



图灵计算机科学丛书



IPv6时代的《TCP/IP详解》!

IPv6详解

卷2：高级协议实现

IPv6 Advanced Protocols Implementation

[美] Qing Li [日] Tatuya Jinmei [日] Keiichi Shima 著
王嘉祯 彭德云 文家福 刘晓芹 译



人民邮电出版社
POSTS & TELECOM PRESS

TURING

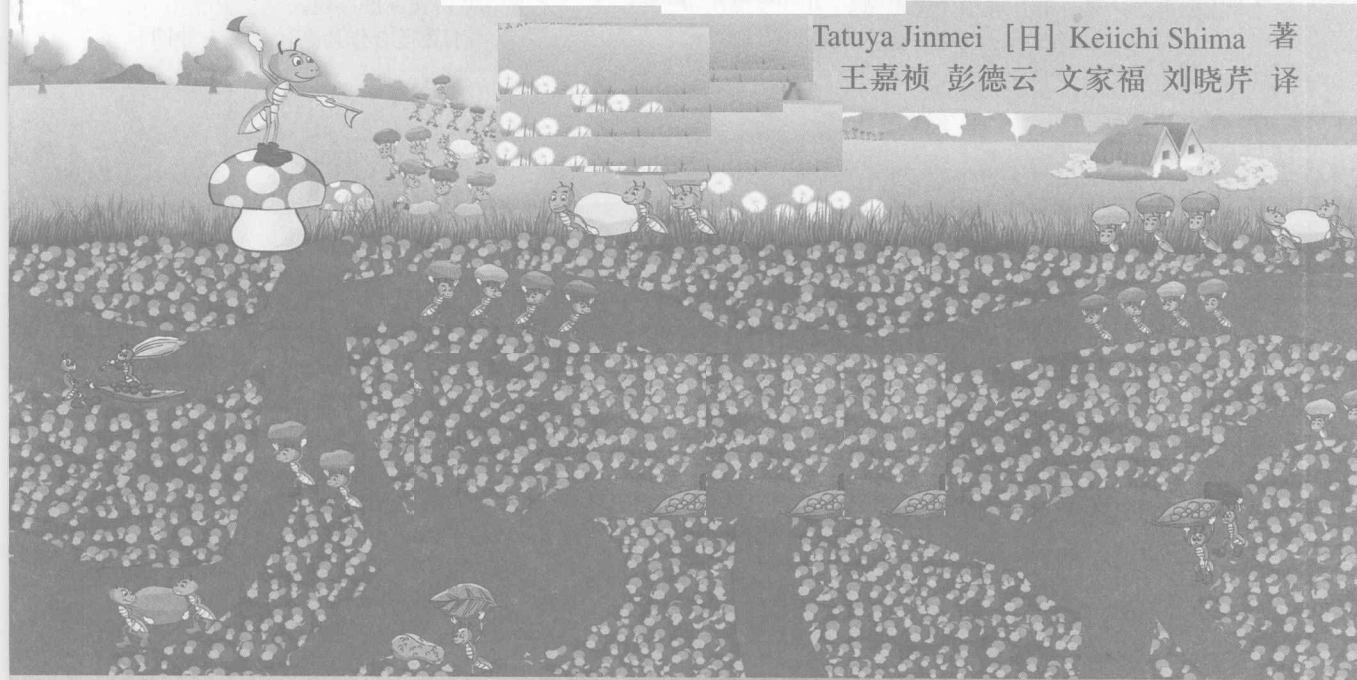
图灵计算机科学丛书

IPv6详解

卷2：高级协议实现

IPv6 Advanced Protocols Implementation

Tatuya Jinmei [日] Keiichi Shima 著
王嘉祯 彭德云 文家福 刘晓芹 译



人民邮电出版社
北京

图书在版编目(CIP)数据

IPv6详解, 第2卷, 高级协议实现/(美)李清, (日)神明达哉, (日)岛庆一著; 王嘉祯等译. —北京: 人民邮电出版社, 2009.8

(图灵计算机科学丛书)

书名原文: IPv6 Advanced Protocols Implementation, First Edition

ISBN 978-7-115-20891-0

I. I… II. ①李… ②神… ③岛… ④王… III. 计算机网络—传输控制协议 IV. TN915.04

中国版本图书馆CIP数据核字(2009)第080223号

内 容 提 要

本书全面讲解IPv6及相关协议实现的事实标准KAME, 从源代码的角度揭示了KAME IPv6协议栈的所有细节, 对每行代码到底做了什么, 以及为什么要这样设计都进行了解释。全书共分6章, 分别介绍IPv6单播路由选择协议、IPv6多播技术、IPv6的DNS DHCPv6、移动IPv6、IPv6与IP安全。每章都包含两个主要部分: 第一部分是相关规范的综述, 第二部分则逐行代码地描述和分析实际的实现。

本书是IPv6的权威参考书, 适合网络设计和开发人员阅读。此外, 本书还适合作为高校相关专业网络课程的教学参考书。

图灵计算机科学丛书

IPv6详解, 卷2: 高级协议实现

◆ 著 [美] Qing Li [日] Tatuya Jinmei [日] Keiichi Shima
译 王嘉祯 彭德云 文家福 刘晓芹
责任编辑 杨海玲

◆ 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>
北京顺义振华印刷厂印刷

◆ 开本: 787×1092 1/16
印张: 55.5

字数: 1529千字 2009年8月第1版
印数: 1-3000册 2009年8月北京第1次印刷

著作权合同登记号 图字: 01-2007-4742号

ISBN 978-7-115-20891-0/TP

定价: 128.00元

读者服务热线: (010) 51095186 印装质量热线: (010) 67129223

反盗版热线: (010) 67171154

版 权 声 明

IPv6 Core Protocols Implementation, First Edition by Qing Li, Tatuya Jinmei, and Keiichi Shima, ISBN: 978-0-12-370479-5.

Copyright © 2007 by Elsevier. All rights reserved.

Authorized Simplified Chinese translation edition published by the Proprietor.

ISBN: 978-981-259-975-9

Copyright © 2009 by Elsevier (Singapore) Pte Ltd. All rights reserved.

Elsevier (Singapore) Pte Ltd.

3 Killiney Road

#08-01 Winsland House I

Singapore 239519

Tel: (65)6349-0200

Fax: (65)6733-1817

First Published 2009

2009年初版

Printed in China by POSTS & TELECOM PRESS under special arrangement with Elsevier (Singapore) Pte Ltd. This edition is authorized for sale in China only, excluding Hong Kong SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书简体中文版由人民邮电出版社与Elsevier (Singapore) Pte Ltd.合作出版。本版仅限于中华人民共和国（不包括香港特别行政区和台湾地区）出版及标价销售。未经许可之出口，视为违反著作权法，将受法律之制裁。

献给Huaying、Jane和Adalia。

——Qing Li

献给我在KAME的同事们：与你们这些天才一起工作是激动人心的一种体验，也正因如此这一派生的项目才得以完成。

——Tatuya Jinmei

献给所有开发KAME的人、开发因特网的人以及将要开发未来因特网的人。

——Keiichi Shima

译 者 序

《IPv6详解》是作者所著的关于IPv6协议实现的2卷本系列书。《卷1》主要讲述IPv6的基本内容和大多数实现支持的主要协议，而本书讲述的是在更高级和更复杂的IPv6网络环境中才可能有协议，如路由选择、多播技术、DNS、DHCPv6、移动IPv6和安全等。

本书每一章以相关规范的综述开始，接着通过大量的代码描述和分析实际的实现，并辅以合理、精确的解释和应用。用作者的话说：“本书的一半篇幅写的是源代码。”书中包含大量图表，这些极为形象的图表对读者深刻理解协议的内容及实现方法无疑有巨大的帮助。如此详尽地阐述IPv6及相关协议的书籍在目前是绝无仅有的。

作者Qing Li等均是KAME项目的开发人员（其中Tatuya Jinmei自项目的开始到结束一直都是核心开发者），而KAME项目是总部设在日本的WIDE组织的重要课题，旨在为BSD系统提供IPv6、IPsec（包括IPv4和IPv6）以及移动IP的参考实现。三位作者都有着深厚的研究背景和丰富的开发经验，本书可以说是他们实际工作的提炼与总结。

本书的读者对象主要是目前正在从事或即将从事IPv6及相关协议实现的开发人员，对于研究和学习IPv6及相关协议的人员也有极大的参考价值。阅读本书需要读者具有IPv6相关协议的基础知识和C语言编程的经验。

本书由王嘉祯、彭德云、文家福、刘晓芹等翻译，由王嘉祯、文家福等审校并统稿。参加翻译工作的还有刘会英、刘爱珍和杨素敏等。中国人民解放军军械工程学院院长米东教授等为翻译工作提出了许多宝贵的意见和建议，对此表示深切的谢意。

由于本书所涉及的内容较新，许多术语在国内还没有统一的译法，加之译者水平有限，书中难免有错误和不妥的地方，敬请广大读者批评指正。

前言

我们的这一系列书通过KAME实现详细论述了IPv6和相关协议，本书为系列中的第2本。对于IPv6和在多种版本的BSD操作系统上开发的IP安全协议来说，KAME是广泛应用的实际参考实现。

系列书的第一本名为《IPv6详解，卷1：核心协议实现》（*IPv6 Core Protocol Implementation*）^①，以下简称为《卷1》，主要讨论了IPv6的基本内容和被大多数实现所支持的主要协议。这些协议在支持IPv6的大小设备上运转。《卷1》也描述了较高层协议（如TCP和UDP）上的IPv6内容，并介绍了与IPv6相关的应用编程接口。

书中所讨论的协议用于更强的IPv6设备，这些设备一般部署在较复杂的IPv6网络环境中，或者部署在不专门针对IPv6但是可扩展支持IPv6的网络环境中。具体地说，它涵盖了更高级的主题，如路由选择、多播技术、DNS、移动和安全等。

本书的结构和风格与《卷1》完全相同。每一章以相关规范的综述开始，接着就是逐行代码地描述和分析实际的实现。

我们希望这两本书能帮助读者牢固地掌握IPv6的基础理论和基本实践知识。这两本书全方位地涵盖了IPv6技术，这是前所未有的。

全书包含如下6章。

- 第1章讨论一般的路由概念和各类单播路由选择协议的基本内容。这一章详细讨论IPv6的简单路由选择协议RIPng，总结了为BGP4+和OSPFv3路由选择协议定义的IPv6扩展，比较了几种协议的复杂性、稳定性和操作问题以及各自的解决方案，还通过描述IPv6的路由选择API和展示KAME的RIPng实现的代码（**routed6d**守护进程），解释了在各种BSD版本上实现IPv6路由选择协议所必需的基本知识。最后这一章给出某些典型场景的**routed6d**配置示例。
- 第2章讨论IPv6的多播技术，特别是多播路由选择机制。这一章首先讲解主机到路由器协议和多播路由选择协议的基本知识，针对IPv6具体讨论了多播监听发现协议第1版（MLDv1）和协议无关多播（PIM）。这一章的后面部分描述MLDv1和IPv6多播转发的KAME内核实现。
- 第3章描述的是将IPv6扩展到DNS（域名系统）协议规范以及实现。这一章从DNS协议的一般描述及其支持IPv6的扩展开始，继而描述KAME的DNS客户（称作解析器）实现，并特别强调对IPv6的支持。这里也完整地介绍了getaddrinfo()库函数，该库函数在《卷1》中也有部分描述。这一章的后半部分说明如何操作支持IPv6的BIND9 DNS服务器，并指出这些服务器带有的普遍缺陷和IPv6相关操作特有的问题。
- 第4章详细讨论DHCPv6（IPv6的动态主机配置协议）的协议规范和KAME实现。虽然该协议的基本概念大多来自于IPv4的DHCP（DHCPv4），但是在设计上还是做了许多改进，并且期望的应用模型与DHCPv4也有所不同，这一章澄清了它们的主要区别。实现的描述包含所有协议的应用程序，即客户、服务器和中继代理，并且深入探讨了协议的工作原理。这一章还讲述了在一般应用场景下如何通过KAME实现来操作DHCPv6。

^① 其中文版及英文影印版已由人民邮电出版社出版。——编者注

- 第5章介绍IPv6主机移动协议（称作移动IPv6）。这一章先描述基本的移动IPv6，然后详细讨论协议规范和数据结构，中间部分讨论实际的实现。KAME移动IPv6实现支持归属代理和移动节点功能。代码描述部分详细讨论所有数据结构和函数。最后基于KAME移动IPv6实现，通过一个示例配置文件简单介绍了移动IPv6的运行。
- 第6章开始介绍IPsec协议和在IKE（因特网密钥交换）协议环境下键控概念，其余部分则集中描述流行的*racoon* IKE守护进程，全面解释它的配置和操作。这一章最后给出了一些使用*racoon*的示例。这一章不包含代码描述，原因是IP安全的基本机制和大多数实现并不是针对IPv6的，展示不针对IPv6的代码不是本书的目的。

目标读者

总的来讲，本书与《卷1》所面向的读者相同，都是实现IPv6和相关协议的开发人员，有志于从事这些协议项目开发的学生，特别是基于KAME/BSD实现的学生。不过，与《卷1》不同的是，本书讨论的是更高级的主题，如最近才被标准化的协议。所以，本书也可以作为这些协议的参考，DHCPv6和移动IPv6就是两个主要协议。

与《卷1》一样，本书假定读者非常熟悉C语言编程。另外，还假定读者已经掌握IPv6的基本概念和《卷1》中描述的相关协议，本书提到的其他参考资料也可以帮助没有阅读《卷1》的读者理解本书的内容。本书的第2章和第5章还要求读者对于BSD内核实现有大致的了解。

与《卷1》不同的是，本书的每一章都是完全独立的。尽管有些章的内容有交叉，但一般情况下读者仍然可以根据自己的兴趣从任何一章开始阅读。

排版约定

本书采用与《卷1》相同的排版约定。

在代码描述中提到的变量名、函数名、结构名、结构成员和编程语言关键字都用等宽字体表示。函数名用等宽字体表示，后跟一个圆括号，如`ip6_mforward()`；结构名用等宽字体表示，后跟一个花括号，如`ip6_mh{}`。

程序名显示为**粗体**，如**route6d**。程序的命令行输入和输出用等宽字体表示。

配套资源

本书英文原版书附带两张CD-ROM^①。第一张是FreeBSD4.8-RELEASE的ISO镜像，它是第1、2、3和6章所使用的基本操作系统，并且是可引导的CD-ROM，包括安装文件。安装过程从打开计算机，载入CD-ROM开始。详细的安装过程在INSTALL.TXT文件中可以找到，该文件在CD-ROM的根目录下。

类似地，第二张CD-ROM也是可引导的FreeBSD4.9-RELEASE的ISO镜像，是第5章使用的基本操作系统。

注意 FreeBSD 4.8和4.9RELEASE有几个安全漏洞，并且目前已经不再由FreeBSD项目支持。因此，这些系统仅仅作为阅读本书时学习KAME实现的参考，不建议在与因特网连接的产品环境中使用这些FreeBSD版本。

^① 中译本未随书附带CD-ROM，文中提到的CD-ROM中的FreeBSD4.8-RELEASE和FreeBSD4.9-RELEASE可以从<http://www.freebsd.org/releases/4.8R/installation.html>和<http://www.freebsd.org/releases/4.9R/installation.html>下载，KAME源代码可以从<http://www.kame.net>下载，以上内容也可从图灵网站（www.turingbook.com）直接下载。

第一张CD-ROM还包含本书所讨论的KAME源代码。可以通过根目录中的appendix目录来访问,该目录有两个子目录:kame-snap和rtadd6。

kame-snap子目录包含以下归档文件。

- kame-20030421-freebsd48-snap.tgz: 在2003年4月21日获取的FreeBSD 4.8的KAME快照。
- kame-20040712-freebsd49-snap.tgz: 在2004年7月12日获取的FreeBSD 4.9的KAME快照。这与第5章的内容有关,并且与第二个CD-ROM中的FreeBSD 4.9系统一起使用。
- kame-dhcp6-20050509.tgz: 包含在2005年5月9日获取的KAME快照中的KAME的DHCPv6实现,与第4章的内容有关。

安装KAME快照时,解压归档文件,在名为kame的顶级目录(在本书中也称作 $\{KAME\}$)中找到INSTALL文件。对于手中有《卷1》的读者来说,书的第1章提供了更详细的使用方法。本书的第4章解释如何安装DHCPv6实现。

另一个子目录rtadd6包含第1章中的rtadd6程序的源代码,这段代码是为本书新写的。

源代码版权

本书有许多源代码,这些代码是由KAME项目及志愿人员开发的。其中还涉及作为FreeBSD发布一部分的系统头文件。所有这些源代码都有版权提示。

错误报告及勘误表

作者希望接收到有关本书内容的错误报告,并计划在网提供勘误表。

致谢

首先,本书作者主要感谢所有的KAME开发者。同《卷1》一样,本书一半的篇幅都是用源代码填充的,这就意味着他们是本书的幕后作者。

我们也深深地感谢技术审稿人,他们有选择地阅读了本书的有关章节,提出了许多有价值的意见和建议,并纠正了一些错误。他们是Mark Andrews、David Borman、Francis Dupont、Daniel Hartmeier、Jeffrey Hsu、Akira Kato、T. J. Kniveton、Ted Lemon、Tsuyoshi Momose、George Neville-Neil、Yasuhiro Ohara、Shawn Routhier、Shoichi Sakane、Shigeya Suzuki、Shinsuke Suzuki、Christian Vogt和Carl Williams。审阅本书需要有关协议规范的详尽知识和高水平的编程技巧。我们知道,只有为数不多的人才具有如此才干,我们很荣幸拥有世界上最好的审稿团队。

书的封面是基于知名的KAME海龟图设计的,它被指定为这一项目的吉祥物,设计者是Manabu Higashida和Chizunu Higashida。

其次,我们还要感谢本书的编辑Rich Adams、Rachel Roumeliotis、Dawnmarie Simpson以及Morgen Kaufmann/Elsevier编辑部的全体成员,感谢他们在本项目3年半的时间中的耐心和鼓励。

最后,我们感谢Gary R. Wright和W. Richard Stevens。他们的工作激励我们开始自己的项目并让我们对于这一工作的价值充满信心。

Qing Li的个人致谢

我很想感谢Rich Adams,是他强烈地意识到本书的重要性,即本书填补了一项市场空白。

他对本书建议的认可是一种无价的动力。我感谢我的妻子Huaying Cheng在编写本书的过程中给予我的理解与支持。还要感谢VMware公司，感谢它捐赠了一个VMwareWorkstation 4软件的许可证，感谢MKS Software捐赠MKS工具包企业开发版8.6软件。

Tatuya Jinmei的个人致谢

我想感谢东芝公司的现任和前任经理们，感谢他们对这项工作的认可和支持。他们是Yukio Kamatani、Toshio Murai、Yasuhiro Katsube和Atsushi Inoue。我还要感谢我的WIDE项目的“管理员”Jun Murai和Hiroshi Esaki。

Keiichi Shima的个人致谢

我感谢为出版本书付出艰苦努力和支持这一工作的所有人，特别是Internet Initiative Japan公司的经理Eiiti Wada，同时感谢所有参与因特网运营、建设和研究的人。

目 录

第1章 IPv6单播路由选择协议	1
1.1 引言	1
1.2 路由选择概念概述	2
1.3 基于向量的算法和链路状态算法概述	4
1.3.1 距离向量算法	4
1.3.2 路径向量算法	5
1.3.3 链路状态算法	6
1.4 RIPng简介	7
1.4.1 RIPng报文格式	8
1.4.2 RIPng操作	11
1.4.3 RIPng存在的问题	12
1.5 BGP4+简介	13
1.5.1 BGP4+操作	14
1.5.2 BGP4+报文	17
1.5.3 路径属性	20
1.5.4 BGP4+的IPv6扩展	22
1.5.5 BGP4+路由选择过程	24
1.6 OSPFv3简介	25
1.6.1 路由器邻接和LSDB同步	25
1.6.2 区域类型和路由器分类	26
1.6.3 链路状态广告和LSA类型	27
1.6.4 LSA格式	28
1.6.5 OSPF树的构造和路由计算	34
1.7 代码介绍	37
1.8 BSD内核中的IPv6路由表	38
1.9 路由选择API	42
1.9.1 路由选择套接字	42
1.9.2 通过sysctl()转储路由表	48
1.10 route6d守护进程概述	51
1.11 公用数据结构、例程和全局变量	51
1.11.1 RIPng报文结构	52
1.11.2 route6d的路由表	52
1.11.3 本地接口的结构	54
1.11.4 route6d路由过滤器条目	56
1.11.5 子例程和全局变量	56
1.12 接口配置	57
1.12.1 ifconfig()函数	58
1.12.2 ifconfig1()函数	60
1.13 RIPng协议操作	63
1.13.1 sendrequest()函数	64
1.13.2 riprecv()函数	65
1.13.3 riprequest()函数	77
1.13.4 ripsend()函数	77
1.13.5 ripalarm()函数	84
1.14 使用route6d的路由选择操作	85
1.14.1 叶子网络	85
1.14.2 简单环状网	87
1.14.3 层次网	90
第2章 IPv6多播技术	92
2.1 引言	92
2.2 IPv6多播地址到第2层多播地址的 映射	92
2.3 多播监听发现协议	93
2.3.1 MLD协议报文格式	93
2.3.2 路由器警告选项	94
2.3.3 源地址选择	94
2.3.4 目的地址选择	94
2.3.5 MLD查询器	95
2.3.6 操作变量	95
2.3.7 MLD的加入进程	95
2.3.8 MLD离开进程	96
2.4 多播路由选择的基本原理	97
2.4.1 反向路径转发	97
2.4.2 多播路由选择模型	97
2.4.3 协议无关多播	101
2.4.4 PIM的IPv6特定问题	104
2.4.5 IPv6多播的未来: MLDv2和 SSM	105
2.5 代码介绍	106
2.6 MLD实现	107
2.6.1 类型和结构	108
2.6.2 mld6_init()函数	109
2.6.3 加入组: mld6_start_ listening()函数	111

2.6.4	离开组: mld6_stop_ listening() 函数	112	3.2.3	DNS事务和分组格式	177
2.6.5	输入处理: mld6_input() 函数	113	3.2.4	名字解析和缓存	179
2.6.6	mld6_fasttimeo() 函数	117	3.3	DNS中与IPv6相关的主题	181
2.6.7	mld6_sendpkt() 函数	118	3.3.1	AAAA资源记录	181
2.6.8	mld_allocbuf() 函数	121	3.3.2	IPv6的DNS反向树	182
2.7	IPv6多播接口: mif6{}结构	122	3.3.3	DNS的IPv6传输	182
2.8	IPv6多播路由选择API	124	3.3.4	分组大小的问题和EDNS0	183
2.8.1	ip6_mrouter_set() 函数	125	3.3.5	DNS服务器对AAAA的不当 操作	186
2.8.2	ip6_mrouter_init() 函数	126	3.3.6	废弃的标准	188
2.8.3	ip6_mrouter_get() 函数	128	3.4	IPv6 DNS解析器的实现	189
2.8.4	set_pim6() 函数	128	3.4.1	_dns_getaddrinfo() 函数	191
2.8.5	add_m6if() 函数	129	3.4.2	getanswer() 函数	196
2.8.6	del_m6if() 函数	131	3.4.3	res_queryN() 函数	204
2.8.7	ip6_mrouter_done() 函数	132	3.4.4	解析器声明结构	206
2.8.8	mrt6_ioctl1() 函数	135	3.4.5	res_init() 函数	208
2.8.9	get_mif6_cnt() 函数	135	3.4.6	res_send() 函数	211
2.9	IPv6多播转发缓存	136	3.4.7	IPv6反向查找: _dns_ ghbyaddr() 函数	219
2.9.1	add_m6fc() 函数	137	3.5	BIND的IPv6 DNS操作	223
2.9.2	del_m6fc() 函数	141	3.5.1	BIND9概述	224
2.9.3	expire_upcalls() 函数	142	3.5.2	获得BIND9	224
2.9.4	get_sg_cnt() 函数	143	3.5.3	构建和安装BIND9	225
2.10	IPv6多播转发	144	3.5.4	IPv6操作的BIND9配置	225
2.10.1	ip6_mforward() 函数	145	3.5.5	实现的相关要点	231
2.10.2	ip6_mdq() 函数	152	3.5.6	完整的配置示例	238
2.10.3	phyint_send() 函数	158	3.5.7	dig和host工具	241
2.10.4	register_send() 函数	160	第4章	DHCPv6	243
2.10.5	socket_send() 函数	162	4.1	引言	243
2.10.6	pim6_input() 函数	163	4.2	DHCPv6协议概述	243
2.11	IPv6多播操作	169	4.2.1	DHCPv6案例	244
2.11.1	ifmcstat命令	169	4.2.2	DHCPv6的定义	246
2.11.2	启动IPv6多播路由选择	169	4.2.3	DHCPv6报文交换	249
2.11.3	pim6dd和pim6sd路由选择 守护进程	170	4.2.4	DHCPv6选项总结	259
2.11.4	pim6stat输出	171	4.2.5	与邻居发现的交互	265
2.11.5	netstat命令	172	4.2.6	与DHCPv4的比较	265
第3章	IPv6的DNS	174	4.3	代码介绍	265
3.1	引言	174	4.4	客户实现	271
3.2	DNS的定义和协议基础	174	4.4.1	客户特定的数据结构	272
3.2.1	DNS、域和区域	174	4.4.2	client6_mainloop() 函数	277
3.2.2	资源记录和区域文件	176	4.4.3	client6_timo() 函数	278
			4.4.4	client6_send() 函数	282

4.4.5	client6_recv() 函数	287	4.8.7	配置控制命令密钥	409
4.4.6	client6_recvadvert() 函数	289	4.8.8	DHCPv6 服务操作	410
4.4.7	client6_recvreply() 函数	294	第5章 移动IPv6		416
4.4.8	处理身份关联	299	5.1	引言	416
4.4.9	update_ia() 函数	302	5.2	移动IPv6概述	416
4.4.10	update_address() 函数	306	5.2.1	节点的类型	417
4.4.11	reestablish_ia() 函数	310	5.2.2	移动IPv6运行的基本过程	418
4.4.12	ia_timo() 函数	315	5.3	首部扩展	421
4.4.13	释放资源	320	5.3.1	对齐要求	422
4.5	服务器实现	323	5.3.2	归属地址选项	422
4.5.1	server6_mainloop() 函数	326	5.3.3	类型2路由首部	423
4.5.2	server6_recv() 函数	327	5.3.4	移动首部	423
4.5.3	process_relayforw() 函数	331	5.3.5	移动选项	429
4.5.4	react_solicit() 函数	334	5.3.6	邻居发现报文	431
4.5.5	react_request() 函数	340	5.3.7	ICMPv6报文	433
4.5.6	make_ia() 函数	345	5.4	移动IPv6的过程	436
4.5.7	react_renew() 函数	355	5.4.1	协议常量和变量	436
4.5.8	react_rebind() 函数	358	5.4.2	归属注册	437
4.5.9	binding_timo() 函数	365	5.4.3	双向隧道技术	439
4.5.10	react_release() 函数	367	5.4.4	截获移动节点分组	440
4.5.11	react_informreq() 函数	371	5.4.5	返回归属	441
4.5.12	server6_send() 函数	373	5.5	路由优化	442
4.6	中继代理实现	377	5.5.1	迂回路由	443
4.6.1	relay6_loop() 函数	378	5.5.2	发送发起报文	443
4.6.2	relay6_recv() 函数	379	5.5.3	响应发起报文	444
4.6.3	relay_to_server() 函数	382	5.5.4	计算共享密钥	445
4.6.4	relay_to_client() 函数	387	5.5.5	验证报文	446
4.7	DHCPv6鉴别实现	390	5.5.6	安全性考虑	446
4.7.1	与DHCPv6鉴别相关的数据 结构	390	5.5.7	注销通信节点绑定	447
4.7.2	set_auth() 函数	392	5.5.8	向后兼容性	447
4.7.3	process_auth() 函数 (客户端)	394	5.6	移动检测	448
4.7.4	process_auth() 函数 (服务器 端)	397	5.7	动态归属代理地址发现	449
4.8	DHCPv6操作	403	5.8	移动前缀请求/广告	451
4.8.1	构建DHCPv6实现	403	5.9	与IPsec的关系	452
4.8.2	配置DUID	404	5.10	代码介绍	454
4.8.3	配置DHCPv6服务器	404	5.11	与移动IPv6相关的结构	455
4.8.4	配置DHCPv6客户	405	5.11.1	文件	455
4.8.5	配置DHCPv6中继代理	408	5.11.2	移动首部报文: ip6_mh{}	456
4.8.6	配置DHCPv6鉴别	409		结构	456
			5.11.3	绑定刷新请求报文: ip6_mh_binding_request{}	456
				结构	456

5.11.4 归属测试发起报文: ip6_mh_home_test_init{}	5.11.19 类型2路由首部: ip6_rthdr2{}
结构457	结构466
5.11.5 转交测试发起报文: ip6_mh_careof_test_init{}	5.11.20 修改的路由器广告报文: nd_router_advert{}
结构457	结构466
5.11.6 归属测试报文: ip6_mh_home_test{}	5.11.21 修改的前缀信息选项: nd_opt_prefix_info{}
结构458	结构467
5.11.7 转交测试报文: ip6_mh_careof_test{}	5.11.22 广告间隔选项: nd_opt_adv_interval{}
结构459	结构467
5.11.8 绑定更新报文: ip6_mh_binding_update{}	5.11.23 归属代理信息选项: nd_opt_homeagent_info{}
结构460	结构468
5.11.9 绑定确认报文: ip6_mh_binding_ack{}	5.11.24 动态归属代理地址发现请求报文: mip6_dhaad_req{}
结构461	结构468
5.11.10 绑定错误报文: ip6_mh_binding_error{}	5.11.25 动态归属代理地址发现应答报文: mip6_dhaad_rep{}
结构462	结构469
5.11.11 移动选项报文的结构462	5.11.26 移动前缀请求报文: mip6_prefix_solicit{}
5.11.12 移动选项报文: ip6_mh_opt{}	结构469
结构463	5.11.27 移动前缀广告报文: mip6_prefix_advert{}
5.11.13 绑定刷新建议选项: ip6_mh_opt_refresh_	结构469
advice{}	5.11.28 绑定缓存条目: mip6_bc{}
结构463	结构470
5.11.14 替换转交地址选项: ip6_mh_opt_altcoa{}	5.11.29 绑定更新列表条目: mip6_bu{}
结构463	结构471
5.11.15 现时索引选项: ip6_mh_opt_nonce_index{}	5.11.30 归属代理条目: mip6_ha{}
结构464	结构473
5.11.16 授权数据选项: ip6_mh_opt_auth_data{}	5.11.31 前缀条目: mip6_prefix{}
结构464	结构474
5.11.17 内部移动选项: mip6_mobility_options{}	5.11.32 归属虚拟接口: hif_softc{}
结构465	结构475
5.11.18 归属地址选项: ip6_opt_home_address{}	5.12 宏和类型定义479
结构465	5.13 全局变量480
	5.14 实用函数480
	5.14.1 文件481
	5.14.2 创建IPv6首部481
	5.14.3 校验和计算482
	5.15 常规移动首部处理484
	5.15.1 文件484
	5.15.2 移动首部输入484

5.15.3	生成绑定错误报文	490	5.17.11	归属代理列表的管理	653
5.15.4	绑定错误报文的速率限制	491	5.17.12	前缀信息的管理	661
5.15.5	绑定错误报文的创建	492	5.17.13	通过路由器广告报文接收前缀信息	672
5.15.6	向原始套接字传递移动首部报文	493	5.17.14	发送移动前缀请求报文	680
5.16	归属代理和通信节点	496	5.17.15	接收移动前缀广告报文	682
5.16.1	文件	496	5.17.16	发送动态归属代理地址发现请求报文	690
5.16.2	绑定更新报文的输入	497	5.17.17	接收动态归属代理地址发现应答报文	693
5.16.3	绑定缓存条目的管理	504	5.17.18	接收ICMPv6错误报文	698
5.16.4	移动选项的处理	512	5.17.19	状态机	700
5.16.5	通信节点绑定更新报文的确认	514	5.17.20	主状态机	701
5.16.6	K _{bm} 和授权数据计算	515	5.17.21	从状态机	718
5.16.7	通信节点对绑定缓存条目的管理	520	5.17.22	虚拟归属接口	725
5.16.8	发送绑定刷新请求报文	522	5.17.23	迂回路由过程和路由优化	737
5.16.9	归属注册的处理过程	526	5.17.24	路由优化通信	753
5.16.10	DAD过程	532	5.17.25	隧道控制	762
5.16.11	代理邻居发现的控制	538	5.17.26	从隧道中接收分组	765
5.16.12	归属注销的过程	542	5.17.27	I/O控制	767
5.16.13	发送绑定确认报文	545	5.18	运行移动IPv6	770
5.16.14	现时和节点关键字的管理	551	5.18.1	重建支持移动IPv6扩展的内核	770
5.16.15	接收归属地址选项	554	5.18.2	重建用户空间程序	770
5.16.16	通过隧道向移动节点发送分组	561	5.18.3	IPsec信令保护	771
5.16.17	恢复临时禁用的代理条目	565	5.18.4	配置节点	773
5.16.18	接收ICMPv6错误报文	566	5.18.5	查看状态信息	774
5.16.19	归属代理列表的管理	570	5.18.6	查看统计信息	775
5.16.20	前缀列表的管理	583	5.19	mip6control手册	776
5.16.21	发送移动前缀广告报文	583	第6章	IPv6与IP安全	778
5.16.22	构造净荷	585	6.1	引言	778
5.17	移动节点	587	6.2	鉴别首部	778
5.17.1	文件	587	6.3	ESP	780
5.17.2	绑定更新列表条目的管理	588	6.4	传输模式和隧道模式	781
5.17.3	移动检测	596	6.5	安全关联数据库	783
5.17.4	配置归属地址	606	6.5.1	安全策略数据库	783
5.17.5	发送绑定更新报文	615	6.5.2	安全关联数据库	784
5.17.6	接收绑定确认报文	629	6.5.3	SAD和SPD示例	784
5.17.7	接收类型2路由首部	641	6.6	IPsec流量处理	785
5.17.8	接收绑定刷新请求报文	644	6.7	SPD和SAD管理	786
5.17.9	接收绑定错误报文	646	6.8	手工配置	787
5.17.10	源地址选择	649	6.8.1	配置文件格式	788
			6.8.2	处理SP条目的示例	791

6.8.3 处理SA条目的示例	793	6.11 场景	803
6.9 ISAKMP概述	794	6.11.1 在3个网络间创建VPN	803
6.9.1 ISAKMP交换	795	6.11.2 创建星型拓扑VPN	808
6.9.2 解释域	797	6.11.3 使用传输模式的IP安全	812
6.9.3 IKE协议	797	6.11.4 从公共接入点到服务器的连接	813
6.10 Racoon操作	798	参考文献	816
6.10.1 配置Racoon	798	索引	821
6.10.2 配置文件格式	798		

IPv6单播路由选择协议

1.1 引言

每当任意一对节点之间通信时，尤其是当该通信涉及驻留在不同网段的节点时，必须确定每个分组的流向。这一决策通常称为分组路由选择（routing）决策，或者分组转发（forwarding）决策。相关的中介网络设备一般称为路由器（router），负责完成路由选择功能，包括根据每个分组的最终目的地址作出路由选择决策。

可以依据每个路由器上手工配置的路由信息作出路由选择决策，但这种方式在大中型复杂网络中显然是行不通的。路由选择协议（routing protocol）提供了使路由器能自动地作出正确路由选择决策所必需的信息。由于分组的目的地址可能是单播或多播的地址（将广播地址看作多播的特例），所以，设计的路由选择协议也要适用于单播路由选择（unicast routing）或者多播路由选择（multicast routing）。本章将重点关注路由选择协议。

在IPv4领域中，RIPv2[RFC2453]、集成IS-IS[RFC1195]以及OSPFv2[RFC2328]都是在诸如企业环境等中小网络中部署的单播路由选择协议，而BGP-4[RFC4271]则是在诸如因特网服务提供商（Internet Service Provider, ISP）等大型组织中部署的通用路由选择协议。总体上讲，由于IPv4和IPv6在路由选择的概念上是一致的，因此这些路由选择协议就自然地扩展以支持IPv6。尽管分组的格式会有所改变，但原理在很大程度上仍然是一致的。

不过IPv6也有其特有的问题。需要特别指出的是，大多数IPv6路由选择协议在很大程度上依赖链路本地地址，原因在于从路由选择的角度看，其范围有限，所以采用这类地址通信是稳定的。另外，链路本地地址的二义性我们在《卷1》的第2章中已经讨论过，在实现这些协议时需要重点关注。因此，理解这些协议的细节以及如何实现它们，即使对于已经熟悉IPv4路由选择协议的人来说，也是非常重要的。

本章将提供理解和实现IPv6单播路由选择协议的所有基本内容。首先介绍的是路由选择的基本概念，然后介绍IPv6单播路由选择协议，包括RIPng[RFC2080]、OSPFv3[RFC2740]和BGP4+[RFC2545]，还将全面介绍RIPng协议。另外，除了IPv6相关的协议分组内容之外，我们还对OSPFv3以及BGP4+的常规协议操作进行了总结，但没有涉及协议的具体细节。不需要理解这些高级问题的读者大可放心地跳过这两节（1.5节和1.6节）内容，因为本书其他部分并不依赖这些内容。

在协议基础知识后面的几节重点讲解实现，这里提供了在BSD系统上开发IPv6路由选择程序的全部重要内容，囊括了从核心架构到路由选择程序代码等各个方面。首先说明在BSD系统上如何处理IPv6路由选择信息，从内核的内部数据结构到API（应用程序接口）等各个方面的内容。同时，也指出了采用这些API进行链路本地地址处理时存在的主要缺陷。然后描述KAME项目RIPng路由选择守护进程route6d的实现，主要讨论其RIPng协议的处理。RIP起源于route6d程序，该程序主流的实现版本在各种平台上都得到了广泛支持。其流行的原因在于实