

多协议标记交换(MPLS)技术研究

赵云栋 (河北省财政局信息中心)

摘要:主要介绍了多协议标记交换(MPLS)技术的基本概念、网络构成和工作原理,在此基础上,提出了实现 MPLS 交换机路由器的关键技术。

关键词:标记 转发等效类 标记分发协议(LDP) 多协议标记交换(MPLS)

1 MPLS 的基本概念

现有互连网是由路由器加专线构成的,IP 数据包在所经过的每个路由器节点都要进行目的 IP 地址的最大匹配和选路,因此要求每个节点的路由器功能及能力都很强。但随着目前线路速率和通信业务量的增长,负责寻路和转发功能的路由器已成为网络的瓶颈。为了解决传统路由器转发速度低这一问题,业界采取了一系列提高路由器转发速度的措施,如采用效率更高的算法和数据结构、速度更高的处理器以及一些专用集成电路等,但这些措施增加了每个节点的成本,因而使得建网费用较高,而多协议标记交换提供了更好的解决方案,它把整个网络的节点设备分为两类:即标记边线路由器(LER)和标记交换路由器(LSR),由 LER 构成 MPLS 网的接入部分,LSR 构成 MPLS 网的核心部分。LER 发起或终止标记交换通道(LSP)连接并完成传统 IP 数据包转发和标记转发功能。入口 LER 完成 IP 包的分类、寻路、转发表和 LSP 表的生成、FEC(转发等效类)至标记的映射。出口 LER 终止 LSP,并根据弹出的标记转发剩余的包。LSR 只是根据交换表完成转发功能。这样所有复杂功能都在 LER 内完成,LSR 只完成高速转发功能。其优点是网络简单、易于扩展、建网费用低。

1.1 MPLS 的封装

为了能够在数据包中携带标记栈信息,需要在封装中引入标记栈字段。MPLS 为底层的链路规定了不同的填充装格式。在 MPLS 专用的环境中,也就是在以太网和点到点这样的链路上,采用了“shim”封装。shim 位于第三层协议和第二层协议头之间,与两层协议无关,因而被称为通用 MPLS 封装。通用 MPLS 封装包括标记栈、TTL 和 CoS(分类、排队、规划)等字段。

1.2 MPLS 的层次化结构

层次化结构对于提高网络的可扩展性是非常重要的。对于 MPLS 而言,要实现层次化结构,有三项工作必须完成:一是标记分发协议必须能够透过其他节点传递标记信息;二是必须在数据包中传递标记栈信息,也就是说数据包必须同

时携带多层标记信息;三是必须让 LSR 知道何时可使用分层标记,也就是必须配置路径的两个端节点,以便它们能够完成“换层”的处理。第一项和第三项工作都由标记分发协议(LDP)完成。

2 MPLS 的网络构成

MPLS 网络由标记边线路由器(LER)和标记交换路由器(LSR)组成。通过标记分发协议(LDP)在节点间完成标记信息的分发。在 LER 之间、LSR 之间、LER 与 LSR 之间依次需要运行路由协议(如 OSPF、BGP 等),它们通过路由协议来获取网络拓扑信息,并根据这些信息决定数据包的下一跳(在第三层转发时)或交换路径如何建立(在第二层转发)。

一个路由域内可以同时包含 MPLS 节点和非 MPLS 节点,标记信息只在 MPLS 的相邻节点间传递。MPLS 的两个节点间若有非 MPLS 的 ATM 或帧中继交换机,则以 PVP 或 PVC 将两个 MPLS 节点连通。

3 MPLS 工作原理

建立 MPLS 通信网的目的就是为连接到该网的用户建立一条交换路径,MPLS 支持四种交换路径:点到点、多点到点、点到多点和多点到多点。建立交换路径涉及标记分配和标记分发两方面的内容。

3.1 标记分配

标记分配分为上游分配、下游分配和按需下游分配三种。LDP 只规定了其中两种:下游分配和下游按需分配。下游分配模式是由下游节点发起标记的分配和分发过程,而下游按需分配模式则是由上游节点发起申请,由下游节点完成标记的分配和分发。在 ATM 环境下通常使用下游按需分配模式。

不论何种分配模式,都存在如何启动其分配的问题。也就是标记分配的驱动力问题。MPLS 使用了三种标记分配的驱动力:拓扑驱动、基于申请的控制驱动和业务流驱动。

拓扑驱动是指依靠路由协议控制信息来驱动标记的分配,进而完成交换路径的建立。因为标记是预分配的,路径是预建立的,所以在数据到达时可以以快的速度转发。使用拓扑驱动的技术有标签交换、ARIS 等。

基于申请的控制驱动是指由基于申请的控制信息驱动标记的分配过程,这种驱动方式与拓扑驱动一样具有路径预建立的优点,但比拓扑驱动耗费更多的标记资源,这种方式非常适用于有特殊质量需求的业务。基于申请的控制业务流有 RSVP。

业务流驱动的优点是可以避免不必要的标记资源的占用,但因为业务流驱动是业务流到达时临时建立交换路径和进行标记分配,所以部分数据包的转发速度会比较低,同时它还要求交换机有较高的处理能力(建立路径的同时还需处理数据转发)。CSR 和 Ipsilon 使用的是业务流驱动的方式。

3.2 标记分发

MPLS 使用了两种标记分发方式,即用控制消息携带和使用专用标记分发协议。虽然使用专用标记分发协议会使路由的分发和标记的分发分离,但是因为使用控制消息存在诸如无法协商标记的可用范围、无法处理无效的标记分配等问题。因此建议采用专用标记分发协议(LDP)。

(1) 标记分发过程。LDP 用四类消息完成标记的分发过程,它包括发现消息、会话消息、公布消息、通知消息。标记分发协议的运行也可以分为三个阶段:

① 发现阶段:在这个阶段里,LSR 可以自动发现它的 LDP 对等,而无需进行人工配置。LDP 将发现机制分为两种:一是基本发现机制,用来发现本地的 LDP 对等;二是扩展的发现机制,用来发现远地的 LDP 对等。

基本发现机制通过在 LDP 键路上周期性地发送“Hello”包来通知相邻节点本地对等关系。这个“Hello”按 UDP 发送,目的协议地址为所有路由器的组播组地址,UDP 端口号为公认的 LDP

发现端口号。“Hello”包的接收方会将这个对等关系标识为链路层可达。

扩展的发现机制则由 LSR 周期性地发送特定 IP 地址的“Hello”包。“Hello”按 UDP 发送,目的协议地址为特定的 IP 地址,UDP 端口号为公认的 LDP 发现端口号。“Hello”包的接收方会将这个对等关系标识为网络局可达。

② 会话建立阶段:在两个 LSR 交互发现“Hello”包后,就会启动 LDP 会话建立阶段。会话建立阶段又划分为两个子阶段——建立传输层连接和初始会话。建立传输层连接就是在两个 LSR 之间建立 TCP 连接,而初始会话就是通过交互 LDP 初始化消息来协商会话的参数,包括 LDP 协议版

本、标记分发的方法、定时器值、标记的范围(当底层使用 ATM 网络时是 VPI/VCI 范围,当底层使用帧中继网络时则是 DLCI 范围)等。

③ 会话的维持和删除:LDP 通过监测会话传输连接上 LDP PDU 的接收情况来判定会话连接是否完整。一个 LSR 为每个会话保持一个“保持存活”定时器,当从该会话上接收到一个 LDPPDU 时,“保持存活”定时器会刷新。如果“保持存活”定时器到时,则 LSR 认为该传输连接中断,该对等失效,并由此关闭传输层连接,从而终止这个对等会话。

(2) 标记分发控制模式。标记分发控制有独立的标记分发控制和有序的标记分发控制两种模式。

① 独立标记分发控制是指每个节点可以在任何时候向其相邻节点分发标记映射信息。使用这种模式可能出现这样的情况,在一条 LSP 上的各个节点还未完成标记分发的时候,该 LSP 就已开始传输数据,这些数据在传输中途被丢弃或以路由的方式继续传递;

② 有序标记分发控制是指 LSR 只在两种情况下才发起标记映射的传输,一种情况是该标记映射已经拥有了到下一跳的标记映射,另一种情况是该 LSR 是出口 LSR。其他情况下,LSR 必须等待从其下游 LSR 收到标记映射后,才能向其上游分发它的标记映射信息。使用有序标记分发控制,可以保证标记分发的完整性和一致性。也就是说,在入口 LSR 收到标记分发信息时,就可以肯定这条 LSR 上的所有节点都已完成标记分配和分发过程。

(3) 标记保持模式。MPLS 将只记录下一跳所分发的标记信息的处理模式称为保守的标记保持模式。因为在这种模式下会丢弃非下一跳分发的标记信息,所以在路由发生改变时,就必须借助下游节点按需分配向新的下一跳 LSR 申请标记。正因如此,保守的标记保持模式通常和标记通知模式中的下游节点按需分配模式一起使用。采用这种模式的优点是可以节省存储器空间。

MPLS 将记录所有对 LSR 分发的标记信息的处理模式称为自由的标记保持模式。采用这种模式,在路由发生转换时,可以在所记录的标记信息中找出新的下一跳节点曾分发的信息,所以可以加快路由转换的速度。

(4) 循环控制。因为标记的分配和交换路径的建立可以逐段完成,也就是说,建立路径时只考虑了相邻节点的情况,所以从整条路径上看,存在出现循环的可能,特别是在路由过渡时期。因此,循环处理对于 MPLS 是非常重要的。循环处理分为三种:减轻循环、防止循环和检测循环。

减轻循环就是利用 TTL 字段等手段将循环所造成的影响降到最小。利用 TTL(生命周期)字段可以将循环的数据包在一定时间内丢弃,避免它们占用过多的网络资源(带宽和

处理器资源)。类似的手段还有动态路由和公平排队等。因为循环数据通常会阻塞通路,从而降低路由控制信息传递的速度,也就同时降低了该条路由的收敛速度,所以动态调整后的路由就会收敛在没有循环的路由上。而公平排队则是将属于不同流的数据分别排队,在某些流出现循环后,不会影响其他业务流的传递。

检测循环就是在建立路径时不加控制,允许有循环的路径建立,但是进行检测,发现循环后再将其删除。检测循环的一种方法是在路径发生改变时,沿着新的路径向目的地发送循环检测控制包。因为要做循环检查,所以沿途各节点不能以通常的第二层交换方式传递这个控制包。在传递过程中,如果出现 TTL 到时或发送者收到这个包,则判定为路径循环,该路径将被舍弃。

防止循环是指境保建立交换路径时不出现循环。防止循环的一种方法是在从下游向上游分发标记信息的控制消息中包含所经历的路径信息表,每个上游节点在收到这个消息时都要做循环检查(检查路径信息表中是否包含本节点),检查到循环则将该控制消息删除,循环路径便不会形成。防止循环的另一种方法是使用显式路由,即在建立路径的发起端就指定路径所经历的每个节点,只要这个发起端具有足够智能,就不会出现循环路径,但是显式路由要求节点了解全网路由信息(而不只是下一跳信息),只有 OSPF 和 BGP 能够支持这种路由方式。

4 实现 MPLS 交换机的关键技术

以 ATM 交换机作为硬件平台,ATMMPLS 多协议标记交换机由两部分组成:一部分是互连网路由模块,另一部分是面向连接的转发模块。路由模块包括:支持 OSPF 和 BGP 的路由协议功能模块、TCP/IP 协议栈、LDP 及其运行结果、用于标记分发和分配的 LIB 转发模块采用 ATM 交换矩阵。

为了实现 MPLS-ATM 交换机,需解决如下关键技术:

- (1) 如何决定标记分配的流的分类粒度;
- (2) 选择何种标记分配的驱动力;
- (3) 选择何种标记分发方式;
- (4) 选择何种标记分发控制方式;
- (5) 选择何种标记保持模式;
- (6) 选择何种循环控制方式;
- (7) 组播的实现方法;
- (8) 如何管理每条 LSP 上的 QoS 特性。

上述各类技术的选择取决于实现的难易程度、网络的可扩展性、与现有网络的兼容程度、能否兼容多种路由协议、能

否易于管理网络资源等。

在 ATM 交换机作为 LSR 的环境中,仍然沿用 ATM 的信元封装,标记栈信息也就相应地有了一定的改变。MPLS 规定了三种对 ATM 信无头的编码:

① ATM SVC 编码:标记栈是用 VPI 和 VCI 字段共同组成的一层标记;

② ATM SVP 编码:标记栈是用 VPI 和 VCI 字段分别构成的两层标记,VPI 位于栈顶,VCI 位于栈底;

③ ATM SVP 多点编码:标记栈是用 VPI 和 VCI 组成的两级栈,但 VCI 只使用了原来 VCI 字段的一部分,VCI 字段的剩余部分用来标识 LSP 的入口。

如果在 ATM-LSR 环境下需要多于两层的标记,则只能将 ATM 的情无封装与通用。MPLS 封装相结合。标记分配的驱动力采用拓扑驱动方式。采用拓扑驱动方式时,由于标记是预分配的,路径是预建立的,所以在数据到达时可以以最快的速度转发。由于标记分配仅受控制信息的控制,而不是数据与控制信息的混合控制,从而简化了整个系统的行为。

标记分发方式采用专用分发协议(LDP),虽然使用专用标记分发协议会使路由的分发和标记的分发分离,但是采用 LDP 易于协商标记的可用范围、处理无效的标记分配等等问题。而且 LDP 协议建议采用受限路由——标记分发协议(CR-LDP),CR-LDP 易于和 ATM 兼容,可以使用非最短路径,以充分利用网络资源,但存在如何寻找最短路径路由问题。

标记分发控制采用有序标记分发控制方式,这样可以保证标记分发的完整性和一致性。

即,在入口 LSR 收到标记分发信息时,就可以肯定这条 LSP 上的所有节点都已完成标记分配和分发过程。同时 VCI/VPI 可以更有效地利用。

标记保持模式采用保守的标记保持模式。因为在这种模式下会丢弃非下一跳分发的标记信息,所以在路由发生改变时,就必须借助下游节点按需分配向新的下一跳 LSR 申请标记。所以保守的标记保持模式通常和标记通知模式中的下游节点按需分配模式一起使用。采用这种模式的优点是可以节省存储器空间,且易于标记管理。

循环控制方式采用防止循环方式的显式路由,即在建立路径的发起端就指定路径所经历的每个节点,只要这个发起端具有足够智能,就不会出现循环路径,但是显式路由要求节点了解全网路由信息,而不只是下一跳信息,只有 OSPF 和 BGP 能够支持路由方式。