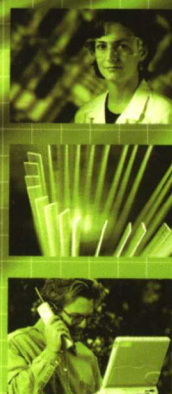


MPLS: Implementing the Technology

MPLS: Implementing the Technology



宽带 **Zone** 丛书

MPLS 实现技术

MPLS: Implementing the Technology

(美) Eric W. Gray 著

罗志祥 朱志实 黄本雄 译



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

宽带 **Zone** 丛书

MPLS 实现技术

MPLS: Implementing the Technology

(美) Eric W. Gray 著

罗志祥 朱志实 黄本雄 译

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书对 MPLS 技术所蕴含的有关思想进行了深入浅出的描述,介绍了 MPLS 的历史、MPLS 与现有交换协议间的关系、MPLS 的系统结构及其特殊的 MPLS 封装;详细论述了 ATM、帧中继、POS 及以太网如何支持 MPLS, MPLS 如何提供增强服务(如服务质量、流量工程和虚拟专用网);本书还详述了 MPLS 与其他技术(包括路由协议、链路层以及网络层技术)的相互影响,以及隧道技术、封装、标签分布、环路缓解、检测与预防等。本书为网络工程师、网络设计者和网络管理者提供了实现与开发 MPLS 技术所必需的所有资料,是一本有关 MPLS 技术的非常理想、完备的参考书。

Authorized Translation from the English language edition, entitled MPLS: Implementing the Technology, 1st Edition by Eric W. Gray ISBN 0201657627, published by Pearson Education, Inc, publishing as Addison-Wesley, copyright © 2001 Addison-Wesley.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

CHINESE SIMPLIFIED language edition published by PUBLISHING HOUSE OF ELECTRONICS INDUSTRY, Copyright © 2003

本书中文简体版专有出版权由 Pearson Education 授予电子工业出版社,未经许可,不得以任何方式复制或抄袭本书的任何部分。

版权贸易合同登记号 图字: 01-2002-2741

图书在版编目(CIP)数据

MPLS 实现技术/格林(Gray, E.W.)等编著;罗志祥,朱志实,黄本雄译. —北京:电子工业出版社, 2003.11(宽带 ZONE 丛书)

书名原文: MPLS: Implementing the Technology

ISBN 7-5053-9239-5

I. M… II. ①格… ②罗… ③朱… ④黄… III. 宽带通信系统—综合业务通信网 IV. TN915.142

中国版本图书馆 CIP 数据核字(2003)第 093054 号

责任编辑:张来盛 宋 梅

印 刷:北京增富印刷有限公司

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:11 字数:240 千字

版 次:2003 年 11 月第 1 版 2003 年 11 月第 1 次印刷

印 数:4000 册 定价:18.00 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。质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

前 言

多协议标签交换 (Multiprotocol Label Switching, MPLS) 是由 Internet 工程任务组 (Internet Engineering Task Force, IETF)^① 发展起来的一种减少网络中转发数据复杂程度的方法, 是一个工业标准。对于不同的人来说, MPLS 就是不同的事务。但是, MPLS 首先是一种实现了第 2 层 (链路层) 交换技术简化的转发特性, 且保留了同等合乎需要的第 3 层 (网络层) 路由灵活性与扩展性的方法。采用 MPLS 时, 在一个网络设备上做出路由判决, 允许相似的一些设备采用一种简化的标签置换模式来转发数据。因为通过减少必须执行路由判决的次数, MPLS 减少了网络的工作, 所以 MPLS 对许多路由问题都可以提供改进的解决方案。

目前, 以这种规律而言, MPLS 最重要的两个用途是流量工程和虚拟专用网 (Virtual Private Network, VPN)。虽然流量工程和虚拟专用网都可以采用现有的标准协议来完成, 但是, 因为 MPLS 有可能利用路由和转发分离的优势来减少或者消除某些路由的局限性, 所以 MPLS 可以使得这些任务变得更加简单。

例如, 以流量工程为例, 下述方法是可能的: 在建立一条通道期间指定某些明确的路径, 从而使得数据可以绕过网络的热点重路由。网络热点 (拥塞点) 发展的原因是寻径趋向于对每个可能的 (集合的) 目的地集中选择一条惟一的花费最少的路径。使用一个明确路由将该业务流的重要部分引导到未被寻径过程选择的网络部分, 通过这种方法, 可以允许分组通过 (部分地) 忽略寻径来旁路网络故障点。

正如在支持虚拟专用网通常所必需的那样, 也可以用多种方法将 MPLS 用于将专用网的业务流以隧道的形式跨越公用或者主干基础设施。网络操作员可以使用 MPLS 在 VPN 站点之间以隧道方式传输分组, 使得地址转换以及更加昂贵的隧道方案变得不必要。

MPLS 在一些其他的应用上也具有潜在优势。举例来说, 在小型 ISP、商务以及为多个计算机提供网络访问的地方, MPLS 可以在 Linux 平台上得到充分支持。如果允许使用 Linux 终端的话, 例如, 用于网络管理、行政和会计应用以及其他目的 (比如玩 “文明” 或者 “Pod 赛跑” 游戏), 则 MPLS 通过软件路由器转发分组的方式就可以使用做路由器的主机在性能上产生明显的差别。

MPLS 并不代表链路层和网络层的合并。实际上, 它以第 2 层和第 3 层技术仲裁程

^①IETF 是由科学家、工程师以及工业界的专家组成的团队, 他们致力于建立网络以及网络组件如何一起协同工作的规范。IETF 内的标准化进程是由 “多数人的意见和工作代码” 来驱动的。有关 IETF 的更加详细的信息, 请登录网站:

<http://www.ietf.org/overview>

序的角色与这两层相互作用。在网络层和链路层封装之间，MPLS 定义了一个封装，但在某些情况下，它也为链路层封装自身中重要的域定位定义值。例如，这种情况发生在 MPLS 与异步传输模式（Asynchronous Transfer Mode, ATM）以及帧中继（Frame Relay）共同使用的情况下。通过分配值到链路层封装中的域，使用位于链路层和网络层封装之间的一个填充片层封装，或者在网络层封装中嵌入一个长度固定的值^①，就可以获得期望的链路层行为特性。MPLS 也包括一些机制的定义，这些机制用于在链路层和 MPLS 填充片层域中建立标签值，以及在处理分组中使用这些值。当然，使用这种技术的最终目的是处理和转发分组。

MPLS 并不是一种期望中的端到端的解决方案。若有主机牵涉在 MPLS 标签分配与利用中，从中只能获得相对较小的利益。另外，MPLS 的可扩展性部分依赖于限制 MPLS 域的范围。虽然标签合并还不能自始至终地扩展到端节点——除非这种情况，例如在 Internet 中仅仅有一个接收端，但是，随着 MPLS 域的扩大，标签的合并将变得必不可少。最终，在转发中利用标签就意味着在分配标签的系统和利用这些标签的系统之间存在一种强烈的信托关系。目前在 Internet 中，这种信托关系的层次从端到端还不存在。

基本的 MPLS 概念可以归纳如下：

- MPLS 采用固定长度的标签来表示关于分组的任何信息。这些信息可能是通常包含于分组中的信息，也可能是已知的关于某些分组的一条流的信息，这样的信息通常并不包含于每一个分组中。
- 标签用于本地识别需要同等转发的数据分组。
- 用固定长度的标签表示转发信息，简化了采用精确匹配算法基础上的判决。这种简化了的转发算法称为标签交换。
- 一个参与标签交换的路由器是一个标签交换路由器（LSR）。
- 标签被加到一个 MPLS 入口点，通过中间 LSR 进行交换，并在一个 MPLS 出口点被删除。对于任意一个特别的开始标签，沿着一条通道发生上述情况，该通道就称为标签交换式通道。

关于本书

这本书被确定为一本关于 MPLS 技术的、独立的、知识完备的参考书。但是，MPLS 作为一种技术，它与大量的其他技术（例如各种各样的路由协议、链路层技术以及（至少是在理论上）网络层技术）相互作用，相互影响。本书将不对单个的路由协议规范进行描述，除非那些与 MPLS 直接相关的路由协议。关于路由与网络以及链路层技术，有许多很好的参考书。

本书对导致 MPLS 发展的各种不同技术进行了概括性的描述，是一个较高层次的

^① 这最后一个选项是一个为了在第 6 版因特网协议（IPv6）中使用的早期提案。

MPLS 发展过程的历史进程总结，它对于理解在 MPLS 发展过程中所做出的一些选择是十分有用的。但是，因为这本书的目标是使得协议本身是可理解的，所以，本书的重点是各种不同提议合并过程中所产生的协议细节，而不是各个提议的详细分析与比较。已经有一些书至少对那些早期的提议进行了很好的比较。

本书中的材料都是在公众可获得的信息基础之上的。本书中的阐述可以由任何一个期望证明它的人来证明。当然，并不是所有的公共信息都是一致的。特别是，对于在期望认真讨论标签交换时完全不可避免的一些话题，要想从那些可获得的信息中提取出一个通用的结论是不可能的——至少可以这么说。各种各样的参与者提供的信息通常包含一些比较模糊的参考书目来说明一些可能是广为人知或者鲜为人知（甚至是不可以公开获得）的信息，并且经常是与其他信息提供者提供的相关信息不一致。说明这个问题的一个典型例子就是对于“在什么环境下哪一个信令协议最有效”这样一个问题持续的争论。因此，试图使得 MPLS 技术容易理解的目的与讨论 MPLS 技术标准发展的结果保持公平与客观的目的稍微有些不一致。太注重公平性以及客观性将导致太多的模糊。我希望我已经做到了在达到易懂性目的的同时而没有丧失太多的公平性与客观性。

哪些读者应该阅读本书

本书假定读者已经听说过 MPLS 技术并且至少稍微知道一些术语。附录 A 和 B 提供的缩略语和术语解释用来填补读者对 MPLS 相关词汇的任何空缺，并且我已经尽最大努力来使得任何一个人通过阅读本书就可以更好地理解 MPLS 技术成为可能。

如果你是一个网络工程师、一个网络规划者或者设计者，或者是一个考虑在你自己的网络中采用 MPLS 技术的其他个人，那么你可能对本书有兴趣。如果你是一个技术管理者，或者是一个考虑在你自己的产品中实现这一技术的工程师，那么你可能会发现这本书很有趣。你也可能是一个学生，自愿地把你的时间和精力投入到这个难啃的科目。或者你仅仅只是曾经听说过这一技术，并对此感兴趣，想通过阅读本书增长知识。虽然你以前从没有听说过这一技术，你也可能从书架上拿起这本书——要么是在一个书店，要么是在一个技术图书馆。毕竟，MPLS 只不过是新一代的四字母缩略语中的一个成员而已，对于那些还没有曾经听说过这一技术的人来说，它将不太可能是一个迫切的兴趣。

在我书写本书的时候，还没有适当、最近、翔实的技术书籍介绍 MPLS，并且对于这一激动人心的新技术还存在很多的疑问和一般的兴趣。如果你已经很久以前就读过这一点，这个事实就说明你就是对这个新技术抱有疑问的很多人中的一员。因此，除非你始终坚信 MPLS 是明尼阿波利斯（Minneapolis，美国城市）的缩写语，你还仍然把这本书放在你面前的事实说明你很可能就是应当阅读这本书的人。

需要具备的知识

为了尽量利用这本书，读者应当对网络技术，如路由、交换、数据传输、封装、信

令、控制功能与结果等有一个基本的了解。这本书包含了许多有关这些基本话题的参考文献和书籍。当某些基本概念不是很清楚时，这些参考文献应当用做补充读物。如果你是一个网络工程师或者一个技术管理者，那么本书中的材料与对数据网络的基本了解一起对于你明智地判决是否实现和采用这一技术应当是足够的。如果你是一个实现者，那么这本书应当可以提供有关这一技术的坚固的基础知识。然而，在你能够利用这些知识实现它之前，你将需要阅读一些特别的相关材料，这些材料将为你所试图实现的产品细节提供更详细的知识。对于特殊的路由协议、服务质量（Quality-of-Service, QoS）以及虚拟专用网而言，上述观点更是特别真实。

如何阅读本书

本书的材料以贴近读者的方式进行组织，使得几乎所有读者都能够理解它并从中获益。举例来说，第 1 章对体现该技术的一些思想提供了一个总体的看法。对于某些读者来说，第 1 章可能就是他们需要阅读的全部。另外，广泛地使用了脚注。这样做的目的是，如果读者对某些背景讨论或者对某个特别的注释或观测来源于何处不感兴趣，那么他们可以完全不理睬这些脚注。使用脚注使得一些读者可以更加轻松地跟随正在讨论的一些思想的发展进程。

我也已经历尽艰辛来以这样一种方式组织各个章节，使得每一章都独自解决 MPLS 功能的一个特别的子集部分，那些对其他章节不那么感兴趣的读者也可能对该章讲述的 MPLS 功能感兴趣。例如，对 MPLS 的历史不感兴趣的读者可以完全跳过第 2 章，而几乎没有丧失对 MPLS 的理解，比方说，对一个特别的网络应用而言，不会丧失对 MPLS 有效性的理解。当然，也将有许多读者，他们对这一技术的各个方面都感兴趣。因此，我的出发点是实现两个目标：使得每一章都成为一个完整的整体；为那些对 MPLS 整体技术都感兴趣的读者提供交叉索引，把那些材料紧密地结合在一起。

本书的实际编排如下。第一部分包括第 1 章到第 5 章，对 MPLS 技术进行了概述；第二部分包括第 6 章和第 7 章，详细探讨了 MPLS 标准的细节。

第 1 章通过一些例子的介绍以及某些技术细节的概括性叙述，描述了 MPLS 技术的基础知识。本章讨论的主要概念包括什么是标签交换，什么是标签置换，它们如何与路由相比较，还有，需要什么来发送标签信令。

第 2 章从演化进程出发，在演化进程中发展了一些创新的概念。这一简要的演化进程从一部分人们试图解决的问题、一部分解决这些问题的早期提案，以及这些问题本身演化发展的方式入手。然后讨论了实实在在推动工业界开发一个标准解决方案的一些主要的提案。最后，总结了标准是如何达到目前的状况的。贯穿全章的时间线以及其他的一些图表展示了各种不同的提案是如何相互影响的。

第 3 章、第 4 章和第 5 章让读者进一步地了解 MPLS 技术以及该技术与网络世界的关系。第 3 章解释了 MPLS 必须如何与路由、网络层和链路层相互作用以便提供至少与

现有技术一样好的转发服务。本章也概括地阐述了在这个框架内 MPLS 应该可以提供的一些优点。第 4 章详细地介绍了 MPLS 系统的体系结构，包括那些组件、功能以及操作模式。第 5 章解释了特殊的 MPLS 封装以及哪些信令方法在哪些地方是最适用的。这些章节为后续一些章节将要更加详细讨论的问题提供了基础知识。

第 6 章将 MPLS 和其他一些解决这些相同问题的可选方法进行了详细的比较。本章举例说明了 MPLS 如何不同于其他方法，并阐明了使用 MPLS 的益处。这一章还说明了在各种不同技术上如何支持 MPLS。这些技术包括 ATM、帧中继、SONET 上分组技术 (Packet over SONET, POS)，以及以太网技术。

最后，第 7 章描述了如何使用 MPLS 来支持一些服务，如 QoS、流量工程以及虚拟专用网。

附录 A 和 B 给出了缩略语的全称，并且定义了本书中使用的一些术语和惯用语。

致 谢

我要感谢以下人员，他们是 Jhilmil Kochar、Loa Andersson、SBC 技术资源有限公司的 Muckai K. Girish、Georgia 技术学院的 Randal T. Abler、New Hampshire 大学互操作实验室 Rob Blais、Ron Bonica、Ross W. Callon、思科（Cisco）系统有限公司的 Thomas D. Nadeau、Tom Herbert 以及 Marconi 通信的 Walt Wimmer。他们非常细心地审阅了本书。他们的工作极大地鼓舞了我，并且帮助我提高了这本书的可读性和正确性。

我也要感谢那些在我写作过程中给予我鼓励和支持的同事。我要特别感谢在朗讯（Lucent）工作时的同事 Barbara Fox、Pramod Kalyanasundaram 和 Vasanthi Thirumalai 以及我在 Zaffire 工作时的同事 Fong Liaw、George Frank、John Yu 和 Michael Yao。

最亲密的人给了我最大的帮助。我要非常感谢在关键时期帮助我的家人。

最后，我要感谢 Addison-Wesley 出版社员工，特别是 Karen Gettman、Mary Hart、Emily Frey 以及 Marcy Barnes 的帮助与耐心。

译者序

20 世纪末，人类社会已进入信息时代，计算机网络技术与应用飞速发展，出现了许多新技术、新成果。多协议标签交换（Multiprotocol Label Switching, MPLS）就是最引人注目的新技术之一。MPLS 是 Internet 工程任务组为了减少网络中的转发复杂度而设计的适合工业标准的一种方法。

为了使广大从事网络应用和开发的人员全面掌握 MPLS 的原理及其应用，我们将 Eric W. Gray 所著的《MPLS 实现技术》一书介绍给广大读者。

Eric W. Gray 是 Zaffire 责任有限公司负责 MPLS 技术的高级设计工程师。在加入 Zaffire 责任有限公司之前，他是朗讯（Lucent）技术有限公司贝尔（Bell）实验室的一名成员。在贝尔实验室，他参与了朗讯公司有关 MPLS 和路由的工作。正因为如此，Eric W. Gray 所著的这本书具有很高的实用价值。

全书编排如下。第 1 部分包括第 1 章到第 5 章，对 MPLS 技术进行了概述，描述了 MPLS 技术的基础知识、发展进程以及 MPLS 技术与网络世界的关系。第 2 部分包括第 6 章和第 7 章，详细探讨了 MPLS 标准的细节，阐述了如何在各种不同的技术上支持 MPLS，以及如何使用 MPLS 来支持一些服务，如 QoS、流量工程和虚拟专用网。

本书中的材料都是以公众可获得的信息为基础的，是一本关于 MPLS 技术的独立的知识完备的参考书，特别适合于大专院校计算机、信息技术及相关专业的本科生、研究生以及网络规划人员、网络技术人员和网络管理人员阅读。本书还可供其他各类专业人员自学使用。

本书由华中科技大学罗志祥教授、武汉电视台朱志实高工和黄本雄翻译。

另外，对在本书翻译过程中给予支持的曹明翠、罗风光、周新军、徐军等老师表示感谢。

由于时间紧迫，书中错误之处在所难免，敬请读者指正。

译者
2003 年 9 月

目 录

第一部分 标签交换概述

第 1 章 思想	(3)
1.1 标签交换	(5)
1.1.1 桥接、交换以及标签交换	(7)
1.1.2 转发的复杂性	(7)
1.1.3 新的交换范畴	(11)
1.2 标签置换	(11)
1.3 信令标签	(12)
参考文献	(15)
第 2 章 历史概述	(17)
2.1 早期的概念	(19)
2.1.1 局域网仿真	(19)
2.1.2 ATM 上的多协议	(19)
2.1.3 信元交换路由器	(22)
2.1.4 Ipsilon 公司的 IP 交换	(23)
2.2 标志交换、ARIS 以及其他提案	(25)
2.3 MPLS 工作组	(26)
2.3.1 信令草案的发展	(27)
2.3.2 封装以及相关草案的发展	(31)
2.3.3 框架、体系结构以及其他通用草案的发展	(33)
2.3.4 虚拟专用网络、流量工程以及多路径优化草案的发展	(33)
参考文献	(49)
第 3 章 整体框架	(53)
3.1 必要条件	(54)
3.1.1 与路由的关系	(54)
3.1.2 与网络层协议的关系	(58)
3.1.3 与链路层协议的关系	(58)
3.2 优点	(58)
3.2.1 简单的转发范畴	(58)

3.2.2	显式路由的 LSP	(59)
3.2.3	流量工程	(59)
3.2.4	服务质量	(59)
3.2.5	工作划分	(59)
3.2.6	路由协议的扩展性	(60)
3.2.7	通用信令	(60)
3.2.8	简化管理	(60)
3.2.9	缩短时延	(61)
	参考文献	(61)
第 4 章	体系结构	(63)
4.1	MPLS 系统的组件	(65)
4.1.1	标签交换路由器	(65)
4.1.2	标签交换式路径	(67)
4.1.3	打标过的分组	(69)
4.2	MPLS 系统的功能	(69)
4.2.1	标签分布	(70)
4.2.2	合并	(71)
4.2.3	标签堆栈操作	(73)
4.2.4	路由选择	(74)
4.3	MPLS 的操作模式	(75)
4.3.1	标签分配模式	(75)
4.3.2	标签分布模式	(75)
4.3.3	标签保持模式	(76)
4.3.4	控制模式	(77)
4.3.5	标签空间	(78)
	参考文献	(80)
第 5 章	应用性	(81)
5.1	概述	(83)
5.2	分组的封装	(84)
5.3	信令	(84)
5.3.1	硬状态协议和软状态协议	(85)
5.3.2	LDP	(85)
5.3.3	CR-LDP	(86)
5.3.4	RSVP-TE	(87)
5.3.5	MPLS-BGP	(87)

参考文献	(90)
------------	--------

第二部分 MPLS 标准详述

第 6 章 执行选择	(93)
6.1 拓扑与流	(94)
6.2 隧道技术	(95)
6.2.1 对等体到对等体隧道	(95)
6.2.2 显式路由隧道	(99)
6.3 封装	(101)
6.3.1 媒介方面的细节	(101)
6.3.2 MPLS 填充片	(104)
6.4 标签分布	(104)
6.4.1 采用 BGP 捎带标签分布	(104)
6.4.2 采用 RSVP 捎带标签分布	(106)
6.4.3 标签分布协议	(112)
6.5 环路以及环路缓解、检测和预防	(119)
6.5.1 环路的影响	(120)
6.5.2 环路缓解	(120)
6.5.3 环路检测	(121)
6.5.4 环路预防	(122)
参考文献	(122)
第 7 章 服务	(125)
7.1 基本的服务	(127)
7.1.1 采用 LDP	(127)
7.1.2 采用 BGP	(128)
7.2 QoS: 保险的服务	(128)
7.2.1 采用集成服务方式	(128)
7.2.2 采用差分服务方式	(128)
7.3 流量工程	(129)
7.3.1 MPLS 的角色	(130)
7.3.2 流量工程的工作方式	(130)
7.3.3 业务流干线的属性	(132)
7.3.4 资源属性	(135)
7.3.5 基于约束的路由	(136)
7.3.6 路径建立和维护	(137)

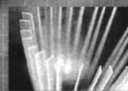
7.3.7 采用 TE 业务流进行负载分布	(138)
7.3.8 故障处理	(138)
7.3.9 方法	(139)
7.4 虚拟专用网	(140)
7.4.1 方法	(141)
参考文献	(142)
附录 A 缩略语	(145)
附录 B 术语解释	(149)

第一部分

标签交换概述

MPLS MPLS MPLS MPLS

MPLS MPLS MPLS MPLS MPLS MPLS MPLS MPLS MPLS



宽带 **Zone** 丛书

第1章 思 想

- ◆ 标签交换
- ◆ 标签置换
- ◆ 信令标签