

一种基于 PIM-SSM 的 MPLS VPN 组播算法

王 渊 明

(浙江万里学院, 浙江 宁波 315101)

摘要:文章提出一种基于指定信源的协议无关组播(Protocol Independent Multicast-Source Specific Multicast, PIM-SSM)协议的多协议标签交换虚拟专用网(MPLS VPN)组播算法,该算法利用处于分支点的核心路由器才需要维护组播表和复制组播数据的特点,在 MPLS VPN 中引入了组播管理点来更好地实现组播。文章描述了算法的具体细节和实现原理,评价了该算法的可行性和优越性。

关键词:组播;指定信源的协议无关组播协议;多协议标签交换虚拟专用网

中图分类号:TN915

文献标识码:A

文章编号:1005-8788(2011)02-0014-03

A PIM-SSM-based multicast algorithm for MPLS VPN

Wang Yuanming

(Zhejiang Wanli University, Ningbo 315101, China)

Abstract: This paper presents a new multicast algorithm based on Protocol Independent Multicast-Source Specific Multicast (PIM-SSM) for MPLS VPN. Utilizing the theory that only the branching core routers need to maintain the multicast routing-table and copy the multicast data, this algorithm introduces multicast managers into MPLS VPN to better implement multicast. This paper describes its specific details and implementation principle and evaluates its feasibility and advantages.

Key words: multicast; PIM-SSM protocol; MPLS VPN

0 引 言

多协议标签交换虚拟专用网(MPLS VPN)技术日趋成熟,已成为 VPN 发展的方向。同时以组播(Multicast)技术为支持的多媒体业务已逐步成为下一代网络的主流应用。在 MPLS VPN 网络中更好地支持组播具有重要意义。

1 MPLS VPN 组播研究现状

目前在 MPLS VPN 中实施组播时使用较多的是组播虚拟专用网(MVPN)算法,该算法有很多优势,但致命缺点是不能提供最优的路由。因为它要求所有连接到该客户边缘路由器 CE 的服务提供商边界路由器 PE 都要接收该客户的组播数据包,而不管那些 PE 分支上是否真的有该组播的接受者,浪费了一定带宽。文献[1]指出了该算法的另外一个缺点,即 PE 路由器维持着每一个邻接 CE 的协议无关组播(PIM)状态,加重了 PE 的负担。MPLS 技术对组播支持的最新研究是基于 PIM 协议的 MPLS 对组播的支持,该算法的优点是:一般的路由器组播表为一个<入标签,入接口>对应一个<出标签,出接口>,而分支点的路由器组播表为一个<入标签,入接口>对应多个<出标签,出接口>。本

文参考了文献[2],把基于 PIM 的 MPLS 组播算法的思想运用到 MPLS VPN 中,提出一种基于 PIM-SSM(指定信源的组播)的 MPLS VPN 组播算法。

2 基于 PIM-SSM 的 MPLS VPN 组播算法

2.1 算法原理

基于 PIM-SSM 的 MPLS VPN 组播算法的原理如图 1 所示。首先 PE₂ 一侧的 VPN₁ 有一个组

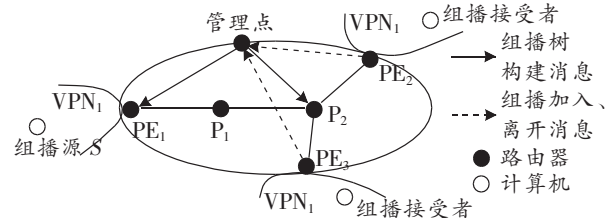


图 1 基于 PIM-SSM 的 MPLS VPN 组播算法原理图

播接受者想接受组播组 G 的组播,它通过某种途径知道组播组 G 所在的组播源 S 的地址,SSM 的特性决定了组播接受者知道组播源地址。组播接受者通过 IGMPv3(Internet 组管理协议)向所连路由器发送成员身份报告消息,所连路由器将成员身份报告消息转换成 PIM 的加入消息,并发送到 PE₂。PE₂

收到加入消息后,在组播虚拟路由转发表(MVRF)中生成 (S,G) 组播表(见图2),并把加入消息、 VPN_1 的唯一标识符(RD)和路由目标(RT)一起发送给组播管理点(Manager)。管理点收到上面的信息后生成相应的表格(见图2),并根据组播源 S 的地址在 VPN_1 的虚拟路由转发表(VRF)中找到组播源 S 所在的PE路由器 PE_1 ,然后发送组播树构建消息给 PE_1 ,告诉 PE_1 发送组播数据到 PE_2 。 PE_1 收到组播树构建消息后在MVRF中生成 (S,G) 组播表(见图2),与 PE_2 生成的组播表不同的是出接口为到 PE_2 的标签交换路径(LSP)。

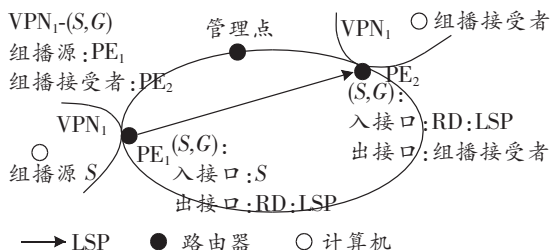


图2 当 PE_2 有组播接受者加入后路由器形成组播表

这时 PE_3 所在的 VPN_1 内也有组播接受者想加入组播组 G ,同样加入消息到达 PE_3 , PE_3 在MVRF中生成 (S,G) 组播表(见图3),并把加入消息、 VPN_1 的RD和RT一起发送给管理点。管理点收到上面的消息后,查找 VPN_1 组播表,发现已有 (S,G) 表项,则在组播接受者表项上加入 PE_3 。管理点同时在MPLS域的路由表上计算 PE_1 到 PE_2 和 PE_3 的最短路径,并计算出两者的重复路径,定位出分支点路由器 P_2 (见图3),然后发送组播树构建消息给 P_2 。 P_2 收到消息后生成组播信息表MFT^[3],当 P_2 收到 PE_1 通过 LSP_1 发送过来的组播数据包时,就会复制两份,分别通过 LSP_2 、 LSP_3 发送给 PE_2 和 PE_3 (见图3)。

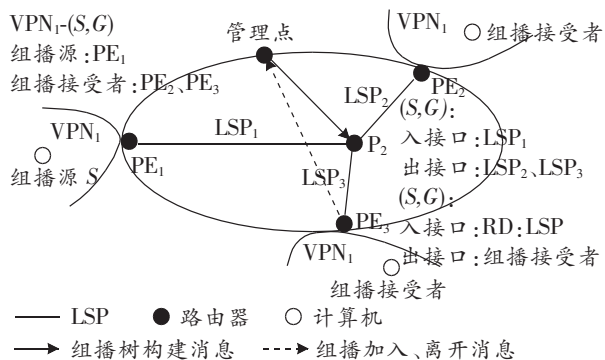


图3 当 PE_3 有组播接受者加入时路由器形成组播表

每次有组播接受者加入或离开导致PE路由器发生变化时,管理点都会重新计算组播树,并向相应的路由器发送组播树构建消息,让网络重新生成新的组播树。例如,若 PE_4 有组播接受者加入,那么管理点重新计算组播树,并选择新的分支点,同时通过消息去除老的分支点,生成新的分支点。

2.2 算法扩展

在实际网络环境中必然存在多个VPN,每个VPN又有多个组播,网络的规模也一定很大,管理点既要保存全局和各个VPN的路由表、组播表,还要进行消息收发和计算组播树,负担繁重。如果管理点运行速度慢、反应不过来或死机,那么整个网络的组播运行就会瘫痪。针对上述情况,本文提出一种扩展算法,该算法一方面提升了管理点处理VPN中有多个组播,以及网络中有多个VPN且每个VPN都有自己的组播的情况的能力;另一方面通过采用备份和多管理点的方法可以增加网络的冗余和健壮性。

(1) 当一个VPN内有多个组播时,管理点需要在 VPN_1 的组播表中生成两个组播项 (S_1, G_1) 和 (S_2, G_2) ,为 VPN_1 生成两棵组播树,见图4中组播树1和组播树2,选择了两个分支点 P_2 和 P_3 ,当 PE_3 从MPLS域中收到组播数据包后,去除其标签,并根据组播源和组播组地址发送给相应的组播接受者。

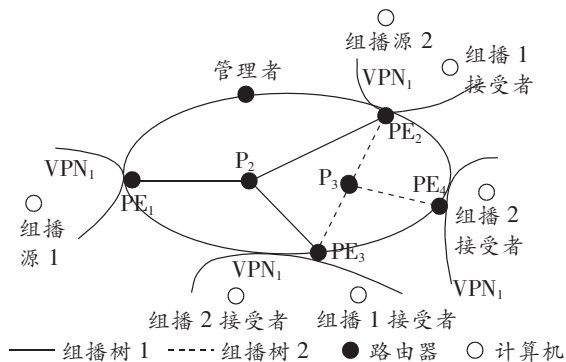


图4 一个VPN内有多个组播

(2) 当多个VPN内有组播时,管理点需要在 VPN_1 和 VPN_2 中分别生成组播表 (S_1, G_1) 和 (S_2, G_2) ,为 VPN_1 生成组播树1,为 VPN_2 生成组播树2,分别为两棵组播树选择了两个分支点 P_2 和 P_3 ,分支点路由器并不需要知道组播属于哪个VPN,它只需根据LSP来复制组播数据就可以了。当 PE_3 从MPLS域中收到组播数据包后,根据边界网关协议(BGP)的标签来决定组播数据包属于哪个VPN,然后发送给相应的组播接受者,如图5所示。

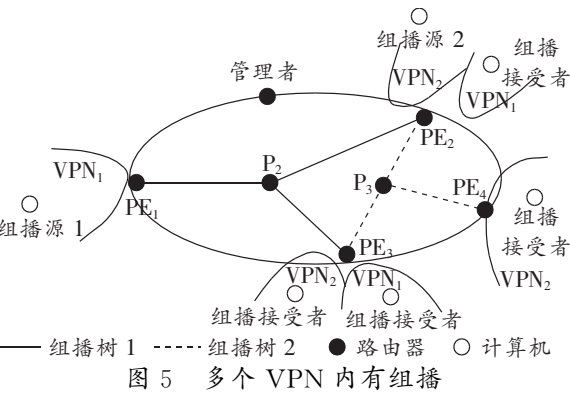


图 5 多个 VPN 内有组播

(3)管理点在算法中起着非常重要的作用,如果管理点死机,则整个 MPLS 域内的组播都将无法进行,所以需要备份管理点。在图 6 中,管理点 2 是备份组播管理点,PE 路由器也会向管理点 2 发送组播加入、离开消息,但管理点 2 在管理点 1 运行正常的情况下不发送组播树构建消息,只有在管理点 1 死机后,才会代替它成为组播管理点。管理点 2 用保活消息(Keepalive)来知道管理点 1 的状态。管理点 1 死机后恢复运行时,可以用高优先级方式抢占组播管理点位置。例如,当 PE₁ 收到所连的一个 VPN 中的一个组播加入消息时,查找自己的配置发现有两个管理点,就分别向两个管理点发送加入消息;管理点 1 和 2 收到加入消息后都进行组播树计算,由于组播数据库一样,它们的计算结果也一样,但只有管理点 1 向组播树上的分支点路由器发送组播树构建消息,如图 6 中的 P₂ 收到了管理点 1 的构建消息,在网络中形成这个 VPN 的组播树。管理点 1 和 2 会每隔 60 s 发送一次保活消息,若管理点 2 在 3 倍保活时间内没收到管理点 1 的保活消息,就认为管理点 1 出现故障,此时若 PE₃ 也发送了这个 VPN 中的一个组播加入消息,那么管理点 2 收到加入消息后不仅进行组播树计算,还向相应的分支点路由器发送组播树构建消息,即管理点 2 代替了管理点 1 成为此时网络的组播管理点。

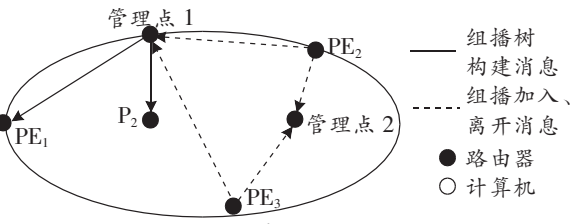


图 6 管理点冗余

(4)当网络中有多个 VPN,一个管理点管理不过来时,就需要多个管理点分别管理各自 VPN 内的组播。如图 7 所示,PE₁、PE₂ 和 PE₃ 在对 VPN₁ 的配置中手工指定组播管理点为管理点 1,在对

VPN₂ 的配置中手工指定组播管理点为管理点 2。当 VPN₁ 中有组播时,就会通过 PE 向管理点 1 发送组播加入消息,管理点 1 计算组播树后发送组播树构建消息到相应的分支点路由器 P₂,形成相应的组播树。而当 VPN₂ 中有组播时,PE 向管理点 2 发送组播加入消息,管理点 2 计算组播树后发送组播树构建消息到相应的分支点路由器 P₃,形成相应的组播树。由于管理点 1 和管理点 2 分别管理不同的 VPN,相互之间并不会产生冲突,它们各自的管理负担也会大大减轻。

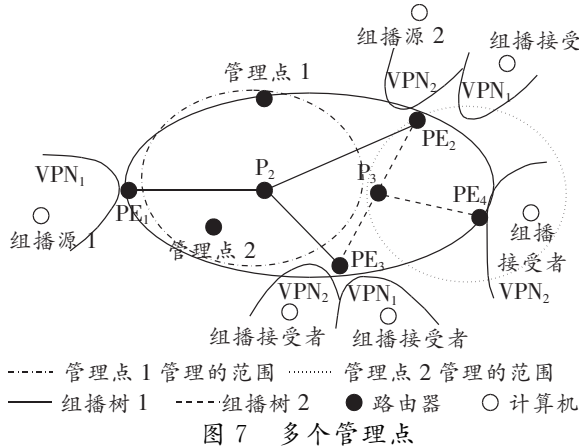


图 7 多个管理点

3 结束语

由于本算法在 MPLS 域内只有分支点核心路由器需要维护组播表和复制组播数据,其他核心路由器并不需要加入组播树等工作,所以分支点核心路由器的个数在总核心路由器个数中所占比例越小,采用本算法的优越性就越大,大部分核心路由器不需要保存组播状态信息,实现组播无状态,从而极大地提高了多播的扩展性,节约了核心路由器资源。

参考文献:

[1] 黄冰, 刘文霞. MPLS VPN 组播的研究与仿真 [J]. 宽带网络与传输, 2006, (5): 67-69.

[2] Striegel A, Bouabdallah A, Bettahar H, et al. A New Approach for Scalable DiffServ Multicasting [DB/OL]. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.2.8301>, 2003-03-30.

[3] 鲍丽红, 王昱, 韩双旺. 可扩展的基于源树的 MPLS 组播算法 [J]. 计算机工程与设计, 2007, 28: 4948-4954.

[4] Stoica I, Eugene NTS, Zhang Hui. REUNITE: A recursive unicast approach to multicast [DB/OL]. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.6895>, 2000-03.