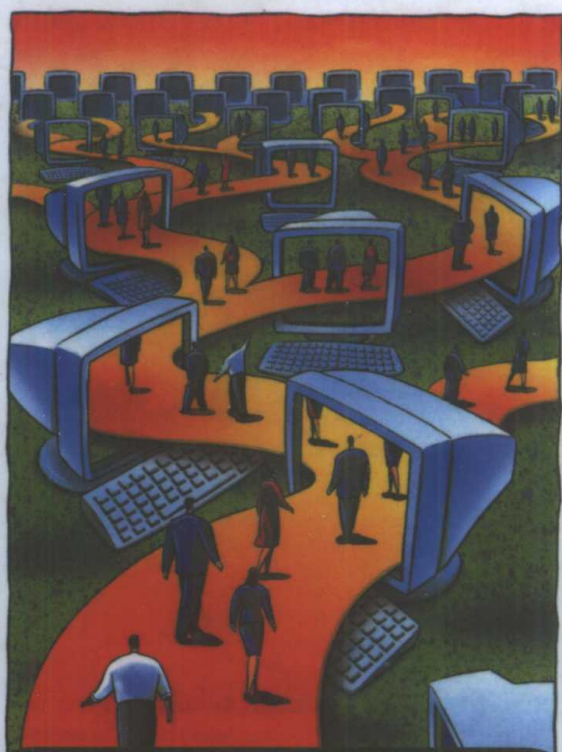


因特网的路由选择技术

(第2版)



网络工程丛书

Internet Routing Architectures
Second Edition

[美] Sam Halabi Danny McPherson 著
彭业飞 译

CISCO SYSTEMS
CISCO PRESS



电子工业出版社
Publishing House of Electronics Industry
URL: <http://www.phei.com.cn>

网络工程丛书

因特网的路由选择技术

(第2版)

Internet Routing Architectures
Second Edition

[美] Sam Halabi Danny McPherson 著

彭业飞 译

电子工业出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

本书对当代因特网的结构、服务提供者和编址技术进行了介绍。讲述了链路状态和距离向量选路协议的特性以及为什么需要域间选路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选路策略。

本书通过在概念上和在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。尽管本书讲解了不同等级的专门技术,但是也逻辑性地涉及了从最简单到最复杂的概念和问题。本书适合于网络技术专业人员和网络爱好者阅读。

Authorized translation from the English language edition published by Cisco Press. Copyright © 2000. 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 the publisher.

Simplified Chinese language edition published by Publishing House of Electronics Industry. Copyright © 2001.

本书中文简体专有翻译出版版权由Pearson教育集团所属的Cisco Press授予电子工业出版社。其原文版权及中文翻译出版版权受法律保护。未经许可,不得以任何形式或手段复制或抄袭本书内容。

图书在版编目(CIP)数据

因特网的路由选择技术(第2版)/(美)哈莱拜(Halabi, S),迈克皮尔森(McPherson, D)著;彭业飞译.-北京:电子工业出版社,2001.6
(网络工程丛书)

书名原文: Internet Routing Architectures

ISBN 7-5053-6763-3

I. 因... II. ①迈...②彭... III. 因特网—路由选择 IV. TP393

中国版本图书馆CIP数据核字(2001)第035380号

丛 书 名: 网络工程丛书

书 名: 因特网的路由选择技术(第2版)

原 书 名: Internet Routing Architectures, Second Edition

著 者: [美] Sam Halabi Danny McPherson

译 者: 彭业飞

责任编辑: 赵红燕

排版制作: 今日电子公司制作部

印 刷 者: 北京东光印刷厂

出版发行: 电子工业出版社 URL: <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: 850 × 1168 1/32 印张: 16 字数: 460千字

版 次: 2001年6月第1版 2001年6月第1次印刷

书 号: ISBN 7-5053-6763-3
TP · 3793

定 价: 28.00元

版权贸易合同登记号 图字: 01-2000-4278

凡购买电子工业出版社的图书,如有缺页、倒页、脱页者,请向购买书店调换;若书店售缺,请与本社发行部联系调换。

译 者 序

当前的时代是一个信息时代，毫无疑问，信息高速公路将成为信息社会的主要传播媒介。因特网作为一项技术奇迹，正在急剧发展。近年来因特网的网络规模和业务量一直遵循每6~9个月翻一番的新摩尔定律在爆炸式增长。因特网的迅猛发展正在引发着整个电信领域的一场极其深刻的变革，一个以因特网为中心的电信市场新格局正在形成和发展，新的因特网时代已经来临。传统电信运营商必须去迎接一场巨大的挑战。

因特网的联网概念是利用装有联网协议的路由器进行对等对话来选择路由，使得信息到达最后的目的地。它的核心技术是编址和路由选择。本书对于联网的核心技术做了详尽的探讨，是一本有关因特网联网设计的经典著作。本书的最大特点是附图多、实例多，读者可以直观地学习到因特网路由选择的基本概念和实用知识。本书对于因特网的各级提供者是特别实用的；对于电信网的工程技术人员来说，学习因特网的联网技术也是必然的趋势。

本书由彭业飞翻译，由于译者水平所限，译文中难免有不妥之处，望读者给予指正。

关于作者

Sam Halabi 是因特网服务提供商行业的最重要的专家之一，最近担任一个 IP 联网开办机构的市场副总裁，他曾在 Cisco 系统公司工作多年，领导 IP 载体市场营销计划工作。Halabi 先生是复杂的选路协议方面的专家，在设计大型 IP 网络方面富有专长。

作为本行业的活跃成员，Halabi 是光互联网论坛的理事和 MPLS 论坛的会员。

Danny McPherson 目前是 Amber Networks 的 CTO 办公室的设计主任。他曾在四个因特网服务提供商处担任技术领导职务（Qwest、GTE Internetworking、Genuity 和 internetMCI），负责网络和产品设计、选路设计、对等化和其他与业务和策略有关的任务。

McPherson 对因特网工程任务组（IETF）以及其他几个标准机构都有积极的贡献。他在因特网体系结构和选路协议方面是一位公认的专家。

前 言

因特网，这种 20 世纪 60 年代末崛起的学术实验，今天已争得了地位并取得了成功。从 ARPANET（美国高级研究项目管理网）、NSFNET（美国国家科学基金网）到任何人的网络，因特网已不再属于某个实体；它属于任何买得起它的存储空间的人。数千万的使用者在寻求网络连接，数万个公司如果不进入因特网会感到孤立。这已使网络设计者和管理者们处于很大的压力之下，那就是要跟上组网和连通性的需求。学习如何联网，尤其是如何选择路由，已经成为必不可少的事情。

当网络出错并瘫痪时有些人会惊讶，而有些人则是当它不出错时才会惊讶。这原因似乎是因为有用的知识太少。直到现在，那些网络设计者和管理者们能得到的关于选路的知识都用处不大：那些知识使你认为已经知道怎样建立自己的网络，但一尝试就会发现并非如此。本书第 1 版是一本有关实用的路由选择问题的书，它以一种综合易懂的方式讲解实际的选路问题。

现在的第 2 版不仅更新了原始材料，还引入了 BGP 协议最新的增强功能，讨论了因特网号码登记和分配的有关改变，并提供了科研和教育网络的信息。

本书目的

本书的目的是使你成为把自己的网络结合到全球因特网中去的专家。通过在概念上和在实例上提供实用的编址、选路和连通的技术，本书旨在培养你对路由选择的理解，以便你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。无论你是因特网的用户还是提供者（或者都是），本书都能预见并讲解那些面对

你的网络的路由选择的难题。

读者对象

本书是为任何一个可能想要进入因特网的组织和个人准备的。无论你要成为服务提供者还是正要接到某个服务提供者，都能找到联网所需的知识。本书通篇考虑了网络管理者、综合者和设计者的观点。尽管本书讲解了不同等级的专门技术，但是也逻辑性地涉及了从最简单到最复杂的概念和问题，而且本书的基本内容是简单易懂的，实际例子是任何人都可参考的。不要求读者有较多的选路或TCP/IP背景知识。任何理解选路技术所需的基础或背景知识会在需要时在正文的讨论中提到，而不是假定为读者必备技能的一部分。

本书的组织

全书由四部分组成：

- 第一部分：当代因特网。第1章到第3章对当代因特网做了关于它的结构、服务提供者和编址技术的基本介绍。即使读者已很熟悉因特网的一般结构，还是应该读一读第1章中关于网络接入点、路由仲裁方案和网络信息服务的部分。加在因特网这些方面的压力对于管理者所面对的选路设计问题有着实际意义。第2章提供了评价因特网服务提供者的有价值的准则。如果你是一个提供者，或已是一个用户，可能你已熟悉这些信息。第3章讨论了无类别域间选路（CIDR）、VLSM（可变长度子网掩码）、IPv6以及因特网编址的其他方面。
- 第二部分：选路协议基础。第4章和第5章涉及了这些基础：链路状态和距离向量选路协议的特性以及为什么需要域间选路协议和它们是如何工作的。这些主题的论述既有一般讲解，也有对现代的域间选路协议BGP（边界网

关协议)的事实标准的特殊表述。对 BGP 的特殊能力和属性有全面的介绍。

- 第三部分：有效的因特网选路设计。第 6 章到第 10 章深入探讨了 BGP 的实用的、面向设计的应用。第二部分中介绍的属性以工作状态展示在多种有代表性的网络方案中。BGP 的属性在实现设计目标时要用到，例如冗余度、对称性和负荷平衡。对域内和域间选路协调工作的难题、正在增长或已很大的系统的管理和稳定性的保持等都有讨论。
- 第四部分：因特网选路手段的配置。第 11 章和第 12 章包括很多 BGP 属性的编码实例和许多选路策略。在读完前面数章之后，就会彻底明白这些编码实例，因为它们涉及了许多的概念和设计目标。所以读者还可以把前面数章的文字讨论和第 11 章、第 12 章的编码实例并列阅读，在前面的章节中加上了称为“配置例子”的指针。当看到它时，就可以快速地翻到提到的那页去阅读正在讨论的属性或策略的配置例子。
- 第五部分：有几个附录提供了进一步阅读的参考资料、最新的 Cisco IOS BGP 命令的列表以及为了提供更直观的 BGP 命令行接口的有关 IOS 的修正信息。

方法

以易懂的风格描写技术知识是非常困难的。去掉太多技术细节的信息会失掉它的本意，而过于完整精细的技术细节会使读者厌烦和难以弄懂概念。本书循序渐进地介绍技术细节，只要可能，总是结合实际情况。最复杂的技术信息，亦即以 Cisco IOS 语言书写的配置例子，只出现在本书的最后两章里，以便使它们的基础预先奠定于前面的概念和样本拓扑中。

尽管你的最终目标是设计和完成选路策略，但在把它们应用到特定网络之前掌握有关概念和原理仍是重要的。本书为了平衡

概念与实际的方案，遵循了逻辑的、渐进的、从一般到特殊、从概念到实施的步骤。甚至在必须采取大量描述方法的章节，亲自做一做的兴趣也可通过用指针找到配置例子、经常提出的问题和基于实例的解释得到满足。

基于实例的方法是本书非常重要的部分：它用有代表性的网络拓扑作为基础来说明几乎所有讨论过的协议属性和选路策略。尽管读者也许还没有看到网络的具体情况说明，但那些实例已足够具体、而且足够普遍，可使自己能推断怎样把那些所说明的概念应用到具体情况中。

特点和文本惯例

本书致力于不去避免叙述协议的细节和面向设计的知识，而同时认识到那些一般基础和概念的理解需要先行。两个用来帮助强调实用性和面向设计的特点作为基础概念被引入：

- 指向配置例子的指针——放在相关的讨论内容附近，这些参考点指向可找到相关配置例子的第 11 章和第 12 章。
- 经常提出的问题——放在每章的最后，这些问题预先估计了你在读完本章后可能会提出的关于特定网络的实际和面向设计的问题。

命令语法惯例

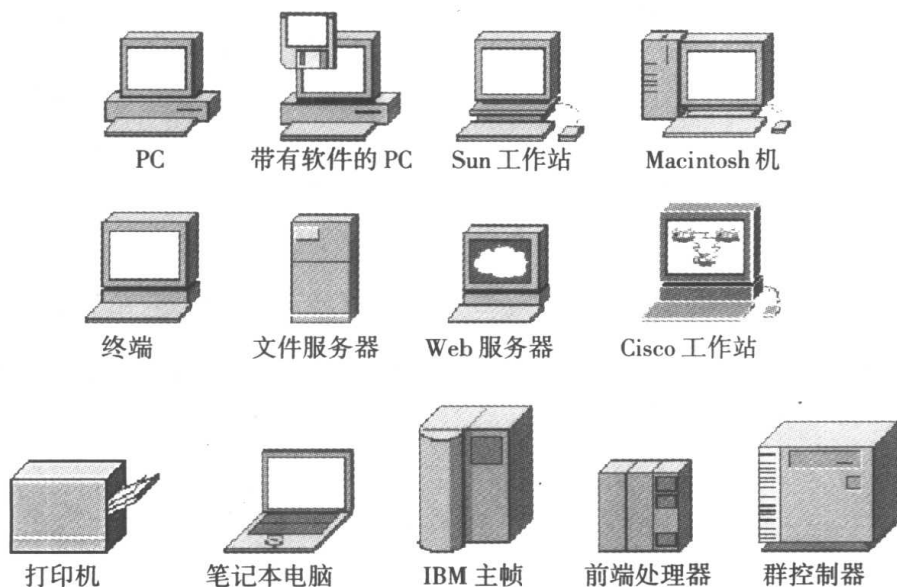
本书用来表达命令语法的惯例与 IOS 命令参考中使用的惯例一致。命令参考说明了如下惯例：

- 竖线 (|) 分隔了不同的、互相排斥的元素。
- 方括号 ([]) 指出了可选的元素。
- 大括号 ({}) 指出了要求的选项。
- 方括号中的大括号 ([{}]) 指出了 n 个可选元素中要求的选项。

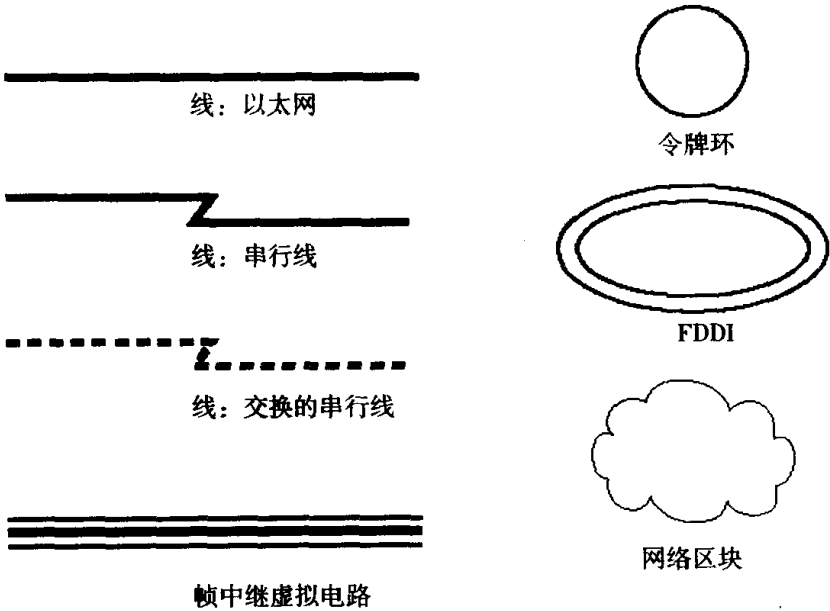
本书使用的图标



在本书中，还可以看到如下图标，用来表示外围设备以及其他设备。



在本书中，还将看到如下图标，用来表示网络和网络连接。



目 录

第一部分 当代因特网

第 1 章 因特网的演进	2
1.1 因特网的起源和近年的历史	3
1.1.1 从 ARPANET 到 NSFNET	4
1.1.2 现代因特网	6
1.1.3 NSFNET 的征求	7
1.2 网络接入点	8
1.2.1 什么是 NAP	8
1.2.2 NAP 管理者的征求	9
1.2.3 联邦因特网交换点	10
1.2.4 商业因特网交换点协会	10
1.2.5 现代 NAP 的物理配置	11
1.2.6 NAP 的备选方案：直接互连	12
1.3 路由仲裁者方案	13
1.4 超高速骨干网服务（vBNS）.....	17
1.5 NSFNET 到区域网络的转移	19
1.6 NSF 征求 NIS 管理者	21
1.6.1 网络信息服务	22
1.6.2 InterNIC 的建立	22
1.6.3 索引和数据库服务	22
1.6.4 注册服务	24
1.6.5 支持 NIC 的服务	25
1.7 其他因特网注册处	25
1.7.1 因特网号码美国注册处	25

1.7.2	IP 网络欧洲协调中心	26
1.7.3	亚太网络信息中心	26
1.8	因特网路由注册处 (IRR)	27
1.9	因特网的过去和未来	28
1.9.1	下一代因特网的起步	28
1.9.2	因特网 2 (第 2 代因特网)	30
1.9.3	Abilene	31
1.10	展望	32
1.11	经常提出的问题	33
1.12	参考资料	34
第 2 章	ISP 服务与特性	35
2.1	ISP 服务	35
2.1.1	专用因特网接入	36
2.1.2	帧中继和 ATM 因特网接入	36
2.1.3	拨号连接服务	38
2.1.4	数字用户线路	38
2.1.5	电缆调制解调器	39
2.1.6	专门代管服务	39
2.1.7	其他 ISP 服务	40
2.2	ISP 的价格、服务等级协定和技术特性	41
2.2.1	ISP 服务价格	41
2.2.2	服务等级协定	42
2.2.3	ISP 骨干网选择准则	42
2.2.4	分界点	48
2.3	展望	51
2.4	经常提出的问题	52
第 3 章	IP 编址和分配技术	53
3.1	因特网编址的历史	54
3.1.1	基本 IP 编址	54
3.1.2	基本子网划分	56

3.1.3	可变长度子网掩码	58
3.2	IP 地址空间的缺乏	61
3.2.1	IP 地址分配	62
3.2.2	无类别域间选路	64
3.2.3	专用地址和网络地址的转换	76
3.2.4	IP 版本 6	79
3.3	展望	83
3.4	经常提出的问题	83
3.5	参考资料	85

第二部分 选路协议基础

第 4 章	域间选路基础	88
4.1	路由器与选路概述	88
4.1.1	基本选路例子	89
4.2	选路协议的概念	91
4.2.1	距离向量选路协议	91
4.2.2	链路状态选路协议	94
4.3	把世界划分成自治系统	97
4.3.1	静态路由、默认路由和动态路由	97
4.3.2	自治系统	98
4.4	展望	103
4.5	经常提出的问题	103
4.6	参考资料	105
第 5 章	边界网关协议版本 4	106
5.1	BGP 是怎样工作的	107
5.1.1	BGP 消息报头格式	110
5.1.2	BGP 相邻体协商	111
5.1.3	有限状态机的前景	113
5.1.4	NOTIFICATION 消息	116
5.1.5	KEEPALIVE 消息	118

5.1.6	UPDATE 消息和选路信息	118
5.2	BGP 能力协商	124
5.3	BGP 的多协议扩展	125
5.4	TCP MD5 签名选项	126
5.5	展望	128
5.6	经常提出的问题	128
5.7	参考资料	129

第三部分 有效的因特网选路设计

第 6 章	BGP 能力的协调	132
6.1	建立对等对话	133
6.1.1	物理的和逻辑的连接	134
6.1.2	取得一个 IP 地址	135
6.1.3	鉴别 BGP 对话	136
6.1.4	在一个 AS 内的 BGP 连续性	136
6.1.5	AS 内部的同步	137
6.2	路由更新的源	140
6.2.1	向 BGP 动态注入信息	140
6.2.2	向 BGP 静态注入信息	143
6.2.3	路由的起点	144
6.2.4	静态对动态选路的一个例子：移动网络	145
6.3	重叠协议：后门	146
6.4	简化的选路过程	148
6.4.1	BGP 路由：通告和存储	149
6.4.2	BGP 选路信息库	150
6.4.3	从对等体收到的路由	151
6.4.4	输入策略机	151
6.4.5	路由器使用的路由	152
6.4.6	输出策略机	152
6.4.7	通告给对等体的路由	152

6.4.8	选路环境的例子	152
6.4.9	BGP 决策过程总结	154
6.5	控制 BGP 路由	155
6.5.1	BGP 路径属性	156
6.5.2	多路访问媒体的 NEXT_HOP 性能	169
6.5.3	在非广播多路访问媒体上的 NEXT_HOP 性能	170
6.5.4	使用 next-hop-self 与通告 DMZ	171
6.5.5	使用专用 AS	172
6.5.6	AS_PATH 和路由聚合问题	174
6.5.7	AS_PATH 的操作	175
6.6	路由过滤和属性操作	177
6.6.1	入站和出站过滤	178
6.6.2	路由过滤和操作过程	178
6.6.3	对等体组	186
6.7	BGP-4 聚合	187
6.7.1	仅聚合, 抑制更具体的路由	188
6.7.2	聚合加上更具体路由	188
6.7.3	带有更具体路由子集的聚合	190
6.7.4	聚合中信息的丢失	191
6.7.5	改变聚合的属性	191
6.7.6	基于更具体路由子集的聚合构成	191
6.8	展望	192
6.9	经常提出的问题	193
6.10	参考资料	195
第 7 章	冗余度、对称性和负载平衡	196
7.1	冗余度	197
7.1.1	地理限制的压力	198
7.1.2	设定默认路由	198
7.2	对称性	203
7.3	负载平衡	204
7.4	冗余度、对称性和负载平衡的设计	206

7.4.1	情况 1: 单宿连接	207
7.4.2	情况 2: 到单一提供者的多宿连接	207
7.4.3	情况 3: 到不同提供者的多宿连接	217
7.4.4	情况 4: 具有一个备用链路的相同提供者的用户	221
7.4.5	情况 5: 具有一个备用链路的的不同提供者的用户	224
7.5	展望	229
7.6	经常提出的问题	229
7.7	参考资料	230
第 8 章	自治系统内部选路的控制	231
8.1	非 BGP 路由器与 BGP 路由器的相互作用	232
8.1.1	把 BGP 注入 IGP	232
8.1.2	在 AS 内部遵循默认路由	233
8.2	BGP 策略与内部默认路由的冲突	234
8.2.1	AS 内的默认路由: 主用 / 备用 BGP 策略	234
8.2.2	AS 内的默认路由: 其他 BGP 策略	240
8.3	策略选路	242
8.3.1	基于业务量源的策略选路	242
8.3.2	基于业务量源 / 目的的策略选路	243
8.3.3	默认为动态选路的策略选路	244
8.3.4	策略选路的其他应用	244
8.4	展望	246
8.5	经常提出的问题	247
第 9 章	大规模自治系统的控制	248
9.1	路由反射器	249
9.1.1	无路由反射器的内部对等体	249
9.1.2	有路由反射器的内部对等体	250
9.1.3	命名惯例和操作规则	251
9.1.4	冗余度问题及一个 AS 内的多个路由反射器	252
9.1.5	路由反射拓扑模型	253
9.1.6	路由反射器和对等体组	257