地址, 在全网范围内有效。包含两种特定的组地址:
 - 232.0.0.0/8: SSM (Source-Specific Multicast, 指定信源组播) 组地址。
 - 233.0.0.0/8: GLOP 组地址。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。

说明

224.0.1.0/24 网段内的一些组播地址也被 IANA 预留给了某些组播应用。譬如, 224.0.1.1 被预留给 NTP (Network Time Protocol, 网络时间协议) 所使用。

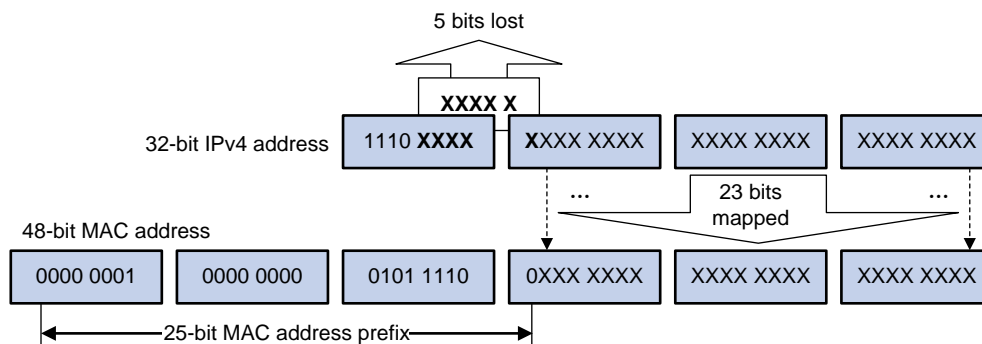
- 239.0.0.0~239.255.255.255: 本地管理组播地址, 仅在本地管理域内有效。使用本地管理组地址可以灵活定义组播域的范围, 以实现不同组播域之间的地址隔离, 从而有助于在不同组播域内重复使用相同组播地址而不会引起冲突。

2.1.2 组播 MAC 地址

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。

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 3 所示。

图3 IP 组播地址与组播 MAC 地址的映射



由于 IPv4 组播地址的高 4 位固定为 1110，而低 28 位中只有 23 位被映射到 IPv4 组播 MAC 地址上，从而导致有 5 位信息丢失。于是，将有 32 个 IPv4 组播地址被重复映射到同一个 IPv4 组播 MAC 地址上。因此设备在进行二层处理时，可能会收到一些本不需要的组播数据，这些多余的组播数据就需要上层进行过滤。

2.2 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大类：

- **ASM (Any-Source Multicast, 任意信源组播) 模型：**在 ASM 模型中，任意一个发送者都可作为组播源向某组播组地址发送组播信息，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，但在任意时间加入或离开组播组。
- **SFM (Source-Filtered Multicast, 信源过滤组播) 模型：**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，组播源经过了筛选。
- **SSM (Source-Specific Multicast, 指定信源组播) 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 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。

2.3 组成员关系管理

组成员关系管理是指在路由器/交换机上建立直联网段内的组成员关系信息，具体说，就是各接口/端口下有哪些组播组的成员。

IGMP (Internet Group Management Protocol, 互联网组管理协议) 运行于三层设备和其直联网段中的用户主机之间，其实现的功能是双向的：

- 主机通过 IGMP 通知路由器希望接收某个特定组播组的信息；

- 路由器通过 IGMP 周期性地查询局域网内的组播组成员是否处于活动状态，实现所连网段组成员关系的收集与维护。

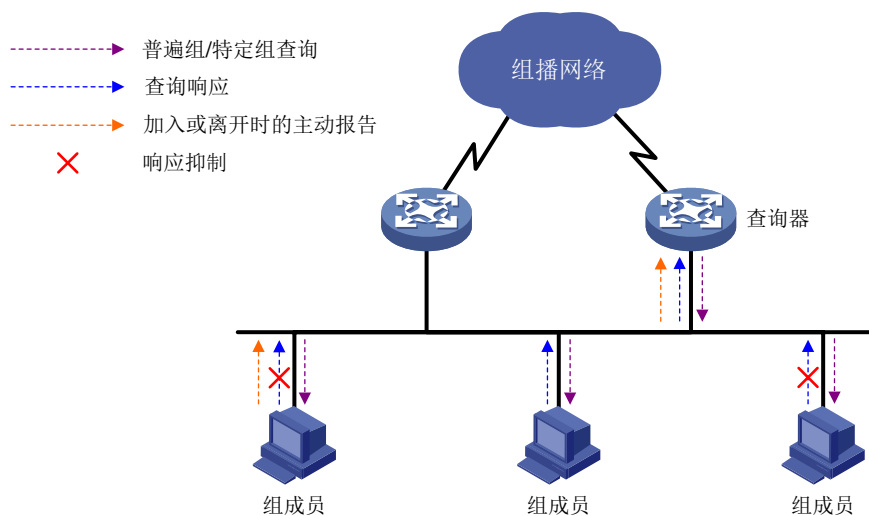
通过 IGMP，可以在路由器上记录某个组播组是否在本本地有组成员，而不是记录组播组与主机之间的对应关系。

目前 IGMP 有以下三个版本：

- IGMPv1 (RFC 1112) 中定义了基本的组成员查询和报告过程；
- IGMPv2 (RFC 2236) 在 IGMPv1 的基础上添加了组成员快速离开的机制等；
- IGMPv3 (RFC 3376) 中增加的主要功能是成员可以指定接收或拒绝来自某些组播源的报文，以实现支持 SSM 模型的支持。

以下着重介绍 IGMPv2 的原理。

图4 IGMPv2 的工作原理



如图 4 所示，当同一个网段内有多台 IGMP 路由器时，IGMPv2 通过查询器选举机制从中选举出唯一的查询器。查询器周期性地发送普遍组查询报文进行成员关系查询，主机通过发送成员关系报告报文来响应查询。而作为组成员的路由器，其行为也与普通主机一样，响应其它路由器的查询。

当主机要加入组播组时，不必等待查询报文，而是主动发送成员关系报告报文；当主机要离开组播组时，也会主动发送离开组报文，查询器收到离开组报文后，会发送特定组查询报文来确定该组的所有组成员是否都已离开。

当关注同一个组的主机成员中，已有某个成员以组播的方式向该组播组发送成员关系报文，由于主机上存在 IGMP 成员关系报告抑制机制，因此关注该组播组的其他主机成员收到报告报文后，将不再发送同样针对该组播组的成员关系报告报文，因为 IGMP 路由器已知道本地网段中有对此组播组感兴趣的主机了。

通过上述机制，在路由器上建立一张 IGMP 组播组信息表，该表记录了路由器每个接口连接的网段上存在的组播组的成员。当路由器收到发往某一组播组的组播数据后，只向那些有该组播组成员的接口转发该数据。至于组播数据在路由器之间如何转发则由组播路由协议决定，而不是 IGMP 的功能。

2.4 组播报文转发

2.4.1 组播转发树

组播报文在网络中沿着树型转发路径进行转发,该路径称为组播转发树。它可分为源树(Source Tree)和共享树(RPT)两大类。

1. 源树

源树是指以组播源作为树根,将组播源到每一个接收者的最短路径结合起来构成的转发树。由于源树使用的是从组播源到接收者的最短路径,因此也称为 SPT (Shortest Path Tree, 最短路径树)。对于某个组,网络要为任何一个向该组发送报文的组播源建立一棵树。

源树的优点是能构造组播源和接收者之间的最短路径,使端到端的延迟达到最小。但付出的代价是,在路由器中必须为每个组播源保存路由信息,这样会占用大量的系统资源,路由表的规模也比较大。

2. 共享树

以某个路由器作为路由树的树根,该路由器称为 RP (Rendezvous Point, 汇集点),由 RP 到所有接收者的最短路径所共同构成的转发树称为 RPT (Rendezvous Point Tree, 共享树)。使用共享树时,对于某个组,网络中只有一棵树,所有的组播源和接收者都使用这棵树来收发报文。组播源先向树根发送数据报文,之后报文又向下转发到达所有的接收者。

共享树的最大优点是路由器中保留的路由信息可以很少,缺点是组播源发出的报文要先经过 RP,再到达接收者,经由的路径通常并非最短,而且对 RP 的可靠性和处理能力要求很高。

2.4.2 组播报文转发机制

当路由器收到组播数据报文时,根据组播目的地址查找组播转发表,对报文进行转发。与单播报文的转发相比,组播报文的转发相对复杂:在单播报文的转发过程中,路由器并不关心报文的源地址,只关心报文的地址,通过其目的地址决定向哪个接口转发;而组播报文是发送给一组接收者的,这些接收者用一个逻辑地址(即组播地址)标识,路由器在收到组播报文后,必须根据报文的源地址确定其正确的入接口(指向组播源方向)和下游方向,然后将其沿着下游方向转发——这个过程称为 RPF (Reverse Path Forwarding, 逆向路径转发)。

RPF 检查的过程为:路由器在单播路由表中查找组播源或 RP 对应的出接口(使用 SPT 时查找组播源对应的出接口,使用 RPT 时查找 RP 对应的出接口)。该出接口即为 RPF 接口。如果组播报文是从 RPF 接口接收的,则 RPF 检查通过,报文向下游接口转发;否则,丢弃该报文。通过 RPF 检查除了可以正确地按照组播路由的配置转发报文外,还可以避免可能出现的环路。

2.5 组播路由协议

与单播路由一样,组播路由协议也分为域内和域间两大类:

- 域内组播路由协议:根据 IGMP 协议维护的组成员关系信息,运用一定的组播路由算法构造组播分发树,在路由器中建立组播路由表项,路由器根据组播路由表进行组播数据包转发;
- 域间组播路由协议:根据网络中配置的域间组播路由策略,在各自治系统间发布具有组播能力的路由信息以及组播源信息,使组播数据能在域间进行转发。

2.5.1 域内组播路由协议

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 可以分为以下几种模式：

- 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）
- PIM-SSM（Protocol Independent Multicast Source-Specific Multicast，协议无关组播—指定源组播）
- BIDIR-PIM（Bidirectional Protocol Independent Multicast，双向协议无关组播，简称双向 PIM）

1. PIM-DM

在 PIM-DM 域中，设备上每个运行了 PIM 协议的接口通过定期向本网段的所有 PIM 设备（224.0.0.13）组播发送 PIM Hello 报文，以发现 PIM 邻居，维护各设备之间的 PIM 邻居关系，从而构建和维护 SPT。同时，PIM 设备通过 PIM Hello 报文在多路访问网络中选举 DR（Designated Router，指定路由器）。



PIM-DM 本身并不需要 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。

PIM-DM 属于密集模式的组播路由协议，使用“推”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到网络中的所有节点。接着 PIM-DM 对没有组播数据转发的分支进行剪枝，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。
- 当被剪枝分支的节点上出现了组播组的成员时，该节点通过主动向其上游发送嫁接报文，从而由剪枝状态恢复成转发状态，以恢复对组播数据的转发。

2. PIM-SM

在 PIM-SM 域中，设备上每个运行了 PIM 协议的接口通过定期向本网段的所有 PIM 设备（224.0.0.13）组播发送 PIM Hello 报文，以发现邻接的 PIM 路由器，并负责在多路访问网络中选举 DR。这里，DR 负责为与其直连的组成员向组播树根节点的方向发送加入/剪枝消息，或是将直连组播源的数据发向组播分发树。

PIM-SM 属于稀疏模式的组播路由协议，使用“拉”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，RPT 选择 PIM 域中某台路由器作为公用的根节点 RP，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，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 负责向 RP 进行注册，把注册报文通过单播方式发送给 RP，该报文到达 RP 后触发建立从组播源到 RP 的 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。

3. PIM-SSM

PIM-SSM 主要应用于接收者已经预先知道组播源的场景。

如果接收者准备加入的组播组属于 SSM 组地址范围(IANA 保留的 SSM 组地址范围为 232.0.0.0/8)，则采用 PIM-SSM 方式来构建组播转发树。

在 PIM-SSM 域中，邻居发现、DR 选举机制和 PIM-SM 相同。

PIM-SSM 也使用“拉”模式传送组播数据，由于接收者已经预先知道组播源的具体位置，因此 SSM 域中无需 RP，无需构建 RPT，无需组播源注册过程，也无需通过 MSDP 来发现其它 PIM 域内的组播源。其基本原理如下：

- 接收者借助 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。
- 连接接收者的路由器发现准备加入的组播组属于某个 SSM 组地址范围，于是向组播源逐跳发送加入报文。沿途所有设备上创建 (S, G) 表项，所经过的路径就形成了 SPT。组播源直接沿着 SPT 路径，将组播报文发送给接收者。

4. 双向 PIM

双向 PIM 主要应用于一个组播组同时对应多个接收者和多个组播源的情况。如果使用传统的 PIM-DM 或 PIM-SM 按 SPT 转发组播数据，需在每台设备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。而双向 PIM 通过建立以 RP 为中心、分别连接组播源和接收者的双向 RPT，在每台设备上只需维护 (*, G) 表项即可，从而节约了系统资源。

在双向 PIM 域中，邻居发现、DR 选举机制和 PIM-SM 相同，但增加了 DF 选举机制。域中的设备在获得 RP 的信息后，会向本网段的所有 PIM 设备 (224.0.0.13) 以组播方式发送 DF 选举报文 (DF Election Message)，按照一定规则选举该网段中唯一的 DF 设备。

接收者侧 RPT 和组播源侧 RPT 的构建过程如下：

- 连接接收者的路由器向某组播组对应的 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接收者侧 RPT 的分支。
- 组播源发向某个组播组 G 的组播数据在途经的每个网段，都被该网段的 DF 无条件地向 RP 转发。所经过的路径形成了组播源侧 RPT 的分支。

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。

2.5.2 域间组播路由协议

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：

- MBGP (Multicast BGP, 组播 BGP)：用于在自治系统之间交换组播路由信息；
- MSDP (Multicast Source Discovery Protocol, 组播源发现协议)：用于在自治系统之间交换组播源信息。

1. MBGP

域间路由的首要问题是路由信息（或者说可达信息）如何在自治系统之间传递，由于不同的 AS 可能属于不同的运营商，因此除了距离信息外，域间路由信息必须包含运营商的策略，这是与域内路由信息的不同之处。

组播的网络拓扑和单播拓扑有可能不同，这里既有物理方面的原因，也有策略方面的原因。网络中的一些路由器可能只支持单播不支持组播，也可能按照策略配置不转发组播报文。为了构造域间组播路由树，除了要知道单播路由信息外，还要知道网络中哪些部分是支持组播的，即组播的网络拓扑情况。简而言之，域间的组播路由信息交换协议应该满足下面的要求：

- 能对单播和组播拓扑进行区分；
- 有一套稳定的对等和策略控制方法。

目前使用最多的域间单播路由协议是 BGP-4，由于 BGP-4 已满足后一个条件，而且已被证明是一个有效的、稳定的单播域间路由协议，因此为了实现域间组播路由信息的传递，合理的解决方案就是对 BGP-4 协议进行增强和扩展，而不是构建一套全新的协议。在 RFC 2858 中规定了对 BGP 进行多协议扩展的方法，扩展后的 BGP 协议（即 MP-BGP，也写作 BGP-4+）不仅能携带 IPv4 单播路由信息，也能携带其它网络层协议（如组播、IPv6 等）的路由信息，携带组播路由信息只是其中一个扩展功能，称为组播 BGP（MBGP）。

有了 MBGP 之后，单播和组播路由信息可以通过同一个进程交换，但是存放在不同的路由表里。由于 MBGP 是 BGP-4 协议的一个增强版，因此 BGP-4 所支持的常见的策略和配置方法都可以用到组播里。

2. MSDP

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。而对于自治系统来说，不希望依靠其它自治系统的 RP 转发组播流量，但同时又要求无论组播源的 RP 在哪里，都能从组播源获取信息发给自己内部的成员。

MSDP 就是为了解决多个 PIM-SM 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。MSDP 通过在网络中选取适当的设备建立 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）报文来共享组播源信息。

尽管 MSDP 是为域间组播开发的，但它在 PIM-SM 域内还有着项特殊的应用——Anycast RP（任播 RP）。Anycast RP 是指在同一 PIM-SM 域内通过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。

2.6 二层组播协议

不管是 IGMP 还是 PIM，都是针对 IP 层设计的，只能记录路由器上的三层接口与 IP 组播地址的对应关系。但在很多情况下，组播报文不可避免地要经过一些二层交换机，如果不能将二层端口和组播报文对应起来，那么组播报文就会泛洪给交换机的所有端口，这显然会浪费大量的系统资源。通过二层组播，可以解决这些问题。

2.6.1 IGMP Snooping

IGMP Snooping（Internet Group Management Protocol Snooping，互联网组管理协议窥探）运行在二层设备上。主机发往 IGMP 查询器的成员关系报告报文经过二层交换机时，交换机对这个报文进行监听并记录下来，为端口和组播 MAC 地址建立起映射关系；当交换机收到组播数据时，根据对应的映射关系，只向连有组播组成员的端口转发组播数据。

2.6.2 PIM Snooping

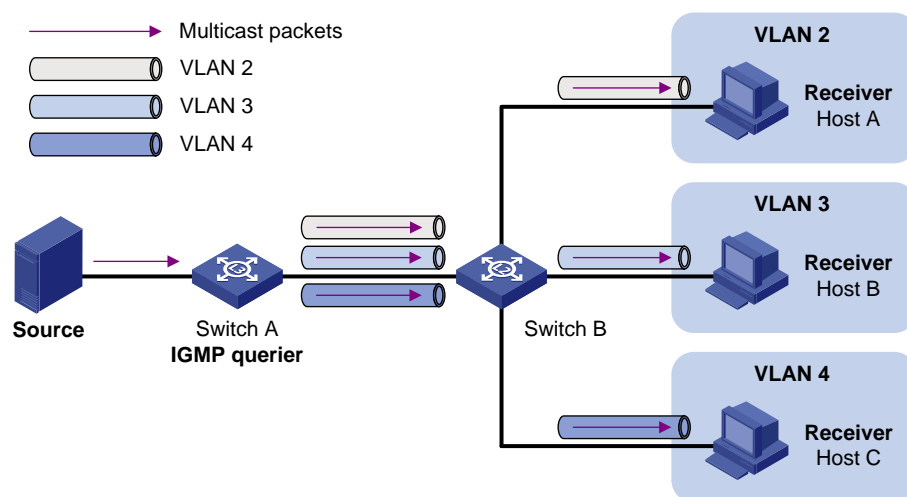
PIM Snooping（Protocol Independent Multicast Snooping，协议无关组播窥探）运行在二层设备上。二层交换机设备对经过的 PIM 协议报文进行监听并且记录下来，将有接收需求的端口添加到 PIM Snooping 路由表的相应表项中，以实现组播报文的精确转发。

- 当二层交换机设备只运行 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，将组播数据报文向 VLAN 内的所有路由器端口转发，而将除 PIM Hello 报文外的其它 PIM 协议报文在 VLAN 内泛洪。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。
- 当二层交换机设备同时运行了 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，将有接收需求的 PIM 路由器所在的端口添加到 PIM Snooping 路由表的相应表项中，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。

2.6.3 组播 VLAN

组播 VLAN 主要用于点播同一个组播组的主机分属于不同 VLAN 的场景，如图 5 所示。在传统的组播点播方式下，三层设备（Switch A）需要将组播数据为每个用户 VLAN（即主机所属的 VLAN）都复制一份后发送给二层设备（Switch B）。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5 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，三层设备只需将组播数据通过组播 VLAN 向二层设备发送一份即可，而不必向每个用户 VLAN 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 有基于子 VLAN 和基于端口两种实现和配置方式。

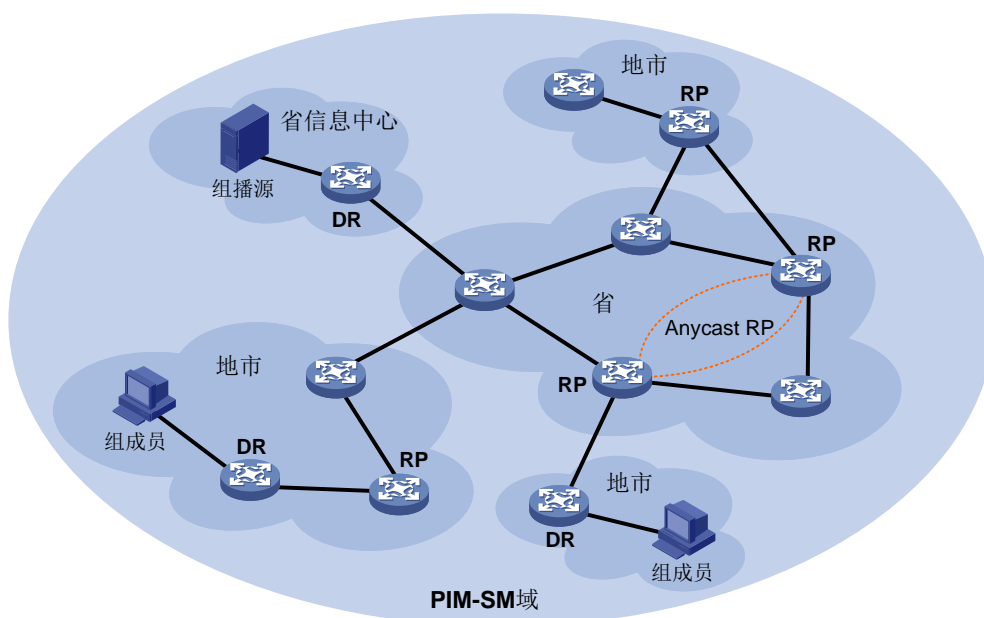
3 典型组网应用

3.1 单域组播组网应用

目前，PIM-SM 是域内组播的公认标准。对于由一个自治系统组成的网络，或者组播仅在域内进行时，仅需在网络中运行 PIM-SM 即可。为了增强 PIM-SM 中 RP 的可靠性，以及对网络中的组播流量进行分担，可在网络中选取若干 RP，运行 Anycast RP，达到冗余备份及负载分担的目的。

使用 PIM-SM 协议的单域组播组网如图 6 所示。

图6 PIM-SM 单域组播组网

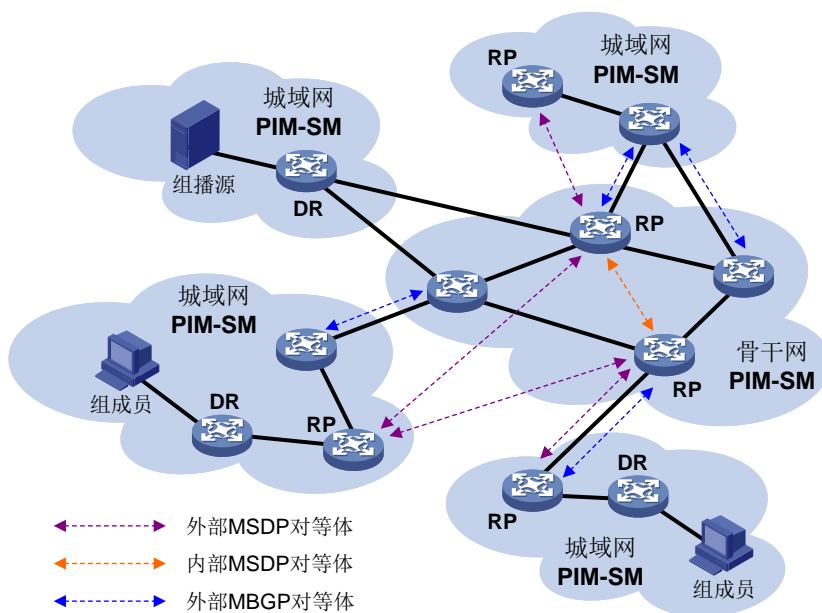


3.2 跨域组播组网应用

3.2.1 PIM-SM&MBGP&MSDP 组合方案

域间组播目前比较成型的解决方案是 PIM-SM&MBGP&MSDP 组合方案，它要求所有的自治系统都支持 PIM-SM、MBGP 和 MSDP。如图 7 所示，在全网各自治系统都运行 PIM-SM，域间运行 MBGP 和 MSDP。

图7 PIM-SM&MBGP&MSDP 组合方案



该方案实际上是 PIM-SM 在域间环境下的扩展,如果把整个 PIM-SM&MBGP&MSDP 组合方案机制看作 PIM-SM,则所有域的 RP 的集合就是 PIM-SM 协议中的“RP”,而该方案无非是增加了以下两个过程:

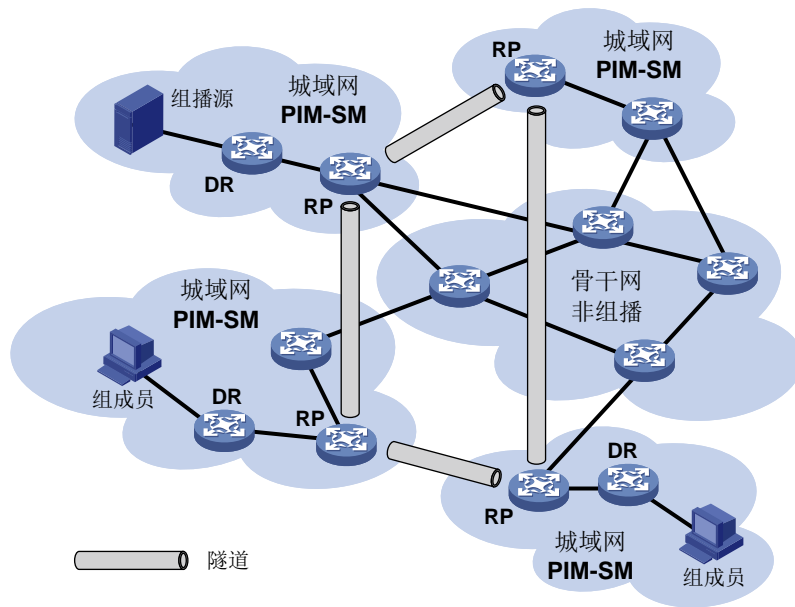
- (1) 组播源信息在 RP 集合中的泛洪,以实现组播源和成员在“RP”点的会合;
- (2) 域间组播路由信息的传递,目的是保证组播报文在域间的顺利转发。在上述过程中,一个 AS 中的 RP 和接收端向另一个 AS 中的远端建立逆向路径的过程中都需要用到 MBGP 传递的组播拓扑信息。

在该方案中,自治系统边界路由器之间配置外部 MBGP 对等体,RP 之间配置外部 MSDP 对等体;自治系统内部路由器之间根据需要配置内部 MBGP 对等体,内部 RP 之间配置内部 MSDP 对等体,运行 Anycast RP;所有的自治系统都运行 PIM-SM。系统内的组播路由和组播源信息收集工作由 PIM-SM 完成,系统间则由 MBGP 传播组播拓扑信息、MSDP 传播组播源信息。

3.2.2 PIM-SM 与隧道 (MBGP&MSDP) 组合方案

如图 8 所示,在骨干网不支持或不运行组播的情况下,在城域网内部运行 PIM-SM,各个城域网的 RP 与其它城域网 RP 之间通过隧道构成虚拟网络,在此虚拟网络中运行 PIM-SM、MBGP 和 MSDP。这种方案的优势是不要求骨干网支持 PIM-SM、MBGP 和 MSDP,组播流量对骨干网络来说是透明的,可以避免组播报文转发对设备性能造成的影响。缺点是要求 RP 之间既要支持 PIM-SM,还要支持 MBGP 和 MSDP,配置和管理繁琐、对设备要求较高。

图8 PIM-SM与隧道（MBGP&MSDP）组合方案



3.3 组播穿越防火墙组网应用

如图9所示，当多个组播网段之间跨越一个不支持组播的网络（如 Internet），或者防火墙上配置了 NAT 或 IPSec VPN 时，防火墙不能与对端设备建立 PIM 邻居关系并生成组播路由。在这种情况下，通过配置 GRE 隧道可以将各分离的组播网段连接成起来，实现组播的应用。

图9 组播穿越防火墙组网

