

组播技术白皮书

关键词: 组播, IGMP, IGMP Snooping, PIM, MBGP, MSDP, SSM Mapping

摘 要:组播技术实现了IP网络中点到多点的高效数据传送,由于组播能够有效地节约网络带宽、降低网络负载,因此在实时数据传送、多媒体会议、数据拷贝、游戏和仿真等诸多方面都有广泛的应用。本文介绍了组播的基本概念和目前通用的组播协议,以及组播组网的基本方案。

缩略语:

H-H /H		
缩略语	英文全名	中文解释
AS	Autonomous System	自治系统
ASM	Any-Source Multicast	任意信源组播
BSR	BootStrap Router	自举路由器
C-BSR	Candidate-BSR	候选BSR
C-RP	Candidate-RP	候选RP
DR	Designated Router	指定路由器
IANA	Internet Assigned Numbers Authority	互联网编号分配委员会
IGMP	Internet Group Management Protocol	互联网组管理协议
MBGP	Multicast Border Gateway Protocol	组播边界网关协议
MP-BGP	MultiProtocol Border Gateway Protocol	多协议边界网关协议
MSDP	Multicast Source Discovery Protocol	组播源发现协议
PIM-DM	Protocol Independent Multicast-Dense Mode	协议无关组播一密集模式
PIM-SM	Protocol Independent Multicast-Sparse Mode	协议无关组播一稀疏模式
RP	Rendezvous Point	汇集点
RPF	Reverse Path Forwarding	逆向路径转发
RPT	Rendezvous Point Tree	共享树
SPT	Shortest Path Tree	最短路径树
SSM	Source-Specific Multicast	指定信源组播 联系
-		

145年常景/

- *1.2 接条犹点
- 2 组播技术实现
- 2.1 组播地址机制
- 2.1.1 IP组播地址
- 2.1.2 IP组播地址到链路层的映射
- 2.2 组成员关系管理
- 2.2.1 IGMP
- 2.2.2 IGMP Snooping
- 2.3 组播报文转发
- 2.3.1 组播转发树
- 2.3.2 组播报文转发机制
- 2.4 组播路由协议
- 2.4.1 域内组播路由协议
- 2.4.2 域间组播路由协议
- 2.5 组播模型分类
- 3 典型组网应用
- 3.1 单域组播组网应用
- 3.2 跨域组播组网应用
- 3.2.1 PIM-SM/MBGP/MSDP方案
- 3.2.2 PIM-SM/隧道(MBGP&MSDP)方案
- 3.2.3 PIM-SM/隧道(PIM-DM)方案
- 3.3 组播穿越防火墙组网应用
- 4 总结和展望

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需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,里复的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。

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,可 以利用网络的组播特性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<u>组播报文转发</u>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<u>组播路由协议</u>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把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。

32位IP地址

1110 XXXX

XXXX XXXX

XXXX XXXX

XXXX XXXX

图1 IP组播地址格式

如图1所示,IP组播地址前四位均为"1110",而整个IP组播地址空间的划分则如图2所示。

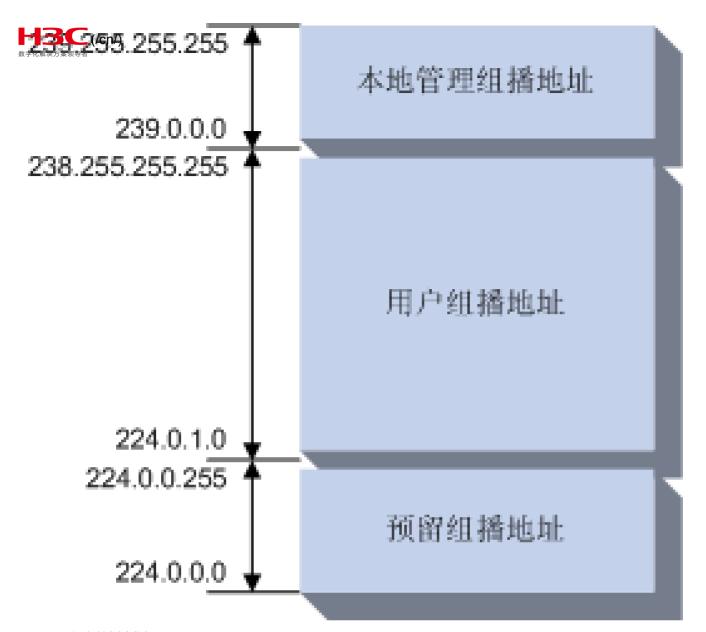


图2 IP组播地址划分

- 224.0.0.0到224.0.0.255被IANA预留,地址224.0.0.0保留不做分配,其它地址供路由协议及拓扑查找和维护协议使用。该范围内的地址属于局部范畴,不论TTL为多少,都不会被路由器转发;
- 224.0.1.0到238.255.255.255为用户可用的组播地址,在全网范围内有效。其中 232.0.0.0/8为SSM组地址,而其余则属于ASM组地址。有关ASM和SSM的详细介绍,请参 见"2.5 组播模型分类"一节;
- 239.0.0.0到239.255.255.255为本地管理组播地址,仅在特定的本地范围内有效,也属于 ASM组地址。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。

□ 说明:

联系我们

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。



2.1.2 IP组播地址到链路层的映射

□ 说明:

本文只讨论以太网链路层协议的组播实现,其它链路层协议的组播实现并不作为本文讨论的重点。

IANA将MAC地址范围01:00:5E:00:00:00~01:00:5E:7F:FF:FF分配给组播使用,这就要求将28位的IP组播地址空间映射到23位的组播MAC地址空间中,具体的映射方法是将组播地址中的低23位放入MAC地址的低23位,如图3所示。

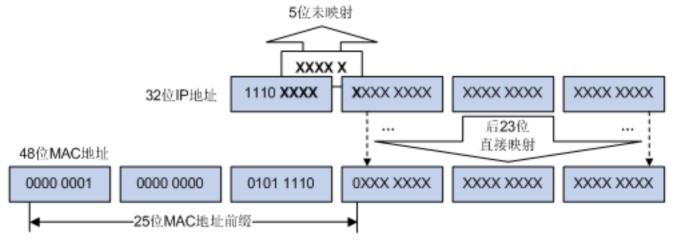


图3 IP组播地址到组播MAC地址的映射

由于IP组播地址的后28位中只有23位被映射到组播MAC地址,这样会有32个IP组播地址映射到同一组播MAC地址上。

2.2 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。

2.2.1 IGMP

IGMP运行于主机和与主机直连的路由器之间,其实现的功能是双向的:一方面,主机通过IGMP通知路由器希望接收某个特定组播组的信息;另一方面,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,在路由器中记录的信息是某个组播组是否在本联系我们地有组成员,而不是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;

▶ 3 (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*** 以实现对SSM模型的支持。

以下着重介绍IGMPv2的原理。

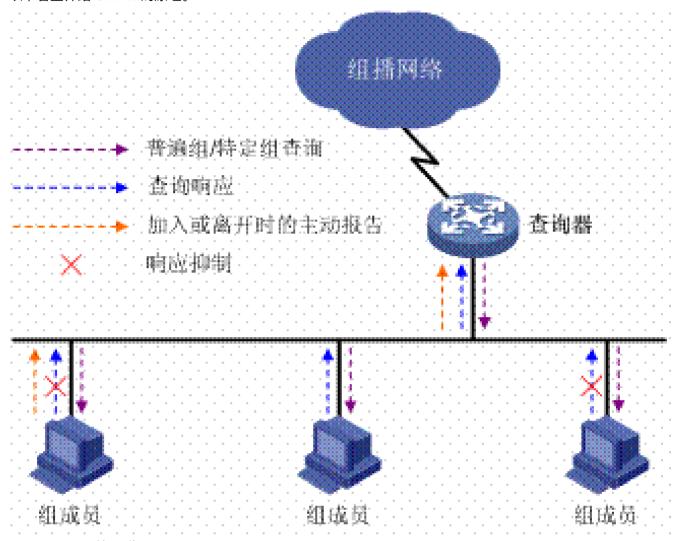


图4 IGMPv2的工作原理

如图4所示,当同一个网段内有多个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消息进行成员关系查询,主机通过发送报告消息来响应查询。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,不必等待查询消息,而是主动发送报告消息;当主机要离开组播组时,也会主动发送离开组消息,查询器收到离开组消息后,会发送特定组查询消息来确定该组的所有组成员是否都已离开。通过上述机制,在路由器里建立起一张表,其中记录了路由器各接口所对应子网上都有哪些组的成员。当路由器收到发往组G的组播数据后,只向那些有G的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。

2.2.2 IGMP Snooping

联系我们

IGMP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交换机,如果没有一种机制将二层端口与组播MAC地址对应起来,组播报文就会转发给交换机的所有端口,这显然会浪费大量的系统资源。

2.3 组播报文转发

2.3.1 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,因此也称为最短路径树(SPT)。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,该路由器称为汇集点(RP),共享树就是由RP到所有接收者的最短路路径 所共同构成的转发树。使用共享树时,对应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。

2.3.2 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,只关心报文的目的地址,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,然后将其沿着远离组播源的下游方向转发——这个过程称为逆向路径转发(RPF)。在RPF执行过程中会利用原有的单播路由表确定上、下游的邻接节点,只有报文从上游节点所对应的接口(称为RPF接口,即路由器上通过单播方式向该地址发送报文的出接口)到达时,才向下游转发。RPF的主体是RPF检查,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。路由器收到组播报文后先对其进行RPF检查,只有检查通过才执行转发。

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RPF接口(使用SPT时查找组播源对应的RPF接口,使用RPT时查找RP对应的RPF接口),如果组播报文是从RPF接口接收下来的,则RPF检查通过,报文向下游接口转发;否则,丢弃该报文。



2.4 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状态,路由器根据这些状态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。

2.4.1 域内组播路由协议

在众多域内组播路由协议中,PIM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。

1. PIM-DM

在PIM-DM域中,运行PIM-DM的路由器周期性地发送PIM Hello消息,以发现邻接的PIM路由器,进行叶子网络、叶子路由器的判断,并负责在多路访问网络中选举DR——尽管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。然后,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"扩散一剪枝"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,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,运行PIM-SM的路由器周期性地发送PIM Hello消息,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: 联系我们

● 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;

● 组播源如果要向某组播组发送组播数据,首先由与组播源侧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。

2.4.2 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: 用于在自治域之间交换组播路由信息;

● MSDP:用于在ISP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ISP来说,不希望依靠其它ISP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通过将某个域内的RP与其它域内的RP建立MSDP对等体关系,以连通各PIM-SM域的RP,并利用这些对等体关系交换组播源信息。

口管是DP是为域间组播开发的,但它在PIM-SM域内还有着一项特殊的应用——Anycast RP(任播RP)。

**Ahrycast 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。

2.5 组播模型分类

根据接收者对组播源处理方式的不同,组播模型分为以下两大类:

- ASM模型:即任意源组播模型。在ASM模型中,任一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组播组地址标识的组播组以获得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组播组。
- SSM模型:即指定信源组播模型。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。

上一节所描述的组播路由协议构架主要针对ASM模型。在ASM模型下,接收者无法选择组播源,只能被动地接收所有组播源的信息,而SSM模型的提出则为指定源组播提供了解决方案。

SSM模型要求使用与ASM模型不同的组地址,通过PIM-SM直接在接收者与组播源之间建立专用的组播转发路径。由于接收者可通过其它渠道(如广告咨询等)事先了解组播源的地址,因此SSM模型无需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-SM域内的组播源。同时,SSM模型还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:

- 如果接收者主机运行的是IGMPv3,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/v2,由于IGMPv1/v2的报告报文中无法指定组播源的地址,因此可以通过在路由器上配置SSM Mapping静态映射规则,将IGMPv1/v2报告报文中所包含的(*, G)信息映射为(G, INCLUDE,(S1, S2…))信息。

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。对于由一个自治域组成的网络,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5所示。

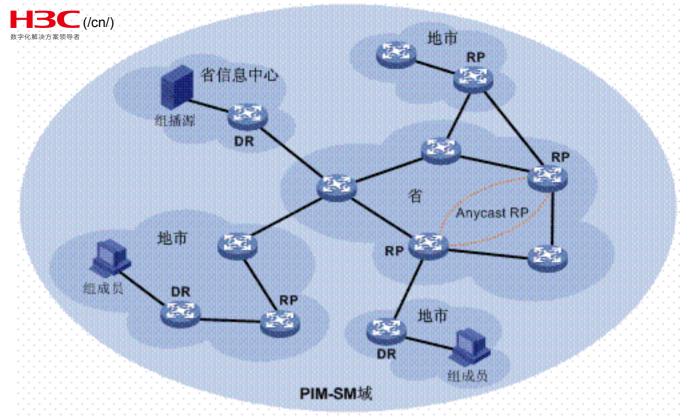


图5 PIM-SM单域组播组网

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,它要求所有的自治域都支持PIM-SM、MBGP和MSDP。如图6所示,在全网各自治域都运行PIM-SM,域间运行MBGP和MSDP。

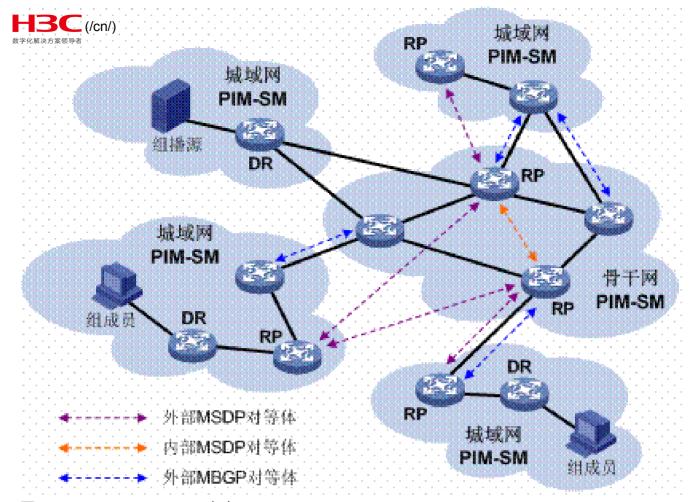


图6 PIM-SM/MBGP/MSDP方案

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集合中的泛滥,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,自治域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;自治域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;所有的自治域都运行PIM-SM。域内的组播路由和组播源信息收集工作由PIM-SM完成,域间则由MBGP传播组播拓扑信息、MSDP传播组播源信息。

3.2.2 PIM-SM/隧道(MBGP&MSDP)方案

如图7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_{联累我们}求骨干支持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隧道,配置和管理繁琐、对设备要求较高。

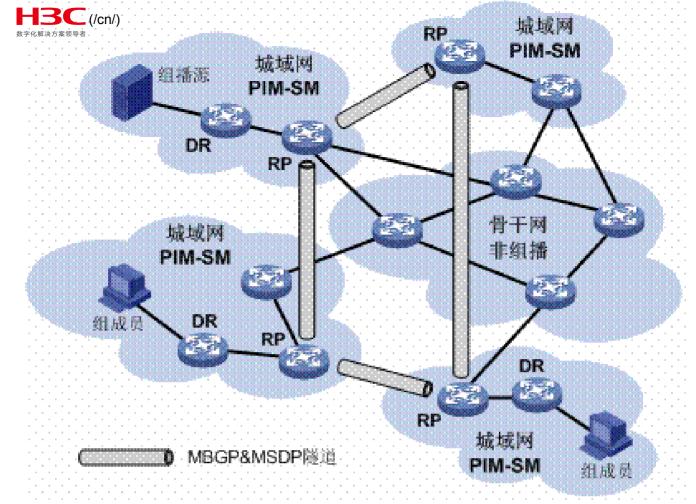


图7 PIM-SM/隧道(MBGP&MSDP)方案

3.2.3 PIM-SM/隧道(PIM-DM)方案

如图8所示,在骨干网不支持或不运行组播的情况下,城域网内部运行PIM-SM,各个城域网的RP与其它城域网的RP之间通过隧道构成虚拟网络,在此虚拟网络中运行PIM-DM。这种方案的优势是不要求骨干支持PIM-SM、MBGP和MSDP,组播流量对骨干网络透明,因此也不需要在骨干上保存大量的组播路由状态;缺点是RP之间运行PIM-DM,组播流量定期扩散可能会造成骨干网的带宽浪费。

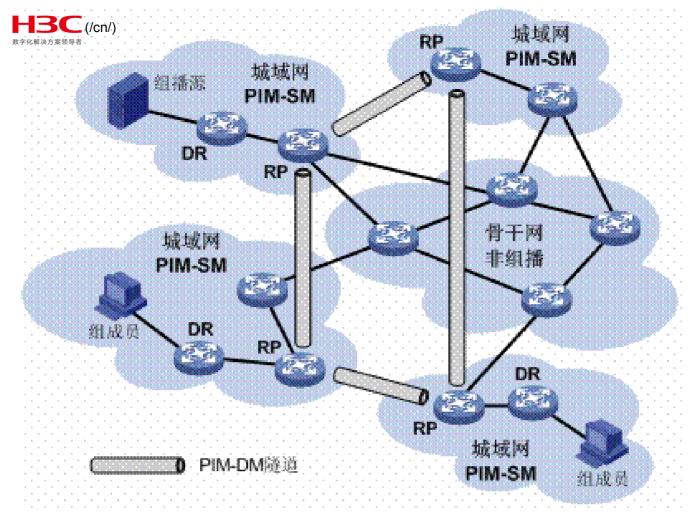


图8 PIM-SM/隧道(PIM-DM)方案

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。

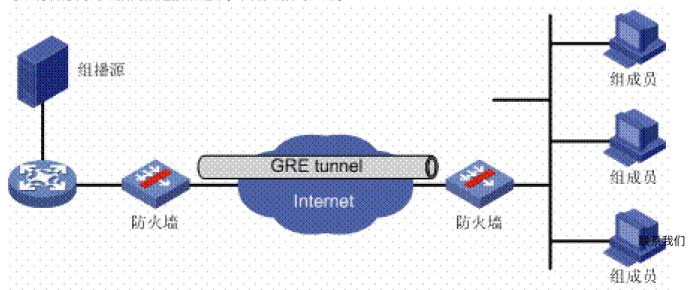


图9 组播穿越防火墙组网

4 总结和展望

组播技术涵盖了从地址方案、成员管理和路由建立等各个方面,其中组播地址的分配方式、域间组播路由以及组播安全等仍是研究的热点。从目前的情况看,组成员管理技术普遍采用IGMPv2;PIM-SM因其良好的扩展性以及从RPT向SPT切换的能力而成为域内组播路由技术的佳选;域间组播路由协议现阶段普遍采用PIM-SM/MBGP/MSDP的组合方案。

组播技术可以提供包括流媒体、视频会议、IPTV等在内的各种宽带增值业务,但这些业务的顺利开展还依赖于有效的业务管理、监控及安全控制。结合在业务运营管理方面的理解和经验积累,H3C公司提供不断完善的可运营、可管理的组播解决方案,我们将继续致力于推动组播技术的发展、组播业务的普及和功能的完善。

Copyright ©2004-2008 杭州华三通信技术有限公司 版权所有,保留一切权利。

非经本公司书面许可,任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部,并不得以任何形式传播。本文档中的信息可能变动,恕不另行通知。

附件下载

组播技术白皮书.pdf (https://download.h3c.com/app/cn/download.do?id=333564)(492.54 KB)



我们非常抱歉发生了这样的系统异常,请 联系管理员,或者在线给我们反馈,非常感谢!

如何购买							
关于新华三							
联系新华三							
常用链接							
6	%	具条	J				

版权所有 2003-2021 新华三技术有限公司.保留一切权利. 浙ICP备09064986号-1 (http://beian.miit.gov.cn/) 🧶 浙公网安备 33010802004416号 (http://www.beian.gov.cn/portal/registerSystemInfo?recordcode=33010802004416)

隐私政策 版权声明 网站地图 联系我们 (/cn/Home/Privacy_Clause/) (/cn/Home/Legal___Privacy/) (/cn/Home/sitemap/) (/cn/About_H3C/Contact_Us/)

联系我们