

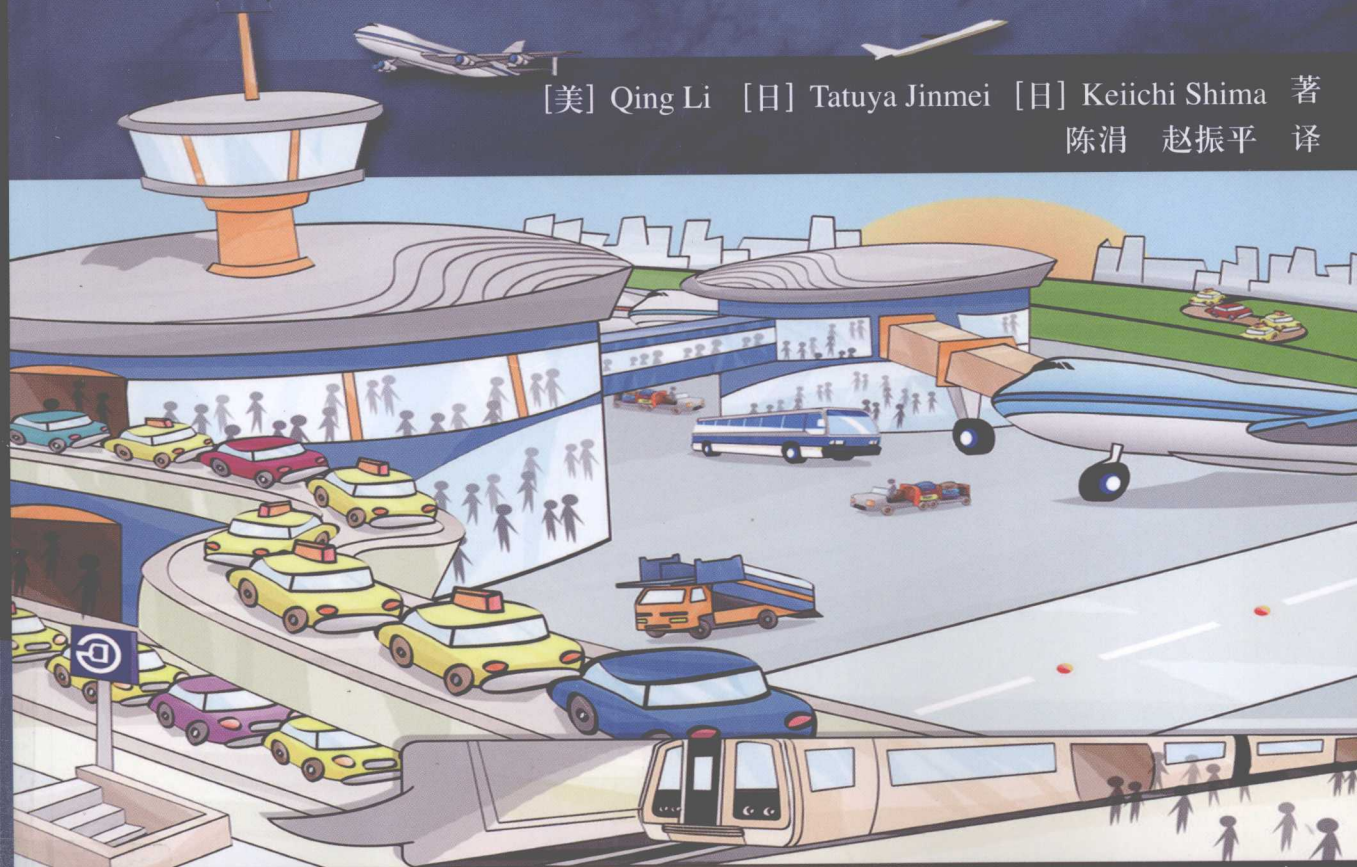
IPv6时代的《TCP/IP详解》!

IPv6详解

卷1：核心协议实现

IPv6 Core Protocols Implementation

[美] Qing Li [日] Tatuya Jinmei [日] Keiichi Shima 著
陈涓 赵振平 译



人民邮电出版社
POSTS & TELECOM PRESS



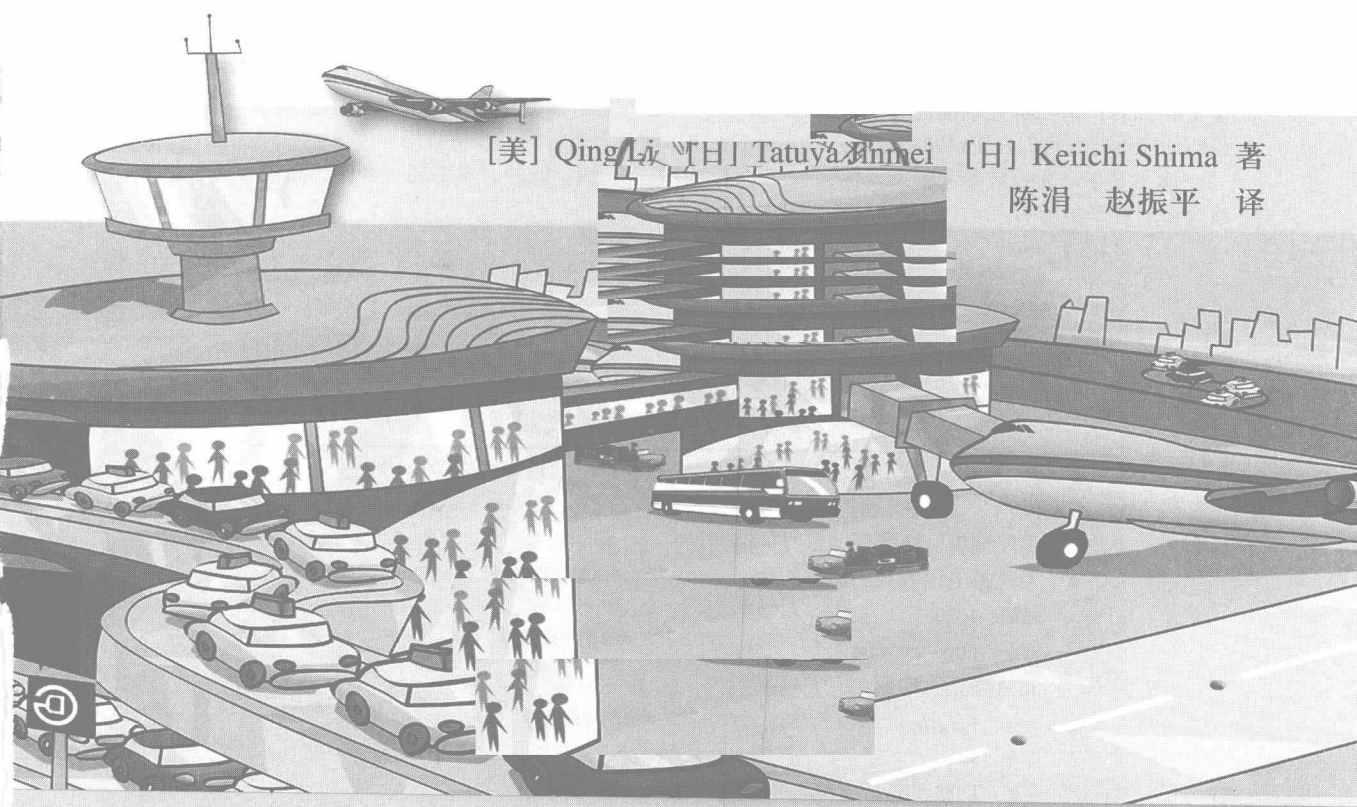
图灵计算机科学丛书

IPv6详解

卷1：核心协议实现

IPv6 Core Protocols Implementation

[美] Qing Li [日] Tatsuhiro Minai [日] Keiichi Shima 著
陈涓 赵振平 译



人民邮电出版社
北京

图书在版编目(CIP)数据

IPv6详解,第1卷,核心协议实现/(美)李清,(日)神明达哉,(日)岛庆一著;陈涓,赵振平译. —北京:人民邮电出版社,2009.1

(图灵计算机科学丛书)

书名原文: IPv6 Core Protocols Implementation

ISBN 978-7-115-18950-9

I. I… II. ①李… ②神… ③岛… ④陈… ⑤赵… 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(2008)第154120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,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7章,分别介绍IPv6地址结构、IPv6、ICMPv6、邻居发现和无状态地址自动配置、传输层协议及套接字API扩展。每章都包含两个主要部分:第一部分提供了作为RFC发布的主要协议规范的摘要,定义各种协议分组,解释每个分组中每个字段的含义和目的;第二部分则描述实现这些RFC的KAME的数据结构及功能。书的最后还提供了一些例子,说明如何编写既可以在IPv4网络上运行、又可以在IPv6网络上运行的可移植应用程序。

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1:核心协议实现

◆ 著 [美] Qing Li [日] Tatuya Jinmei [日] Keiichi Shima
译 陈 涓 赵振平
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>
北京鑫正大印刷有限公司印刷

◆ 开本: 787×1092 1/16

印张: 54

字数: 1 488 千字

印数: 1—3 500 册

2009年1月第1版

2009年1月北京第1次印刷

著作权合同登记号 图字: 01-2007-4741号

ISBN 978-7-115-18950-9/TP

定价: 128.00元

读者服务热线: (010) 88593802 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: 978-0-12-447751-3.

Copyright © 2007 by Elsevier. All rights reserved.

Authorized Simplified Chinese translation edition published by the Proprietor.

ISBN: 978-981-259-947-6

Copyright © 2008 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 2008

2008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Elsevier (Singapore) Pte Ltd.授权人民邮电出版社独家出版。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

献给所有参与、支持KAME项目的人：项目发起者、管理者、开发者、为此作出贡献的人和用户。

——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

序

早在1994年IETF接受这个现在被称为IPv6的建议时，我就确信，如果没有可以免费获得的高质量的开源参考实现，这样一个新的基础协议即便不是不可能，也是很难被网络群体接受、采纳和部署的。

这种信念源于我对最初的TCP/IPv4协议，以及加州大学伯克利分校完成的这些协议的BSD实现的密切参与。我亲眼目睹了BSD实现是如何对TCP/IPv4（通常统称为因特网协议）的成功作出巨大贡献的。我们需要能够为IPv6扮演同样角色的新的成就。

从开发者的视角为因特网世界作出贡献是我们的责任，但那时，因特网的繁荣带来的经济冲击已经使我伯克利的同事们忙得不可开交了。我知道这是我们的使命，为此，我们在WIDE项目中成立了IPv6工作组，这个工作组最终发展成了KAME项目。

对KAME项目组开发的软件提出的要求之一就是：说明IPv6协议是怎样工作的，以及这些协议在实际环境中运行得如何。这是一项艰难且具有挑战性的任务。在所有支持者长期、耐心的帮助下，KAME项目组成员以他们的勤勉和不懈努力完成了这个目标。所有主要的BSD变体都已经将KAME实现当作事实上的IPv6实现。在IETF会议上讨论IPv6时也经常提到KAME。我坚信，在IPv6技术被广泛接受并不断被采纳的过程中，KAME项目的成功起了非常重要的作用。

现在，我将部分注意力转移到了采纳这项技术后的知识传递方面。我们必须将我们的经验和智慧与那些追随KAME脚步以及将成为KAME主要用户的新一代工程师们分享。虽然一个成熟的工程师可以直接从源代码中进行学习，但一本由“知道怎样做”的专家写成的关于“如何做”的好书所能提供的深入理解和指导，要远比标准文档和源代码所能提供的多得多。

尽管现在有很多关于IPv6的学习资料都很不错，但那些资料通常只是一些更易读的协议规范。我一直渴望能有些与众不同的东西，为下一代网络高手提供一个创造性的舞台，使其在这个新技术舞台上做更多的事情。本书填补了这个空白。

本书揭示了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。通过阅读本书，读者可以深刻理解这一实现，并从整体上理解IPv6。

本书是关于KAME和IPv6的权威参考书，涵盖了很多其他书中都没有的内容。希望本书能够成为你在IPv6方面最主要而且最好的参考书。

IPv6正向着无处不在的应用发展，我相信本书将会成为当今“因特网协议”的教科书。

Jun Murai博士

WIDE项目

庆应义塾（KEIO）大学副校长

前言

无论是在学术界还是在商业界，KAME的IPv6实现都已经作为IPv6实现的事实参考。世界上无数研究人员都将他们对IPv6及下一代因特网的研究建立在实现了KAME的系统上。对KAME的接受已经超出了纯学术研究的范畴，而成为了很多商业产品的一个组成部分。

本系列书通过KAME实现对IPv6及相关协议进行了详细的描述。书中采用的方式与W. Richard Stevens和Gary R. Wright编写《TCP/IP详解》所采用的方式类似，《TCP/IP详解》为20世纪90年代的开发人员提供了对TCP/IP4网络本质的真知灼见。我们希望在本书中与读者分享我们在KAME和IPv6方面的深刻理解和开发经验，帮助读者投入到这个全新的、普遍连通的美妙世界中。

现在关于IPv6的书有很多，在Amazon上搜一下书名中包含IPv6的书就能得到不下几十个结果。但我们的书并不仅仅是另一套关于IPv6的书。据我们所知，我们的书是同类书中的第一套也是唯一一套，书中对一个已经被很多开源操作系统以及商业操作系统采纳的实际IPv6实现进行了深入研究，并对代码及其设计进行了剖析，说明了我们是如何根据标准对IPv6及其相关协议进行解释和实现的。书中会阐明标准中那些具有二义性的地方，这些地方可以用各种方法解释，并且在应用中存在问题，因而为实现带来挑战。读者将从本质上理解早期实现经验能够为一种技术的开发和成长提供什么样的帮助。

本书的组织

本书是关于IPv6及相关协议的两卷本中的第一本。本书通过完整的代码描述以及对KAME本质的阐释，从规范到操作介绍了核心IPv6协议。

本书对IPv6协议基础及特征、ICMPv6协议、对在IPv6上的TCP或UDP这样的传输层问题的考虑以及开发支持IPv6的应用程序所需的套接字扩展进行了讨论。下面对本书每章的内容进行简单的说明。

- 第1章介绍KAME项目的简要历史，描述KAME发行版的组件，解释构建一个实现了KAME的BSD系统的步骤，介绍BSD联网系统的基本知识以及KAME对基本系统的增强。
- 第2章讨论各种类型IPv6地址的格式和结构，随后是与地址配置及操作有关的代码描述。
- 第3章讨论IPv6的协议结构、扩展首部及在IPv6层处理IPv6分组输入及输出的KAME实现。本章还会介绍用来传输IPv6分组的默认地址选择算法。
- 第4章讨论ICMPv6的协议结构，这是IPv6操作的一个组成部分，描述KAME实现中基本ICMPv6操作的输入函数和输出函数。作为ICMPv6的一个有趣应用，本章还会详细描述节点信息查询协议。
- 第5章对邻居发现（ND）协议进行了描述。邻居发现运行于ICMPv6之上，并提供了各种基础服务，例如链路层地址解析、路由器发现以及路由重定向等。自动地址配置服务也是通过邻居发现分组提供的。本章会非常详细地阐述邻居发现协议和无状态地址自动配置过程。本章会描述与邻居发现协议操作有关的主要KAME内核函数。
- 第6章描述为使传输层协议运行于IPv6之上而对TCP和UDP所做的修改，还会说明为使传

输层协议成为双协议栈，而对BSD内核实现所做的修改。

- 第7章讨论对套接字API所做的IPv6扩展和增强。本章还会讨论几个已扩展或创建以支持IPv6的标准用户库。本章最后会提供一些例子，说明如何编写既可以在IPv4网络上运行又可以在IPv6网络上运行的可移植应用程序。

本书的姊妹篇为《IPv6详解，卷2：高级协议实现》（英文版书名为*IPv6 Advanced Protocols Implementation*，文中简称为《卷2》），书中讲述了下列主题。

- 第1章——IPv6单播路由选择协议。
- 第2章——IPv6多播。
- 第3章——IPv6的DNS。
- 第4章——IPv6的动态主机配置协议（DHCPv6）。
- 第5章——移动IPv6。
- 第6章——IPv6与IP安全。

图0-1以图形方式描述了这两本书涵盖的IPv6的各个方面。我们将本书称为“卷1”，将高级协议那本称为“卷2”。

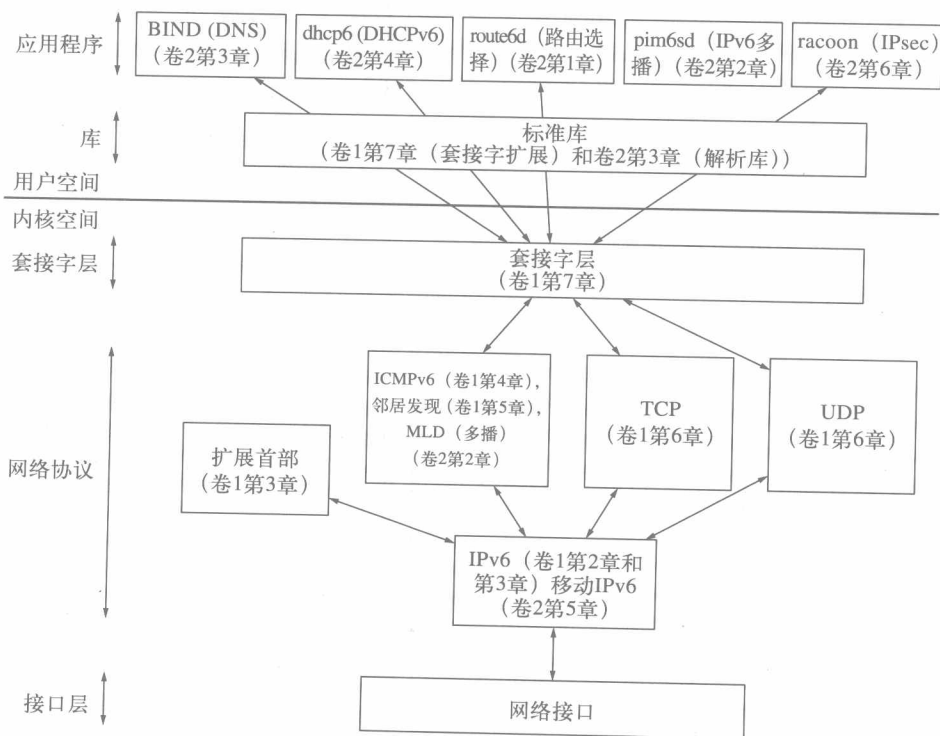


图0-1 两本书涵盖的IPv6的各个方面

这两本书涵盖了很多协议规范，这些规范主要是作为RFC（请求评论）发布的，并且是互相关联的。图0-2总结了本系列图书涵盖的主要RFC和描述特定规范的各章之间的关系。箭头表示两个特定规范之间的引用关系。比如，从RFC3041到RFC2461的箭头就说明前者引用了后者。

每章都包含两个主要部分，每部分都由一个或两个连续的小节组成。第一部分提供了对那一章介绍的、作为RFC发布的主要协议规范的摘要。对第3章至第5章这样的协议章来说，第一部分描述的是特定协议的基本操作。第一部分定义各种协议分组，解释每个分组中每个字段的

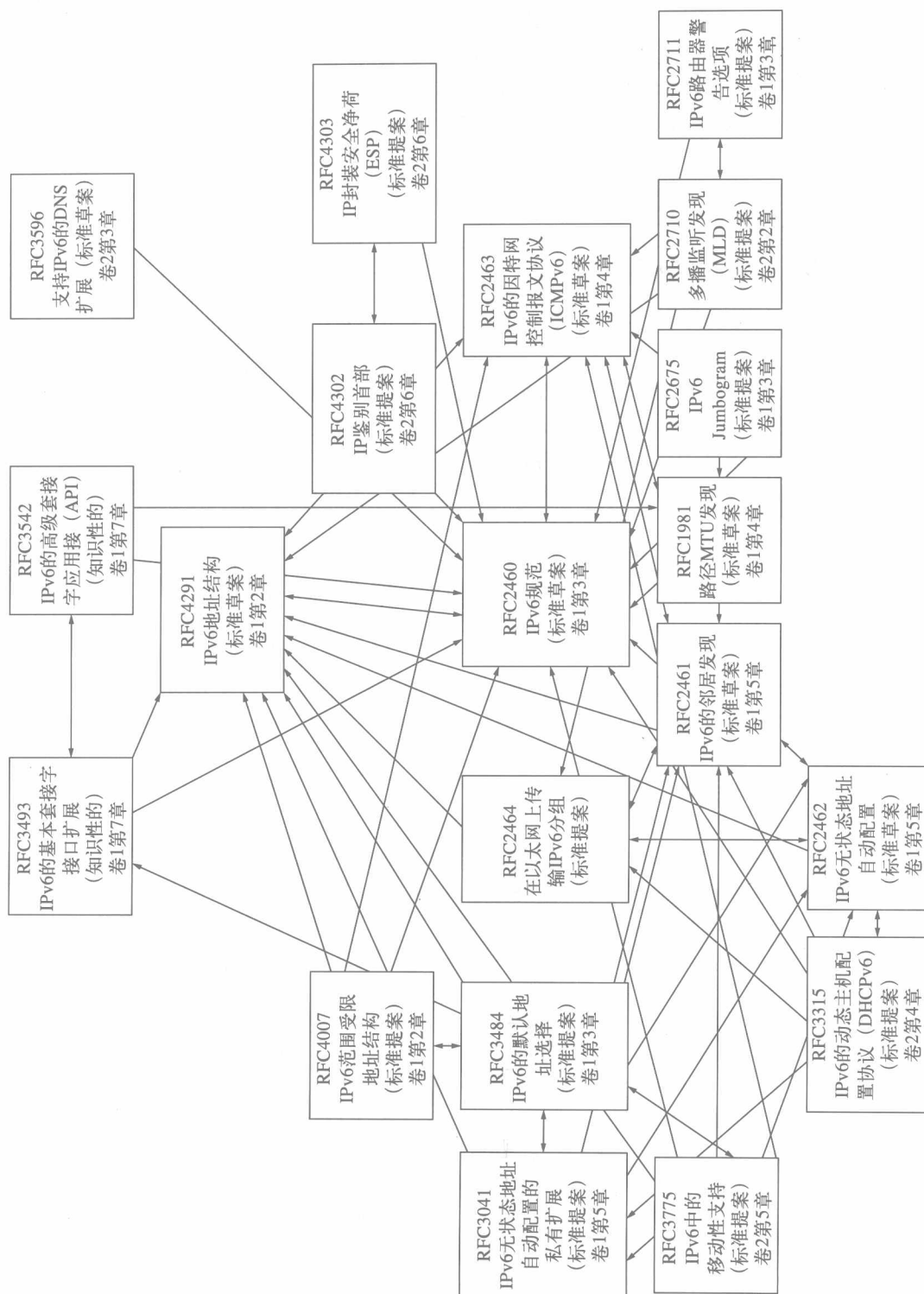


图0-2 本系列图书涵盖的主要RFC和描述特定规范的各章之间的关系

含义和目的，描述有关协议是如何处理分组的。这部分还尝试着在必要的地方对RFC中有歧义的文本进行了澄清。

每章的第二部分则描述实现这些RFC的KAME的数据结构和函数。书中提供了一些图，以图形方式说明了各种数据结构之间的关系，还提供了函数调用图来说明代码的来龙去脉。在这一部分中，代码叙述也能澄清规范中存在的一些有歧义的地方，并标识出RFC中一些难以实现的部分。

有些章还有一个附加的尾部。这部分描述一些有用的工具，并说明其用途，从而强调这个实现在实际操作中是怎样工作的。

本书的目标读者及使用方法

对刚接触IPv6或KAME的开发人员来说，可以将本书当作IPv6方面的第一本书。对需要掌控或维护基于KAME的系统的开发人员，以及需要在不同的BSD变体上采用KAME代码并维护及增强KAME的BSD开发人员来说，可以将本书当作一本参考教程。

本书也适合作为大学本科四年级关于IPv6及相关协议的计算机科学课程的教材，也可以作为一年级研究生IPv6方面基于项目的课程的教学参考书。指导教师必须设置一些计算机网络、网络编程方面的计算机科学基础课程以及一门关于TCP/IP的课程（选修）作为先修课程。学生还应该熟练掌握C语言。

通常我们鼓励读者按顺序来阅读书中各章的内容，但贯穿全书的大量交叉引用使读者可以直接从一章跳到另一章感兴趣的主题，而不必担心迷失在这个代码迷宫中。对那些熟悉现有TCP/IP技术的读者来说，第6章和第7章对如何将现有应用程序转换到IPv6中去，以及如何编写IPv6应用程序提供了有用的过渡指导。

对各种设计决策的详细描述揭示了规范及实现中存在的偏差。这些深入的理解可以帮助研究人员找出新的研究课题。从KAME中获取的知识使研究人员可以将KAME设为一个平台，并在这个平台上建立一些实验解决方案。

配套资源

本书英文原版带有2张CD-ROM^①。第一张CD-ROM是FreeBSD4.8-RELEASE的ISO映像，这是本书讨论的基础操作系统。这是一张可引导CD-ROM，其中包含安装文件。用CD-ROM开机就可以启动安装过程。CD-ROM根目录中的INSTALL.TXT文件中有详细的安装过程。第二张CD-ROM中包含FreeBSD4.8-RELEASE的源文件和本书提到的KAME快照发行版。在CD-ROM的根目录中有两个子目录。

- FreeBSD4.8-RELEASE——包含FreeBSD4.8-RELEASE的整个源树。
- kame-snap-20030421——包含2003年4月21日创建的KAME快照的源树。

本书描述的所有源文件都可以在第二张CD-ROM中找到。读者阅读本书时可以在CD-ROM中找到完整的源代码，也可以用目录kame-snap-20030421下的源树来安装KAME实现，以查看KAME栈的实际表现。1.3.2节对KAME快照的安装过程进行了说明。

源代码版权

本书提供了KAME项目开发的很多源代码，也引用了作为FreeBSD系统的一部分发布的其

^① 中译本未随书附带光盘，光盘中的内容可以从图灵网站www.turingbook.com本书网页免费注册下载。

他源代码。所有这些源代码都是有版权声明的，这些版权声明在第二张CD-ROM包含的源代码副本中。

报告错误及勘误表

尽管在本书的审阅和更新阶段，我们尽了最大努力找出并修正错误，但在这么大部头的一本书中，还是不可避免地会有一些疏漏的错误。作者很高兴收到关于本书内容的错误报告，并在因特网上提供一个错误更正页面。这个页面位于<http://books.elsevier.com/companions/0124477518>。

致谢

Qing Li的个人致谢

2003年1月，我刚刚在风河（Wind River）公司完成了一项历时3年的嵌入式IPv6项目，正在整理我的笔记和文章以响应ACM SIGCOMM对新通信协议的征稿通知。

我厚厚的笔记本里记录了代码路径、对各种数据结构及代码片断的描述、大量的图表，以及数不清的关于KAME实现的注解。阅读笔记时，我想，如果有一本关于KAME的书该多好啊。这个想法马上打动了：我可以把所有这些注解都放到一本书中去，这样就可以帮助其他人理解IPv6，并推动他们的项目的发展了。这个想法一发不可收，并逐渐变成一种讨论整个KAME实现的大胆尝试。

我在2003年3月向KAME核心开发者Jinmei先生提出了关于写这本书的想法，并邀请他加入。Jinmei先生是IPv6及多播路由选择、DHCPv6和DNS等各种重要因特网技术领域一位知名且受人尊敬的研究者。他欣然接受了我的邀请。

另一名核心KAME开发者Shima先生在3月底加入了我们的行列。Shima先生是IPv6及移动领域中一位知名且受人尊敬的研究者。

我们在2003年5月就一份写作计划达成了一致，但书在差不多3年之后才完成。

Jinmei先生为整本书提供了技术指导。他对内容、细节及技术精确性的洞察力和细心的关注确保了本书的质量。他在内容上为我提供了很多帮助，即使很忙，甚至在24小时工作的时候，也总是积极地回答我的一些技术问题和质询。

Shima先生在内容和修订方面都为我提供了很多帮助。尤其是在项目快结束，我承受着工作和个人生活的双重压力时，Shima先生和Jinmei先生帮助我冲过了“终点线”。Shima先生作出的贡献显然比他应该承担的责任要多。

和我的合作者一起工作是一个绝妙的学习过程。他们出色的专业技能、他们的勤奋以及他们的工作理念都为我设定了一个新的学习标准。

我要感谢Rick Adams与我分享了对书的看法，感谢他在过去3年里付出的耐心和时间。在我们的谈话中，大部分时间都是我在解释为什么时间表再次推后了，而在很多次谈话中，他都是以很委婉的方式来表达他的沮丧，对此我非常感激。

我要感谢我的妻子Huaying Cheng，感谢她对我经常在电脑前工作到深夜并在电脑前度过那么多个周末的理解和支持。

Tatuya Jinmei的个人致谢

除了在下边的小组致谢中感谢的那些为本书作出贡献的人之外，我个人还想感谢东芝公司的经理Yukio Kamatani、Toshio Murai、Yasuhiro Katsube和Atsushi Inoue，他们为我在KAME项目组的工作提供了支持，并批准了这项派生出来的工作。要是开始没有他们的理解和支持，我就不会成为一名KAME的开发者，更不要说写一本关于KAME的书了。

同样，如果没有Jun Murai和Hiroshi Esaki，我也不会取得今天的成就。在WIDE和KAME项目中，他们总是鼓励我，为我提供新的机会，还教会了我无私奉献的精神。Murai先生还友好地接受了为本书作序的邀请。

最后，我要感谢我的合作者。很显然，如果Qing没有邀请我，我永远也想不到要写一本关于KAME的书。对于我编写书稿时的延误他总是能宽容对待，并在我为种种细节斟酌不休，延误交稿达半年之久的情况下，保证了我们没有超过出版合同约定的时限。当我为在截稿日期之前修改完书稿感到绝望时，Shima先生帮助了我——这远远超出了他的责任。

Keiichi Shima的个人致谢

首先，我要感谢所有创建了高质量IPv6参考代码的KAME开发者。没有这些代码，就不会有这本书，也就是说，没有那些开发者不断润饰代码的努力，没有他们用代码来实现IPv6技术的坚定决心，就不会有这本书。幸运的是，我加入了这个项目，这是我的荣耀，我可以为这个伟大的项目作出我的贡献。对我来说，这是一段宝贵的时光，是一种无法替代的体验。我开始做与IPv6相关的工作时还是个学生。Jun Murai先生和Kazuhiko Yamamoto先生给我提供了在WIDE项目中为IPv6工作的机会。我无法想象如果没有遇到他们我会怎么样。Kazumasa Utashiro先生和Eiiti Wada先生为我在KAME项目中的活动提供了支持。没有因特网技术，我所有的工作都无法完成。我要感谢所有为因特网的运行做出不懈努力的人们。

小组致谢

作者要感谢所有的KAME开发者。这也是“他们的书”，因为他们编写的KAME代码占本书一半的篇幅。尤其要感谢Shinsuke Suzuki，他阅读了本书的每一章，并提出了很多宝贵的注解和建议。Kenjiro Cho、Tsuyoshi Momose和Ryuji Wakikawa仔细地审阅了部分章节，并以大量细致的注解帮助我们提高了本书的质量。我们还要感谢那些项目之外的审稿人，特别是Rob Austein、George Neville-Neil、Andre Oppermann、Shawn Routhier、Randall Stewart、Shinta Sugimoto和Kazumasa Utashiro。他们自愿花费了大量的个人时间。本书的封面图片基于知名的KAME形象，这个形象是（小鼯鼠画室的）Manabu Higashida和Chizuru Higashida设计的KAME项目标志。

我们还要感谢Gary R. Wright和W. Richard Stevens，是他们的书帮助我们建立了基本的TCP/IP知识。他们描述BSD网络代码的模式非常有效，所以我们决定在本书中借用这种模式来描述KAME的IPv6实现。我们的贡献应该部分归功于他们，归功于他们对网络工程界的贡献。我们希望能够为IPv6技术中的下一代工程人员作出同样的贡献。

最后，我们要向本书的编辑Rick Adams、Rachel Roumeliotis、Dawnmarie Simpson以及Morgan Kaufmann/Elsevier的编辑们表达我们诚挚的谢意，感谢他们为将我们的书稿出版为一本专业书籍所做的努力。

目 录

第1章 绪论	1	2.9 对范围区域的处理	39
1.1 引言	1	2.9.1 范围区域的初始化	39
1.2 IPv6和KAME的简史	1	2.9.2 范围区域ID	40
1.3 KAME发行版概述	5	2.9.3 地址结构中的区域ID	41
1.3.1 源代码树结构	5	2.9.4 与范围相关的实用函数	43
1.3.2 构建过程	6	2.10 接口地址结构	49
1.4 BSD网络实现概述	8	2.10.1 ifaddr{}和in6_ifaddr{} 结构	50
1.5 源代码描述	10	2.10.2 in6_ifreq{}和in6_aliasreq{} 结构	52
1.5.1 排版约定	10	2.10.3 多播地址结构	53
1.5.2 源代码描述示例	11	2.11 IPv6前缀结构	56
1.5.3 预处理器变量	12	2.12 地址操纵程序概述	57
1.5.4 网络设备与体系结构假设	12	2.13 IPv6的接口初始化	60
1.6 mbuf与IPv6	13	2.13.1 in6_if_up()函数	60
1.6.1 常见的mbuf操纵宏和函数	14	2.13.2 in6_ifattach()函数	61
1.6.2 mbuf标签	14	2.13.3 in6_ifattach_loopback{} 函数	64
1.6.3 IPv6的mbuf要求	18	2.13.4 in6_ifattach_linklocal() 函数	65
1.6.4 诊断mbuf链	20	2.13.5 get_ifid()函数	69
第2章 IPv6编址体系结构	22	2.13.6 get_hw_ifid()函数	71
2.1 引言	22	2.13.7 get_rand_ifid()函数	75
2.2 IPv6地址	22	2.13.8 in6if_do_dad()函数	76
2.3 IPv6地址的文本表示	23	2.14 IPv6接口地址配置	77
2.4 地址范围	24	2.14.1 in6_control()函数	77
2.4.1 范围区域	25	2.14.2 in6_update_ifa()函数	83
2.4.2 区域索引	27	2.14.3 in6_joingroup()和 in6_leavegroup()函数	95
2.4.3 范围受限地址的文本表示	29	2.14.4 in6_addmulti()和 in6_delmulti()函数	96
2.4.4 单播站点本地地址的废弃	30	2.14.5 in6_ifinit()函数	98
2.5 IPv6地址格式	31	2.14.6 in6_ifaddloop()和in6_ ifloop_request()函数	100
2.5.1 接口标识符的产生	32	2.15 删除IPv6地址	102
2.5.2 有关地址格式的注解	33	2.15.1 in6_purgeaddr()函数	103
2.5.3 多播地址格式	33	2.15.2 in6_ifremloop()函数	104
2.6 节点地址要求	35		
2.7 IPv6地址空间管理	35		
2.8 代码介绍	36		
2.8.1 IPv6地址结构: in6_addr{}与 sockaddr_in6{}	36		
2.8.2 宏与变量	37		

2.15.3 in6_unlink_ifa() 函数	105
2.16 用地址配置工具进行的操作	106
第3章 IPv6	110
3.1 引言	110
3.2 IPv6首部格式	110
3.3 IPv6扩展首部	112
3.3.1 扩展首部的顺序	113
3.3.2 逐跳选项首部	113
3.3.3 目的选项首部	114
3.3.4 路由首部	114
3.3.5 片首部	116
3.3.6 IPv6选项	118
3.4 源地址选择	120
3.4.1 默认地址选择	120
3.4.2 源地址选择	121
3.4.3 目的地址选择	123
3.5 代码介绍	124
3.5.1 统计数据	125
3.5.2 首部结构	126
3.5.3 ip6protosw{} 结构	130
3.6 mbuf中的IPv6分组地址信息	132
3.6.1 ip6_setdstifaddr() 函数	134
3.6.2 ip6_getdstifaddr() 函数	134
3.6.3 ip6_setpktaddrs() 函数	134
3.6.4 ip6_getpktaddrs() 函数	136
3.7 输入处理: ip6_input() 函数	136
3.8 逐跳选项首部的处理:	
ip6_hopopts_input() 函数	150
3.8.1 对每个选项进行处理: ip6_	
process_hopopts() 函数	151
3.8.2 处理未知选项:	
ip6_unknown_opt() 函数	155
3.9 对目的选项首部的处理:	
dest6_input() 函数	156
3.10 已分片分组的重装	157
3.10.1 用于分组重装的结构	157
3.10.2 frag6_input() 函数	160
3.11 对路由首部的处理:	
route6_input() 函数	172
3.12 转发: ip6_forward() 函数	177
3.13 输出处理	186
3.13.1 源地址选择:	
in6_selectsrc() 函数	187
3.13.2 路由选择:	
ip6_selectroute() 函数	199
3.13.3 ip6_output() 函数	206
3.13.4 构建扩展首部:	
ip6_copyexthdr() 函数	237
3.13.5 分割首部: ip6_splthdr()	
函数	237
3.13.6 插入特大净荷选项: ip6_	
insert_jumboopt() 函数	238
3.13.7 分片: ip6_insertfraghdr()	
函数	241
3.13.8 路径MTU判断: ip6_getpmtu()	
函数	242
3.13.9 多播环回: ip6_mloopback()	
函数	245
第4章 ICMPv6	247
4.1 引言	247
4.2 ICMPv6报文	247
4.2.1 目的不可达报文	248
4.2.2 分组太长报文	249
4.2.3 超时报文	250
4.2.4 参数问题报文	251
4.2.5 回送请求报文	252
4.2.6 回送应答报文	252
4.2.7 ICMPv6报文处理规则	253
4.3 PMTU发现机制	253
4.4 节点信息查询	254
4.4.1 节点信息报文的格式	255
4.4.2 NOOP查询	257
4.4.3 支持的Qtype查询	257
4.4.4 节点名查询	257
4.4.5 节点地址查询	258
4.4.6 IPv4地址查询	259
4.5 代码介绍	260
4.5.1 统计数据	260
4.5.2 ICMPv6首部	262
4.6 ICMPv6输入处理	263
4.6.1 icmp6_input() 函数	263
4.6.2 错误通知: icmp6_notify_error()	
函数	272
4.7 PMTU发现的实现	280

4.8 ICMPv6输出处理	283
4.8.1 发送错误: icmp6_error() 函数	284
4.8.2 错误速率限制: icmp6_ratelimit()函数	290
4.8.3 icmp6_reflect()函数	291
4.9 节点信息查询的实现	297
4.9.1 类型和变量	297
4.9.2 ping6命令: 发送查询	298
4.9.3 ping6命令: 接收应答	303
4.9.4 ping6命令: 打印支持的Qtype	309
4.9.5 ping6命令: 打印节点地址	312
4.9.6 查询处理: ni6_input()函数	314
4.9.7 节点名的操作	322
4.9.8 创建节点地址应答: ni6_store_addrs()函数	330
4.10 节点信息操作	333
第5章 邻居发现和无状态地址自动配置	335
5.1 引言	335
5.2 邻居发现协议概述	336
5.3 无状态地址自动配置概述	336
5.4 邻居发现协议报文	337
5.5 邻居发现协议报文的交换示例	338
5.6 邻居发现协议分组类型及格式	340
5.6.1 路由器请求报文	340
5.6.2 路由器广告报文	341
5.6.3 邻居请求报文	344
5.6.4 邻居广告报文	344
5.6.5 重定向报文	346
5.7 邻居发现选项类型及格式	347
5.7.1 链路层地址选项	347
5.7.2 前缀信息选项	348
5.7.3 重定向首部选项	349
5.7.4 MTU选项	349
5.7.5 路由信息选项	350
5.8 下一跳判定和地址解析	351
5.9 邻居不可达检测算法	351
5.10 无状态地址自动配置	352
5.10.1 地址的构成和地址状态	353
5.10.2 重复地址检测算法	354
5.10.3 处理路由器广告	355
5.10.4 隐私扩展	356
5.11 路由器特有的操作	357
5.11.1 发送未经请求的路由器广告	359
5.11.2 处理路由器请求	359
5.11.3 处理路由器广告	360
5.12 主机特有的操作	360
5.12.1 发送路由器请求	361
5.12.2 处理路由器广告	361
5.12.3 默认路由器选项	362
5.13 代码介绍	362
5.13.1 邻居发现报文定义	362
5.13.2 邻居缓存: llinfo_nd6{}结构	365
5.13.3 操作变量: nd_ifinfo{}结构	367
5.13.4 默认路由器: nd_defrouter{} 结构	368
5.13.5 前缀: nd_prefix{}结构	369
5.13.6 前缀控制: nd_prefixctl{} 结构	370
5.13.7 邻居发现报文选项: nd_opts{} 结构	371
5.13.8 DAD队列条目: dadq{}结构	372
5.13.9 IPv6地址: in6_ifaddr{}结构	372
5.13.10 目的地缓存	372
5.13.11 操作常量	372
5.14 初始化函数	373
5.14.1 nd6_init()函数	373
5.14.2 nd6_ifattach()函数	374
5.15 邻居缓存管理函数	375
5.15.1 nd6_rtrequest()函数	375
5.15.2 nd6_cache_lladdr()函数	385
5.15.3 nd6_lookup()函数	395
5.15.4 nd6_free()函数	398
5.15.5 nd6_timer函数	401
5.16 邻居发现协议报文处理函数	408
5.16.1 nd6_ns_output()函数	408
5.16.2 nd6_ns_input()函数	414
5.16.3 nd6_na_input()函数	422
5.16.4 nd6_na_output()函数	432
5.16.5 nd6_rs_input()函数	437
5.16.6 nd6_ra_input()函数	440
5.16.7 icmp6_redirect_input() 函数	447
5.16.8 icmp6_redirect_output() 函数	454

5.17 邻居发现协议报文选项处理函数	461
5.17.1 nd6_option_init() 函数	461
5.17.2 nd6_option() 函数	462
5.17.3 nd6_options() 函数	463
5.18 默认路由器管理函数	465
5.18.1 defrouter_addreq() 函数	465
5.18.2 defrouter_delreq() 函数	466
5.18.3 defrouter_addifreq() 函数	467
5.18.4 defrouter_delifreq() 函数	469
5.18.5 defrouter_lookup() 函数	470
5.18.6 defrouter_select() 函数	471
5.18.7 defrtrlist_del() 函数	475
5.18.8 defrtrlist_update() 函数	477
5.19 前缀管理函数	479
5.19.1 nd6_prelist_add() 函数	479
5.19.2 prelist_remove() 函数	481
5.19.3 prelist_update() 函数	482
5.19.4 find_pfxlist_reachable_router() 函数	491
5.19.5 与在链条件有关的前缀和 地址状态	491
5.19.6 pfxlist_onlink_check() 函数	493
5.19.7 nd6_prefix_onlink() 函数	497
5.19.8 nd6_prefix_offlink() 函数	500
5.20 无状态地址自动配置函数	503
5.20.1 in6_ifadd() 函数	503
5.20.2 in6_tmpifadd() 函数	506
5.20.3 regen_tmpaddr() 函数	509
5.21 重复地址检测函数	511
5.21.1 nd6_dad_find() 函数	511
5.21.2 nd6_dad_starttimer() 函数	512
5.21.3 nd6_dad_stoptimer() 函数	512
5.21.4 nd6_dad_start() 函数	512
5.21.5 nd6_dad_stop() 函数	515
5.21.6 nd6_dad_timer() 函数	516
5.21.7 nd6_dad_duplicated() 函数	519
5.21.8 nd6_dad_ns_output() 函数	520
5.21.9 nd6_dad_ns_input() 函数	521
5.21.10 nd6_dad_na_input() 函数	522
5.22 其他函数	523
5.22.1 nd6_is_addr_neighbor() 函数	523
5.22.2 nd6_output() 函数	525
5.22.3 rt6_flush() 函数	531
5.22.4 nd6_rtmsg() 函数	533
第6章 传输层的实现	534
6.1 简介	534
6.2 IPv6上的TCP和UDP	534
6.3 IPv6的伪首部	535
6.4 IPv4校验和与IPv6校验和的区别	536
6.5 IPv4映射的IPv6地址的用法	536
6.6 代码介绍	536
6.7 对PCB和套接字的一般操作	544
6.7.1 IPv6 PCB的分配: in_pcballoc() 函数	544
6.7.2 绑定本地地址: in6_pcbbind() 函数	546
6.7.3 固定远程地址: in6_pcbconnect() 函数	554
6.7.4 in6_pcbldaddr() 函数	556
6.7.5 搜索PCB条目: in6_ pcblookup_local() 函数	558
6.7.6 搜索IPv4映射的PCB: in_pcblookup_local() 函数	561
6.7.7 搜索PCB条目: in6_pcblookup_hash() 函数	563
6.7.8 搜索IPv4映射的PCB: in_pcblookup_hash() 函数	565
6.7.9 分离一个IPv6 PCB: in6_pcbdetach() 函数	567
6.7.10 控制报文信令: in6_pcbnotify() 函数	569
6.7.11 清空PCB缓存路由: in6_rtchange() 函数	573
6.7.12 获取对等实体地址: in6_setpeeraddr() 函数	573
6.7.13 获取本地地址: in6_setsockaddr() 函数	575
6.8 IPv6上的TCP	576
6.8.1 ip6protosw{} 的IPv6上的 TCP实例	576
6.8.2 TCP输出	577

6.8.3 初始化首部: tcp_fillheaders() 函数	582	7.3.1 一些高级的定义	667
6.8.4 TCP输入: tcp6_input() 和 tcp_input() 函数	582	7.3.2 IPv6原始套接字	668
6.8.5 TCP控制输入: tcp6_ctlinput() 函数	587	7.3.3 辅助数据介绍	670
6.8.6 TCP用户请求	590	7.3.4 IPv6分组信息	672
6.9 IPv6上的UDP	596	7.3.5 处理IPv6扩展首部	674
6.9.1 ip6protosw{ }的IPv6上的UDP 实例	596	7.3.6 路径MTU的API	678
6.9.2 UDP输出: udp6_output() 函数	597	7.3.7 用于一批“r”命令的套接字 扩展	679
6.9.3 UDP输入: udp6_input() 函数	602	7.3.8 列表总结套接字选项	679
6.9.4 UDP控制输入: udp6_ctlinput() 函数	609	7.4 IPv6套接字API的内核实现	681
6.9.5 UDP用户请求的处理	612	7.4.1 代码介绍	682
6.10 原始IPv6	618	7.4.2 ip6_pktopts{ }结构	684
6.10.1 原始IPv6统计数据	619	7.4.3 IPv6套接字选项处理: ip6_ctloutput() 函数	688
6.10.2 原始IPv6输出: rip6_output() 函数	619	7.4.4 获取套接字选项: ip6_getpcbopt() 函数	703
6.10.3 原始IPv6输入: rip6_input() 函数	623	7.4.5 设置套接字选项与辅助数据	705
6.10.4 ICMPv6输入: icmp6_rip6_input() 函数	627	7.4.6 清理: ip6_freepcbopts() 函数	720
6.10.5 原始IPv6控制输入: rip6_ctlinput() 函数	632	7.4.7 IPv6多播套接字选项	721
6.10.6 原始IPv6控制输出: rip6_ctloutput() 函数	633	7.4.8 IPv6原始套接字选项: ip6_raw_ctloutput() 函数	731
6.10.7 原始IPv6用户请求处理	637	7.4.9 ICMPv6套接字选项: rip6_ctloutput() 函数	734
6.11 对IPv4映射的IPv6地址操作的 总结	644	7.4.10 传送输入信息: ip6_savecontrol() 函数	736
6.12 用netstat查看IPv6连接	648	7.5 套接字选项与辅助数据示例	743
6.13 配置IPv4映射的IPv6地址支持	650	7.5.1 发送路径示例	743
第7章 套接字API扩展	652	7.5.2 接收路径示例	745
7.1 简介	652	7.6 库函数的实现: libinet6	746
7.2 基本套接字API	652	7.6.1 inet_pton() 函数和inet_pton6() 函数	747
7.2.1 基本定义	652	7.6.2 inet_ntop() 函数和inet_ntop6() 函数	752
7.2.2 接口标识	653	7.6.3 getaddrinfo() 函数	756
7.2.3 在AF_INET6套接字上进行 IPv4通信	654	7.6.4 地址排序示例	781
7.2.4 地址和名字转换函数	656	7.6.5 freeaddrinfo() 函数	786
7.2.5 基本套接字选项	663	7.6.6 gai_strerror() 函数	786
7.3 高级套接字API——[RFC3542]	667	7.6.7 getnameinfo() 函数	787
		7.6.8 其他库函数	795
		参考文献	797
		索引	801