

CISCO SYSTEMS



Cisco Press

CCIE



TCP/IP 路由技术 (第 2 卷)

CCIE™ Professional Development
Routing TCP/IP
Volume II

A detailed examination of exterior routing
protocols and advanced IP routing issues

[美] Jeff Doyle Jennifer DeHaven Carroll 著
毕立波 魏亮 刘述 译

人民邮电出版社
POSTS & TELECOMMUNICATIONS PRESS

TCP/IP 路由技术

(第2卷)

[美] Jeff Doyle Jennifer DeHaven Carroll 著

毕立波 魏亮 刘述译

人民邮电出版社

图书在版编目(CIP)数据

TCP/IP 路由技术.第 2 卷 / (美)多伊尔(Doyle, J.)著;毕立波,魏亮,刘述译.

—北京:人民邮电出版社,2002.8

ISBN 7-115-10096-9

I. T... II. ①多...②毕...③魏...④刘... III. 计算机网络—通信协议—路由选择 IV. TN915.04

中国版本图书馆 CIP 数据核字(2002)第 029269 号

内 容 提 要

本书深入系统地阐述了 TCP/IP 路由技术,内容包括几种重要的网络协议,如外部网关协议(EGP)、边界网关协议(BGP4),以及相应的高级 IP 路由技术与应用——网络地址翻译、IP 多播路由、IPv6 技术、路由器管理等。

本书内容全面,可读性强,含有协议配置、网络实施、故障排除等方面的大量实例,非常适合 CCIE 开发人员、网络与通信系统工程技术人员阅读。

TCP/IP 路由技术 (第 2 卷)

-
- ◆ 著 [美] Jeff Doyle Jennifer DeHaven Carroll
 - ◆ 译 毕立波 魏 亮 刘 述
责任编辑 陈万寿
 - ◆ 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 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
读者热线 010-67180876
北京汉魂图文设计有限公司制作
北京鸿佳印刷厂印刷
新华书店总店北京发行所经销
 - ◆ 开本: 787×1092 1/16
印张: 43.5
字数: 1061 千字 2002 年 8 月第 1 版
印数: 1-4 000 册 2002 年 8 月北京第 1 次印刷

ISBN 7-115-10096-9/TN · 1841

定价: 72.00 元

本书如有印装质量问题,请与本社联系 电话:(010) 67129223

版 权 声 明

Jeff Doyle and Jennifer DeHaven Carroll:

Routing TCP/IP, Volume II

AuthoriZed translation from English language edition published by Cisco Press.

Copyright © 2001 Cisco Systems, Inc.

All rights reserved. For sale in mainland China only.

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

图字：01-2000-3065

版权所有，侵权必究。

惯用的命令语法

这本书中所用的命令语法与 IOS 命令参考中所用的相同。命令参考中描述的惯例如下：

- 竖线(|)分开可选的、互相排斥的成分。
- 中括号指示可选的成分。
- 大括号指示必选项。
- 中括号中的大括号指示可选项中的必选项。
- 黑体字表示合法输入的命令和关键字。在实际的配置和输出(不是通用的命令语法)中, 黑体字表示由用户手工输入的命令(例如 **show** 命令)。
- 斜体字表示需要提供实际值的参量。

前 言

自从《TCP/IP 路由技术(第 1 卷)》发行以来,在 Cisco 出版社的 CCIE 开发系列中又出版了许多新书。CCIE 编程本身也扩展到包括专业方面的许多不同的领域。但是 IP 路由协议还保留着基本功能,CCIE 的报考者必须在这些基本知识的基础上掌握自己的专业技能。这就如同地基不牢,房子早晚会坍塌是同一个道理。

在第 1 卷的简介中我曾声明:“……随着互联网在规模以及复杂性方面的增长,路由问题会立即变得巨大而且错综复杂。”由于我们从内部网关协议转移到既检查自治系统之间的路由又要解决更多特殊的路由问题,例如多播和 IPv6,因此,扩展性以及对此种增长的管理将还是本书论述的中心问题。

通过 CCIE 实验室考试后,你的名字后面会增加一个经过考评的,有价值的 CCIE 号,但是这本书的主要目的是不仅要帮助你从这种 CCIE 实验室考试中走出来,而且要增加你的知识和技能从而使你能够无愧于 CCIE 资格。正如我在第 1 卷中所讲的,我的目的是使你成为一个真正的 CCIE,而不是简单地通过 CCIE 考试。在这本书中,你会发现更多的信息而不只是那些通过实验室考试所需要的知识,但是所有的材料对于一个被认知的互联网专家来讲都是非常重要的。

当我学习 CCIE 的时候,实验室大部分都是由 AGS+路由器组成的。当然,实验室和考试的本质相对于以前来讲都有了重大的改变。如果还有什么不同的话,就是现在实验室考试更难了。CCIE 过程增加的另外一个部分就是重新认证要求。甚至在我参加重新认证之前,就有人告诉我,第 1 卷为他们准备考试提供了很大的帮助——尤其是在准备 IS-IS 方面,因为除了业务供应商之外很少有人使用该协议。因此,我在写第 2 卷的时候,不仅考虑了 CCIE 的投考者,还考虑到已经是 CCIE 的人在进行重新认证时复习的需要。关于多播和 IPv6 的章节就直接起到这种指导作用。

我努力使本书与第 1 卷的结构相同。在第 1 卷中,通过普通的术语介绍一种协议,然后讲述通过 Cisco IOS 软件来进行协议配置的例子,最后讲述的是故障排除用到的 Cisco IOS 软件工具。在讲解 BGP 和 IP 多播的时候,这种结构对于一个单一的章节来讲太长了,因此我把它们分布到了多个章节之中。

我希望你们在读这本书的时候,能像我写这本书时所获得的知识一样多。

作者简介

Jeff Doyle, CCIE #1919, 是位于美国丹佛的 Juniper Networks 公司的专业业务顾问。他精通 IP 路由协议以及 MPLS 业务量工程。在北美洲、欧洲和亚洲, Jeff 帮助设计并实现了许多大型的 Internet 业务供应商的网络。Jeff 还针对更高级的网络技术在业务供应商论坛例如 NANOG 和 APRICOT 做了讲座。在加入 Juniper Networks 之前, Jeff 是国际网络业务方面的资深网络顾问。Jeff 的联系方式是 jeff@juniper.net。

Jennifer DeHaven Carroll, 是朗讯科技的主要顾问, 也是 Cisco 认证的互联网方面的专家(CCIE #1402)。在过去的 13 年中, 她规划、设计并组建了许多大型的网络, 同时还开发和讲授全 IP 路由协议的理论以及 Cisco 的应用。Jenny 的联系方式是 jennifer.carroll@ieee.org。

审稿人简介

Henry Benjamin, CCIE #4695, CCNA, CCDA, 工学学士, 是 Cisco 认证的互联网专家并且是 Cisco 公司的 IT 网络设计工程师。他有 8 年多在 Cisco 的工作经验, 包括规划、设计、组建大型的运行 IGRP、EIGRP 和 OSPF 协议的 IP 网络。现在 Henry 在澳大利亚悉尼的 Cisco 内部的 IP 设计组工作。Henry 获得了悉尼大学的工学学士学位。

Peter J.Moyer, CCIE#3286, 是 Juniper Networks 的专业业务顾问, 他曾经在该公司设计并组建了大型的 ISP 网络。除了他的顾问工作, Peter 还为 Juniper 的客户以及合作者建立了高级 IP 培训课程并创办了 IP 网络设计研讨会。在网络讨论会上他负责讲述 MPLS 等先进的技术。在加入 Juniper 之前, Peter 是国际网络服务公司(INS)的资深网络顾问, 他在这里设计并组建了大型的企业网。Peter 获得了马里兰大学计算机与信息科学的工学学士学位。

声 明

Jeff Doyle: 一本技术书籍的作者就像是一个由一群有才气以及有献身精神的人组成的小部队的前锋，这本书也不例外。听起来就好像我在进行接受学术奖的演讲，但是我还是要感谢许多人。

首先并且最重要的，我要感谢 Jenny Carroll，作为一名技术编辑，她在第 1 卷中的努力令人惊异。作为技术编辑，Jenny 不仅继续在本书中贡献她的专业知识，而且当我感觉到无法预期完成该书的时候，她作为一名合著者加入进来，在我的请求下，她写了最后两章。如果没有她无价的建议和对细节的注意，这本书将无法成功。

同样我要感谢 Pete Moyer，他是我的朋友和助手，他作为一名技术编辑加入到第 2 卷的工作当中。Pete 对该项目之外的我的生活有着重要的影响，我将永远感激他。

我还要感谢 Laurie McGuire 以及 Chris Cleveland，感谢他们作为开发编辑而给予我的专业指导。他们使该书更加完美同时也使我成了一名更好的作者。

感谢 Brett Bartow 以及 Cisco Press 的所有朋友们，他们在我努力完成本书的过程中给了我极大的关心，从而使我能够在最终期限之前完成本书。我敢肯定，他们对我的第一本书不是很满意，但是在整个项目进行的过程中，他们还是非常善意地对待我。

最后，我要感谢你们，亲爱的读者，是你们让我的第一本书如此成功，并且如此耐心地等待第 2 卷的完成。我希望这本书值得你们的等待。

Jennifer Dehaven Carroll: 我要感谢 Jeff Doyle，是他给了我机会，让我为他的书做了一些贡献，这是一个有趣且富有挑战性的工作。

目 录

第一部分 外部网关协议(EGP)

第 1 章 外部网关协议	2
1.1 EGP 的起源	2
1.2 EGP 的操作	3
1.2.1 EGP 拓扑问题	3
1.2.2 EGP 的功能	5
1.2.3 EGP 消息格式	12
1.3 EGP 的不足	18
1.4 配置 EGP	19
1.4.1 案例研究: 一个 EGP 末梢网关	19
1.4.2 案例研究: 一个 EGP 核心网关	22
1.4.3 案例研究: 间接邻居	25
1.4.4 案例研究: 缺省路由	27
1.5 EGP 的故障排除	28
1.5.1 解释邻居表	29
1.5.2 案例研究: 聚合到 Syrup 的速度	30
1.6 尾注	31
1.7 展望	32
1.8 复习问题	32
1.9 配置练习	33
1.10 故障排除练习	36
第 2 章 BGP4 简介	38
2.1 无类域间路由	38
2.1.1 归纳摘要	39
2.1.2 无类路由	40
2.1.3 路由总结: 优势、劣势以及不对称性	43
2.1.4 Internet: 经过多年后还保持着分层结构	45
2.1.5 CIDR: 减轻了路由表的爆炸性增长	48
2.1.6 CIDR: 降低了 B 类地址空间的消耗	51
2.1.7 CIDR 遇到的问题	51
2.2 谁需要 BGP	54

2.2.1	一个单宿主自治系统	55
2.2.2	多宿主到一个单一的 AS	57
2.2.3	多宿主到多个自治系统	60
2.2.4	“负载均衡”中应当注意的一个问题	62
2.2.5	BGP 的危险	63
2.3	BGP 基础知识	64
2.3.1	BGP 消息类型	66
2.3.2	BGP 有限状态机	67
2.3.3	路径属性	70
2.3.4	管理权值	78
2.3.5	AS_SET	79
2.3.6	BGP 决策过程	80
2.3.7	路由抑制	82
2.4	IBGP 和 IGP 的同步	83
2.5	管理大型 BGP 对等关系	88
2.5.1	对等组	88
2.5.2	团体	88
2.5.3	路由反射器	88
2.5.4	联盟	93
2.6	BGP 消息格式	94
2.6.1	Open 消息	95
2.6.2	Update 消息	96
2.6.3	Keepalive 消息	97
2.6.4	Notification 消息	97
2.7	尾注	99
2.8	展望	99
2.9	推荐的读物	99
2.10	复习题	99
第 3 章	BGP4 的配置以及故障排除	105
3.1	基本的 BGP 配置	105
3.1.1	案例研究: 建立 BGP 路由器之间的对等	105
3.1.2	案例研究: 向 BGP 中注入 IGP 路由	110
3.1.3	案例研究: 向 IGP 注入 BGP 路由	115
3.1.4	案例研究: 没有 IGP 的 IBGP	120
3.1.5	案例研究: IGP 上的 IBGP	126
3.1.6	案例研究: EBGP 多跳	132
3.1.7	案例研究: 聚合路由	135
3.1.8	管理 BGP 连接	150
3.2	路由策略	153

3.2.1	重置 BGP 连接	153
3.2.2	案例研究: 通过 NLRI 过滤路由	155
3.2.3	案例研究: 通过 AS_PATH 过滤路由	161
3.2.4	案例研究: 通过路由图过滤路由	164
3.2.5	案例研究: 管理权值	166
3.2.6	案例研究: 管理距离以及后门路由	173
3.2.7	案例研究: 使用 LOCAL_PREF 属性	178
3.2.8	案例研究: 使用 MULTI_EXIT_DISC 属性	182
3.2.9	案例分析: 附加 AS_PATH	187
3.2.10	案例分析: 路由标记	190
3.2.11	案例分析: 路由抑制	194
3.3	大型 BGP	197
3.3.1	案例分析: BGP 对等组	198
3.3.2	案例分析: BGP 团体	201
3.3.3	案例分析: 专用 AS 号	212
3.3.4	案例分析: BGP 联盟	215
3.3.5	案例分析: 路由反射器	225
3.4	展望	230
3.5	推荐的读物	230
3.6	命令归纳	231
3.7	配置练习	235
3.8	故障排除练习	240

第二部分 高级 IP 路由问题

第 4 章	网络地址翻译	250
4.1	NAT 的操作	250
4.1.1	NAT 的基本概念	250
4.1.2	NAT 和 IP 地址的保存	252
4.1.3	NAT 和 ISP 的变更	254
4.1.4	NAT 和多宿主 AS	255
4.1.5	端口地址翻译	257
4.1.6	NAT 和 TCP 负载分配	258
4.1.7	NAT 和虚拟服务器	259
4.2	NAT 的问题	260
4.2.1	信头校验和	260
4.2.2	分段	260
4.2.3	加密	260
4.2.4	安全性	261
4.2.5	具体协议涉及到的问题	261

4.3	配置 NAT	268
4.3.1	案例研究: 静态 NAT	268
4.3.2	案例研究: 动态 NAT	274
4.3.3	案例研究: 网络合并	278
4.3.4	案例研究: 用 NAT 实现 ISP 多宿	281
4.3.5	端口地址翻译	286
4.3.6	案例研究: TCP 负载均衡	287
4.3.7	案例研究: 服务分配	288
4.4	NAT 故障排除	290
4.5	尾注	292
4.6	展望	292
4.7	命令归纳	292
4.8	配置练习	293
4.9	故障排除练习	295
第 5 章	IP 多播路由介绍	297
5.1	对 IP 多播的要求	299
5.2	组成员概念	303
5.2.1	加入和退出组	304
5.2.2	因特网组管理协议(IGMP)	308
5.2.3	Cisco 组员资格协议(CGMP)	313
5.3	多播路由的问题	320
5.3.1	多播的前转	320
5.3.2	多播路由	321
5.3.3	稀疏与密集拓扑的比较	322
5.3.4	隐式加入与显式加入的比较	323
5.3.5	基于源的树与共享树的比较	325
5.3.6	多播的范围	326
5.4	距离向量多播路由协议(DVMRP)的操作	329
5.4.1	对邻居的发现和维持	330
5.4.2	DVMRP 路由表	330
5.4.3	DVMRP 包的前转	332
5.4.4	DVMRP 消息的格式	332
5.5	MOSPF 的操作	338
5.5.1	MOSPF 基础	339
5.5.2	区域间的 MOSPF	340
5.5.3	AS 间的 MOSPF	342
5.5.4	MOSPF 扩展的格式	343
5.6	基于核心的树(CBT)的操作	345
5.6.1	CBT 基础	345

5.6.2	寻找核心	346
5.6.3	CBT 指定路由器	347
5.6.4	成员与非成员的多播源	348
5.6.5	CBT 消息格式	349
5.7	与协议无关的多播(PIM)的介绍	353
5.8	与协议无关多播, 密集模式(PIM-DM)的操作	354
5.8.1	PIM-DM 基础	354
5.8.2	Prune 消息的覆盖	359
5.8.3	单播路由的改变	361
5.8.4	PIM-DM 指定路由器	361
5.8.5	PIM 前转器的选举	361
5.9	与协议无关的多播, 稀疏模式(PIM-SM)的操作	364
5.9.1	PIM-SM 基础	364
5.9.2	查找会聚点	365
5.9.3	PIM-SM 和共享树	367
5.9.4	源的注册	369
5.9.5	PIM-SM 与最短路径树	375
5.9.6	PIMv2 消息格式	379
5.10	尾注	385
5.11	展望	386
5.12	推荐读物	386
5.13	命令归纳	386
5.14	复习问题	388
第 6 章	IP 多播路由的配置和故障排除	394
6.1	配置 IP 多播路由	394
6.2	案例研究: 配置与协议无关多播, 密集模式(PIM-DM)	395
6.3	配置与协议无关多播, 稀疏模式 (PIM-SM)	403
6.3.1	案例研究: 静态配置 RP	403
6.3.2	案例研究: 配置 Auto-RP	409
6.3.3	案例研究: 配置稀疏——密集模式	416
6.3.4	案例研究: 配置自举协议	419
6.4	案例研究: 多播负荷分担	423
6.5	IP 多播路由的故障排除	429
6.5.1	使用 mrinto	430
6.5.2	mtrace 与 mstat 的使用	432
6.6	展望	436
6.7	配置练习	436
6.8	排错练习	438
第 7 章	大范围 IP 多播路由	441

7.1	多播范围控制	441
7.2	案例学习: 多播穿过非多播域	443
7.3	连接到 DVMRP 网络	445
7.4	AS 间多播	448
7.4.1	BGP 的多协议扩展(MBGP)	450
7.4.2	多播源发现协议(MSDP)运行	451
7.4.3	MSDP 消息格式	453
7.5	案例学习: 配置 MBGP	456
7.6	案例学习: 配置 MSDP	460
7.7	案例学习: MSDP 全连接组	464
7.8	案例学习: 泛播 RP	466
7.9	案例学习: MSDP 缺省对等实体	470
7.10	命令归纳	473
7.11	尾注	474
7.12	展望	474
7.13	复习问题	474
第 8 章	IPv6	476
8.1	IPv6 的设计目标	476
8.1.1	提高可扩展性	477
8.1.2	易于配置	477
8.1.3	安全性	478
8.2	当前 IPv6 状态	478
8.2.1	IPv6 规范(RFC)	478
8.2.2	厂商支持	479
8.2.3	实现	479
8.3	IPv6 包格式	480
8.3.1	IPv6 地址	480
8.3.2	地址空间	481
8.3.3	地址的文字表示	481
8.3.4	地址前缀的文字表示	482
8.3.5	地址类型分配	482
8.4	地址结构	484
8.4.1	可聚合全球地址格式	484
8.4.2	IPv6 头	493
8.5	IPv6 功能	497
8.5.1	在 Cisco 路由器上使能 IPv6 能力	497
8.5.2	ICMPv6	498
8.5.3	邻居发现	499
8.5.4	自动配置	506

8.5.5	路由	509
8.5.6	泛播处理过程	521
8.5.7	多播	522
8.5.8	服务质量	526
8.6	从 IPv4 向 IPv6 过渡	526
8.6.1	双协议栈	527
8.6.2	DNS	527
8.6.3	IPv4 中的 IPv6 隧道	528
8.6.4	网络地址翻译—协议翻译	530
8.7	尾注	530
8.8	展望	530
8.9	推荐书目	531
8.10	复习问题	531
8.11	参考文献	533
第 9 章	路由器管理	535
9.1	规则和程序定义	536
9.1.1	服务等级协议	536
9.1.2	改变管理	536
9.1.3	扩大提交过程程序	538
9.1.4	更新规则	538
9.2	简单网络管理协议	538
9.2.1	SNMP 概述	538
9.2.2	CiscoWorks	540
9.2.3	路由器的 SNMP 配置	540
9.3	RMON	545
9.3.1	RMON 概述	545
9.3.2	路由器的 RMON 配置	546
9.4	记录日志	548
9.5	系统日志(Syslog)	551
9.5.1	Syslog 概述	551
9.5.2	路由器上 Syslog 的配置	552
9.6	网络时间协议(NTP)	553
9.6.1	NTP 概述	553
9.6.2	路由器的 NTP 配置	554
9.7	记账	557
9.7.1	IP 记账	558
9.7.2	NetFlow	559
9.8	配置管理	564
9.9	故障管理	565

9.10	性能管理	567
9.11	安全管理	567
9.11.1	口令类型和加密	568
9.11.2	控制交互式访问	568
9.11.3	减少拒绝服务攻击的危险	569
9.11.4	TACACS+	570
9.11.5	RADIUS	575
9.11.6	安全的命令解释器	576
9.12	设计支持管理程序的服务器	577
9.13	网络健壮性	577
9.13.1	HSRP	577
9.13.2	多组 HSRP	578
9.13.3	配置 HSRP	579
9.13.4	配置 MHSRP	582
9.14	实验室	583
9.15	推荐书目	584
9.16	尾注	585
9.17	展望	585
9.18	命令归纳	585
9.19	复习问题	589
9.20	配置练习	590
9.21	参考文献	590

第三部分 附 录

附录 A	show ip bgp neighbors 的显示	594
附录 B	正则表达式指南	599
附录 C	保留的多播地址	603
附录 D	复习问题的答案	619
附录 E	配置练习的答案	631
附录 F	故障排除练习答案	664