

国外计算机教材系列

ANSI C版

用 TCP/IP 进行网际互联

第二卷：设计、实现与内核 (第三版)

Internetworking With TCP/IP Vol II:

Design, Implementation, and Internals

Third Edition

[美] Douglas E. Comer 著
David L. Stevens

张娟 王海 黄述真 译
谢希仁 审校



电子工业出版社

Publishing House of Electronics Industry
URL: <http://www.phei.com.cn>

国外计算机科学教材系列

用 TCP/IP 进行网际互联

第二卷：设计、实现与内核

(第三版)

Internetworking with TCP/IP Vol II :
Design, Implementation, and Internals
Third Edition

[美] Douglas E. Comer 著
David L. Stevens

张 娟 王 海 黄述真 译

谢希仁 审校

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内 容 简 介

本书是一部有关计算机网络的经典教科书。它是目前美国大多数大学里所开设的计算机网络课程的主要参考书。目前国内外能见到的各种有关TCP/IP的书籍,其主要内容均出自本书。本书的特点是:强调原理、概念准确、深入浅出、内容丰富且新颖。

全书共分为三卷。第一卷介绍了TCP/IP的基本概念,第二卷在第一卷的基础上,进一步详细讨论了TCP/IP的实现过程,这一卷的突出特点是非常注重实际。本书作者利用程序代码实现了TCP/IP的每一个具体细节,并且所有的代码在书中均可找到,这对于读者深入了解并掌握TCP/IP的细节内容大有帮助。各章之后附有很多习题,内容全面且结合实际。全书最后还有三个附录,分别给出了过程调用交叉参考表、程序代码中用到的C数据结构交叉参考表以及Xinu函数和常量。本书可供计算机和通信专业的研究生、高年级本科生作为教科书和学习参考书,也可供各种从事科研的人员参考。

Authorized translation from the English language edition published by Prentice-Hall, Inc. Copyright © 1999.
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Prentice-Hall, Inc.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

用TCP/IP进行网际互联第二卷:设计、实现与内核(第三版)/(美)科默(Comer, D. E.)
史蒂文(Stevens, D. L.)著;张娟等译. —北京:电子工业出版社, 2001. 4
书名原文: Internetworking With TCP/IP Vol II: Design, Implementation, and Internals Third Edition
国外计算机科学教材系列
ISBN 7-5053-6630-0

I. 用... II. ①科... ②张... III. 计算机网络-通信协议-教材 IV. TN915.04

中国版本图书馆CIP数据核字(2001)第22857号

丛 书 名: 国外计算机科学教材系列

书 名: 用TCP/IP进行网际互联第二卷:设计、实现与内核(第三版)

原 书 名: Internetworking With TCP/IP Vol II: Design, Implementation, and Internals Third Edition

著 者: [美] Douglas E. Comer David L. Stevens

译 者: 张 娟 王 海 黄述真

审 校 者: 谢希仁

责任编辑: 马 岚

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 电话: 68279077

经 销: 各地新华书店

开 本: 787 × 1092 1/16 印张: 34 字数: 849千字

版 次: 2001年4月第1版 2001年4月第1次印刷

书 号: ISBN 7-5053-6630-0
TP · 3686

定 价: 52.00元

版权贸易合同登记号 图字: 01-2000-3470

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。

出版说明

随着 21 世纪的到来, 计算机技术的发展更加迅猛, 在各行各业的应用更加广泛, 越来越多的高等院校增设了有关计算机科学的课程内容, 或对现有计算机课程设置进行了适当调整, 以紧跟前沿技术。在这个教学体系和学科结构变革的大环境下, 对适合不同院系、不同专业、不同层次的教材的需求量与日俱增。此时, 如果能够借鉴、学习国外一流大学的先进教学体系, 引进具有先进性、实用性和权威性的国外一流大学计算机教材, 汲取其精华, 必能更好地促进中国高等院校教学的全面改革。

美国 Prentice Hall 出版公司是享誉世界的高校教材出版商, 自 1913 年成立以来, 一直致力于教材的出版, 所出版的计算机教材为美国众多大学采用, 其中有不少是专业领域中的经典名著, 已翻译成多种文字在世界各地的大学生中使用, 成为全人类的共同财富。许多蜚声世界的教授、学者都是该公司的资深作者, 如道格拉斯·科默 (Douglas E. Comer)、威廉·斯大林 (William Stallings) 等。早在 1997 年, 电子工业出版社就从 Prentice Hall 引进了一套计算机英文版专业教材, 并将其翻译出版, 同时定名为《国外计算机科学教材系列》(下称: 第一轮教材)。截至 2000 年 12 月, 该系列教材已出版 23 种, 深受读者欢迎, 被许多大学选为高年级学生和研究生教材或参考书。

4 年过去了, 已出版的教材中多数已经有了后续版本。因此, 我们开始设计新一轮教材 (第二轮教材) 的出版, 成立了由我国计算机界著名专家和教授组成的“教材出版委员会”, 并结合第一轮教材的使用情况和师生反馈意见, 组织了第二轮《国外计算机科学教材系列》出版工作。

第二轮教材的出版原则为:

1. 引进 Prentice Hall 出版公司 2000 年和 2001 年推出的新版教材, 作为替换版本。
2. 在著名高校教授的建议下, 除了从 Prentice Hall 新选了一些教材之外, 还从 McGraw-Hill 和 Addison Wesley Longman 等著名专业教材出版社、麻省理工学院出版社和剑桥大学出版社等著名大学出版社引进了一些经典教材, 作为增补版本。
3. 对于第一轮中无新版本的优秀教材, 我们将其作为延用版本, 直接进入第二轮使用。
4. 对于第一轮中翻译质量较好且无新版本的教材, 我们将其进行了修订, 也作为延用版本, 进入第二轮使用。

这次推出的教材覆盖学科范围广、领域宽、层次多, 既有本科专业课程教材, 也有研究生课程教材, 以适应不同院系、不同专业、不同层次的师生对教材的需求。广大师生可自由选择 and 自由组合使用。

按照计划, 本轮教材规划出版 37 种, 其中替换版本 8 种, 新增版本 14 种, 延用版本 15 种。教材内容涉及的学科方向包括网络与通信、操作系统、计算机组织与结构、算法与数据结构、数据库与信息处理、编程语言、图形图像与多媒体、软件工程等。本轮教材计划于 2001 年 7 月前全部出版。教材的使用年限平均为 3 年。我们还将陆续推出一些教材的参考课件, 希望能为授课老师提供帮助。

为了保证本轮教材的选题质量和翻译质量, 我们约请了清华大学、北京大学、北京航空航天大学、复旦大学、上海交通大学、南京大学、浙江大学、哈尔滨工业大学、华中科技大学、西安交通

大学、国防科学技术大学、解放军理工大学等著名高校的教授和骨干教师参与了本轮教材的选题、翻译和审校工作。他们中既有讲授同类教材的骨干教师和博士,也有积累了几十年教学经验的教授和博士生导师。

在本轮教材的选题、翻译和编辑加工过程中,为提高教材质量,我们做了大量细致的工作,包括:

1. 对于新选题和新版本进行了全面论证。
2. 对于延用版本,认真审查了前一版本教材,修改了其中的印刷错误。
3. 对于译者和编辑的选择,达到了专业对口。
4. 对于从英文原书中发现的错误,我们通过与作者联络、从网上下载勘误表等方式,一一做了修改。
5. 对于翻译、审校、编辑、排版、印刷质量进行了严格的审查把关。

通过这些工作,保证了本轮教材的质量较前一轮有明显的提高。相信读者一定能够从字里行间体会到我们的这些努力。

今后,我们将继续加强与各高校教师的密切联系,为广大师生引进更多的国外优秀教材和参考书,为我国计算机科学教学体系与国际教学体系的接轨做出努力。

由于我们对国际计算机科学、我国高校计算机教育的发展存在认识上的不足,在选题、翻译、出版等方面的工作中还有许多有待提高之处,恳请广大师生和读者提出批评和建议。

电子工业出版社
2001 年春

教材出版委员会

主 任	杨芙清	北京大学教授 中国科学院院士 北京大学信息与工程学部主任 北京大学软件工程研究所所长
委 员	王 珊	中国人民大学信息学院院长、教授
	胡道元	清华大学计算机科学与技术系教授 国际信息处理联合会通信系统中国代表
	钟玉琢	清华大学计算机科学与技术系教授 中国计算机学会多媒体专业委员会主任
	谢希仁	中国人民解放军理工大学教授 全军网络技术研究中心主任、博士生导师
	尤晋元	上海交通大学计算机科学与工程系教授 上海分布计算技术中心主任
	施伯乐	中国计算机学会常务理事、上海市计算机学会理事长 上海国际数据库研究中心主任、复旦大学教授
	邹 鹏	国防科学技术大学计算机学院教授、博士生导师 教育部计算机基础教学课程指导委员会副主任委员
	张昆藏	青岛大学信息工程学院教授

前言

很荣幸 Doug Comer 给我这个机会，让我能在他的这本书第三版付印时与诸位探讨一些个人的想法。在过去的十年期间，Internet 以惊人的速度发展着。Internet 上主机的数目从 1989 年的 100 000 台发展到 1998 年的 30 000 000 台。在本书第二版出版的时候，连接到 Internet 的网络大约有 26 000 个。到 1998 年，这个数字大约在 200 000 到 350 000 之间，这还不包括那些利用互联网技术但未连接到公共网络上的专用内联网。

除了规模上的变化，Internet 在应用上也发生了惊人的变化。Internet 技术与正在蓬勃发展的 WWW 作为一个完整的、不可缺少的系统已被人们所认可。而 WWW 在商业、学术和政府等部门都掀起了一场革命。在 WWW 上每天出现的“网页”可达 3.2 亿之多，并且还在不断涌现。有些学校大约四分之一的入学申请都是通过电子邮件或其他网页申请形式进行的。Dell 公司透露，他们通过网站每天销售价值大约 600 万美元的 PC 机。Amazon 网上书店以季度盈利达 6600 万美元而成为历史上发展最快的公司，这也是第一个在一年之内销售额超过 2.5 亿美元的公司。

目前，至少有两千多家无线电台将其音频服务推向了 Internet，许多网站也开始提供音频甚至低质量的视频服务。随着访问速度提高到兆比特的范围以及主干网容量的增加，视频服务质量也将得到提高。

随着 1996 年网络电视和 1997 年 Nokia 推出的支持网络功能的手机的问世，其他支持网络功能的产品也不断地涌现。更多的实用产品实现了相似的网络功能，例如水暖加热器既可以由家用电脑来控制，也可以由电力公司来控制，以满足适度的调峰需求。尽管智能代理还没有成为主流，但是 XML 近来的发展也表明 Internet 上有关事务处理的应用正在急剧增多。SML 的标准“文档”表示方式以及一致的常用解释格式构成了一种可传输对象，该对象已形成所有商业、金融事务、数据库事务及其他需要标准表示方式和解释格式的交易的基础。

再回头看看电话产品的发展，无论传统的服务商还是现代企业都在将支持网络的电话产品推向市场。“Soft PBX”系统利用 LAN 和 Internet 实现传统的专用交换机的功能，基于微处理器的电话机也正在改变远程通信的经济情况和可操作性。具有 IP 功能的传真机已经出现。网关把原有的模拟世界与现在的 Internet 世界联系在一起，起着非常重要的作用。

对 Internet 服务需求量的增长速度赶上甚至超过了网络本身的增长。主干网络的速度达到 622Mb/s 已是很平常的事。利用硬件 IP 交换机的新一代路由器可望以 10Gb/s (OC192) 以上的速度处理 Internet 分组。下一个挑战是利用单模光纤处理每秒兆兆字节的信息量。

网络的安全性过去总是处于后台操作状态，随着更多领域依赖于 Internet，安全性已逐渐走向前台。防火墙技术、端-端加密、密钥管理、证书系统和鉴别系统已成为成功管理 Internet 的关键因素。

再向未来展望，网络协议和结构已经朝着星际互联的方向发展，也许会形成和互联网一样的形式。域名系统如果还存在，还要考虑将不同的行星计入命名体系。“互联的 Internet”协议将能够处理更高的延迟，传统的 TCP 概念也将被更多的单向过程取代。未来的网际遨游者回看这十年，一定会觉得这是个充满挑战但技术相对落后的时代。在他们看来，这段前言中的观点可能已经过时又难以理解，因为他们所面对的是经历了另外二十年变革的 Internet。

Vinton Cerf
Camelot
Northern Virginia

序 言

用TCP/IP进行网际互联第二卷提供了第一卷所没有包含的关于TCP/IP协议的一些细节问题。第二卷如同将TCP/IP置于放大镜下，考察每个协议的具体细节。它讨论了协议的实现方案，并着重于介绍协议软件的内部机制。第三版包含了对某些协议的修改和更新的内容。其中的代码改用C++语言的ANSI标准C子集，包含了函数原型和参数的声明。另外还纠正了一些错误。我们将SNMP更新为SNMPv2，其中包括替换了地址转换表，增加了UDP listener 列表。在本书的最后，还增加了附录2，给出了书中代码所用到的主要数据结构的声明及变量的交叉参考表，并扩充了附录1中的过程调用交叉参考的内容。

书中的范例代码用Gnu C++编译器在Intel体系平台编译通过，也在Pentium™系统上经过测试。所有的代码均可在以下网址得到：

<ftp://ftp.cs.purdue.edu/pub/comer/TCPIP-vol2.dist.tar.Z>

虽然本书受到版权保护，但书中的代码可供读者使用，而且已经在许多商业产品中实现。这些代码使用的惟一限制是不得在公开出版物中出版。

我们鼓励读者利用计算机工具来查看、修改、编译和测试这些代码。事实上，尽管附录1和附录2中提供了定位代码条目的有效途径，但在查看大段代码时，UNIX grep 程序的价值更是不可估量。

对于各种正式的协议规范，以及对协议的实现和使用的讨论，可参见请求评论文档（RFC）。尽管一些RFC文档对初学者来讲难以理解，但这些文档是信息详尽的权威性资源，没有哪个作者能够做到在自己编写的书中包含RFC文档中的所有内容。尽管RFC文档涉及了每一个协议，但有时它们对协议之间的交互问题并未加以说明。例如，选路信息协议（RIP或OSPF）之类的选路协议规定了网关如何将路由置入IP路由表中，以及如何将表中的路由通知其他网关。RIP还规定路由必须设立定时机制，一旦某条路由超时，就将其删除。但是，我们在RFC文档中并不容易看出RIP和其他协议之间是如何交互作用的，随之而来的问题是：“路由超时机制将如何影响路由表中那些由ICMP设置的路由呢？”我们可能还会考虑这样一个问题：“当RIP更新路由时，应不应该推翻那些由管理员直接输入的路由呢？”

为了有助于解释协议之间的交互作用，并确保我们的方案能协调工作，我们设计并构造了一个工作系统，作为全书的中心范例。该系统提供了TCP/IP协议族中的大部分协议，包括：TCP、IP、ICMP、IGMP、UDP、ARP、RIP、SNMPv2以及OSPF的主要部分。另外，该系统还有一个finger服务的客户和服务器的范例。由于本书包括了每一个协议的程序代码，读者可以研究其实现方法并了解其内部结构。最重要的是，由于范例系统将所有协议软件集成为一个工作整体，读者可以清楚地了解协议之间的交互作用。

范例的程序代码试图做到一方面遵守协议标准，另一方面引入一些新的思想。例如，我们的TCP程序代码中包含了“糊涂窗口预防”技术、Jacobson-Karels的“慢启动”和“拥塞预防”等优化技术，诸如此类的性能可能在商业实现中被忽略。但同时我们也清楚地认识到商业领域并不总是遵从已经公布的标准，因此我们也努力将系统调整到能够在现实环境中使用。例如，程序代码中包含了

一个配置参数，使得它既可以采纳 Internet 标准，也可以采纳 BSD UNIX 中“TCP 紧急数据指针”的实现方法。

我们并不认为书中所提供的程序代码都是准确无误的，甚至不能断言它肯定比其他实现方法要好。事实上，经过多年使用，我们仍然在不断完善这套软件，同时，也希望读者跟我们一起继续改进它。

本书可以作为网络工程人员的高级教程或者作为研究生教材使用。在作为本科教程使用时，应将重点放在前几章，而忽略有关 OSPF、SNMP 和 RIP 这几章内容。研究生可能会在有关 TCP 的章节中发现一些最为有趣和最难理解的概念。为保证其高性能而采用的自适应重发和相关的试探法尤为重要，应当仔细加以研究。纵观全书，绝大部分习题都向大家提示了其他可选择的实现方案及其大致情况，这些内容并不要求死记硬背，学生们可能需要阅读本书以外的其他资料，才能解答这些习题。

正如任何耗费甚巨的工程一样，本书中包含了许多人的心血，对此我们表示衷心的感谢。作者之一 David Stevens 完成了大部分软件的编制工作，其中包括一个完整的 TCP 版本。Shawn Ostermann 为本书付出了许多努力。Shawn 将 TCP/IP 代码集成到 Xinu 版本 8 中，并将其从最初的 Sun 3 平台移植到 DECstation 3100 上。在这一版本中，他还对上一版中由 Vic Norman 编写的 SNMP 软件进行了大量的修改，以符合第二版的标准并有助于代码的调试。John Lin 对书中的技术细节进行了校对，纠正了一些错误。

珀杜（Purdue）大学的网际互联研究小组（Internetworking Research Group）的许多成员对程序代码的早期版本做过很多贡献。Andy Muckelbauer 和 Steve Chapin 建立了一个 UNIX 兼容库，并与 Shawn Ostermann 和 Scott Mark 合作，使用 TCP 代码运行一个 X window 服务器。他们对 TCP 做了大量的测试工作，并指出其性能上存在的几个问题。Scott M. Ballew 参与了一些软件的开发工作。

我的妻子 Christine 对本书手稿进行了编辑，并提出了许多建议。最后，我们感谢珀杜大学的计算机科学系和计算中心所提供的帮助。

Douglas E. Comer

David L. Stevens

目 录

第 1 章 引言与概述	1
1.1 TCP/IP 协议	1
1.2 了解细节的必要性	1
1.3 协议间交互作用的复杂性	1
1.4 本书采用的方法	2
1.5 研究代码的重要性	2
1.6 Xinu 操作系统	2
1.7 本书其余部分的组织	3
1.8 小结	3
深入研究	4
第 2 章 操作系统中的 TCP/IP 软件结构	5
2.1 引言	5
2.2 进程的概念	5
2.3 进程的优先级	6
2.4 进程的同步通信	6
2.5 进程间通信	8
2.5.1 端口	8
2.5.2 消息传送	9
2.6 设备驱动程序和输入、输出程序	9
2.7 网络的输入和中断	10
2.8 向高层协议传递分组	11
2.9 IP 协议与传输协议之间的数据报传递	12
2.9.1 将传入的数据报发送给 TCP	12
2.9.2 将传入的数据报发送给 UDP	12
2.10 向应用程序的传递操作	13
2.11 输出时的信息流	13
2.12 从 TCP 经过 IP 到网络输出	14
2.13 UDP 输出处理	15
2.14 小结	15
深入研究	16
习题	16
第 3 章 网络接口层	18
3.1 引言	18

3.2 网络接口抽象模型	18
3.2.1 接口数据结构	18
3.3 以太网的基本定义	21
3.3.1 应用中的统计数据	24
3.4 接口的逻辑状态	24
3.5 本地主机接口	24
3.6 缓冲区管理	25
3.6.1 大缓冲区方案	25
3.6.2 链表方案 (mbuf)	26
3.6.3 方案举例	26
3.6.4 有关缓冲区的其他议题	26
3.7 传入分组的多路分解	27
3.8 小结	28
深入研究	29
习题	29
第 4 章 地址的发现及绑定 (ARP)	30
4.1 引言	30
4.2 ARP 软件在理论上的结构	30
4.3 ARP 设计方案举例	30
4.4 ARP 高速缓存的数据结构	31
4.5 ARP 输出处理	34
4.5.1 搜索 ARP 高速缓存	34
4.5.2 ARP 请求分组的广播	35
4.5.3 输出过程	36
4.6 ARP 输入处理	38
4.6.1 向表中增加已转换的表项	38
4.6.2 发送等待发送的分组	39
4.6.3 ARP 输入过程	40
4.7 ARP 高速缓存的管理	42
4.7.1 高速缓存表项的分配	42
4.7.2 高速缓存的定期维护管理	43
4.7.3 释放队列中的分组	44
4.8 ARP 初始化	45
4.9 ARP 参数配置	46
4.10 小结	46
深入研究	47
习题	47
第 5 章 IP: 软件的总体结构	48
5.1 引言	48

5.2 中心环节	48
5.3 IP 软件设计思想	48
5.4 IP 软件结构和数据报流程	49
5.4.1 选择传入数据报的策略	49
5.4.2 允许 IP 进程被阻塞	51
5.4.3 IP 使用的常量的定义	54
5.4.4 校验和的计算	57
5.4.5 处理定向广播	57
5.4.6 识别一个广播地址	59
5.5 IP 首部中的字节顺序	60
5.6 向 IP 发送数据报	61
5.6.1 发送本地生成的数据报	61
5.6.2 发送传入数据报	63
5.7 表格的维护	63
5.8 小结	65
深入研究	65
习题	65
第 6 章 IP: 选路表和选路算法	67
6.1 引言	67
6.2 路由维护和查找	67
6.3 选路表结构	67
6.4 选路表数据结构	68
6.5 路由的生成源及保持时间	70
6.6 为数据报选择路由	70
6.6.1 实用过程	70
6.6.2 获得一个路由	73
6.6.3 数据结构初始化	74
6.7 选路表的定期维护	75
6.7.1 增加路由	77
6.7.2 删除路由	80
6.8 IP 选项处理	82
6.9 小结	83
深入研究	83
习题	84
第 7 章 IP: 分片与重组	85
7.1 引言	85
7.2 数据报的分片	85
7.2.1 为一个数据报片再次分片	85
7.3 分片的实现	85

7.3.1	发送一个数据报片	87
7.3.2	复制数据报首部	89
7.4	数据报的重组	90
7.4.1	数据结构	90
7.4.2	互斥操作	91
7.4.3	在链表中加入一个数据报片	91
7.4.4	溢出时的丢弃	93
7.4.5	测试一个完整的数据报	94
7.4.6	将数据报片组装成完整的数据报	96
7.5	数据报片链表的维护管理	97
7.6	初始化	99
7.7	小结	99
	深入研究	100
	习题	100
第 8 章	IP：差错处理 (ICMP)	101
8.1	引言	101
8.2	ICMP 报文格式	101
8.3	ICMP 报文的实现	101
8.4	传入 ICMP 报文的处理	103
8.5	ICMP 重定向报文的处理	105
8.6	设置子网掩码	107
8.7	为一个 ICMP 分组选择源地址	108
8.8	生成 ICMP 差错报文	109
8.9	避免出现关于差错报文的差错报文	111
8.10	为 ICMP 报文分配缓冲区	112
8.11	ICMP 报文中的数据部分	114
8.12	ICMP 重定向报文的生成	116
8.13	小结	117
	深入研究	117
	习题	117
第 9 章	IP：组播处理 (IGMP)	119
9.1	引言	119
9.2	维护组播主机群的成员信息	119
9.3	主机群表	119
9.4	查找一个主机群	121
9.5	向主机群表中增加一个表项	122
9.6	为一个组播地址设置网络接口	124
9.7	IP 组播地址和硬件组播地址之间的转换	125
9.8	从主机群表中删除一个组播地址	126

9.9 加入一个主机群	127
9.10 维持与一个组播路由器的联系	129
9.11 IGMP 成员关系报告的实现	130
9.12 计算随机时延	131
9.13 发送 IGMP 报告的进程	132
9.14 处理输入的 IGMP 报文	133
9.15 脱离主机群	134
9.16 IGMP 数据结构的初始化	136
9.17 小结	137
深入研究	137
习题	137
第 10 章 UDP: 用户数据报	138
10.1 引言	138
10.2 UDP 端口和多路分解处理	138
10.2.1 成对通信使用的端口	138
10.2.2 多对一通信使用的端口	138
10.2.3 操作模式	139
10.2.4 多路分解处理中的细节问题	139
10.3 UDP 的输入处理	141
10.3.1 UDP 数据结构的说明	141
10.3.2 传入数据报队列的说明	142
10.3.3 UDP 端口号与队列的映射	144
10.3.4 分配空闲队列	144
10.3.5 网络字节顺序与本机字节顺序之间的相互转换	145
10.3.6 处理一个已到达的数据报	146
10.3.7 UDP 校验和的计算	148
10.4 UDP 输出的处理	149
10.4.1 一个 UDP 数据报的发送	150
10.5 小结	151
深入研究	152
习题	152
第 11 章 TCP: 数据结构和输入处理	153
11.1 引言	153
11.2 TCP 软件概览	153
11.3 传输控制块	153
11.4 TCP 报文段格式	158
11.5 序列空间中的比较	159
11.6 TCP 有限状态机	160
11.7 状态变迁举例	160

11.8 有限状态机的说明	161
11.9 TCB 的分配及初始化	163
11.9.1 分配一个 TCB	163
11.9.2 释放一个 TCB	164
11.10 有限状态机的实现	165
11.11 处理一个输入报文段	165
11.11.1 将 TCP 首部转换为本地字节顺序	167
11.11.2 计算 TCP 的校验和	168
11.11.3 为报文段查找对应的 TCB	169
11.11.4 检查报文段的有效性	171
11.11.5 为当前状态选择一个过程	172
11.12 小结	173
深入研究	173
习题	173
第 12 章 TCP：有限状态机的实现	175
12.1 引言	175
12.2 CLOSED 状态处理	175
12.3 从容关闭	175
12.4 关闭后的延迟计时	176
12.5 TIME-WAIT 状态处理	177
12.6 CLOSING 状态处理	178
12.7 FIN-WAIT-2 状态处理	179
12.8 FIN-WAIT-1 状态处理	180
12.9 CLOSE-WAIT 状态处理	181
12.10 LAST-ACK 状态处理	182
12.11 ESTABLISHED 状态处理	183
12.12 处理报文段中的紧急数据	184
12.13 处理报文段中的其他数据	186
12.14 经常注意已接收的八位组	188
12.15 终止一个 TCP 连接	190
12.16 建立 TCP 连接	191
12.17 初始化 TCB	191
12.18 SYN-SENT 状态处理	193
12.19 SYN-RECEIVED 状态处理	194
12.20 LISTEN 状态处理	196
12.21 为一个新 TCB 初始化窗口变量	197
12.22 小结	199
深入研究	199
习题	199

第 13 章 TCP: 输出处理	200
13.1 引言	200
13.2 TCP 输出控制的复杂性	200
13.3 TCP 输出的四种状态	200
13.4 作为一个进程的 TCP 输出	201
13.5 TCP 输出报文	201
13.6 对输出状态和 TCB 编号的编码	202
13.7 TCP 输出进程的实现	202
13.8 互斥操作	203
13.9 IDLE 状态的实现	203
13.10 PERSIST 状态的实现	204
13.11 TRANSMIT 状态的实现	205
13.12 RETRANSMIT (重发) 状态的实现	206
13.13 发送一个报文段	206
13.14 计算 TCP 数据长度	210
13.15 序号计算	210
13.16 其他 TCP 过程	211
13.16.1 发送复位	211
13.16.2 转换成网络字节顺序	213
13.16.3 等待输出缓冲空间	213
13.16.4 唤醒等待 TCB 的进程	214
13.16.5 选择初始序号	216
13.17 小结	217
深入研究	217
习题	217
第 14 章 定时器管理	218
14.1 引言	218
14.2 定时事件的通用数据结构	218
14.3 TCP 事件使用的数据结构	219
14.4 定时器、事件和报文	220
14.5 TCP 定时器进程	220
14.6 删除 TCP 定时器事件	222
14.7 删除一个 TCB 的所有事件	223
14.8 确定出现一个事件的尚需时间	224
14.9 插入 TCP 定时器事件	225
14.10 启动无时延的 TCP 输出	227
14.11 小结	227
深入研究	228
习题	228

第 15 章 流量控制和自适应重发	229
15.1 引言	229
15.2 自适应重发中的难题	229
15.3 自适应重发的调整	229
15.4 重发定时器和退避	230
15.4.1 Karn 算法	230
15.4.2 重发输出状态的处理	230
15.5 基于窗口的流量控制	231
15.5.1 糊涂窗口综合症	232
15.5.2 接收方预防糊涂窗口	232
15.5.3 零窗口之后的性能优化	233
15.5.4 调整发送方的窗口	233
15.6 最大报文段长度的计算	235
15.6.1 发送方的最大报文段长度	235
15.6.2 选项处理	236
15.6.3 通告一个最大输入报文段长度	238
15.7 网络拥塞预防与控制	239
15.7.1 成倍递减法	239
15.8 慢启动和拥塞预防	239
15.8.1 慢启动	239
15.8.2 超过上限后减速递增	240
15.8.3 递增拥塞窗口尺寸的实现	240
15.9 平均往返时间估值及超时	242
15.9.1 一种快速平均值更新算法	242
15.9.2 传入确认的处理	243
15.9.3 为窗口外的数据创建确认报文段	245
15.9.4 接收到一个确认后改变输出状态	246
15.10 技巧和注意事项	247
15.11 小结	247
深入研究	248
习题	248
第 16 章 紧急数据处理和推功能	250
16.1 引言	250
16.2 带外信令	250
16.3 紧急数据	250
16.4 标准的解释	250
16.4.1 带外数据的解释法	251
16.4.2 数据标记解释法	252
16.5 为 Berkeley 紧急指针解释法而进行的配置	252
16.6 通知应用程序	253