



CCIE职业发展系列
CCIE Professional Development

ciscopress.com



TCP/IP路由技术

(第二卷) (英文版)

Routing TCP/IP

Volume II

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

[美] **Jeff Doyle, CCIE #1919** 著
Jennifer DeHaven Carroll, CCIE #1402

 **人民邮电出版社**
POSTS & TELECOM PRESS

CCIE职业发展系
CCIE Professional Development

TCP/IP路由技术

(第二卷) (英文版)

Routing TCP/IP
Volume II

[美] Jeff Doyle, CCIE #1919 著
Jennifer DeHaven Carroll, CCIE #1402

人民邮电出版社
北京

图书在版编目 (CIP) 数据

TCP/IP路由技术. 第2卷: 英文 / (美) 多伊尔 (Doyle, J.), (美) 卡罗尔 (Carroll, J. D.) 著. —2版. —北京: 人民邮电出版社, 2009. 9
(CCIE职业发展系列)
ISBN 978-7-115-21099-9

I. T… II. ①多…②卡… III. 计算机网络—通信协议—路由选择—英文 IV. TN915.04

中国版本图书馆CIP数据核字 (2009) 第121296号

版权声明

Authorized translation from the English language edition, entitled Routing TCP/IP, Volume II, 9781578700899 by Jeff Doyle and Jennifer DeHaven Carroll, published by Pearson Education, Inc, publishing as Cisco Press, Copyright © 2001 Pearson Education, Inc.

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.

仅限于中华人民共和国境内 (不包括中国香港、澳门特别行政区和中国台湾地区) 销售。
本书封面贴有 Pearson Education (培生教育出版集团) 激光防伪标签。无标签者不得销售。

CCIE 职业发展系列

TCP/IP 路由技术 (第二卷) (英文版)

-
- ◆ 著 [美] Jeff Doyle, CCIE#1919
Jennifer DeHaven Carroll, CCIE#1402
- 责任编辑 李 际
执行编辑 赵 轩
- ◆ 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>
北京铭成印刷有限公司印刷
- ◆ 开本: 787×1092 1/16
印张: 60.5
字数: 1371 千字 2009 年 9 月第 2 版
印数: 3 001—5 500 册 2009 年 9 月北京第 1 次印刷
著作权合同登记号 图字: 01-2009-3161 号

ISBN 978-7-115-21099-9

定价: 128.00 元

读者服务热线: (010)67132705 印装质量热线: (010)67129223

反盗版热线: (010)67171154

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——网络地址转换、IP 组播路由技术、IPv6 技术、路由器管理等。本书共分 9 章, 首先介绍并讨论了有关主题的基本原理, 接着是设计用来展示实际网络环境中相关概念的一系列配置范例, 最后提供给读者经实践验证过的故障排除方法以解决网络可能会出现的问题。重点介绍了自治系统之间的路由选择策略以及诸如组播和 IPv6 等更具挑战性的路由选择和实施技术。

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, 是备战 CCIE 认证考试的经典之作, 适合准备参加 CCIE 考试的人员、网络与通信系统工程技术人员阅读。

About the Authors

Jeff Doyle, CCIE #1919, is a Professional Services Consultant with Juniper Networks, Inc. in Denver, Colorado. Specializing in IP routing protocols and MPLS Traffic Engineering, Jeff has helped design and implement large-scale Internet service provider networks throughout North America, Europe, and Asia. Jeff has also lectured on advanced networking technologies at service provider forums such as the North American Network Operators' Group (NANOG) and the Asia Pacific Regional Internet Conference on Operational Technologies (APRICOT). Prior to joining Juniper Networks, Jeff was a Senior Network Systems Consultant with International Network Services. Jeff can be contacted at jeff@juniper.net.

Jennifer DeHaven Carroll, is a principal consultant with Lucent technologies and is a Cisco Certified Internetwork Expert (CCIE # 1402). She has planned, designed, and implemented many large networks over the past 13 years. She has also developed and taught theory and Cisco implementation classes on all IP routing protocols. Jenny can be reached at jennifer.carroll@ieee.org.

About the Technical Reviewers

Henry Benjamin, CCIE #4695, CCNA, CCDA, B. Eng., is a Cisco certified Internet Expert and an IT Network Design Engineer for Cisco Systems, Inc. He has more than eight years of experience in Cisco networks, including planning, designing, and implementing large IP networks running IGRP, EIGRP, and OSPF. Currently Henry is working for the IT design team internally at Cisco in Sydney, Australia. Henry holds a Bachelor of Engineering degree from Sydney University.

Peter J. Moyer, CCIE #3286, is a Professional Services Consultant for Juniper Networks, where he designs and implements large-scale ISP networks. In addition to his consulting work, Peter has developed and delivered advanced IP training courses and IP network design seminars to Juniper customers and partners. He has presented at networking conferences on such advanced topics as MPLS. Before joining Juniper, Peter was a Senior Network Consultant for International Network Services (INS), where he designed and implemented large-scale enterprise networks. Peter holds a Bachelor of Science degree in Computer and Information Science from the University of Maryland.

Dedications

Jeff Doyle: This book is dedicated to my wife, Sara, and my children, Anna, Carol, James, and Katherine. They are my refuge, and they keep me sane, humble, and happy.

Jennifer DeHaven Carroll: To my husband, Mike, and son, Mitchell, who continue to encourage me.

Acknowledgments

Jeff Doyle: An author of a technical book is just a front man for a small army of brilliant, dedicated people. This book is certainly no exception. At the risk of sounding like I'm making an Academy Award acceptance speech, I would like to thank a number of those people.

First and foremost, I would like to thank Jenny Carroll, whose efforts as a technical editor on *Volume I* were amazing. Not only has Jenny again contributed her technical expertise to this second volume as a technical editor, but when I became hopelessly behind schedule, she stepped in as a coauthor, at my request, and wrote the last two chapters. Neither volume would be what they are without her invaluable advice and attention to detail.

I would also like to thank Pete Moyer, my friend and associate, who came aboard as a technical editor for this second volume. Pete has had a profound influence on my life beyond this project, and I will always be indebted to him.

My gratitude goes to Laurie McGuire and Chris Cleveland for their expert guidance as development editors. They have made the book a better book and me a better writer.

Thanks to Brett Bartow and all the folks at Cisco Press for their enormous patience with me as I struggled to finish the book and let deadline after deadline slip. They continued to show me great kindness throughout the project when I'm sure they would have preferred to bash me on the head with a copy of my first book.

Finally, I would like to thank you, good reader, for making the first book such a success and for waiting so patiently for me to finish this second volume. I hope the book proves to be worth the wait.

Jennifer DeHaven Carroll: I'd like to thank Jeff Doyle for giving me the opportunity to contribute to his books. It has been fun and challenging.

Introduction

Since the publication of *Volume I of Routing TCP/IP*, many volumes have been added to the Cisco Press CCIE Professional Development series. And the CCIE program itself has expanded to include various areas of specialization. Yet the IP routing protocols remain the essential foundation on which the CCIE candidate must build his or her expertise. If the foundation is weak, the house will tumble.

I stated in the introduction to *Volume I* that "...as internetworks grow in size and complexity, routing issues can become at once both large and subtle." Scalability and management of growth continues to be a central theme in this second volume, as we move beyond the interior gateway protocols to examine both interautonomous system routing and more exotic routing issues such as multicasting and IPv6.

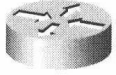
My objective in this book is not only to help you walk away from the CCIE lab exam with one of those valued and valuable numbers after your name, but also to help you develop the knowledge and skills to live up to the CCIE title. As with the first volume, I want to make CCIEs, not people who can pass the CCIE lab. In this vein, you will find in this book more information than you will need to pass the lab, but certainly all of the material is important in your career as a recognized internetworking expert.

When I earned my CCIE, the lab still consisted mostly of AGS+ routers. Certainly the lab and the nature of the exam have changed substantially since that ancient time. If anything, the lab is more difficult now. Another addition to the CCIE program has been the recertification requirement. Even before I took the recertification exam for the first time, people were telling me how much *Volume I* had helped them prepare for the test—particularly for IS-IS, a protocol that few outside of service provider environments are exposed to. I have therefore written this second volume with not only CCIE candidates in mind, but also existing CCIEs who need to review for their recertification. The chapters on multicasting and IPv6 are directed to this audience.

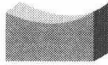
I have endeavored to follow the same structure that I followed in *Volume I*, in which a protocol is introduced in generic terms, followed by examples of configuring the protocol using Cisco IOS Software, and finally by examples of Cisco IOS Software tools for troubleshooting the protocol. In the case of BGP and IP multicast, this structure is far too lengthy for a single chapter and therefore spans multiple chapters.

I hope you learn as much from reading this book as I have from writing it.

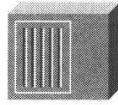
Icons Used in This Book



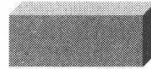
Router



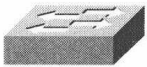
Bridge



Hub



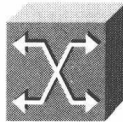
DSU/CSU



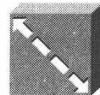
Catalyst switch



Multilayer switch



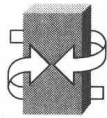
ATM switch



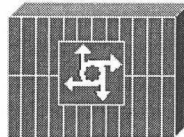
ISDN switch



Communication server



Gateway



Access server



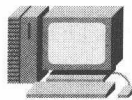
PC



PC with software



Sun Workstation



Mac



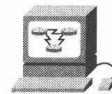
Terminal



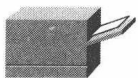
File server



Web server



CiscoWorks Workstation



Printer



Laptop



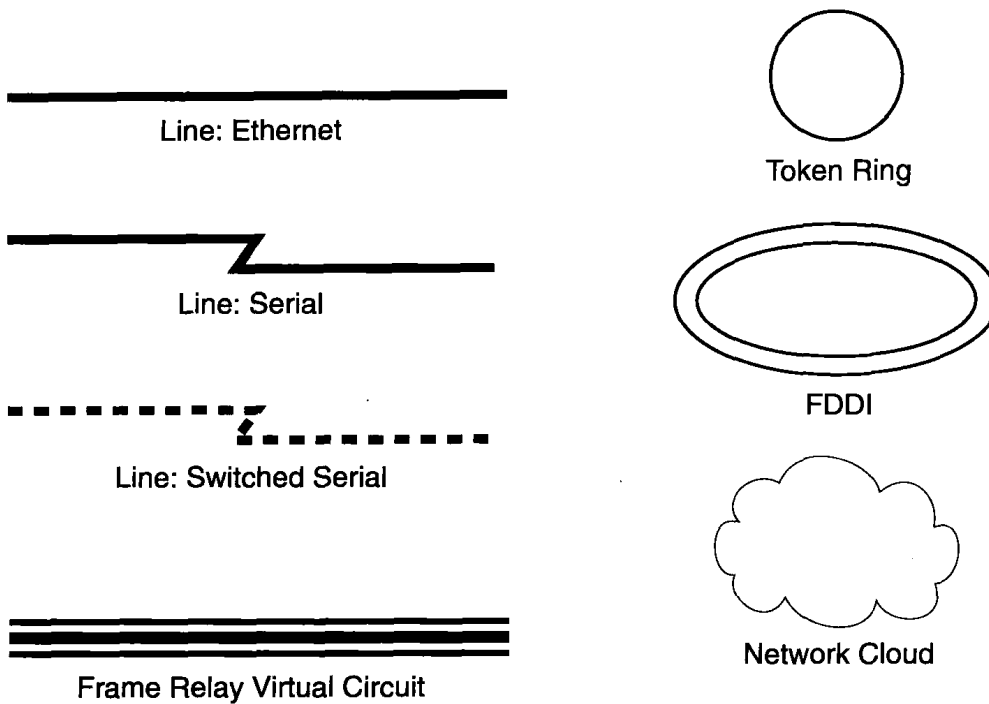
IBM mainframe



Front-End Processor



Cluster Controller



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the IOS Command Reference. The Command Reference describes these conventions as follows:

- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets [] indicate optional elements.
- Braces { } indicate a required choice.
- Braces within brackets [{ }] indicate a required choice within an optional element.

- **Boldface** indicates commands and keywords that are entered literally as shown. In actual configuration examples and output (not general command syntax), boldface indicates commands that are manually input by the user (such as a **show** command).
- *Italics* indicates arguments for which you supply actual values.

Table of Contents

Part I Exterior Gateway Protocols 3

Chapter 1 Exterior Gateway Protocol 5

- The Origins of EGP 5
- Operation of EGP 7
 - EGP Topology Issues 7
 - EGP Functions 9
 - EGP Message Formats 18
- Shortcomings of EGP 26
- Configuring EGP 26
 - Case Study: An EGP Stub Gateway 27
 - Case Study: An EGP Core Gateway 31
 - Case Study: Indirect Neighbors 35
 - Case Study: Default Routes 38
- Troubleshooting EGP 39
 - Interpreting the Neighbor Table 39
 - Case Study: Converging at the Speed of Syrup 41
- End Notes 43
- Looking Ahead 43
- Review Questions 44
- Configuration Exercises 45
- Troubleshooting Exercise 51

Chapter 2 Introduction to Border Gateway Protocol 4 55

- Classless Interdomain Routing 55
 - A Summarization Summary 56
 - Classless Routing 57

- Summarization: The Good, the Bad, and the Asymmetric 61
- The Internet: Still Hierarchical After All These Years 64
- CIDR: Reducing Routing Table Explosion 69
- CIDR: Reducing Class B Address Space Depletion 71
- Difficulties with CIDR 72

Who Needs BGP? 76

- A Single-Homed Autonomous System 78
- Multihoming to a Single Autonomous System 80
- Multihoming to Multiple Autonomous Systems 84
- A Note on “Load Balancing” 88
- BGP Hazards 88

BGP Basics 90

- BGP Message Types 93
- The BGP Finite State Machine 95
- Path Attributes 98
- Administrative Weight 111
- AS_SET 112
- The BGP Decision Process 114
- Route Dampening 116

IBGP and IGP Synchronization 118

Managing Large-Scale BGP Peering 124

- Peer Groups 124
- Communities 124
- Route Reflectors 125
- Confederations 131

BGP Message Formats 132

- The Open Message 133
- The Update Message 134
- The Keepalive Message 137
- The Notification Message 137

End Notes 138

Looking Ahead 139

Recommended Reading 139

Review Questions 139

Chapter 3 Configuring and Troubleshooting Border Gateway Protocol 4 149

Basic BGP Configuration 149

- Case Study: Peering BGP Routers 149

- Case Study: Injecting IGP Routes into BGP 155
- Case Study: Injecting BGP Routes into an IGP 161
- Case Study: IBGP without an IGP 167
- Case Study: IBGP Over an IGP 174
- Case Study: EBGP Multihop 182
- Case Study: Aggregate Routes 185

Managing BGP Connections 204

Routing Policies 208

- Resetting BGP Connections 209
- Case Study: Filtering Routes by NLRI 211
- Case Study: Filtering Routes by AS_PATH 219
- Case Study: Filtering with Route Maps 224
- Case Study: Administrative Weights 226
- Case Study: Administrative Distances and Backdoor Routes 234
- Case Study: Using the LOCAL_PREF Attribute 241
- Case Study: Using the MULTI_EXIT_DISC Attribute 246
- Case Study: Prepending the AS_PATH 252
- Case Study: Route Tagging 256
- Case Study: Route Dampening 261

Large-Scale BGP 265

- Case Study: BGP Peer Groups 265
- Case Study: BGP Communities 270
- Case Study: Private AS Numbers 284
- Case Study: BGP Confederations 287
- Case Study: Route Reflectors 300

Looking Ahead 307

Recommended Reading 307

Command Summary 307

Configuration Exercises 312

Troubleshooting Exercises 318

Part II Advanced IP Routing Issues 331

Chapter 4 Network Address Translation 333

Operation of NAT 333

- Basic NAT Concepts 333
- NAT and IP Address Conservation 336
- NAT and ISP Migration 338
- NAT and Multihomed Autonomous Systems 340
- Port Address Translation 342

NAT and TCP Load Distribution 343

NAT and Virtual Servers 344

NAT Issues 345

Header Checksums 346

Fragmentation 346

Encryption 346

Security 347

Protocol-Specific Issues 347

Configuring NAT 356

Case Study: Static NAT 356

Case Study: Dynamic NAT 364

Case Study: A Network Merger 369

Case Study: ISP Multihoming with NAT 374

Port Address Translation 379

Case Study: TCP Load Balancing 381

Case Study: Service Distribution 383

Troubleshooting NAT 386

End Note 387

Looking Ahead 388

Command Summary 388

Configuration Exercises 389

Troubleshooting Exercises 392

Chapter 5 Introduction to IP Multicast Routing 395

Requirements for IP Multicast 398

Multicast IP Addresses 399

Group Membership Concepts 404

Internet Group Management Protocol (IGMP) 411

Cisco Group Membership Protocol (CGMP) 419

Multicast Routing Issues 427

Multicast Forwarding 428

Multicast Routing 429

Sparse Versus Dense Topologies 431

Implicit Joins Versus Explicit Joins 432

Source-Based Trees Versus Shared Trees 435

Multicast Scoping 437

Operation of the Distance Vector Multicast Routing Protocol (DVMRP) 441

Neighbor Discovery and Maintenance 442

The DVMRP Routing Table	443
DVMRP Packet Forwarding	444
DVMRP Message Formats	445
Operation of Multicast OSPF (MOSPF)	453
MOSPF Basics	454
Inter-Area MOSPF	456
Inter-AS MOSPF	459
MOSPF Extension Formats	460
Operation of Core-Based Trees (CBT)	462
CBT Basics	463
Finding the Core	465
CBT Designated Routers	465
Member and Nonmember Sources	466
CBT Message Formats	468
Introduction to Protocol Independent Multicast (PIM)	473
Operation of Protocol Independent Multicast, Dense Mode (PIM-DM)	475
PIM-DM Basics	475
Prune Overrides	481
Unicast Route Changes	484
PIM-DM-Designated Routers	484
PIM Forwarder Election	484
Operation of Protocol Independent Multicast, Sparse Mode (PIM-SM)	487
PIM-SM Basics	488
Finding the Rendezvous Point	489
PIM-SM and Shared Trees	492
Source Registration	496
PIM-SM and Shortest Path Trees	502
PIMv2 Message Formats	507
End Notes	517
Looking Ahead	518
Recommended Reading	518
Command Summary	519
Review Questions	520

Chapter 6	Configuring and Troubleshooting IP Multicast Routing	531
	Configuring IP Multicast Routing	531

Case Study: Configuring Protocol-Independent Multicast, Dense Mode (PIM-DM) 533

Configuring Protocol-Independent Multicast, Sparse Mode (PIM-SM) 542

Case Study: Multicast Load Sharing 568

Troubleshooting IP Multicast Routing 576

Using mrinfo 577

Using mtrace and mstat 580

Looking Ahead 584

Configuration Exercises 585

Troubleshooting Exercises 587

Chapter 7 Large-Scale IP Multicast Routing 593

Multicast Scoping 593

Case Study: Multicasting Across Non-Multicast Domains 595

Connecting to DVMRP Networks 598

Inter-AS Multicasting 602

Multiprotocol Extensions for BGP (MBGP) 606

Operation of Multicast Source Discovery Protocol (MSDP) 607

MSDP Message Formats 609

Case Study: Configuring MBGP 614

Case Study: Configuring MSDP 619

Case Study: MSDP Mesh Groups 624

Case Study: Anycast RP 627

Case Study: MSDP Default Peers 633

Command Summary 636

End Notes 637

Looking Ahead 637

Review Questions 638

Chapter 8 IP Version 6 643

Design Goals of IPv6 643

Improve Scalability 644

Ease of Configuration 645

Security 645