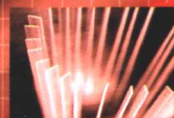


MPLS MPLS MPLS MPLS

MPLS MPLS MPLS MPLS MPLS MPLS MPLS MPLS MPLS



宽带 **Zone** 丛书

MPLS 技术与实现

李晓东 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

www.phei.com.cn

宽带 **Zone** 丛书

MPLS 技术与实现

李晓东 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

运营商骨干网络从时分复用的专线网络发展到了 ATM 网络,又正在从 ATM 网络过渡到 MPLS。虽然 MPLS 技术还没有最后标准化,但是在业界已经颇有些山雨欲来风满楼的气势。

本书分为三个部分。第一部分是全书的“引子”,即多协议标记交换 MPLS 的基础,包括 MPLS 概览、MPLS 体系结构;第二部分是全书的“内核”,即多协议标记交换 MPLS 的控制,MPLS 是如何完成诸多功能,如何成为运营商网络的核心技术的;最后一部分,也是本书的画龙点睛之笔,集中了目前多协议标记交换 MPLS 在运营商网络的主要应用,包括服务质量(QoS)保证,流量工程(Traffic Engineering),虚拟专用网(VPN)以及新的 MPLS 在光网络上的应用和在以太网快速恢复中所起的作用。

本书适合于运营商、服务提供商、网络技术人员、市场人员阅读,也可作为大专院校的选修课教材。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

MPLS 技术与实现 / 李晓东编著. —北京: 电子工业出版社, 2002.12

(宽带 Zone 丛书)

ISBN 7-5053-8274-8

I. M… II. 李… III. 宽带通信系统—计算机通信网—通信技术 IV. TN915.142

中国版本图书馆 CIP 数据核字(2002)第 093241 号

责任编辑: 周晓云

印 刷: 北京四季青印刷厂

出版发行: 电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×980 1/16 印张: 17.75 字数: 360 千字

版 次: 2002 年 12 月第 1 版 2002 年 12 月第 1 次印刷

印 数: 5000 册 定价: 29.00 元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。联系电话: (010) 68279077

前 言

虽然 MPLS 技术仍然还没有最后标准化,但是在业界已经颇有些山雨欲来风满楼的气势。人们期待着又一个全新运营商网络核心技术的出现,期望这一关键技术可以改善目前网络的运营状况,提高网络的效率,甚至希望通过这一关键技术开发出一些新的增值业务,如 IP VPN, IP CoS (Class Of Service) 等,以赚取更大的利润,满足用户更多的需求。

那么,什么是 MPLS? 它可以为运营商带来哪些好处? MPLS 是否真如人们所期望的那样可以解决所有的问题? 本书将围绕这些读者关心的问题对 MPLS 做了一个较为全面的剖析。当读到本书末尾的时候,相信读者可以对这些问题有一个比较全面的认识。

自从 1991 年万维网 (WWW) 推出,促使 Internet 商用化之后,Internet 便以惊人的速度发展,然而随着 Internet 的蓬勃发展以及使用人数的增加,网络拥塞的情况日趋严重,如果不加以改善,网络质量终将恶劣到无法接受的地步,因此各国政府以至业内人士都在集中精力研究解决之道。

1996 年克林顿政府发表了下一代互联网 (Next Generation Internet, NGI), 其他 Internet 相关技术标准更是层出不穷。众所周知,异步传输模式 (Asynchronous Transfer Mode, ATM) 技术可以提供很大的带宽以及各种不同媒体的服务,而且具有端到端的服务质量保证。另一方面,目前在网络层,IP 已经是最流行的网络层协议,更是一个面向应用的协议,垄断了 90% 以上的应用市场,因此 ATM 与 IP 的结合,就成了一个重要的课题。

长期以来,人们一直在 ATM 与 IP 层之间寻找一种合适的技术来解决两层间所固有的问题(由于“重叠”所带来的资源浪费)。例如在 ATM 上承载多协议包 (RFC1483)、局域网仿真 (LANE)、Classical IP (RFC1577)、MPOA 以及如何提高在 ATM 上承载 IP 包的效率。

在广域网上,第三层交换技术被采用,以改善在 ATM 上传送数据包的效率。Ipsilon 的 IP 交换、Cisco 的标记交换 (Tag Switching)、IBM 合成的基于路由的 IP 交换 (Aggregate Route-Based IP Switching) 以及 Nortel Networks 的虚拟网络交换 (VNS), Toshiba 的信元交换路由器 (CSR)、Cascade 的 IP Navigator 等,都是那一时期的产物。

多协议标记交换 MPLS 既融合了第三层乃至所有应用层的业务,也获得了所有底层媒体的支持,所以 MPLS 得到了所有运营商和厂家的青睐,被认为是下一代运营商网络的核心技术。ITU-T 也介入了 MPLS 标准化的研究,为电信运营商和厂商定制相关的标准和规范,为进一步的应用做好准备。

本书分为三个部分。第一部分是全书的“引子”,即多协议标记交换 MPLS 的基础,

包括 MPLS 概览、MPLS 体系结构；第二部分是全书的“内核”，即多协议标记交换 MPLS 的控制，MPLS 是如何完成诸多功能，如何成为运营商网络的核心技术的；最后一部分，也是本书的画龙点睛之笔，集中了目前多协议标记交换 MPLS 在运营商网络的主要应用，包括服务质量（QoS）保证，流量工程（Traffic Engineering），虚拟专用网（VPN）以及新的 MPLS 在光网络上的应用和在以太网快速恢复中所起的作用。

本书的出版，得到了很多人士的大力支持和帮助，在此，我要特别感谢著名电信设备提供商 Nortel Networks 公司的高级网络顾问 Lucian Ma 先生、Mark William 先生以及数据网络厂商 Cisco 公司的高级网络顾问 David zhou 先生等。同时，我也要感谢我的家人石泰、候玉纷、石琦，在本书的写作过程中所给予我的无微不至的照顾。另外也要特别感谢电子工业出版社的大力支持，使得本书能够在最短的时间内与广大读者见面。

本书希望通过对 MPLS 的体系结构、分配协议、流量管理以及 VPN 等方面的介绍，给读者一个较为全面的视角（从实现方式到应用）来看待 MPLS 这样一个新一代运营商核心网络的关键技术，并且为运营商在对这样一个新兴技术的选择上提出了一些可供参考的建议。

编著者

2002 年 11 月

目 录

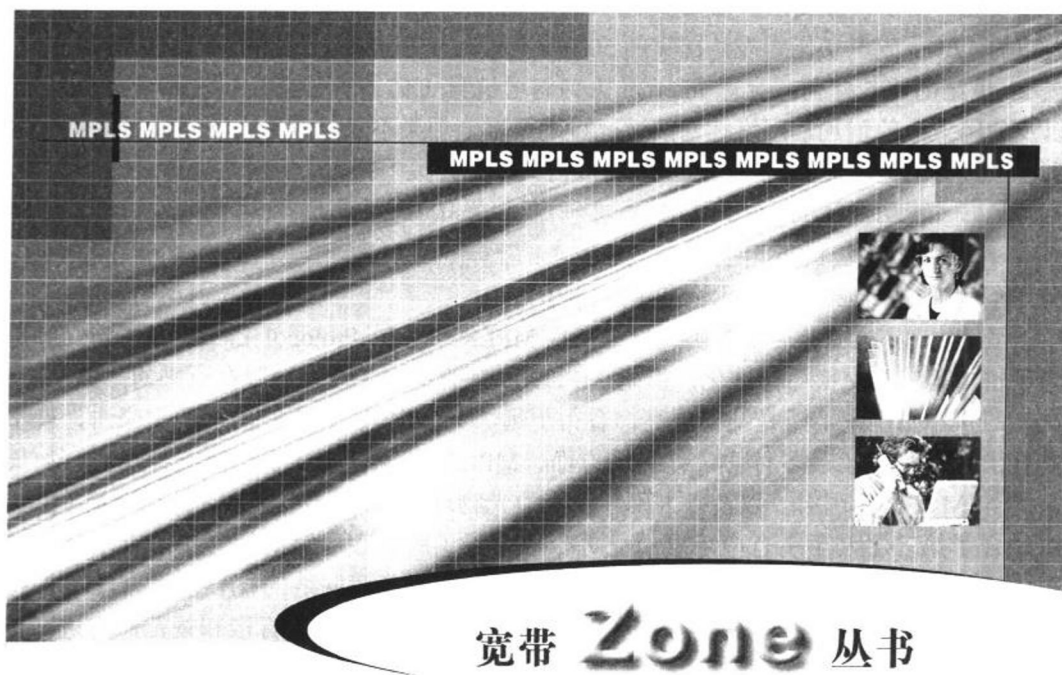
第 1 章 MPLS 概述	(1)
1.1 MPLS “演义”	(2)
1.1.1 运营商核心网络的演进	(2)
1.1.2 “第三层交换”	(5)
1.1.3 MPLS 大事记	(18)
1.2 为什么需要 MPLS	(19)
1.3 什么是 MPLS	(20)
1.4 MPLS 的术语及基本概念	(22)
1.5 MPLS 是怎样工作的	(22)
1.5.1 网络的边缘行为	(23)
1.5.2 网络的核心行为	(23)
1.5.3 如何建立标记交换路径	(24)
1.6 MPLS 的应用	(25)
1.7 总结	(27)
第 2 章 MPLS 的体系结构	(29)
2.1 MPLS 的数据平面	(31)
2.1.1 标记交换路由器和标记	(31)
2.1.2 标记的封装	(32)
2.2 标记的控制平面	(37)
2.2.1 转发等价类	(38)
2.2.2 标记分发协议	(39)
2.2.3 标记的分配和分发	(39)
2.2.4 路由的选择	(40)
2.2.5 环和 TTL	(43)
2.2.6 标记的分发	(45)
2.2.7 标记的合并	(46)
2.3 总结	(47)
第 3 章 MPLS “信令”——标记分配协议	(49)
3.1 LDP——普通的标记分配协议	(50)
3.1.1 普通标记分配协议 LDP 的对等实体	(51)

3.1.2	普通的标记分配协议 LDP 的“消息”	(51)
3.1.3	标记的分发和管理	(52)
3.2	CR-LDP	(55)
3.2.1	限制路由标记分配协议 CR-LDP 的要求	(56)
3.2.2	限制路由中的“限制”	(56)
3.2.3	CR-LDP 的消息	(57)
3.2.4	CR-LDP 的服务质量参数	(58)
3.2.5	如何创建 ER-LSP 隧道	(61)
3.2.6	CR-LDP 的优势	(63)
3.3	RSVP-TE	(63)
3.3.1	RSVP 到扩展的 RSVP	(64)
3.3.2	RSVP 的扩展	(65)
3.3.3	扩展 RSVP 的“消息”(Message)	(66)
3.3.4	RSVP-TE 的服务质量参数	(70)
3.3.5	如何创建 ER-LSP 隧道	(72)
3.3.6	RSVP-TE 的优势	(74)
3.4	几种标记分配协议的比较	(75)
3.5	总结	(79)
第 4 章	MPLS 的主要应用	(81)
4.1	IP 与 ATM 的集成	(82)
4.2	服务质量保证 (QoS)	(84)
4.3	流量工程	(85)
4.4	虚拟专网 VPN	(85)
4.5	IP 与光的融合	(86)
4.6	在城域网中的应用	(87)
第 5 章	MPLS 的服务质量保证	(89)
5.1	ATM 的 QoS 保证	(90)
5.1.1	业务级业务量管理	(91)
5.1.2	路由级业务量管理	(94)
5.1.3	平台级业务量控制机制	(96)
5.2	帧中继 FR 的服务等级保证	(100)
5.2.1	拥塞的通知	(101)
5.2.2	拥塞缓解策略	(102)
5.3	IP QoS	(104)
5.3.1	“Diff-Serv”差分业务结构	(105)

5.3.2	“Inter-Serv”集成业务结构	(107)
5.3.3	IP 网络中的业务等级协议 (SLA)	(109)
5.3.4	IP 网络的服务等级保证体系	(110)
5.4	用 MPLS 来实现 QoS	(113)
5.4.1	MPLS QoS 的实现过程	(113)
5.4.2	MPLS 与 Diff-Serv	(115)
5.4.3	CR-LDP 实现端到端的 QoS	(115)
5.4.4	RSVP-TE 实现端到端的 QoS	(117)
5.5	总结	(117)
第 6 章	MPLS 的流量工程	(119)
6.1	为什么需要流量工程	(120)
6.1.1	以路由器为核心网络中的流量工程	(121)
6.1.2	IP 与 ATM “重叠”网络中的流量工程	(123)
6.2	什么是流量工程	(124)
6.3	MPLS 实现流量工程	(124)
6.3.1	为什么要用 MPLS 实现流量工程	(125)
6.3.2	如何实现 MPLS 流量工程	(125)
6.3.3	MPLS 流量工程的应用	(128)
6.4	MPLS 流量工程实施中的问题	(132)
6.4.1	何时使用 MPLS 流量工程	(132)
6.4.2	MPLS 流量工程的流量如何与传统 IP 流量共存	(132)
6.5	总结	(133)
第 7 章	MPLS VPN	(135)
7.1	VPN 的演进	(138)
7.2	VPN 的意义	(139)
7.3	IP VPN 的分类	(142)
7.3.1	按照层次分	(142)
7.3.2	按照实现方式分	(143)
7.4	MPLS IP VPN	(144)
7.4.1	BGP/MPLS VPN (Cisco)	(145)
7.4.2	北电网络多虚拟路由器解决方案	(157)
7.4.3	Lucent (原 Ascend) IP Navigator 解决方案	(164)
7.4.4	基于 CPE 的 MPLS VPN (Juniper)	(167)
7.5	相关 RFC 介绍	(169)
7.6	几种主要方案的选择	(170)

7.7	MPLS IP VPN 在运营商的应用情况	(171)
7.8	总结	(171)
第 8 章	MPLS 在光层上的应用	(173)
8.1	为什么要在光的智能交换中采用 GMPLS	(175)
8.1.1	网络演进 - 网络革命	(176)
8.1.2	IP 网络比 IP 协议有更多的要求	(176)
8.1.3	MPLS 并不等于第三层交换	(177)
8.1.4	将 ATM 的 QoS 引入到 IP 网络中	(178)
8.1.5	将 MPLS 扩展到光网络	(180)
8.1.6	支持 GMPLS 的光纤业务	(181)
8.1.7	引入 MPLS 交换机	(181)
8.1.8	GMPLS	(182)
8.2	ASTN——智能交换传输网络	(183)
8.3	GMPLS——“通用的”MPLS	(185)
8.3.1	光网络基础	(185)
8.3.2	GMPLS 的需求	(187)
8.3.3	在光网络中的标记	(188)
8.3.4	标记选择的限制	(190)
8.3.5	光的用户网络接口	(191)
8.4	GMPLS 的优势与应用	(193)
8.5	ASTN 的未来	(195)
8.6	总结	(197)
第 9 章	MPLS 的保护	(199)
9.1	链路保护的基本知识	(200)
9.1.1	潜在的资源故障	(200)
9.1.2	修复故障的目的	(200)
9.1.3	检测故障	(201)
9.1.4	克服故障的基本办法	(203)
9.1.5	被保护的链路资源	(204)
9.2	本地修复	(205)
9.2.1	在 MPLS 网络中的本地修复	(206)
9.2.2	流量工程的影响	(207)
9.2.3	恢复的速度	(207)
9.2.4	Crankback	(207)
9.2.5	返回最优路由	(208)

9.3	保护交换 (Protection Switching)	(208)
9.3.1	备份模式以及选项	(209)
9.3.2	通告故障情况	(210)
9.3.3	可选的修复点	(211)
9.3.4	恢复速度	(211)
9.3.5	共享资源	(212)
9.3.6	目前 IETF 的状况	(214)
9.4	快速重路由	(214)
9.4.1	链路保护	(214)
9.4.2	节点保护	(216)
9.4.3	节点及链路保护总结	(217)
9.4.4	用“便道”的方式自动保护	(217)
9.4.5	IETF 中目前的现状	(219)
9.5	几种保护机制的比较	(219)
9.6	MPLS 保护的应用	(220)
9.7	总结	(221)
第 10 章	MPLS 的实现	(223)
10.1	在 ATM 设备上的实现	(224)
10.1.1	MPLS 与 ATM 的关系	(227)
10.1.2	Passport MPLS 基本组件	(229)
10.1.3	Passport MPLS 基本概念	(230)
10.1.4	Passport MPLS 的高级业务	(232)
10.1.5	LDP 与 CR-LDP 的互通	(239)
10.1.6	MPLS IP VPN	(240)
10.2	在路由器上的实现	(241)
10.2.1	MPLS 的流量工程	(242)
10.2.2	MPLS QoS	(247)
10.2.3	MPLS IP VPN	(254)
10.3	测试仪表	(256)
10.4	总结	(257)
第 11 章	MPLS 的未来与展望	(259)
11.1	MPLS 是否会像 ATM 一样, 雷声大, 雨点小	(260)
11.2	MPLS 能否作为下一代运营商网络的核心技术	(260)
11.3	MPLS 准备好了	(261)



第1章 MPLS 概述

- ◆ MPLS “演义”
- ◆ 为什么需要 MPLS
- ◆ 什么是 MPLS
- ◆ MPLS 的术语及基本概念
- ◆ MPLS 是怎样工作的
- ◆ MPLS 的应用
- ◆ 总结

自从 1991 年万维网 (WWW) 推出, 促使 Internet 商用化之后, Internet 便以惊人的速度发展, 目前仅在我国, 在家中上网的人数已经高达 5 660 万人, 而这仅占了我国总人口的 4.6%, 这与美国等发达国家的 50% 以上相比, 还有很大的潜力。而且, 根据预测, 在 21 世纪的前 5 年内, 全球 Internet 人口将增长至 3 亿。然而随着 Internet 的蓬勃发展以及使用人数的增加, 网络拥塞的情况日趋严重, 如果不加以改善, 网络质量终将恶劣到无法接受的地步, 因此各国政府以至业内人士都在集中精力研究解决之道。

目前较受瞩目的宽带网络计划包括由美国研究单位主导的第二代互联网 (Internet2), 以及 1996 年克林顿政府发表的下一代互联网 (Next Generation Internet, NGI), 其他 Internet 相关技术标准更是层出不穷。

虽然, 最初 IETF 提出多协议标记交换 MPLS (Multi Protocol Label Switching) 是为了将“第三层交换”标准化, 但是 MPLS 本身的定位并没有局限在原来的“第三层”交换的狭小范围里, 而是站得更高, 看得更远。一方面从 MPLS 的字面上来讲, 就不难看出, 它可以承载多种协议; 另一方面, MPLS 又可以凌驾于任何一种媒质, 如最初的 ATM、FR (帧中继) 以及以太网和现在的光物理层等。这样, 多协议标记交换 MPLS 既融合了第三层乃至所有应用层的业务, 也获得了所有底层媒质的支持, 所以 MPLS 得到了所有运营商和厂家的青睐, 被认为是下一代运营商网络的核心技术。ITU-T 也介入了 MPLS 标准化的研究, 为电信运营商和厂商定制相关的标准和规范, 为进一步的应用做好准备。

1.1 MPLS “演义”

MPLS 的根源可以追溯到 20 世纪 90 年代中叶 ATM 上承载 IP 的众多技术努力上。前面我们提到了在局域网和广域网方面的努力, 而在这里, 我们仅将目光放在广域网方面, 因为后来的多协议标记交换技术更多的是得利于广域网方面的努力。

在广域网方面, 第三层交换给业界带来了一次新的革命, 说是“革命”, 一点也没有夸张的意味。清楚当年情况的人, 一定还会记得当时给业界带来的冲击。一个小公司似乎在一夜之间站到了已经占据霸主地位的 Cisco 对面, 直接向这个不可一世的霸主挑战。这个公司就是 Ipsilon。虽然在今天, 当时的硝烟早已散尽, 我们看起来仿佛很可笑, 但相信经历过这次变革的局内人, 并不会像今天的读者这么轻松, 能够笑起来。

要想看清这一段历史, 就不得不先来回顾一下从 20 世纪 80 年代末 90 年代初的运营商核心网络的演变, 了解一下这次变革的前因后果。

1.1.1 运营商核心网络的演进

随着 Internet 在全球大规模的快速发展, 以及网络科学技术的进步, 运营商网络经

历了几个主要的历史时期。如果按照在不同历史时期中扮演角色的重要性来划分的话，可以大概地将这段历史划分成三个部分：以路由器为核心的运营商网络阶段，以 ATM 交换机为核心与“IP 重叠”的运营商网络阶段，现今以大规模交换路由器为核心的运营商网络阶段。

1.1.1.1 以路由器为核心的网络

在 90 年代早期，ISP（Internet 服务提供商）网络由路由器组成，路由器之间通过 T1 和 T3 专线连接。在 Internet 开始迅猛增长的时候，Internet 服务提供商们所能提供的链路就已经开始不能完全满足业务量的需要了。ISP 们通过增加更多的链路，扩大带宽来迎接这一挑战。当时，流量工程开始变得对 ISP 重要起来。因为通过流量工程，ISP 可以在多条可选、并行的路径存在时，有效地使用它来汇聚网络带宽。

简单的基于开销的流量工程方案有效地解决了当时 ISP 网络中遇到的问题，但随着 ISP 网络的飞速增长，这种解决方案已经越来越不能满足 ISP 发展的需要了。

显然，这种以传统路由器为核心的网络在支持流量工程方面存在着很多有关扩展性方面的问题。而且，传统的基于软件的路由器由于经常需要软件的介入，所以无法提供高速的汇聚接口和很高的包处理转发能力，从而成为网络的瓶颈。

1.1.1.2 IP 重叠的网络

大约在 1994 年到 1995 年间，Internet 的流量已经接近一个分界点，ISP 需要将他们的网络平滑升级到大于 T3（45 Mbps）的中继容量上。幸运的是，那时 OC3（155 Mbps）ATM 接口在很多交换机和路由器上都可以提供了。为了得到这个速率，ISP 被迫重新设计他们的网络，使他们可以使用 ATM 或帧中继（FR）交换机的高速端口。

一些 ISP 开始有步骤地进行过渡。在边缘，ISP 用带有 OC-3 ATM SAR（拆装子层）接口的路由器取代原来的 T3 点到点链路接口的路由器；同时在 ISP 网络内部，使用交换机构建网络的核心骨干；然后，在大约 6~9 个月之后，在核心 ATM 交换机之间地链路升级到 OC-12（622 Mbps）。

另外一些 ISP 利用帧中继网络将目前的点到点链路扩展到网状链路，并且普遍将链路的带宽升级至 T3（45 Mbps）。当他们开始从帧中继到 ATM 过渡时，虽然在边缘他们也使用 OC-3（155 Mbps），但是在核心他们同样会马上将 ATM 核心交换机之间的链路升级到 OC-12（622 Mbps）。

与传统的基于软件的路由器相比，ATM 交换机提供了更高速率的接口以及更大的汇聚带宽，从而消除了网络 and 核心路由器的瓶颈。总之，在路由器的性能还没有达到人们所期望的目标之前，ISP 在速度、带宽以及性能等方面的要求只能通过 ATM 交换机来实现。

1.1.1.3 以大规模交换路由器为核心的网络

在过去的几年里, ATM 交换机使 ISP 强大起来, 扩展了它们的市场份额, 也增加了它们的利润。在 90 年代中叶, ATM 交换机由于它无可比拟的高速接口、恒定的高性能以及出色的流量管理能力被 ISP 广泛采用。但是今天, ATM 交换机曾经独一无二的特点已经被大规模的 Internet 骨干路由器所取代。Internet 骨干路由器已经具备上述的 ATM 交换机主要的优点。这些最新的在路由技术上的进展使 ISP 不得不重新评价他们是否还要继续容忍这种重叠模式的限制——管理的开销, 昂贵的设备、运营的稳定性以及可扩展性。

基于 ATM 为核心的最根本限制是它需要两套不同的网络管理, 一个 ATM 的基础设施和一个逻辑的 IP 重叠。在 ATM 上承载 IP 网络, ISP 不但增加了他们网络的复杂性, 而且也双倍地加重了 ISP 的负担, 因为他们需要管理和协调两个独立网络的运营。另外, 路由和流量的管理发生在不同的系统之上——路由实施在路由器上, 而流量的管理运行在 ATM 交换机上, 所以它就很难完全将流量管理和路由结合起来。近期的技术发展已经允许 Internet 骨干路由器可以提供高速链路以及足够的性能, 这些因素正是以前 ISP 选择 ATM 交换机的主要因素。当 ISP 需要将网络再升级到 OC-48 (2.5Gbps) 的时候, 他们就不得不重新考虑是否仍然需要继续沿用原来的高开销、复杂的设计, 而此时可以由一套设备, 即集成的基于路由器的核心网络实现同样的功能。

另外, 原有路由器的 ATM 接口已经不能与光接口的发展同步了。目前最快的商用的 ATM SAR 路由器接口是 OC-12 (622 Mbps)。这里我们说的是“商用”, 不包括实验室中的产品。确实令所有的人都很惊奇, 但事实就是如此。在所有的 ISP 和运营商网络中, 使用最快的 ATM SAR 路由器接口是 OC-12 (622 Mbps), 而不是 OC-48 (2.5 Gbps), 或者 OC-192 (10 Gbps)。

当然, ISP 或者运营商的决定不只由于这一个原因, 在单纯的 ATM 网络中, ATM 交换机用 OC-48 (2.5 Gbps), 甚至 OC-192 (10 Gbps) ATM 接口作为中继。注意到这只是作为骨干网使用的纯 ATM 接口, 而且目前大多数运营商和 ISP 只使用了 OC-48 (2.5 Gbps)。是否会向 OC-192 (10 Gbps) ATM 接口过渡还很难说, 因为目前在这个网络上的业务还不足以需要 ATM 网络扩展。促使这个网络扩展的主要动力——Internet 数据业务也正在逐渐地转向另外的网络中去。看看国内外的网络发展, 这一点就不难理解。国内的 Internet 流量由“163”承载, 原来在底层大部分省间和省内中继采用了 OC-12 (622 Mbps) ATM 接口, 而通过 2001 年和 2002 年的扩容工程, 已经基本用 OC-48 (2.5 Gbps) POS 接口取代了原来的 ATM 接口, 换句话说, 已经从 ATM 网上挪到了 SDH 传输网上了。Internet 流量已经跨越 ATM 层, 直接向最底层——光物理层融合。由于光通信方面的发展, 越来越多的业务直接跨越链路层, 承载在光物理层上面, 提高了使用效率, 降低了开销。

综上所述, 在 20 世纪 90 年代中叶, 由于路由器的一些问题, 需要 ATM 网通过网状的 PVC 为其提供互联, 但是在今天, IP 承载在 ATM 之上的模式已经不能满足目前网络业务发展的需要了。以大规模骨干路由器(所谓“骨干路由器”就是每秒钟包转发在千万级以上, 而可以提供 STM-16、STM-48 等高速大容量端口)为核心, 构建新一代运营商核心 IP 网络是市场乃至技术发展的产物, 是历史的必然。

尽管有了大规模的骨干路由器, 但是今天对于运营商来说 MPLS 同样是一个核心网络的关键技术。虽然 MPLS 最初是为了解决在 ATM 网上承载 IP 的效率问题的技术, 但是 MPLS 也可以解决在 IP 网上的 IP 地址问题、服务质量(QoS)保证问题, 可以跨越多个链路层媒质上传送业务数据, 可以提供新的基于 IP 的虚拟专用网(VPN)业务, 它并不简单地只局限在第三层交换的范围内。如果我们将构建下一代运营商核心网络的主要设备——大规模骨干路由器作为网络的“躯体”的话, 我们就可以将 MPLS 比做是躯体的“灵魂”, 如果没有灵魂, 躯体将只能完成最基本的功能, 就像一个“植物人”一样, 没有思想, 没有主动的行为; 而有了灵魂的躯体, 才可以完成今天运营商所希望网络完成的诸多功能。

下面, 我们先从另外一个视角——MPLS 的发展史中来看一下, MPLS 为什么与以往的解决方案不同, 并且可以成为今天运营商骨干网络中的核心技术。

1.1.2 “第三层交换”

尽管, 在“第三层交换”的领域里有很多厂家作出了贡献, 如 IBM 的合成的基于路由的 IP 交换(Aggregate Route-based IP Switching)、Cisco 的标记交换(Tag Switching)、Toshiba 的信元交换路由器(CSR)、Cascade 的 IP 导航(IP Navigator), 以及北电网络的虚拟网络交换(Virtual Network System), 但是我们仍然无法忘记 Ipsilon 在第三层交换发展史上所起的作用。

前面我们曾经回顾过 90 年代中期, 随着 Internet 的迅猛增长, 原有的传统路由器架构已经不能满足人们日益增长的需求了。ATM 必须要有一个很好的技术来解决其承载 IP 方面的效率问题。

下面, 就让我们先来看一下 ATM 论坛在这一方面所做的努力——MPOA。

1.1.2.1 ATM 论坛的 MPOA

1997 年 7 月 ATM 论坛公布其制定的 MPOA(Multi-Protocol Over ATM)协议, 其主要目的在于多种常用的网络层协议及应用能够在 ATM 网络上传送, 并以网络层路由方式解决传统路由器能力的不足之处。一般路由器执行的功能有地址解析、路由选择、包过滤等, 整个 MPOA 架构就有如一个大路由器, 执行这些功能。

MPOA 使用 ATM 论坛的 LANE(LAN Emulation)、IETF 的 NHRP(Next Hop Resolution

Protocol) 以及虚拟路由器 (Virtual Router) 三种技术的观念来构成。LANE 可使 ATM 网络有如一个内部的局域网络; NHRP 则是基于网络层地址, 在 ATM 骨干网络建立“捷径” (Shortcut); 虚拟路由器的观念则能针对网络不同的元件, 提供不同功能的能力, 以增加效率、降低成本。

1. MPOA 的基本元件

MPOA 定义逻辑元件 (Logical Component) 能够安装于多种硬件结构。

(1) 边缘设备 (Edge Devices): 不是很贵的产品, 它是基于目的地的网络层地址和 MAC 层地址, 在局域网和 ATM 界面之间传递数据包。

(2) MPOA 客户端 (MPOA Client, MPC): MPOA 客户端位于边缘设备或 ATM 主机中, 主要功能是判断及进行“捷径”的建立。当数据包进入边缘设备时, 入口 MPOA 客户端会要求 MPOA 服务器 (MPS) 检查其目的地, 判断是否使用“捷径”。如果是, 则 MPOA 客户端会建立一条交换虚拟通道 (SVC) 来传送数据包。MPOA 客户端和 MPOA 服务器之间通过 NHRP 沟通, 前者会将“捷径”的信息隐藏起来以增加处理速度。另一方面, 交换虚拟通道只要有一段时间没用就会被取消掉。

(3) MPOA 路由器: MPOA 路由器执行网络层和 ATM 层之间的映射, 可以安装于现有路由器或交换机, 也可以独立安装, 其保存 MAC 层和 ATM 层的地址以及路由表, 经由 NHRP 彼此沟通找到目的地地址, 如此 MPOA 客户端才能建立“捷径”。MPOA 路由器也安装一般的路由协议 (如 RIP 和 OSPF), 用以与现存的局域网或广域网的路由器沟通。

(4) MPOA 服务器: 是 MPOA 路由器的一个逻辑元件, 负责给 MPOA 客户端提供第三层转发信息, 它同时也包含 NHRP 服务器 (NHS) 的功能, MPOA 服务器会依据路由功能和 NHRP 服务器来寻找 ATM 地址的第二层路径, 此路径信息将返回给提出请求的 MPOA 客户端。

2. MPOA 的实现方式

(1) 当出口 MPOA 客户端 (在边缘设备) 开始接收信息要转发时, 会检查信息判断是否有“捷径”存在, 如果存在, 则直接利用“捷径”的 VPI/VCI 转发; 如果不存在, 则利用 LANE 将信息送至 MPOA 路由器 (入口 MPOA 服务器)。

(2) 如果“捷径”不存在, 入口 MPOA 客户端就会计算数据流的数量, 当数据流量达到某个阈值时, 入口 MPOA 客户端会送出“MPOA Resolution Request”给入口 MPOA 服务器, 以获得目的地的 ATM 地址来建立“捷径”。

(3) 入口 MPOA 服务器收到“MPOA Resolution Request”后, 如目的地的 ATM 地址不在其管辖范围内, 则将“MPOA Resolution Request”改成“NHRP Resolution Request”送到出口 MPOA 服务器。

(4) 出口 MPOA 服务器收到“NHRP Resolution Request”后, 送出“Cache Imposition Request”到出口 MPOA 客户端。

(5) 依相反路径, 出口 MPOA 客户端返回 “Reply” 信息到入口 MPOA 客户端。

(6) 入口 MPOA 客户端根据收到的目的地的 ATM 地址, 使用 ATM 信令 (Signaling) 建立 “捷径”, 直接利用 “捷径” 的 VPI/VCI 传送数据包。

MPOA 虽然已是 ATM 论坛制定的一种标准规范, 具有服务质量 (QoS) 保证的能力, 且除了 IP 外, 其他不同的协议也可以运行在 MPOA 之上, 但是其软件复杂, 需要安装很多协议, 整个网络的管理与运作维护很复杂, 虽有产品上市, 不过实际应用的已经很少了。

1.1.2.2 Ipsilon 的 IP 交换

由于基于 B-ISDN 信令控制的 ATM 基础是面向连接的技术, 与无连接的 IP 技术从根本上不能匹配, 所以, 一个又一个的解决方法提出, 又一个个地被否决, LANE、Classical IP、MPOA 都渐渐地淡出历史的舞台。

Ipsilon 网络公司就是在这一背景下以一个全新的面目脱颖而出, 给沉闷的业界带来一股清新的空气。

与所有以往的 ATM 承载 IP 技术不同, Ipsilon 网络公司倡导一个全新的 “IP 交换” 的概念。首先他们通过使用 Ipsilon 的软件控制 ATM 交换引擎, 完善的 IP 路由功能可以直接叠加到 ATM 交换硬件上; 其次, 他们还根据一定的属性, 将相似的 IP 包划分成一个 IP 流, 这样无须对每一个包都进行复杂的第三层操作, 而真的是一个 “流”。这就可以大大加快路由的速度, 甚至达到无阻塞路由。

下面, 我们来全面看一下 Ipsilon 的 IP 交换解决方案。

将 ATM 集成到 IP 网络中, 需要解决技术上的衔接问题。无论是 ATM 上承载 IP (IP over ATM), 还是局域网仿真 (LANE), 传统 IP (Classical IP), 甚至后来的 MPOA, 都向第三层路由协议隐藏了网络真实的拓扑结构。对于 IP 而言, 物理的网络成为一个很大的不透明的云图。这样的做法首先导致了功能上的重复, 在 ATM 云里需要通过第二层的路由协议 PNNI 计算路由; 而在第三层上, 也同样需要通过第三层的路由协议计算最佳路径。除此之外, 由于两个网络功能的重复, 还会导致更深层次的管理和维护上的问题。这就是今天我们已经所熟知的 “重叠的” 网络结构的由来。

1996 年 Ipsilon 公司提出 IP Switching 技术, 引起工业界广大的兴趣及讨论。其 IP Switch 主要的观念是结合 IP 和 ATM 硬件, 在 ATM 方面只使用 ATM 的 AAL5 层以下的功能, AAL5 以上的 ATM 软件全部拿掉由 IP 的软件取代。这样易于与现存 IP 网络衔接, 不需要使用 ATM 控制层, 而只利用 ATM 快速交换的好处。

Ipsilon 认为 IP 作为第三层协议是必要的而且是足够的。它能扩展到全球互联网就足以证明这一点。IP 是一个强壮的、独立于技术的协议, 几乎所有的操作系统都支持 IP。

Ipsilon 新的方案围绕着两个重点, 一是通过使用 Ipsilon 的软件控制 ATM 交换引擎, 完善的 IP 路由功能可以直接叠加到 ATM 交换硬件上; 二是根据一定的属性, IP 包可以