

国外计算机科学教材系列

用 TCP/IP 进行网际互连

第 2 卷：设计、实现和内部构成（第 2 版）

Internetworking With TCP/IP

Vol II : Design, Implementation, and Internals (Second Edition)

DOUGLAS E. COMER
DAVID L. STEVENS 著

张 娟 王 海 译

谢希仁 校



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY



PRENTICE HALL 出版公司

内 容 提 要

本书是一部有关计算机网络的经典教科书(现在是第二版)。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 2 卷是在第 1 卷介绍了 TCP/IP 的基本概念的基础上,进一步详细讨论了 TCP/IP 的实现细节。这一卷有一个突出的特点,就是非常实际。作者把每一个细节都用程序实现了,并且把每一部分的程序都写在书上。这对读者深入掌握 TCP/IP 的细节将会有很大的帮助。各章之后还附有很多很好的练习题。书后还有两个附录,给出了过程调用交叉参考表和程序代码中使用到的 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。

©1994 by Prentice-Hall, Inc.

本书中文简体版由电子工业出版社和美国 Prentice Hall 出版公司合作出版。未经许可,不得以任何手段和形式复制或抄袭本书内容。版权所有,侵权必究。

丛 书 名: 国外计算机科学教材系列

原 书 名: Internetworking With TCP/IP Vol. II: Design, Implementation, and Internals (Second Edition)

书 名: 用 TCP/IP 进行网际互连第 2 卷:设计、实现和内部构成(第 2 版)

著 者: DOUGLAS E. COMER & DAVID L. STEVENS

译 者: 张 娟 王 海

审 校 者: 谢希仁

责任编辑: 范官清

特约编辑: 苏子栋

印 刷 者: 北京市天竺颖华印刷厂

出版发行:电子工业出版社出版、发行 URL:<http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036 发行部电话(68214070)

经 销:各地新华书店经销

开 本:787×1092 1/16 印张:33.75 字数:810 千字

版 次:1998 年 7 月第 1 版 1998 年 7 月第 1 次印刷

印 数:7000 册

书 号:ISBN 7-5053-4604-0/TP.2185

定 价:52.00 元

著作权合同登记号 图字:01-97-1868

凡购买电子工业出版社的图书,如有缺页、倒页、脱页者,本社发行部负责调换

版权所有·翻印必究

出版说明

计算机科学的迅速发展是 20 世纪科学发展史上最伟大的事件之一。从 1946 年第一台笨重而体积庞大的计算机的发明至今, 仅仅半个多世纪, 计算机已经变得小巧无比却又能力非凡。它的应用已经渗透到了社会的各个方面, 成为当今所谓的信息社会的最显著的特征。

处于世纪之交科技进步的大潮中, 我国正在加强计算机科学的高等教育, 着眼于为下一世纪培养高素质的计算机人才, 以适应信息社会加速度发展的需要。当前, 全国各类高等院校已经或计划在各专业基础课程规划中增加计算机科学的课程内容, 而作为与计算机科学密切相关的计算机、通信、信息等专业, 更是在酝酿着教学的全面革新, 以期规划出一整套面向 21 世纪的、具有中国高校计算机教育特色的课程计划和教材体系。值此, 我们不妨借鉴并引进国外具有先进性、实用性和权威性的大学计算机教材, 洋为中用, 以更好地服务于国内的高校教育。

美国 Prentice Hall 出版公司是享誉世界的高校教材出版商, 自 1913 年公司成立以来, 即致力于教育图书的出版。它所出版的计算机教材在美国为众多大学所采用, 其中有不少是专业领域中的经典名著。许多蜚声世界的教授学者成为该公司的资深作者, 如: 道格拉斯·科默 (Douglas Comer), 安德鲁·坦尼伯姆 (Andrew Tanenbaum), 威廉·斯大林 (William Stallings)……几十年来, 他们的著作教育了一批批不同肤色的莘莘学子, 使这些教材同时也成为全人类的共同财富。

为了保证本系列教材翻译出版的质量, 电子工业出版社和 Prentice Hall 出版公司共同约请北京地区的清华大学、北京大学、北京航空航天大学, 上海地区的上海交通大学、复旦大学, 南京地区的南京大学、解放军通信工程学院等全国著名的高等院校的教学第一线的几十位教师参加翻译工作。这中间有正在讲授同类教材的年轻教师和博士, 有积累了几十年教学经验的教授和博士生导师, 还有我国著名的计算机科学家。他们的辛勤劳动保证了本系列丛书得以高质量地出版面世。

如此大规模地引进计算机科学系列教材, 在我们还是第一次。除缺乏经验之外, 还由于我们对计算机科学的发展, 对中国高校计算机教育特点认识的不足, 致使在选题确定、翻译、出版等工作中, 肯定存在许多遗憾和不足之处, 恳请广大师生和其他读者提出批评、建议。

电子工业出版社

URL: <http://www.phei.com.cn>

Prentice Hall 出版公司

URL: <http://www.prenhall.com>

前 言

这是一本受到广泛关注的教科书,很荣幸能借其再版之机,与诸位探讨一些我个人的想法。从第一版发行到第二版面世,这期间 Internet 的发展速度可以用指数计算,同样呈指数增长趋势的是人们对它的了解与关心程度。诸位也许会对一个事实感兴趣,那就是仅限制在狭小的地理区域内的专用网与连接全球 Internet 的网络相比要多出一倍。就在我撰写本文时,连接全球 Internet 的网络有 26000 个,而与此同时有一倍以上的网络使用了 TCP/IP,但它们都只是专用网。

TCP/IP 及其前一代协议,即 ARPANET 主机到主机协议,其最初设计原则都是基于某种观念,认为对于那些分布在 Internet 上的计算机来说,绝大多数的通信应当发生在不同计算机的进程之间。更进一步说,一个主机中可能有多个进程,而其中每个进程又可能与互连网络中其他计算机上的一个或更多个进程相互联系。我们有充分的证据表明,这种模型已变得越来越重要。用户只要对一些商业新闻媒体,或者,事实上只要对一些大众新闻媒体稍加留意就会发现,智能代理(intelligent agent)已经成为 Internet 环境中的一个主要的新成员。对这些进程的控制,以及对进程之间的交互作用的支持,对进程与 Internet 上分布式数据库之间的交互作用的支持,都有可能随着网络用户的增加而带来可观的数据流量。

随着 Internet 继续保持其非同凡响的发展速度,TCP/IP 在其结构问题上,面临一些非常严峻的挑战。处理以 100% 的速度逐年增长的网络数量将给路由选择系统带来很大压力。从长远观点来看,地址资源本身就会被耗竭。目前,选择 B 类地址空间已经比较紧张,而使用无类型域间路由选择(CIDR ;Classless Inter-Domain Routing)技术仅仅是一个权宜之计。我们需要新版的 IP 协议以支持更大的地址空间,并改善对一般性等级安排问题的处理。

与此同时,在一些新兴的、雄心勃勃的应用中,已经要求 Internet 能够支持规模越来越大的分组语音和图像通信。许多人希望新的信元式交换技术,例如异步传输模式(ATM)能够被用来支持大量的此类应用。

安全性也是重点问题。特别是在目前网络不断地扩展,更多地面向商业应用的情况下,安全性显得更加重要。既要寻求一个统一的设计方案来支持 Internet,同时还要能够与世界范围内各种不同的技术有效地合作,而这些技术中有一些又是被控制出口的技术,这也是我们将要面临的主要挑战之一。

在这样一个有利因素聚集、更新速度快、不断变革的环境中,理解并深入研究 Internet 协议的基本功能是很有必要的,特别是为了迎接令人振奋的未来世界,为了使我们能够正确引导这一正在流行的标准进一步发展,并开创它在 21 世纪应用的天地。愿本书能够让你深入了解并更加喜爱 Internet 及其中的技术。

弗吉尼亚州 Annandale

Internet 协会主席

Vinton Cerf

作者介绍

Douglas Comer 博士是普渡大学(Purdue University)的教授, 讲授操作系统和计算机网络的课程。他已撰写了大量的科研论文和教科书, 目前负责一些网络方面的研究课题。早在七十年代后期, 他就参与了 TCP/IP 和互连网的研究, 并成为世界公认的权威。由他设计实现了 X25NET 和 Cypress 网络, 以及 Xinu 操作系统。Douglas Comer 博士是普渡大学网际互连研究小组的负责人, Journal of Internetworking 杂志的主编, Software - Practice and Experience 杂志的编辑, 以及 Internet 体系结构委员会的前任委员。

David Stevens 是普渡大学计算中心的程序员, 自 1986 年以来一直致力于 UNIX 和互连网的开发工作。目前他是普渡大学数据网研究小组的成员, 曾与他人合作完成了多本有关计算机网络的教科书, 并且是普渡大学网际互连研究组的成员。他获得普渡大学计算机科学硕士学位。

序 言

为了响应那些不满足于第一卷所介绍的内容,希望知道更多有关 TCP/IP 协议细节的读者们的要求,我们出版了 *Internetworking With TCP/IP* 第二卷的第一版。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。此次再版,对某些协议进行了修改,并增加了两章新的内容,分别讨论为 IP 组播处理而设立的 Internet 组管理协议(IGMP)和开放最短路径优先(OSPF)路由选择协议。为了更清楚地说明对数据掩码的解释,我们对 TCP 紧急数据的实现做了一些改动,本书将会讨论其结果。

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,可参见要求评论文档(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涵 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,像选路信息协议(RIP)这样的路由选择协议规定了网关如何将路由安装到 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并非由 RIP 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, SNMP, 以及 OSPF 的主要部分。另外,还具有一个 finger 服务的客户机和服务器范例。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及理解其内部结构。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“防拥塞”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 中的“TCP 紧急数据指针”实现方法。

我们并不认为书中所提供的程序代码都是正确无误的,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。为了方便读者,出版商已同意将书中提供的所有程序代码制作成机器可阅读的拷贝,用户可以利用计算机工具来查看、修改和测试它。读者可通过匿名 FTP 访问 `ftp.cs.purdue.edu` 计算机上的 `pub/comer/v2.dist.tar.Z` 文件得到该拷贝。

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性尤为重要,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,它们很少要求死记硬背,因此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。

就像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。Victor Norman 编写了 SNMP 软件,并对其做了数次修正。Shawn Ostermann 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。Andy Muckelbauer 和 Steve Chapin 建立了一个 UNIX 兼容库,并与 Shawn Ostermann, 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,并不辞辛苦地审核了所有章节及程序代码。珀杜(Purdue)大学的网际互连研究小组(Internetworking Research Group)的许多成员对程序代码的早期版本做过很多贡献。Christine Comer 阅读了本书手稿并提出许多建议。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 引言和概述	(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	(3)
1.7 本书其余部分的组织	(3)
1.8 小结	(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 进程通信处理	(6)
2.5 进程间通信	(8)
2.6 设备驱动程序和输入、输出程序	(10)
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)
2.10 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	(16)
第 3 章 网络接口层	(20)
3.1 引言	(20)
3.2 网络接口抽象模型	(20)
3.3 接口的逻辑状态	(23)
3.4 本地主机接口	(23)
3.5 缓冲区管理	(24)
3.6 输入分组的分路	(26)

3.7 小结.....	(28)
第 4 章 地址的发现及绑定 (ARP)	(29)
4.1 引言.....	(29)
4.2 ARP 软件在理论上的结构	(29)
4.3 ARP 设计方案举例	(29)
4.4 ARP 高速缓存的数据结构	(31)
4.5 ARP 输出处理	(33)
4.6 ARP 输入处理	(37)
4.7 ARP 高速缓存的管理	(41)
4.8 ARP 初始化	(45)
4.9 ARP 参数配置	(46)
4.10 小结	(46)
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)
5.5 IP 首部中的字节顺序	(61)
5.6 向 IP 发送一个数据报	(62)
5.7 表格的维护.....	(65)
5.8 小结.....	(66)
第 6 章 IP:选路表和选路算法	(68)
6.1 引言.....	(68)
6.2 路由维护和查找.....	(68)
6.3 选路表结构.....	(68)
6.4 选路表数据结构.....	(69)
6.5 路由的生成源及保持时间.....	(71)
6.6 为数据报选择路由.....	(71)
6.7 选路表的定期维护.....	(77)
6.8 IP 选项处理	(84)
6.9 小结.....	(85)
第 7 章 IP:分片与重组	(87)
7.1 引言.....	(87)
7.2 数据报的分片.....	(87)
7.3 分片的实现.....	(88)

7.4	数据报的重组	(92)
7.5	数据报片链表的维护管理	(100)
7.6	初始化	(102)
7.7	小结	(102)
第 8 章	IP:差错处理 (ICMP)	(104)
8.1	引言	(104)
8.2	ICMP 报文格式	(104)
8.3	ICMP 报文的实现	(104)
8.4	输入 ICMP 报文的处理	(106)
8.5	ICMP 改变路由报文的处理	(109)
8.6	设置子网掩码	(110)
8.7	为一个 ICMP 分组选择源地址	(112)
8.8	生成 ICMP 差错报文	(113)
8.9	避免出现错误的差错报文	(115)
8.10	为 ICMP 报文分配缓冲区	(116)
8.11	ICMP 报文中的数据部分	(118)
8.12	ICMP 改变路由报文的生成	(120)
8.13	小结	(121)
第 9 章	IP:组播处理 (IGMP)	(123)
9.1	引言	(123)
9.2	维护组播主机群的成员信息	(123)
9.3	主机群表	(123)
9.4	查找一个主机群	(125)
9.5	向主机群表中增加一个表项	(126)
9.6	为一个组播地址设置网络接口	(128)
9.7	IP 组播地址和硬件组播地址之间的转换	(129)
9.8	从主机群表中删除一个组播地址	(131)
9.9	加入一个主机群	(132)
9.10	维持与一个组播路由器的联系	(133)
9.11	IGMP 成员关系报告的实现	(135)
9.12	计算随机时延	(136)
9.13	发送 IGMP 报告的进程	(137)
9.14	处理输入的 IGMP 报文	(138)
9.15	脱离主机群	(139)
9.16	IGMP 数据结构初始化	(141)
9.17	小结	(142)

第 10 章 UDP:用户数据报	(143)
10.1 引言	(143)
10.2 UDP 端口和去复用处理	(143)
10.3 UDP	(146)
10.4 UDP 输出的处理	(155)
10.5 小结	(158)
第 11 章 TCP:数据结构和输入处理	(160)
11.1 引言	(160)
11.2 TCP 软件概览	(160)
11.3 传输控制块	(160)
11.4 TCP 报文段格式	(165)
11.5 序列空间中的比较	(166)
11.6 TCP 有限状态机	(167)
11.7 状态变迁举例	(167)
11.8 有限状态机的说明	(169)
11.9 TCB 的分配及初始化	(170)
11.10 有限状态机的实现	(172)
11.11 处理一个输入报文段	(173)
11.12 小结	(181)
第 12 章 TCP:有限状态机的实现	(183)
12.1 引言	(183)
12.2 CLOSED 状态处理	(183)
12.3 文明关闭	(183)
12.4 关闭后的延迟计时	(184)
12.5 TIME-WAIT 状态处理	(185)
12.6 CLOSING 状态处理	(186)
12.7 FIN-WAIT-2 状态处理	(187)
12.8 FIN-WAIT-1 状态处理	(188)
12.9 CLOSE-WAIT 状态处理	(190)
12.10 LAST-ACK 状态处理	(191)
12.11 ESTABLISHED 状态处理	(192)
12.12 处理报文段中的紧急数据	(193)
12.13 处理报文段中的其他数据	(195)
12.14 经常注意已接收的八位组	(197)
12.15 中止一个 TCP 连接	(200)
12.16 建立 TCP 连接	(200)

12.17	初始化 TCB	(201)
12.18	SYN-SENT 状态处理	(202)
12.19	SYN-RECEIVED 状态处理	(203)
12.20	LISTEN 状态处理	(205)
12.21	为一个新 TCB 初始化窗口变量	(207)
12.22	小结	(209)
第 13 章	TCP:输出处理	(210)
13.1	引言	(210)
13.2	TCP 输出控制的复杂性	(210)
13.3	TCP 输出的四个状态	(210)
13.4	作为一个进程的 TCP 输出	(211)
13.5	TCP 输出消息	(212)
13.6	对输出状态和 TCB 号的编码	(212)
13.7	TCP 输出进程的实现	(212)
13.8	互斥操作	(213)
13.9	IDLE 状态的实现	(214)
13.10	PERSIST 状态的实现	(214)
13.11	TRANSMIT 状态的实现	(215)
13.12	RETRANSMIT 状态的实现	(217)
13.13	发送一个报文段	(217)
13.14	计算 TCP 数据长度	(220)
13.15	序号计算	(221)
13.16	其他 TCP 过程	(222)
13.17	小结	(228)
第 14 章	定时器管理	(230)
14.1	引言	(230)
14.2	定时事件的通用数据结构	(230)
14.3	TCP 事件使用的数据结构	(231)
14.4	定时器、事件和消息	(232)
14.5	TCP 定时器进程	(232)
14.6	删除 TCP 定时器事件	(234)
14.7	删除一个 TCB 的所有事件	(235)
14.8	确定出现一个事件的尚需时间	(236)
14.9	插入 TCP 定时器事件	(237)
14.10	启动无时延的 TCP 输出	(239)
14.11	小结	(240)

第 15 章 TCP:流量控制和自适应重发	(241)
15.1 引言	(241)
15.2 自适应重发中的难题	(241)
15.3 自适应重发的调整	(241)
15.4 重发定时器和退避	(242)
15.5 基于窗口的流量控制	(244)
15.6 最大报文段长度的计算	(247)
15.7 网络拥塞预防与控制	(251)
15.8 慢启动和拥塞预防	(252)
15.9 平均往返时间估值及超时	(254)
15.10 技术和注意事项	(260)
15.11 小结	(260)
第 16 章 TCP:紧急数据处理和急迫功能	(263)
16.1 引言	(263)
16.2 带外信令	(263)
16.3 紧急数据	(263)
16.4 标准的解释	(264)
16.5 为 Berkeley 紧急指针解释法而设的配置	(266)
16.6 通知应用程序	(266)
16.7 从 TCP 中读取数据	(267)
16.8 发送紧急数据	(269)
16.9 TCP 的急迫功能	(270)
16.10 在失序交付时对急迫数据的解释	(270)
16.11 输入时急迫功能的实现	(271)
16.12 小结	(272)
第 17 章 插口级的接口	(274)
17.1 引言	(274)
17.2 通过设备形成的接口	(274)
17.3 作为设备的 TCP 连接	(276)
17.4 TCP 客户程序举例	(276)
17.5 TCP 服务器程序举例	(277)
17.6 TCP 主设备的实现	(279)
17.7 TCP 从设备的实现	(286)
17.8 从设备的初始化	(299)
17.9 小结	(300)

第 18 章	RIP:主动路由传播和被动获取	(302)
18.1	引言	(302)
18.2	主动和被动模式的参与者	(302)
18.3	基本的 RIP 算法和费用度量	(303)
18.4	不稳定性及解决方案	(303)
18.5	报文类型	(306)
18.6	协议特性	(306)
18.7	RIP 的具体实现	(307)
18.8	基本 RIP 进程	(310)
18.9	对输入请求的响应	(315)
18.10	生成更新报文	(317)
18.11	初始化一个更新报文的拷贝	(318)
18.12	生成定期的 RIP 输出	(323)
18.13	RIP 的局限性	(324)
18.14	小结	(324)
第 19 章	OSPF:使用 SPF 算法的路由传播	(325)
19.1	引言	(325)
19.2	OSPF 配置和选项	(325)
19.3	OSPF 的图解理论模型	(326)
19.4	OSPF 的说明	(328)
19.5	邻接关系和链路状态信息的传播	(334)
19.6	用 Hello 发现相邻网关	(335)
19.7	Hello 分组的发送	(336)
19.8	指定路由器	(341)
19.9	选出一个指定路由器	(341)
19.10	变动之后重建邻接关系	(345)
19.11	处理到达的 Hello 分组	(348)
19.12	在相邻网关表中增加一个网关	(349)
19.13	相邻网关状态的变迁	(351)
19.14	OSPF 定时器事件和重发	(352)
19.15	判断是否允许邻接关系	(354)
19.16	OSPF 输入的处理	(355)
19.17	链路状态处理中的说明和过程	(358)
19.18	数据库描述分组的生成	(360)
19.19	创建一个模板	(362)
19.20	传送数据库描述分组	(363)
19.21	处理到达的数据库描述分组	(365)

19.22	处理链路状态请求分组	(370)
19.23	建立链路状态概要信息	(372)
19.24	OSPF 实用过程	(374)
19.25	小结	(377)
第 20 章	SNMP:MIB 变量、表示形式以及变量映射	(378)
20.1	引言	(378)
20.2	服务器的组织和名字的映射	(378)
20.3	MIB 变量	(379)
20.4	MIB 变量名	(380)
20.5	名字之间的字典顺序	(381)
20.6	除去前缀	(381)
20.7	在 MIB 变量上执行的操作	(382)
20.8	表格名	(382)
20.9	名字体系概念上的线索	(383)
20.10	MIB 变量的数据结构	(384)
20.11	用于快速查找的数据结构	(385)
20.12	散列表的实现	(386)
20.13	MIB 绑定的描述	(387)
20.14	绑定中使用的内部变量	(392)
20.15	散列表的查找	(393)
20.16	SNMP 的结构和常量	(395)
20.17	ASN.1 表示形式的处理	(399)
20.18	小结	(409)
第 21 章	SNMP:客户机与服务器	(410)
21.1	引言	(410)
21.2	服务器中数据的表示形式	(410)
21.3	服务器的实现	(410)
21.4	对 SNMP 报文的分析	(412)
21.5	绑定链表中 ASN.1 名字的转换	(416)
21.6	解析一个请求	(417)
21.7	对 get-next 操作的解释	(419)
21.8	操作的间接执行	(420)
21.9	表格的间接寻址	(422)
21.10	回答报文的反向生成	(423)
21.11	从内部格式转换为 ASN.1 表示形式	(426)
21.12	服务器使用的实用函数	(428)
21.13	一个 SNMP 客户机的实现	(428)

21.14	变量初始化	(430)
21.15	小结	(432)
第 22 章	SNMP:表格访问函数	(434)
22.1	引言	(434)
22.2	表格访问	(434)
22.3	表格的对象标识符	(434)
22.4	地址入口表函数	(435)
22.5	地址转换表函数	(441)
22.6	网络接口表的函数	(450)
22.7	选路表函数	(457)
22.8	TCP 连接表函数	(466)
22.9	小结	(473)
第 23 章	实现的回顾	(474)
23.1	引言	(474)
23.2	程序代码统计分析	(474)
23.3	各个协议的程序代码行数	(474)
23.4	每个协议所需的函数和过程	(476)
23.5	小结	(477)
附录 1	过程调用交叉参考表	(479)
附录 2	程序代码中使用到的 Xinu 函数和常量	(499)
参考文献		(515)

第1章 引言和概述

1.1 TCP/IP 协议

TCP/IP Internet 协议族已经成为计算机工业中开放系统互连的事实上的标准。世界各地的计算机系统之所以使用 TCP/IP Internet 协议来通信,是由于它提供了最好的交互操作性能,可以在绝大多数销售商的系统中使用,并且它可兼容多种网络技术,在这一点上,它比其他任何一个协议族要强。科研和教育机构将 TCP/IP 作为基本的数据通信平台。此外,工业上已将 TCP/IP 应用到航空工程、汽车制造、电子技术、宾馆服务、石油化工、出版印刷、药剂制造等众多领域中。

除了常规性地在专用工业网上使用了 TCP/IP 外,许多院校、政府和军事单位也使用 TCP/IP 协议,通过与 Internet 连接,相互通信。这些院校比起没有连接 Internet 的院校来说,在信息交流和研究成果交换方面要优越得多,从而为这些科研院所的研究者们提供了有利的竞争环境。

1.2 了解细节的必要性

尽管 TCP/IP 非常流行并得到了广泛应用,对于大多数计算机专家来说,TCP/IP 协议的细节,以及实现它们的软件的结构仍然是一个谜。从表面上看,是否理解 TCP/IP 的内部细节并不重要,但是,使用 TCP/IP 的程序员们都知道,如果能够完全了解协议是如何运作的,他们将会编写出更健壮的程序代码。譬如说,一个了解 TCP 紧急数据处理过程的程序员可以在他们的应用程序中增加更多的功能,而那些不了解此细节的程序员则无法做到。

即使了解一些简单的概念,诸如 TCP 是如何对数据进行缓冲的,也会对程序员在设计、实施和调试应用程序时有所帮助。举个例子,如果程序员不了解输出缓存、报文段传送、输入缓存和 TCP 急迫数据操作之间的关系的话,那么他就无法成功地编制出利用 TCP 协议的程序。研究 TCP 输入输出的细节,能够使程序员形成一个概念模型,用于解释各部分是如何交互作用的,并有助于理解如何使用底层机制。

1.3 协议间交互作用的复杂性

人们普遍认为,TCP/IP 技术难以理解的一个主要原因是,有关的文章往往单独地讨论每一个协议,而没有去考虑多个协议是如何相互配合协同工作的。例如,各个协议标准文档往往只描述一个协议应当如何工作,它脱离系统的其他部分,孤立地讨论某个协议的行为以及对报文的响应。其实,这些协议最难理解的部分,正是它们之间的交互作用。当我们将所有协议的运作联系在一起考虑时,他们之间的交互会产生复杂的,有时甚至是不可预料的后