



CCIE 职业发展系列
CCIE Professional Development

ciscopress.com



Internet 路由结构

(第二版)(英文版)

Internet Routing Architectures

Second Edition

The definitive BGP resource

[美] Sam Halabi 著
Danny McPherson

人民邮电出版社
POSTS & TELECOMMUNICATIONS PRESS

图书在版编目 (CIP) 数据

Internet 路由结构: 第 2 版 / (美) 哈拉比 (Halabi, S.), (美) 麦克弗森 (Mcpherson, D.) 著. —北京: 人民邮电出版社, 2003.5

(CCIE 职业发展系列)

ISBN 7-115-11216-9

I. I... II. ①哈... ②麦... III. 因特网—路由选择—结构—英文 IV. TP393.4

中国版本图书馆 CIP 数据核字 (2003) 第 025014 号

版 权 声 明

Original edition, Internet Routing Architectures, 2nd Edition, 157870233X, by Sam Halabi and Danny McPherson, published by Pearson Education, Inc., publishing as Cisco Press, Copyright © 2001 Cisco Press.

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.

English reprint published by Pearson Education North Asia Limited and Posts and Telecommunication Press, Copyright © 2003.

This edition is manufactured in the People's Republic of China, and is authorized for sale and distribution in the People's Republic of China exclusively (except Taiwan, Hong Kong SAR and Macao SAR).

本书封面贴有 **Pearson Education** 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

Internet 路由结构 (第二版) (英文版)

◆ 著 [美] Sam Halabi
Danny McPherson
责任编辑 杨长青

◆ 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 010-67132705
北京汉魂图文设计有限公司制作
北京顺义振华印刷厂印刷
新华书店总店北京发行所经销

◆ 开本: 800×1000 1/16
印张: 32.75
字数: 729 千字 2003 年 6 月第 1 版
印数: 1-3 500 册 2003 年 6 月北京第 1 次印刷

著作权合同登记 图字: 01-2003-1485 号

ISBN 7-115-11216-9/TP · 3415

定价: 60.00 元

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

内 容 提 要

本书对当代 Internet 的结构、服务提供商和编址技术进行了介绍。讲述了链路状态和距离向量路由选择协议的特性以及为什么需要域间路由选择协议和它们是如何工作的，并深入探讨了 BGP 的实用的、面向设计的应用。书中包括很多 BGP 属性的编码实例和许多路由选择策略。本书通过在概念上和实例上提供实用的编址、选路和连通的技术，旨在培养读者对路由选择的理解，以便能以一种客观而又有根据的方法计划并完成主要的网络设计。学习本书，可以帮助你成为数据路由选择操作方面的专家。

本书适合于网络技术专业人员和网络爱好者阅读。

About the Authors



Sam Halabi is one of the industry's foremost experts in the Internet Service Provider line of business. Mr. Halabi was recently Vice President of Marketing at an IP networking startup and has spent many years at Cisco Systems where he led the IP Carrier Marketing effort. Mr. Halabi is an expert in complex routing protocols and has specialized in the design of large-scale IP networks.

An active member in the industry, Halabi serves as a board member of the Optical Internet-working Forum and a member of the MPLS Forum.

Danny McPherson is currently Director of Architecture, Office of the CTO, at Amber Networks. Formerly, he held technical leadership positions with four Internet service providers (Qwest, GTE Internetworking, Genuity, and internetMCI), where he was responsible for network and product architecture, routing design, peering, and other business- and policy-related issues. McPherson is an active contributor to the Internet Engineering Task Force (IETF), as well as several other standards bodies. He is an acknowledged expert in Internet architecture and routing protocols.

About the Technical Reviewers

Alexei Roudnev is currently a Software System Engineer for Genesys Labs/Alcatel group in, San Francisco, California. He worked for 10 years as a Network Engineer at Relcom Network, one of the creators of the Russian Internet, in Moscow, Russia. Alexei was also a UNIX based systems Software Developer in Moscow for 9 years.

Abha Ahuja is currently a Senior Network Engineer at Internap Network Services. She works on network design, architecture and operational issues. Previous to Internap, she worked at Merit Network, a leading network research institution where she worked on the Route Server Next Generation project, a nationwide deployment of routing servers at exchange points, and the Internet Performance Measurement and Analysis (IPMA) project. She continues to play an active role in the Internet community and pursues research interests including inter-domain routing behavior and protocols, network operations and performance statistics, and network security. She is a skilled network engineer, certified troublemaker and a classic Scorpio.

Dedications

Danny McPherson: To my wife, Heather, and my two daughters, Kortney and Ashli. You are my infrastructure.

Acknowledgments

This book would not have been possible without the help of many people whose comments and suggestions significantly improved the end result. First, we would like to thank Abha Ahuja, Shane Amante, Johnson Liu, Alvaro Retana, and Alexander Rudenev for their exceptional technical review of this manuscript. We would also like to explicitly acknowledge Henk Smit, Bruce Cole, Enke Chen, Srihari Ramachandra, Rex Fernando, Satinder Singh, and Ravi Chandra, as well as the entire Cisco “BGP Coders” group, and everyone else who provided any amount of input for the second edition. Also, we would like to acknowledge the overwhelming support and patience of Danny McPherson’s present employer, Amber Networks, and previous employer, Qwest Communications, both of which had a significant impact on the value of the content. Finally, we would like to thank Christopher Cleveland, Tracy Hughes, Marc Fowler, Gayle Johnson, and the rest of the Cisco Press folks for keeping us on track and getting the book published.

Introduction

The Internet, an upstart academic experiment in the late 1960s, struggles with identity and success today. From the ARPANET to the NSFnet to ANYBODYSNET, the Internet is no longer owned by a single entity; it is owned by anybody who can afford to buy space on it. Tens of millions of users are seeking connectivity, and tens of thousands of companies are feeling left out if they do not tap into the Internet. This has put network designers and administrators under a lot of pressure to keep up with networking and connectivity needs. Understanding networking, and especially routing, has become a necessity.

Some people are surprised when networks fail and melt down, but others are surprised when they don't. This seems to be the case because there is so little useful information out there. Much of the information on routing that has been available to designers and administrators up until now is doubly frustrating: The information makes you think you know how to build your network—until you try, and find out that you don't. The first edition of this book addressed real routing issues, using real scenarios, in a comprehensive and accessible way.

In addition to providing a thorough update to the original material, this edition introduces recent enhancements to the BGP protocol, discusses changes surrounding registration and allocation of Internet numbers, and provides additional information on research and educational networks.

Objectives

The purpose of this book is to make you an expert on integrating your network into the global Internet. By presenting practical addressing, routing, and connectivity issues both conceptually and in the context of practical scenarios, this book aims to foster your understanding of routing so that you can plan and implement major network designs in an objective and informed way. Whether you are a customer or a provider (or both) of Internet connectivity, this book anticipates and addresses the routing challenges facing your network.

Audience

This book is intended for any organization that might need to tap into the Internet. Whether you are becoming a service provider or are connecting to one, you will find all you need to integrate your network. The perspectives of network administrators, integrators, and architects are considered throughout this book. Even though this book addresses different levels of expertise, it progresses logically from the simplest to the most challenging concepts and problems, and its common denominator is straightforward, practical scenarios to which anyone can relate. No major background in routing or TCP/IP is required. Any basic or background knowledge needed to understand routing is developed as needed in text discussions, rather than assumed as part of the reader's repertoire.

Organization

The book is organized into four parts:

- **Part I: The Contemporary Internet**—Chapters 1 through 3 cover essential introductory aspects of the contemporary Internet with respect to its structure, service providers, and addressing. Even if you are already familiar with the general structure of the Internet, you are encouraged to read the portions of Chapter 1 concerning Network Access Points, the Routing Arbiter Project, and Network Information Services. The pressures that precipitated these components of the Internet have continuing practical implications for routing design problems faced by administrators. Chapter 2 provides valuable criteria by which to evaluate Internet service providers. If you represent such a provider, or are already a customer of one, some of the information might be familiar to you already. Chapter 3 discusses classless interdomain rout-

ing (CIDR), VLSM (variable-length subnet masks), IPv6, and other aspects of Internet addressing.

- **Part II: Routing Protocol Basics**—Chapters 4 and 5 cover the basics: properties of link-state and distance vector routing protocols and why interdomain routing protocols are needed and how they work. These topics are covered both generally and in the specific context of BGP (Border Gateway Protocol)—the de facto standard interdomain routing protocol used in the Internet today. BGP’s particular capabilities and attributes are thoroughly introduced.
- **Part III: Effective Internet Routing Designs**—Chapters 6 through 10 delve into the practical, design-oriented applications of BGP. The BGP attributes introduced in Part II are shown in action, in a variety of representative network scenarios. BGP’s attributes are put to work in implementing design goals such as redundancy, symmetry, and load balancing. The challenges of making intradomain and interdomain routing work in harmony, managing growing or already-large systems, and maintaining stability are addressed.
- **Part IV: Internet Routing Device Configuration**—Chapters 11 and 12 contain numerous code examples of BGP’s attributes and of various routing policies. The code examples will make the most sense to you after you have read the earlier chapters, because many of them address multiple concepts and design goals. So that you can juxtapose textual discussions from earlier chapters with the code examples in Chapters 11 and 12, pointers called “Configuration Examples” appear in the earlier chapters. When you see one, you might want to fast-forward to the referenced page to see a configuration example of the attribute or policy being discussed.

Finally, several appendixes provide additional references for further reading, an up-to-date Cisco IOS™ BGP command reference, and information regarding IOS™ modifications intended to provide a more intuitive BGP command-line interface.

Approach

It is very hard to write about technical information in an accessible manner. Information that is stripped of too much technical detail loses its meaning, but complete and precise technical detail can overwhelm readers and obscure concepts. This book introduces technical detail gradually and in the context of practical scenarios whenever possible. The most heavily technical information—configuration examples in the Cisco IOS language—is withheld until the final two chapters of this book so that it is thoroughly grounded in the concepts and sample topologies that precede it.

Although your ultimate goal is to design and implement routing strategies, it is critical to grasp concepts and principles before applying them to your particular network. This book balances conceptual and practical perspectives by following a logical, gradual progression from general to specific, and from concepts to implementation. Even in chapters and sections that necessarily take a largely descriptive approach, hands-on interests are addressed through pointers to configuration examples, frequently asked questions, and scenario-based explanations.

The scenario-based approach is an especially important component of this book: it utilizes representative network topologies as a basis for illustrating almost every protocol attribute and routing policy discussed. Even though you might not see your exact network situation illustrated, the scenario is specific enough to facilitate learning by example, and general enough that you can extrapolate how the concepts illustrated apply to your situation.

Features and Text Conventions

This book works hard not to withhold protocol details and design-oriented information, while at the same time recognizing that building general and conceptual understanding necessarily comes first. Two features are included to help emphasize what is practical and design-oriented as underlying concepts are developed:

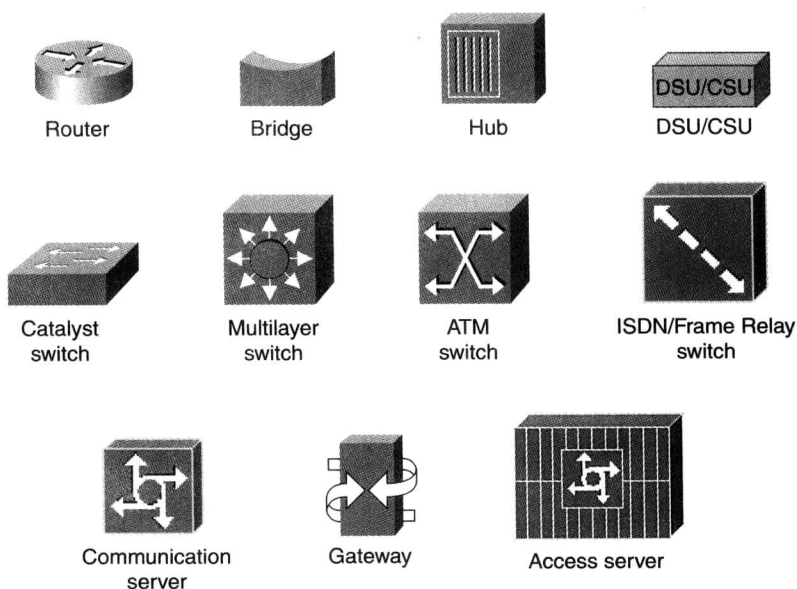
- **Pointers to configuration examples**—Located close to pertinent text discussions, these references point forward to places in Chapters 11 and 12 where related configuration examples can be found.
- **Frequently Asked Questions**—Located at the end of every chapter, these questions anticipate practical and design-oriented questions you might have, for your particular network, after having read the chapter.

Command Syntax Conventions

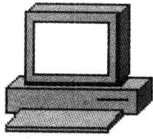
The conventions used to present command syntax in this book are the same conventions used in the IOS Command Reference. The Command Reference describes these conventions as follows:

- 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.

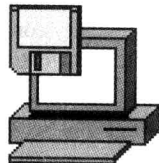
Icons Used in This Book



Throughout the book, you will see the following icons used for peripherals and other devices.



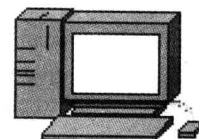
PC



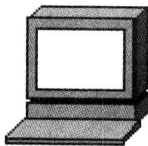
PC with
software



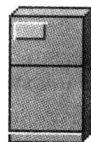
Sun
workstation



Macintosh



Terminal



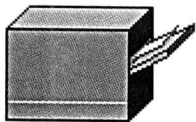
File
server



Web
server



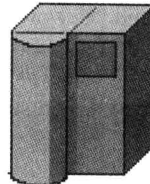
Cisco Works
workstation



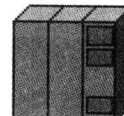
Printer



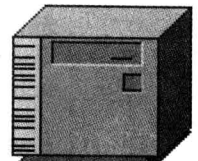
Laptop



IBM
mainframe



Front end
processor



Cluster
controller

Throughout the book, you will see the following icons used for networks and network connections.



Line: Ethernet



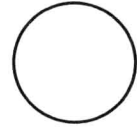
Line: serial



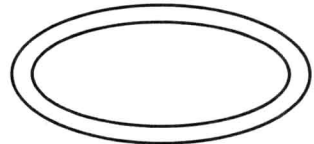
Line: switched serial



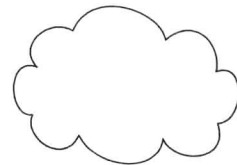
Frame Relay virtual circuit



Token ring



FDDI



Network cloud

Contents at a Glance

Part I	The Contemporary Internet
Chapter 1	Evolution of the Internet
Chapter 2	ISP Services and Characteristics
Chapter 3	IP Addressing and Allocation Techniques
Part II	Routing Protocol Basics
Chapter 4	Interdomain Routing Basics
Chapter 5	Border Gateway Protocol Version 4
Part III	Effective Internet Routing Designs
Chapter 6	Tuning BGP Capabilities
Chapter 7	Redundancy, Symmetry, and Load Balancing
Chapter 8	Controlling Routing Inside the Autonomous System
Chapter 9	Controlling Large-Scale Autonomous Systems
Chapter 10	Designing Stable Internets
Part IV	Internet Routing Device Configuration
Chapter 11	Configuring Basic BGP Functions and Attributes
Chapter 12	Configuring Effective Internet Routing Policies
Part V	Appendixes
A	BGP Command Reference
B	References for Further Study
C	BGP Outbound Route Filter (ORF)
D	Multiprotocol BGP (MBGP)

Contents

Part I The Contemporary Internet 3

Chapter 1 Evolution of the Internet 5

Origins and Recent History of the Internet	5
From ARPANET to NSFNET	7
The Internet Today	8
NSFNET Solicitations	10
Network Access Points	10
What Is a NAP?	11
NAP Manager Solicitation	11
Federal Internet eXchange	12
Commercial Internet eXchange	12
Current Physical Configurations at the NAP	13
An Alternative to NAPs: Direct Interconnections	14
Routing Arbiter Project	14
The Very High-Speed Backbone Network Service	18
Transitioning the Regional Networks from the NSFNET	21
NSF Solicits NIS Managers	22
Network Information Services	23
Creation of the InterNIC	23
Directory and Database Services	23
Registration Services	25
NIC Support Services	25
Other Internet Registries	25
ARIN	26
RIPE NCC	26
APNIC	27
Internet Routing Registries	27
The Once and Future Internet	28
Next-Generation Internet Initiative	28
Internet2	30
Abilene	31
Looking Ahead	32
Frequently Asked Questions	34

References 35

Chapter 2 **ISP Services and Characteristics 37**

ISP Services 37

Dedicated Internet Access 37

Frame Relay and ATM Internet Access 38

Dialup Services 39

Digital Subscriber Line 40

Cable Modems 41

Dedicated Hosting Services 41

Other ISP Services 42

ISP Service Pricing, Service-Level Agreements, and Technical Characteristics 42

ISP Service Pricing 42

Service-Level Agreements 43

ISP Backbone Selection Criteria 43

Demarcation Point 50

Looking Ahead 53

Frequently Asked Questions 54

Chapter 3 **IP Addressing and Allocation Techniques 57**

History of Internet Addressing 57

Basic IP Addressing 57

Basic IP Subnetting 60

VLSMs 62

IP Address Space Depletion 65

IP Address Allocation 66

Classless Interdomain Routing 67

Private Addressing and Network Address Translation 79

IP Version 6 82

Looking Ahead 86

Frequently Asked Questions 87

References 89

Part II **Routing Protocol Basics 91**

Chapter 4 **Interdomain Routing Basics 93**

Overview of Routers and Routing 93

Basic Routing Example 94

Routing Protocol Concepts 96

	Distance Vector Routing Protocols	96
	Link-State Routing Protocols	99
	Segregating the World into Autonomous Systems	101
	Static Routing, Default Routing, and Dynamic Routing	101
	Autonomous Systems	102
	Looking Ahead	107
	Frequently Asked Questions	108
	References	109
Chapter 5	Border Gateway Protocol Version 4	111
	How BGP Works	112
	BGP Message Header Format	115
	BGP Neighbor Negotiation	116
	Finite State Machine Perspective	118
	NOTIFICATION Message	120
	KEEPALIVE Message	122
	UPDATE Message and Routing Information	122
	BGP Capabilities Negotiation	127
	Multiprotocol Extensions for BGP	128
	TCP MD5 Signature Option	129
	Looking Ahead	131
	Frequently Asked Questions	132
	References	133
Part III	Effective Internet Routing Designs	135
Chapter 6	Chapter Tuning BGP Capabilities	137
	Building Peer Sessions	137
	Physical Versus Logical Connections	139
	Obtaining an IP Address	140
	Authenticating the BGP Session	140
	BGP Continuity Inside an AS	141
	Synchronization Within an AS	142
	Sources of Routing Updates	144
	Injecting Information Dynamically into BGP	144
	Injecting Information Statically into BGP	147
	ORIGIN of Routes	148
	An Example of Static Versus Dynamic Routing: Mobile Networks	150

Overlapping Protocols: Backdoors	150
The Routing Process Simplified	152
BGP Routes: Advertisement and Storage	153
The BGP Routing Information Bases	154
Routes Received from Peers	155
Input Policy Engine	155
Routes Used by the Router	155
Output Policy Engine	156
Routes Advertised to Peers	156
Sample Routing Environment	156
BGP Decision Process Summary	158
Controlling BGP Routes	159
BGP Path Attributes	160
NEXT_HOP Behavior on Multiaccess Media	172
NEXT_HOP Behavior Over Nonbroadcast Multiaccess Media	173
Use of next-hop-self versus Advertising DMZ	174
Using Private ASs	175
AS_PATH and Route Aggregation Issues	177
AS_PATH Manipulation	178
Route Filtering and Attribute Manipulation	180
Inbound and Outbound Filtering	181
The Route Filtering and Manipulation Process	182
Peer Groups	190
BGP-4 Aggregation	192
Aggregate Only, Suppressing the More-Specific Routes	192
Aggregate Plus More-Specific Routes	193
Aggregate with a Subset of the More-Specific Routes	195
Loss of Information Inside Aggregates	196
Changing the Attributes of the Aggregate	196
Forming the Aggregate Based on a Subset of the More-Specific Routes	196
Looking Ahead	197
Frequently Asked Questions	199
References	201
Chapter 7 Redundancy, Symmetry, and Load Balancing	203
Redundancy	203
Geographical Restrictions Pressure	204
Setting Default Routes	205
Symmetry	210

Load Balancing	210
Specific Scenarios: Designing Redundancy, Symmetry, and Load Balancing	212
Scenario 1: Single-Homing	213
Scenario 2: Multihoming to a Single Provider	213
Scenario 3: Multihoming to Different Providers	223
Scenario 4: Customers of the Same Provider with a Backup Link	228
Scenario 5: Customers of Different Providers with a Backup Link	231
Looking Ahead	236
Frequently Asked Questions	237
References	239

Chapter 8 Controlling Routing Inside the Autonomous System 241

Interaction of Non-BGP Routers with BGP Routers	241
Injecting BGP into the IGP	241
Following Defaults Inside an AS	242
BGP Policies Conflicting with Internal Defaults	244
Defaults Inside the AS: Primary/Backup BGP Policy	244
Defaults Inside the AS: Other BGP Policies	250
Policy Routing	252
Policy Routing Based on Traffic Source	252
Policy Routing Based on Traffic Source/Destination	253
Policy Routing Defaults to Dynamic Routing	254
Other Applications of Policy Routing	255
Looking Ahead	257
Frequently Asked Questions	258

Chapter 9 Controlling Large-Scale Autonomous Systems 261

Route Reflectors	261
Internal Peers Without Route Reflectors	262
Internal Peers with Route Reflectors	263
Naming Conventions and Rules of Operation	264
Redundancy Issues and Multiple Route Reflectors in an AS	265
Route Reflection Topology Models	266
Route Reflectors and Peer Groups	269
Confederations	271
Confederation Drawbacks	273
Route Exchange and BGP Decisions with Confederations	274
Recommended Confederation Design	274
Confederations Versus Route Reflectors	275