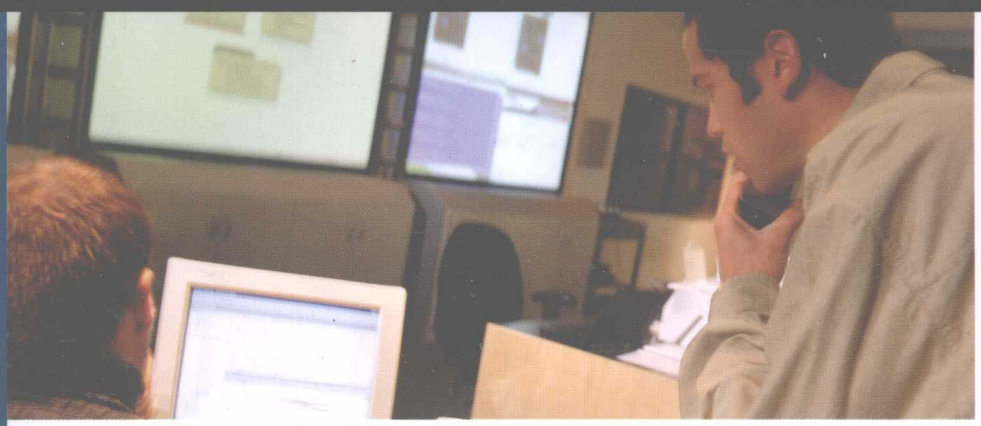
CISCO

ciscopress.com



BGP 设计与实现

BGP Design and Implementation

Practical guidelines for designing and deploying a scalable BGP routing architecture

〔美〕 Randy Zhang, CCIE #5659
Micah Bartell, CCIE #5069 著

黄博 葛建立 译
黄博 审校

人民邮电出版社
POSTS & TELECOM PRESS

BGP 设计与实现

BGP Design and Implementation

〔美〕 **Randy Zhang**, CCIE #5659
Micah Bartell, CCIE #5069 著

黄博 葛建立 译
黄博 审校

人 民 邮 电 出 版 社
北 京

图书在版编目（CIP）数据

BGP 设计与实现/（美）张（Zhang, R.），（美）巴特尔（Bartell, M.）著；黄博，葛建立译. —2 版. —北京：人民邮电出版社，2008.9
ISBN 978-7-115-18441-2

Ⅰ. B… Ⅱ. ①张…②巴…③黄…④葛… Ⅲ. 互联网—路由器 Ⅳ. TN915.05

中国版本图书馆 CIP 数据核字（2008）第 096740 号

版 权 声 明

BGP 设计与实现

◆ 著 [美] Randy Zhang, CCIE#5659
Micah Bartell, CCIE#5069
译 黄 博 葛建立
审 校 黄 博
责任编辑 李 际

◆ 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 ciscobooks@ptpress.com.cn
网址 http://www.ptpress.com.cn
北京艺辉印刷有限公司印刷

◆ 开本：787×1092 1/16
印张：31
字数：774 千字 2008 年 9 月第 2 版
印数：3 501 - 7 000 册 2008 年 9 月北京第 1 次印刷
著作权合同登记号 图字：01-2004-0567 号
ISBN 978-7-115-18441-2/TP

定价：69.00 元

读者服务热线：(010)67132692 印装质量热线：(010)67129223
反盗版热线：(010)67171154

内容提要

本书详细介绍了 BGP 特性及应用。全书共分 5 个部分 12 章。第一部分为理解高级 BGP，其中第 1 章讲解了 BGP 的基本特性，并比较了 BGP 和 IGP 的特性。第 2 章回顾了 BGP 的路径属性，在此基础上讲解了 BGP 的路径选择算法；同时较为深入地介绍了 BGP 进程和内存使用、路由选择信息库以及 IOS 的交换特性。第 3 章主要阐述了 BGP 性能调整的内容，包括有关 TCP 的考虑、队列优化、BGP 更新报文生成、性能调整的相互依赖性、BGP 网络性能特性等方面的内容。第 4 章详细阐述了 BGP 若干策略控制技巧，包括正则表达式、加强 BGP 策略的过滤列表、路由映射、策略列表、过滤处理的顺序等。第二、三部分介绍了设计企业和服务提供商 BGP 网络，这两部分的第 5 章至第 9 章是本书的核心，详细分析了企业的和运营商的 BGP 网络设计，内容包括若干 BGP 架构及其相互比较、企业网络的 Internet 连接性、可扩展的 iBGP 设计和实施指南、路由反射和联盟迁移策略、服务提供商网络架构。第四部分介绍了实施 BGP 多协议扩展，这部分的第 10 章到第 12 章跳出了传统的 BGP 领域，扩展地讲述了多协议 BGP 在其他领域的新应用，包括 MPLS VPN、域间多播、IPv6、CLNS 等方面的知识。第五部分为附录，提供了与本书内容关系密切的资料。

本书层次分明、阐述清晰、分析透彻、理论与实践并重，既深入讲解了传统的 BGP 知识，又讨论了 BGP 的新特性及 IOS 的新发展，非常适合于 ISP 网络管理员、BGP 网络的设计及实施者以及希望深入研究 BGP 的读者。

关于作者

Randy Zhang，博士，他的 CCIE 编号是 5659。Randy Zhang 是 Cisco Systems 公司高级服务组（Advanced Services，AS）的网络顾问工程师，为 Cisco 公司战略性的服务提供商和企业客户提供技术支持。他帮助过许多这样的客户进行大规模的 BGP 和 MPLS 网络的设计、迁移和实施。在加入高级服务组之前，他是 Cisco Systems 公司的高级软件 QA 工程师，研究领域是 Cisco 6x00 系列的 IP DSL 交换机中的 IP 路由选择和 MPLS，他也参与了其他很多项目。Randy Zhang 在不同的领域已经撰写了超过 30 篇的著作。

Micah Bartell，他的 CCIE 编号是 5069。Micah Bartell 是 Cisco Systems 公司的网络顾问工程师，是高级服务组里的 ISP 专家（ISP Experts）组的成员之一，为 Cisco 公司战略性的服务提供商和企业客户提供技术支持。在大规模 IP 网络设计领域，特别是在 BGP、IS-IS 和 IP 多播方面，他是一位公认的专家。此外，Micah Bartell 通过国际标准化组织（International Standards Organization，ISO）和 Internet 工程任务组（Internet Engineering Task Force，IETF），也涉及一些标准化的工作。他最近是 ISO/IEC IS 10589 的编辑。

关于技术审稿人

Juan Alcaide，1999 年加盟 Cisco Systems 公司，然后以联合研究员的身份在杜克大学（Duke University）研究 BGP 的可扩展性。自那时以后，他就一直在 Cisco Systems 技术支援中心（Technical Assistance Center，TAC）的路由选择协议团队中工作。目前，他是一名网络顾问，为大型 ISP 提供支持。

Jonathan Looney，他的 CCIE 编号是 7797。Jonathan Looney 是 Navisite 公司的一位高级网络工程师，在那里，他设计和实现了公司客户的定制网络解决方案和公司拥有的 15 个数据中心。在企业网和服务提供商的网络环境方面，他拥有超过 5 年的实施和维护 BGP 的经验。在 Navisite 公司供职之前，他在一个 ISP——也是一所大型的大学里工作，在那里，他设计和维护了公司的网络。

Vaughn Suazo，他的 CCIE 编号是 5109。他是一位在技术领域工作了 12 年的技术行家，在服务器技术、局域网/广域网和网络安全方面经验丰富。他取得了路由选择和交换、安全的双 CCIE 证书。Vaughn Suazo 在 Cisco Systems 公司的职业生涯开始于 1999 年，并且一开始就为网络服务提供商客户提供技术支撑和工程支持服务。在加盟 Cisco Systems 公司之前，他为一些技术公司工作，并为 Tulsa 和 Oklahoma 城区的很多企业和商业公司客户提供网络设计咨询服务、网络部署前后的技术支持服务以及网络审计服务等。

致　谢

本书是我们需要感谢的许多人共同努力的结果。我们要对很多同事表示深深的谢意，他们在时间紧迫的情况下对本书提供了详细的技术评审——特别是 Rudy Davis、Tony Phelps、Soumitra Mukherji、Eric Louzau 和 Chuck Curtiss。我们也要感谢 Mike Sneed 和 Dave Browning，感谢他们的鼓励和支持。

我们非常感谢 Cisco Press 的一些友善的人们，是他们使本书的出版成为现实。John Kane 在项目的每一个阶段都耐心地指导我们。他的鼓励和指导使得该出版项目减少了一些挑战性。Dayna Isley 和 Amy Moss 是两位很有才华的编辑，他们帮助我们采用正确的方法对书稿进行编辑并加入一些评论，并在手稿修订过程中给我们提供了详细的注释和建议。我们也要感谢 Brett Bartow、Chris Cleveland 和 Tammi Ross，感谢他们在项目初期给予的支持和配合。我们还要对 3 位技术审稿者——Juan Alcaide、Jonathan Looney 和 Vaughn Suazo 表达谢意，他们提出了有益的评论和建议，给本书带来了很多改进。

Randy Zhang：我特别要感谢我的家人、朋友、同事和所有其他这些年来一直帮助和鼓励我的人。

Micah Bartell：我想感谢我的家人和朋友——特别是 Adam Sellhorn 和 Jeff McCombs，感谢他们对该图书出版项目的支持。我也要感谢 Tom Campbell 和全球 Internet 网络运行中心（Global Internet NOC）的其他朋友，他们使网络技术从一开始就变得十分有趣。最后，也是最重要的，我要感谢上帝给了我编写本书的智慧和机会。

献　辞

Randy Zhang：献给 Susan、Amy 和 Ally，感谢她们永远的关爱、支持和耐心。

Micah Bartell：献给我的父母，Merlin 和 Marlene，感谢他们这些年来的支持。

译者序

这是一本内容非常精彩的书！

熟悉网络的人几乎都知道Cisco Press 2000年推出了Sam Halabi与Danny McPherson合著的《Internet Routing Architectures, second edition》（本书英文影印版已由人民邮电出版社出版）一书，这本书一度被业界视为BGP的Bible，而且直到现在仍备受推崇。2001年Cisco Press又推出了Jeff Doyle与Jennifer DeHaven Carroll合著的《Routing TCP/IP, Volume II》（本书英文影印版已由人民邮电出版社出版），这本书的前半部分主要以案例的形式来讲解BGP的策略工具的应用，深入浅出。本书从工程设计和实践的角度出发来讲解BGP，毫不夸张地说，这本书就是新时代、新发展环境下的BGP的Bible。

本书不是BGP的入门指南，而是BGP的高级理解，BGP工具的高级使用技巧。书中大量的图例、例题、案例有助于读者深入地理解书中的内容。内容新是本书的另一大特色。例如，第2章讲解了IOS的交换技术，并对BGP的内存使用做出了估算，为以后的内容打下了基础。第3章BGP性能调整的主题对于优化BGP来说是绝对必要的，但对于绝大多数网络工作者来说又是相对陌生的，作者在这里做了全面的讲解。又如，在第9章“公共对等安全考虑”一节中关于带宽盗用的问题确实对很多人来说十分新奇。

美国Cisco Networkers 2004大会上关于BGP的技术交流文档中推荐了这本书（见http://www.cisco.com/warp/public/732/Tech/routing/docs/deployingbgp.pdf文档第60页）。这在一定程度上反映了业界对本书的认可程度。

本书的译者中，一名工作于四川省创意技术发展有限责任公司，这是一家优秀的网络集成商，主要承接四川电信的多媒体数据网及MPLS VPN网络的建设及维护，也参与了中国电信核心网络改造的区域性的工程施工；另一名工作于江苏网通。由于译者的工作背景，在各自的工作中都遇到过若干关于BGP的难题，设计方面和施工方面的都有。有幸翻译此

书，我们受益匪浅。我们同时也将书中的内容应用到实际工作中。相信读者一定会和我们一样感觉开卷有益！

我们希望尽可能地表达出作者的原意，因此在一些不清楚或模糊的地方加入了注解，而在某些地方我们却保留了书中的原貌，这些内容是需要读者自己去理解的。由于种种原因，部分翻译显得仓促，尽管我们在翻译上力求术语、文字、风格上的统一，但仍恐有出入之处，敬请读者谅解。本书前言及 1～5 章由葛建立翻译，6～12 章、附录 A、B 由黄博翻译，全书由黄博统稿。

译　者

2005 年 1 月

前　言

边界网关协议（Border Gateway Protocol，BGP）是今天的网络中最广泛部署的协议之一，也是 Internet 事实上的路由选择协议。BGP 是一种灵活的协议，这在于它具有很多网络设计者和工程师可用的选项。此外，对它的扩展和软件实现增强也使 BGP 成为一种有力而复杂的工具。

本书的目的超出了基本协议概念和配置，而着重于提供实用的设计和实现的解决方案。在设计和实现复杂的网络方面，BGP 被当做一种有用的工具。通过实际的手法，本书提供了 Cisco IOS 软件的实现细节，以及广泛的例子和案例研究。

读者对象

本书希望涵盖设计和实现 BGP 网络的高级课题。虽然书中也回顾了 BGP 的基本概念，但是本书的重点不在于 BGP 本身或基本的 BGP 配置，而是提供了实用的设计和实现方面的指导建议，以帮助网络工程师、网络管理员以及网络设计者们搭建一个可扩展的 BGP 路由选择架构。本书也可以供任何希望理解 Cisco IOS 中可用的 BGP 高级特性的人使用，此外，对于准备 Cisco 认证考试的考生也会有所帮助。

本书组织结构

本书的章节大致可以分成 5 个部分。

第一部分“理解高级 BGP”，讨论和回顾了 BGP 的一些基础组件和工具。

- 第 1 章“高级 BGP 介绍”，讲述了 BGP 的特性，并比较了 BGP 和 IGP。
- 第 2 章“理解 BGP 的构件块”，通过回顾与 BGP 有关的多种组件来为本书打下基础。

- 第 3 章“调整 BGP 性能”，详细讲述了怎样调整 BGP 性能，并着重讨论了 IOS 的最新发展。
- 第 4 章“有效的 BGP 策略控制”，描述了常用的 BGP 策略控制技巧，这些技巧使 BGP 变得如此灵活。

第二部分“设计企业 BGP 网络”，着重介绍在设计企业网络时怎样运用 BGP 的特性。

- 第 5 章“企业级 BGP 核心网络设计”，讨论使用 BGP 来设计企业核心网络时的多种选择。
- 第 6 章“企业网络的 Internet 连接性”，描述了一个企业网络与 Internet 服务提供商（ISP）相连，以获得 Internet 连接性的设计方法。

第三部分“设计服务提供商 BGP 网络”，着重讨论服务提供商的 BGP 网络设计。

- 第 7 章“可扩展的 iBGP 设计和实施指南”，详细讨论了可用来增强 iBGP 扩展性的两种方法：路由反射和联盟。
- 第 8 章“路由反射和联盟迁移策略”，提供了全连接的 BGP 网络和基于路由反射或基于联盟的网络之间相互迁移的策略，并讲述了几个迁移过程的操作步骤。
- 第 9 章“服务提供商网络架构”，讲述了可用于服务提供商的多种 BGP 设计方法。

第四部分“实施 BGP 多协议扩展”，关注于对 BGP 的多协议扩展。

- 第 10 章“多协议 BGP 和 MPLS VPN”，讨论了为 MPLS VPN 的 BGP 多协议扩展，以及设计和实施复杂 VPN 的多种解决方案。
- 第 11 章“多协议 BGP 和域间多播”，提供了 BGP 怎样被用于域间多播的设计方法。
- 第 12 章“多协议 BGP 对 IPv6 的支持”，讲述了对 IP 版本 6 的 BGP 扩展。

第五部分“附录”，提供了以下信息。

- 附录 A，多协议 BGP 扩展对 CLNS 的支持。
- 附录 B，BGP 特性和 Cisco IOS 软件版本列表。
- 附录 C，其他信息源。
- 附录 D，术语表。

本书使用的图标

Cisco 使用下列标准图标来表示不同的网络设备。在本书中，你可能会碰到一些这样的图标。

命令语法约定

本书中，用来表示命令的语法习惯与 IOS 命令参考（IOS Command Reference）中的命令语法的表示法相同。命令参考描述了以下表示方法：

- 竖线（|）用于分开可选的、互斥的选项；
- 方括号（[ ]）表示可选项；
- 大括号（{}）表示一个必选项；
- 方括号中的大括号（[{}]）表示在可选项中的必选项；
- **粗体字**表示按照文字所显示的内容，而必须被输入的命令和关键字；在实际的配置例子和输出（不是通常的命令语法）中，**粗体字**表示由用户手工输入的命令（例如，**show** 命令）；
- *斜体字*表示需要用实际数值替换的参数。

编址约定

为了简化描述，本书通常分配私有 IP 地址（RFC 1918），相应地，也使用了简单的子

网划分方法。任何这样的地址分配和子网划分机制仅仅用于演示，而不应该被理解为推荐的方法。

AS 号的分配机制代表性地以百计，例如 100，200，300，等等。在适合的时候，本书也使用私有自治系统号。除非特别指出，否则这些 AS 号仅仅用于演示，而不应该被理解为推荐的方法。

Cisco bug 经常被用做记录 IOS 新特性的一种工具。在某些适当和相关的地方，本书会提供 Cisco bug 的标识号。如果要访问这些 bug 信息，你需要注册访问 Cisco Systems 网站（www.cisco.com）的权限。

目　录

第一部分　理解高级 BGP

第二部分　设计企业 BGP 网络

第三部分 设计服务提供商 BGP 网络