

MPLS 环境下组播路由研究^①

Multicast Routing in MPLS

杨振启¹ 张 敏² (1.南京信息工程大学 江苏 南京 210043; 2.曲阜师范大学 山东 日照 276826)

摘要: 组播和 MPLS 可看作是两种互补的技术。通过在 MPLS 中建立组播树将两者结合起来, 不仅有效的解决了 IP 组播的可扩展性问题, 还能获得 QoS 和带宽节约两方面的好处。本文基于 PIM-SM 路由协议, 重点研究了实现 MPLS 组播的关键, 即如何建立 MPLS 组播树和 MPLS 组播数据包的传输, 并在次基础上对 MPLS 环境下的组播发展进行了展望。

关键词: MPLS 组播 组播树 标记

随着 Internet 的迅猛发展, 各种音频和视频等多媒体业务也已得到广泛应用。这些实时多媒体应用往往占用大量的网络资源并在服务质量 (Quality of Service QoS) 方面有了更高要求。例如, 网络视频应用 VoD 点播 (Video on Demand) 要求网络能保证一定的带宽; 视频会议、IP TV、IP 电话对端到端延迟和抖动有严格的限制。因此, 如何保证这些业务所需带宽的同时保证它们所需的服务质量成为目前研究的热点。以上应用都涉及到组播技术, 作为一种实用技术, 全球主要的网络设备厂商, 电信运营公司和 Internet 服务提供商都积极地参与和推动它的发展。组播是通过共享带宽将单一的 IP 数据一次发送到一组客户, 从本质上实现带宽的有效节约, 同时还具有匿名性、限制网络延迟等优点。多协议标记交换 MPLS (Multi-protocol Label Switching) 体系结构, 作为一种成功的电信网络技术与计算机网络技术的融合, 充分利用了网络边缘的 IP 路由和网络核心的交换技术, 不再查找路由表而是仅靠简单的标记进行数据的转发处理, 支持了多种混合的协议并能为用户提供端到端的 QoS 保障。因此组播和 MPLS 可以看作两种互补的技术, 将 MPLS 和组播很好的融合起来具有重要意义。

1 MPLS组播概述

1.1 多协议标记交换 MPLS

20 世纪 90 年代, 针对如何将终端用户的 IP 数据

在 ATM 上传输这一问题, Cisco、Ipsilon、IBM、Toshiba 等几家公司均提出了自己的新技术及新产品。MPLS 技术是 IETF 小组融合各公司的不同技术, 将他们进行统一而提出的。MPLS 技术是一种交换技术, 它的实质是边缘路由, 核心交换。将复杂的路由工作放置到网络的边缘, 在网络内部只进行简单的标记交换实现数据的转发处理。IP 数据在边缘入口路由器依据所属的转发等价类 (Forwarding Equivalence Class, FEC) 被打上标记进入 MPLS 网络, 在 MPLS 域内每到达一个标记交换路由器, 路由器就剥离数据自身携带的标记, 为它打上新的标记并转发到下一跳。通过标记建立起来的从网络入口到网络出口的路径就是标记交换路径 (Label Switching Path, LSP)。当数据到达出口路由器时, 数据所携带的标记被剥离, 之后进行传统的第三层路由转发。

1.2 组播技术

组播技术在 Internet 发展中一直有很好的应用, 一方面网络运营者们希望通过优化网络资源的利用率, 利用有限的网络资源获得更大的经济利益; 另一方面人们对实时多媒体业务的需求也越来越大。使用组播技术, 可以有效节约带宽, 因为组播中使用了组的概念, 把用户分成了不同的组, 每次只把数据传送到组成员的用户而不是所有的用户; 对于同一组的多用户, 源端也只需发送一份数据而不是复制多份数

^① 收稿时间:2008-06-25

据,从而大大提高了网络数据的传送效率,并能改善网络的性能。

历经 20 多年的研究和发展,IP 组播已经形成了较为完整的组播协议体系,包括组播地址分配,组管理协议,组播路由协议等。其中 IGMPT (Internet Group Management Protocol) 协议运行于主机和与之直接相连的组播路由器之间,用于建立、维护组成员关系,实现对组成员的管理。组播路由协议运行于组播路由器之间,负责构造组播树和路由组播数据包。组播路由协议可以分为密集模式协议(距离向量组播路由协议 DVMRP、最短路径优先组播协议 MOSPF、协议无关组播-密集模式 PIM-DM)和稀疏模式协议(协议无关组播-稀疏模式 PIM-SM、核心树协议 CBT)。每种协议都在 Internet 上有所应用,其中 PIM-SM 协议使用最为广泛。

组播路由协议的不同在于建立的组播树类型不同。有两中不同的组播树:有源树和共享树。有源树 (S,G) 由一个源 S 和一个组 G 共同决定;共享树 (*,G) 仅由一个组 G 决定,可以有多个不同的源。有源树的成员(主机与路由器)都保留有所属树的一个 (S,G) 状态,其适用于组播成员密集的情况。共享树则适用于组播成员稀疏的情况,每一个共享树都选有或指定一个或多个 RP(汇聚点)。所有属于该共享树的结点在相应的接口上都创建 (*,G) 状态,所有加入此共享树的组播源都把组播数据发往此共享树的 RP,由 RP 再通过共享树发往接收节点。

1.3 MPLS 组播

MPLS 网络中实现组播不仅能够减少组播流量,而且还可以在速度以及服务质量等方面提高性能。实现的关键是在 MPLS 中应用组播路由,通过建立组播树将两者结合起来即把组播树映射为 LSP 树,以提高网络性能并有效的解决组播可扩展性等问题。因为 PIM-SM 协议是应用最广泛的组播协议,所以我们重点研究了 MPLS 网络中使用 PIM-SM 协议的情况,这些工作对于其他的组播协议同样有益。

2 MPLS组播技术分析

2.1 MPLS 组播树的生成

有源树的生成过程是:源节点首先假定所有的路由器或主机都要加入该组,将组播数据以扩散的方式发送,这样所有要加入该源组的主机和路由器都会在

相应接口建立 (S, G) 转发状态,如果某个网段没有该组成员,路由器将向上行路由器发送剪枝消息,之后该路由器将不再转发任何该源组播树的分组数据,但由于泛洪和剪枝是周期性的,所以间隔一定时间后又会重新自动嫁接上去。

共享树的生成要求所有要加入指定组 G 的节点都向 RP 节点发送显式加入消息:通过使用特定的 IP 组播地址 G,发送加入消息到直连路由器,如果该路由器已属于 (*,G) 树,则在相应接口上建立一个 (*,G) 转发状态;如果该路由器还没有加入 (*,G) 树,则一般是利用 RPF(反向路径转发)接口向 RP 发送组 G 的加入信息,直到加入组播树,这样一棵共享树 (*,G) 最终就被生成了。

有源树中一个 (S,G) 对就要用掉一个标记,而共享树中一个组对应一个标记。所以使用共享树比有源树消耗的标记要少。

以图 1 为例介绍 MPLS 组播树的生成过程如下:假设组播源 S1 与 LSR1 直连,LSR1 为组播树的根。当用户 H4 要接收组播 G 的数据时,H4 要向 LSR3 发送加入 G 的消息,该加入消息在 LSR3 的 if2 上收到,LSR3 在 if2 上建立 (S1,G) 状态,并在本地 FIB 中添加一条项目。此时 LSR3 并未加入 MPLS 组播树中,LSR3 从接口 if1 向根发送加入消息,并在本地生成唯一的标记 1 作为输入标记对应于 (S1,G) 状态,LSR1 在接口 if2 上收到 LSR3 发来的加入请求后,以标记 1 为输出标记并在 LIB 中添加一条相应条目。同理,H1 或 H2 向 LSR4 发送加入请求,LSR4 再向 LSR2 发送加入请求,假设 LSR1 在接口 if1 收到 LSR2 发来的请求绑定标记为 2,则 LSR1 把标记 2 作为接口 if1 对应 (S1,G) 的输出标记,并在 LIB 中添加一条相应条目。这样,一颗 MPLS 组播树 (S1,G) 的 LSP 即被建立了。

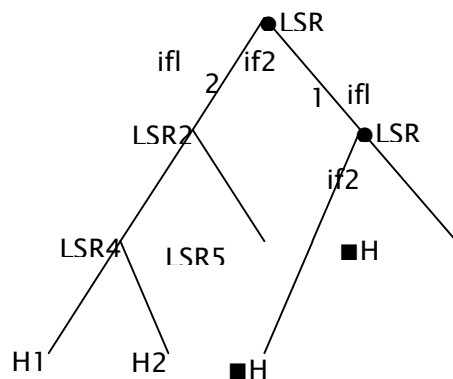


图 1 MPLS 组播树生成示例图

2.2 MPLS 组播 LSP 的建立方式

MPLS 单播的 LSP 建立驱动方式有三种：拓扑驱动、数据流驱动和需求驱动。拓扑驱动是在数据流到达前就建立好 LSP；数据流驱动只在某数据流到达时，LSR 才为其建立一个状态；需求驱动是指 LSP 的建立由控制消息驱动。由于 MPLS 组播的组播组成员加入和离开都是动态的，生成的组播树也是动态的，所以无法象单播一样采用拓扑驱动的方式提前建立好 LSP。对于采用泛洪剪枝模式的组播协议，由于路由是数据流驱动的，所以相应的 LSP 建立也是数据流驱动的，PIM-SM 协议亦是如此。MPLS 组播与 MPLS 单播有很大不同，会消耗大量的标记和信令及时间开销。

2.3 节点的信息表

一个 MPLS 节点保存有三张信息表：LIB (标记信息表 Label Information Base)，PFT (局部转发表 Partial Forwarding Table)，ERB (显式路由信息表 Explicit Routing information Base)。在 MPLS 中实现 PIM-SM 协议，需要做两种映射：(S, G) 树与 (入标记, 入接口) 的映射和 (入标记, 入接口) 与多个 (出标记, 出接口) 的映射。对于第一种映射我们需要建立一个新的信息表 LSG 完成对给定源和组的标记的定义。LSG 表包含四部分：入标记, 入接口, 源端, 组。组播数据包与单播数据包有同样的结构但是 MPLS 节点会使用 LSG 表从 IP 目的地址中确定是否为组播数据包。当有组成员加入或者离开组播组 (S, G) 的时候相应的 LSG 条目填充或者删除；对于第二种映射则不必建立新的表，只需对 LIB 表用同样的 (入标记, 入接口) 来完成不同的 (出标记, 出接口) 的多次填充。

2.4 MPLS 组播包的传输

MPLS 组播包的传输在非分支节点与单播的传输一样，LIB 表都是一对一映射，一个 (入标记, 入接口) 映射到一个 (出标记, 出接口)，过程如图 2 所示；而在组播树的分支节点，LIB 表是一对多的映射，一个 (入标记, 入接口) 映射到多个 (出标记, 出接口)，需要在每个接口对数据包进行复制，并且对每个副本都进行一次标记交换，然后依据相应的出接口把副本数据包发送到下一跳，过程如图 3 所示

3 总结与展望

虽然 IP 组播可以提供高效的尽力而为的数据传送服务，但它的可扩展性比较差，MPLS 技术与组播技术的结合不仅可以有效的解决 IP 组播的可扩展性问题，还能获得 QoS 和带宽节约两方面的好处。本文在 PIM-SM 路由协议的基础上，重点研究了实现 MPLS 组播的关键，即如何建立 MPLS 组播树和 MPLS 组播数据包的传输。今后，诸如将单播数据流与组播数据

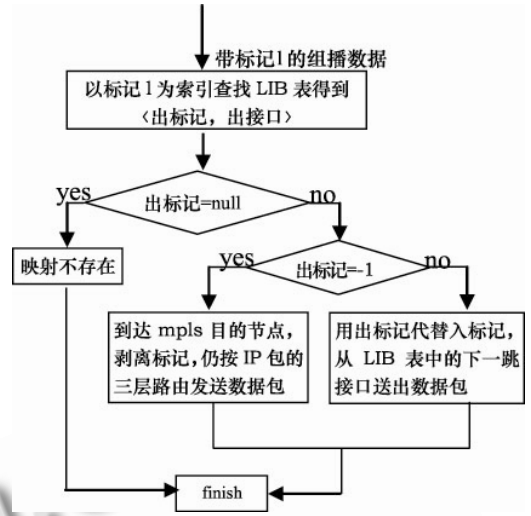


图 2 非分支节点的 MPLS 组播包传输过程

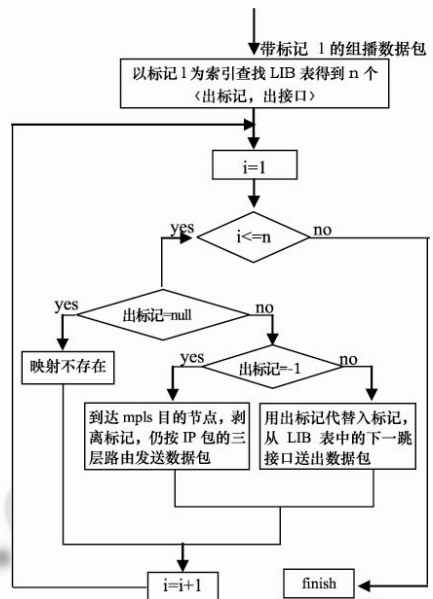


图 3 分支节点的 MPLS 组播包传输过程

流聚合以及组播数据流间的聚合、组播 vpn 在 mpls 中的应用等还有待进一步的研究。

参考文献

- 1 赵宜成,王耀勇.多协议标记交换(MPLS)技术研究.计算机工程与应用,2003,(24).
- 2 Behrouz A.Forouzan,Sophia Chuang Fegan. TCP/ IP Protocol Suite(Second Edition) 2003.
- 3 RFC3353-Overview of IP Multicast in a Multi-Protocol Label Switching (MPLS) Environment. 2002.