



基于MPLS的流量工程

(修订版)

Traffic Engineering with MPLS

Design, configure, and manage MPLS TE
to optimize network performance

[美] Eric Osborne, CCIE #4122 著
Ajay Simha, CCIE #2970 译
张辉 卢炜

基于MPLS的流量工程 (修订版)

Traffic Engineering with MPLS

[美] **Eric Osborne**, CCIE #4122 著
Ajay Simha, CCIE #2970
张辉 卢炜 译

人民邮电出版社
北 京

图书在版编目 (C I P) 数据

基于MPLS的流量工程 / (美) 奥斯巴恩
(Osborne, E.), (美) 辛巴 (Simha, A.) 著; 张辉, 卢
炜译. — 2版. — 北京: 人民邮电出版社, 2012. 12
ISBN 978-7-115-29615-3

I. ①基… II. ①奥… ②辛… ③张… ④卢… III.
①宽带通信系统—综合业务通信网 IV. ①TN915.142

中国版本图书馆CIP数据核字(2012)第235499号

版 权 声 明

Traffic Engineering with MPLS (ISBN: 1587055392)

Copyright © 2002 Pearson Education, Inc.

Authorized translation from the English language edition published by Cisco Press.

All rights reserved.

本书中文简体字版由美国 Cisco Press 授权人民邮电出版社出版。未经出版者书面许可, 对本书任何部分不得以任何方式复制或抄袭。

版权所有, 侵权必究。

基于 MPLS 的流量工程 (修订版)

-
- ◆ 著 [美] Eric Osborne, CCIE #4122
Ajay Simha, CCIE #2970
译 张 辉 卢 炜
责任编辑 傅道坤
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
大厂聚鑫印刷有限责任公司印刷
 - ◆ 开本: 800×1000 1/16
印张: 31.75
字数: 706 千字 2012 年 12 月第 2 版
印数: 4 001—7 000 册 2012 年 12 月河北第 1 次印刷

著作权合同登记号 图字: 01-2012-6904 号

ISBN 978-7-115-29615-3

定价: 89.00 元

读者服务热线: (010)67132692 印装质量热线: (010)67129223

反盗版热线: (010)67171154

广告经营许可证: 京崇工商广字第 0021 号

内容提要

本书向用户提供了关于如何使用 MPLS TE 和相关特性使网络带宽的利用率最大化的信息。全书共分 11 章，分别介绍了理解基于 MPLS 的流量工程、MPLS 转发基础、信息发布、路径的计算和建立、流量在隧道中的转发、基于 MPLS TE 的服务质量、保护与恢复、MPLS TE 管理、支持 MPLS TE 的网络设计、MPLS TE 使用技巧、MPLS TE 故障排查等内容。附录部分介绍了 MPLS TE 命令索引、CCO 和其他参考。

本书适合那些对网络进行配置、故障排查和管理的网络工程师使用，还适合那些设计网络承载不同类型流量和支持不同服务等级协议（SLA）的网络架构师阅读。

关于作者

Eric Osborne, CCIE #4122, 从 1995 年以来一直从事 Internet 工程方面的工作, 并且历经了其中的风风雨雨。在 1998 年, 他进入了 Cisco TAC 部门, 并在 Ajay (本书另一作者) 之后不久也转入了 ISP Expert Team (ISP 专家小组), 并且自从 Cisco 发布其 IOS 11.1CT 版本以来, 就一直参与 MPLS 的工作。他持有心理学学士学位 (心理学要比你想象的有用得多)。在北美举办的 Cisco Networkers 大会中, 他经常作为发言人出现, 自 2000 年以来已经发表过一次关于 “Deploying MPLS Traffic Engineering (部署 MPLS 流量工程)” 的演讲。

Ajay Simha, CCIE #2970, 在印度获得计算机工程学士学位和计算机科学硕士学位。在作为一名数据通信软件开发人员工作了 6 年之后, 他在 1996 年加入了 Cisco TAC 部门, 然后他又作为 Cisco 的 ISP Expert Team (ISP 专家小组) 成员对 Tier 1 ISP (一级 ISP) 和 Tier 2 ISP (二级 ISP) 提供支持。1999 年前期, 他首次接触 MPLS TE, 对其产生了浓厚的兴趣并全身心投入到 MPLS 的工作中。从 1999 年 10 月份起, 他成为一名 MPLS 部署工程师。他在故障排错、设计和部署 MPLS 方面有亲身经验。

关于技术审稿人

Jim Guichard, CCIE #2069, 是 Cisco 公司的一名 MPLS 部署工程师。在最近几年中, 他一直在从事许多大规模 WAN 和 LAN 网络的设计、实现和规划工作。他具有广泛的行业知识、亲身经验, 而且对负责网络体系结构有着深刻的理解, 这使得他能够对 MPLS 以及部署场景有详细的洞察力。

Alexander Marhold, CCIE #3324, 持有工业电子学专业的理科硕士学位, 还持有 MBS 学位。他是 PROIN 公司核心 IP 服务团队的资深顾问和领导人, 这家公司是欧洲一流的致力于向服务提供商网络提供培训和咨询的公司。他关注的领域是包括 MPLS、高级路由、BGP、网络设计与实现在内的核心技术。除了是一名顾问之外, Marhold 还是一名 CCSI, 他开发和维护了与他专业领域相关的培训课程。他的上一份工作经历还包括在一所理工大学讲授通信课程, 以及在化工行业担任过 CIM 项目经理。

Jean-Philippe Vasseur 获得了法国的工学学位, 同时还获得了 SIT (位于新泽西) 的理学硕士学位。他具有 10 年的电信和网络技术工作经验, 在加入 Cisco 前曾为几家 ISP 工作。在同 EMEA 技术咨询小组共事两年之后——致力于向服务提供商提供 IP/MPLS 路由、VPN、流量工程和 GMPLS 设计——他加入了 Cisco 工程团队。他还参与了几个 IETF 草案的制定工作。

献 辞

Ajay Simha: 我要把本书献给我亲爱的妻子 Anitha 和可爱的孩子 Varsha 和 Nikhil，他们不得不忍受我因为写作本书而占用了比以往更多的时间。我还要把本书献给我的父母，正是他们毕生提供的教育基础和价值观念才帮助我达到了今天这个水平。

Eric Osborne: 我要把本书献给在我家附近的很多咖啡馆，没有它们，我就没有足够的动力完成此书。我还要把本书献给我的母亲（她教会了我制定计划表）、我的父亲（他教会我“提高就是积累”的道理）。我还要把本书献给教我网络、写作和思考的人们。这里都有你们的一份贡献。

致 谢

我要感谢很多人。首先，我们要发自内心地感谢整个 MPLS 开发团队，感谢他们经常在我们束手无策时，与我们分享他们的技术经验。特别要感谢 Carol Itturalde、Rob Goguen 和 Bob Thomas，不管我们提出的问题多么琐碎、多么含糊、多么频繁，他们都能给我们答案。

我们还要感谢我们最重要的技术审稿人员——Jim Guichard、Jean-Philippe Vasseur 和 Alexander Marhold。他们的指导和反馈使我们没有偏离方向。

我们还要感谢为保证本书的正确性和相关性而评审了本书部分内容的人们。按字母顺序依次为：Aamer Akhter、Santiago Alvarez、Vijay Bollapragada、Anna Charny、Clarence Filsfils、Jim Gibson、Carol Iturralde、Francois Le Faucheur、Gwen Marceline、Trevor Mendez、Stefano Previdi、Robert Raszuk、Khalid Raza、Mukhtiar Shaikh、George Swallow、Dan Tappan、Mosaddaq Turabi、Siva Valliappan、Shankar Vemulapalli 和 Russ White。Eric 希望感谢 PSU Local #1 的成员，感谢他们确认和纠正了 Eric 就大型 Internet 主干到底怎样运转所作的假设。

最后还要感谢 Cisco Press 的编辑团队——特别是 Chris Cleveland 和 Brett Bartow，他们带领我们经过了这个过程。这花费了一年多的时间，如果没有他们，本书将无法完工。

标记交换 (tag switching) 作为 Cisco 的专有技术在 1996 年 3 月开始演变为 MPLS。当时, 为了管理自己网络中的流量, 几家主要的 ISP 都在运行两级网络。由此可见, IP 总是使用去往目的地的最短路径。这个特性对 Internet 的可扩展性非常重要, 因为它允许路由最大限度地成为一个自动的过程。然而, 最短路径并不总是最快的路径或者负载最轻的路径。此外, 在任何没有流量工程的网络中, 你会发现链路利用情况会这样分布: 少数几条链路的负载非常重, 而其他很多的链路负载则很轻。最终导致许多网络用户竞争繁忙链路的资源, 而其他的链路却没有被利用。这导致服务水平和运行成本都没有得到优化。实际上, 一个 ISP 会说: 有了流量工程, 只使用没有流量工程时 60% 的链路, 就可以提供同等的服务水平。

因此, 流量工程在经济上成为必需, 且其需要程度足以建立一个完整而独立的二层网络。对于流量工程来说, 一个 ISP 需要在其 IP 网络的主要站点之间建立一个全互连的链路 (虚电路), 然后使用二层网络——要么是帧中继要么是 ATM——通过指定虚电路的方式来显式地对流量进行路由。

到 1996 年 4 月, Cisco 公司内部认识到, 标记交换提供了一种在 IP 网络内部创建显式路由的方式, 从而不需要一个二级网络。因为这对减少 ISP 的主要成本有潜在的帮助, 所以人们旋即投入了很大的热忱。几家 ISP 和设备制造商一起制定出了详细的需求和技术方法。

Eric Osborne 和 Ajay Simha 都在 Cisco 负责实现流量工程的开发组工作。他们积极参与了流量工程在很多网络中的部署工作。而且通过这些部署, 他们也成为具有流量工程实践经验的人员。这本书就是他们经验的结晶。它深入并且实用地阐述了构成流量工程应用的各种元素——路由选择 (routing)、路径选择 (path selection) 和通告 (signaling)。在本书之中, 这些

解释都和实际的配置命令以及示例联系在一起。这使得本书成为对流量工程感兴趣，以及打算部署流量工程的人员的宝贵资料和指南。

George Swallow

Cisco 公司

流量工程架构师，IETF MPLS 工作组副主席

前言

这本书集中讨论 MPLS TE 在现实世界中的使用。我们花费了大部分时间来讨论配置情况、用来进行故障排错和管理 MPLS TE 的工具，以及设计场景。

本书不是 MPLS 的介绍性指南。Cisco 公司和其他公司都已经有很多著作了，所以我们没有必要在 MPLS 的介绍上花费时间。虽然第 2 章回顾了 MPLS 的基础知识，但是我们通常还是假设读者已经熟悉了 3 种基本的标签操作（入站、出栈和交换）以及 MPLS 分组在网络中是如何转发的。如果你知道的比这还多，那么，本书就适合你阅读。

或许你已经在你的网络中使用了 MPLS TE。如果是这样的话，本书也适合你。本书有很多细节内容，我们希望你继续使用和探究 MPLS TE 能有所帮助。

或许你正在设计一个新的主干网，并且正考虑在网络中使用 MPLS TE。如果是这样，本书也同样适合你。你不仅需要理解协议机制来合理地设计一个网络，还需要理解选择不同的设计带来的结果。

本书的读者对象

每一个人！你、你的朋友、你的祖母和她周围的朋友、你的孩子和他们幼儿园的同学——每一个人！实际上，我们并不在乎谁会购买本书以及谁会阅读本书，但是购买本书而不阅读是一件相当愚蠢的事。

严肃地说，本书适合两种人。

- **网络工程师**——那些对网络进行配置、故障排错和管理的人。
- **网络设计师**——那些设计网络承载不同类型流量（语音和数据）和支持不同服务等级协议（SLA）的人。

有些朋友在工作中会同时具备这两种角色。对他们来说（如果你也是这样的情况，那么也是对你来说），我们建议：“不错！买两本吧！”

本书的内容组织

本书可以一个部分一个部分地来读，也可以逐章来读。本书可粗略地分为四部分：

- 第 1 章和第 2 章讨论了 MPLS 和 MPLS TE 的历史、动机和基本操作。
- 第 3、4、5 章介绍了在网络中设置和建立 TE 隧道的基本过程。
- 第 6、7 章介绍了高级的 MPLS TE 应用：MPLS TE 和 QoS、使用快速重路由（FRR）进行保护。

• 第 8、9、10 和 11 章介绍了网络管理、设计、部署和故障排错——这些是在现实世界应用 MPLS TE 需要理解的内容。

下面是每章的主要内容。

• **第 1 章，“理解基于 MPLS 的流量工程”**——这一章讨论了基本数据网络的历史和作为网络发展的下一步的 MPLS 和 MPLS TE 的动机。

• **第 2 章，“MPLS 转发基础”**——这一章是对 MPLS 转发如何工作的一个快速回顾。虽然本书不是 MPLS 的介绍性指南，但是你也也许会发现复习一些细节也是有好处的，而这就是这一章提供的内容。

• **第 3 章，“信息发布”**——这一章是本书 3 个核心章节的第一章。MPLS TE 的协议和机制有 3 个部分，第一部分就是在 IGP 中发布 MPLS TE 信息。

• **第 4 章，“路径的计算和建立”**——这一章是 3 个核心章节的第二章。这章介绍了通过 IGP 发布信息后，对信息做了哪些处理。在本章突出介绍的两个部分是约束 SPF（CSPF）和资源预留协议（RSVP）。

• **第 5 章，“流量在隧道中的转发”**——这一章是 3 个核心章节的最后一章。这章介绍了在隧道建立后，对 TE 隧道的处理。这章还介绍了各种环境下的负荷分担、把 TE 隧道作为转发邻接通告到 IGP 以及使用一种称为自动带宽（auto bandwidth）的 Cisco 的机制实现自动的隧道带宽调整。

• **第 6 章，“基于 MPLS TE 的服务质量”**——这一章介绍了 MPLS 和 MPLS TE 与 DiffServ 体系的集成。它还介绍了支持 DiffServ 的流量工程（DS-TE）。

• **第 7 章，“保护和恢复”**——这一章介绍了在 Cisco 的 FRR 下的各种流量保护和恢复机制——如何配置这些服务、它们如何工作以及在网络发生故障时如何大大减少分组丢失。

• **第 8 章，“MPLS TE 管理”**——这一章介绍了管理一个 MPLS TE 网络的工具和机制。

• **第 9 章，“基于 MPLS TE 的网络设计”**——这一章主要介绍了可扩展性。它介绍了在网络中实现 MPLS TE 的几种不同的方式，以及在网络增长时各种解决方案的扩展性如何。

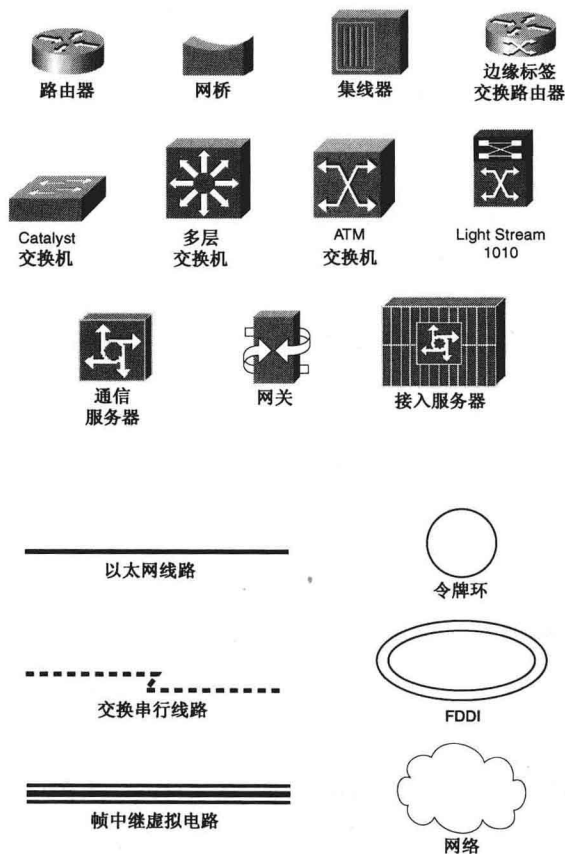
• **第 10 章，“MPLS TE 实现技巧”**——这一章介绍了各种技巧、最佳练习以及与网络中实现 MPLS TE 相关的案例研究。

• **第 11 章, “MPLS TE 故障排查”**——这一章讨论了一个在运行的网络上排查 MPLS TE 故障所需的工具和技术。

同时还提供了两个附录。附录 A 列出了和 MPLS TE 相关的所有主要的命令。

附录 B 列出了诸如 URL 和一些其他参考书的资源。附录 B 在 www.ciscopress.com/1587050315 也可获得。

本书使用的图标



命令语法定义

本书命令语法遵循的惯例与 IOS 命令手册使用的惯例相同。命令手册对这些惯例的描述如下。

- **粗体字**表示照原样输入的命令和关键字, 在实际的设置和输出 (非常规命令语法) 中, 粗体字表示命令由用户手动输入 (如 **show** 命令)。

- *斜体字*表示用户应提供的具体值参数。
- 竖线（|）用于分隔可选的、互斥的选项。
- 方括号（[]）表示任选项。
- 花括号（{}）表示必选项。
- 方括号中的花括号（[{}]）表示必须在任选项中选择一个。

目 录

第 1 章 理解基于 MPLS 的流量工程	1
1.1 基本网络概念	1
1.1.1 时分复用 (TDM)	2
1.1.2 统计复用	2
1.2 什么是流量工程	4
1.3 MPLS 之前的流量工程	5
1.4 进入 MPLS	7
1.4.1 分离路由和转发的关系	8
1.4.2 IP 和 ATM 世界的良好结合	9
1.4.3 使用 MPLS 建立服务	10
1.4.4 MPLS 的误区	11
1.5 现实中使用的 MPLS	13
1.6 小结	13
第 2 章 MPLS 转发基础	17
2.1 MPLS 术语表	17
2.2 转发的基本原理	19
2.2.1 什么是标签	19
2.2.2 控制平面和数据平面	21
2.2.3 分类	22
2.2.4 MPLS 网络中的控制平面	23
2.2.5 转发机制	24
2.2.6 标签分配的概念	33

2.3 标签分配协议	39
2.3.1 LDP PDU 头部	40
2.3.2 LDP 消息格式	40
2.3.3 LDP 的主要功能	42
2.3.4 环路检测	56
2.4 标签分配协议配置	56
2.4.1 配置 CEF	57
2.4.2 全局配置 MPLS 转发	59
2.4.3 接口级配置	59
2.5 小结	65
第 3 章 信息发布	67
3.1 MPLS 流量工程的配置	67
3.2 发布了什么信息	70
3.2.1 可用带宽信息	70
3.2.2 隧道优先级	72
3.2.3 属性标志	78
3.2.4 管理权重	80
3.3 什么时候发布信息	81
3.3.1 对重大变化立即进行泛洪	82
3.3.2 对不重要的变化周期性泛洪, 但是比 IGP 刷新频率更高	85
3.3.3 如果一个没有泛洪的变化导致了错误, 马上进行泛洪	85
3.4 信息如何发布	86
3.4.1 OSPF 中的 MPLS 流量工程	86
3.4.2 IS-IS 中 MPLS 流量工程的泛洪	88
3.5 小结	91
第 4 章 路径的计算和建立	93
4.1 SPF 如何工作	94
4.2 CSPF 如何工作	98
4.2.1 CSPF 中的最高仲裁方法	101
4.2.2 影响 CSPF 的其他因素	102

4.2.3 CSPF 的调节	106
4.3 隧道的重新优化	112
4.3.1 定期重新优化	113
4.3.2 手工重新优化	113
4.3.3 事件驱动的重新优化	113
4.3.4 防范禁闭	114
4.4 资源预留协议 (RSVP)	115
4.4.1 RSVP 基础	116
4.4.2 RSVP 分组	119
4.4.3 RSVP 操作	140
4.4.4 现实世界的 RSVP	150
4.5 区间隧道	155
4.5.1 IGP 术语	155
4.5.2 区间隧道的作用	156
4.5.3 区间隧道如何工作	157
4.5.4 区间隧道不能做什么	163
4.6 链路管理	164
4.6.1 show mpls traffic-eng link-management admission-control 命令	165
4.6.2 show mpls traffic-eng link-management advertisements 命令	166
4.6.3 show mpls traffic-eng link-management bandwidth-allocation 命令	167
4.6.4 show mpls traffic-eng link-management igp-neighbors 命令	169
4.6.5 show mpls traffic-eng link-management interfaces 命令	170
4.6.6 show mpls traffic-eng link-management statistics 命令	171
4.6.7 show mpls traffic-eng link-management summary 命令	171
4.7 小结	172
第 5 章 流量在隧道中的转发	175
5.1 使用静态路由沿隧道 转发流量	175
5.2 使用基于策略的路由沿隧道转发流量	176
5.3 使用自动路由沿隧道转发流量	177
5.4 负荷分担	183