计 算 机 科 学 丛 书

TCP/IP详解

卷2：实现

TCP/IP Illustrated Volume 2: The Implementation

(美) Gary R. Wright
W. Richard Stevens 著
陆雪莹 蒋 慧 等译
谢希仁 校

机械工业出版社
China Machine Press
Addison-Wesley

计算机科学丛书

TCP/IP详解
卷2：实现

(美) Gary R. Wright
W. Richard Stevens 著

陆雪莹 蒋 慧 等译

谢希仁 校

机械工业出版社
China Machine Press

本书完整而详细地介绍了TCP/IP协议是如何实现的。书中给出了约500个图例，15 000行实际操作的C代码，采用举例教学的方法帮助你掌握TCP/IP实现。本书不仅说明了插口API和协议族的关系以及主机实现与路由器实现的差别。还介绍了4.4BSD-Lite版的新的特点，如多播、长肥管道支持、窗口缩放、时间戳选项以及其他主题等等。读者阅读本书时，应当具备卷1中阐述的关于TCP/IP的基本知识。

本书适用于希望理解TCP/TP协议如何实现的人，包括编写网络应用程序的程序员以及利用TCP/IP维护计算机网络的系统管理员。

本书版权登记号：图字：01-1999-2856

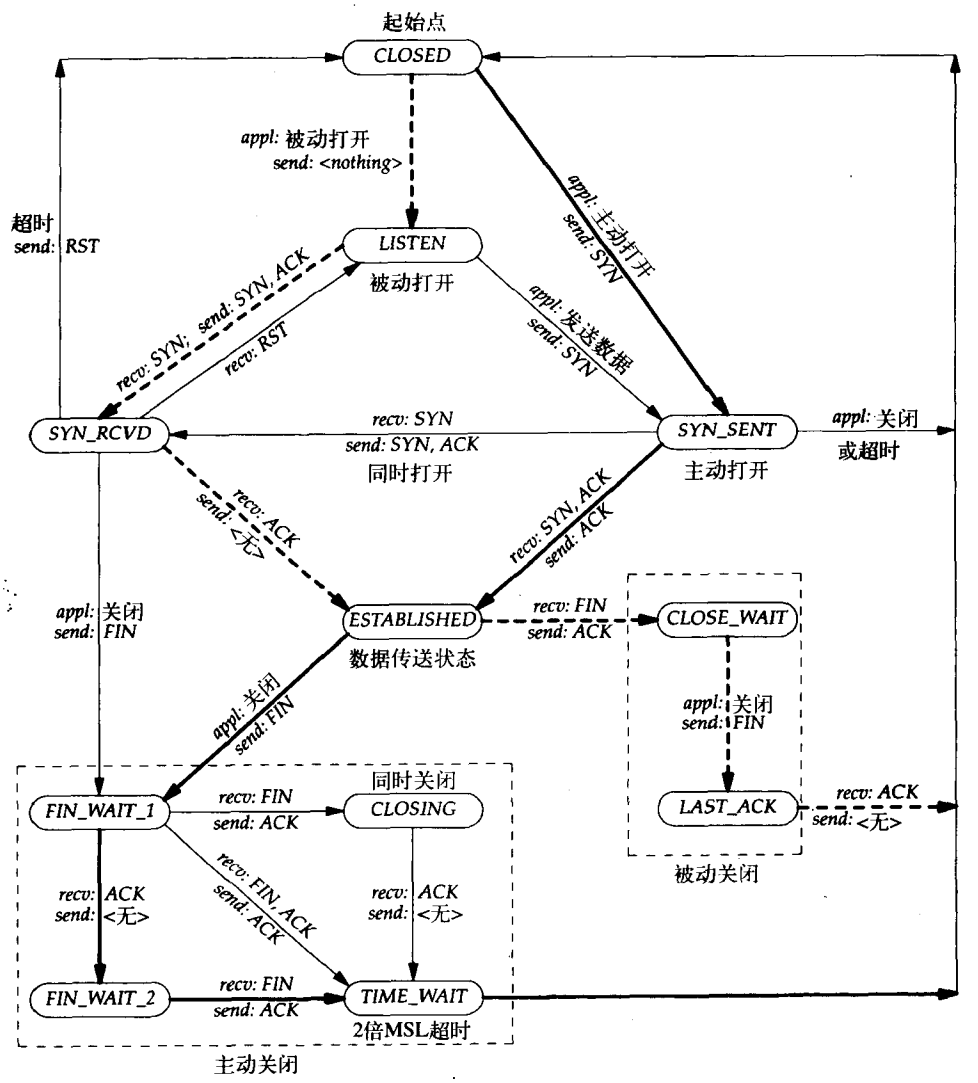
图书在版编目 (CIP) 数据

TCP / IP详解 卷2：实现/ (美) 莱特 (Wright, G. R.)，(美) 史蒂文斯 (Stevens, W. R.) 著；陆雪莹等译. -北京：机械工业出版社，2000.7
(计算机科学丛书)
书名原文：TCP/IP Illustrated, Volume 2: The Implementation
ISBN 7-111-07567-6

Ⅰ. T… Ⅱ. ① 莱… ② 史… ③ 陆… Ⅲ. 传输控制协议 Ⅳ. TN915.04

中国版本图书馆CIP数据核字 (2000) 第21745号

机械工业出版社(北京市西城区百万庄大街22号 邮政编码 100037)
责任编辑：陈贤舜
北京忠信诚胶印厂印刷 · 新华书店北京发行所发行
2000年7月第1版 · 2000年8月第2次印刷
787mm × 1092mm 1/16 · 57.75印张
印数：7 001 - 15 000册
定价：78.00元



客户端状态的正常变迁

服务器端状态的正常变迁

appl: 应用程序产生操作时发生的状态变迁

recv: 收到报文段时发生的状态变迁

send: 状态变迁时发送的内容

TCP状态变迁图

结 构 定 义

译 者 序

我们愿意向广大的读者推荐W. Richard Stevens关于TCP/IP的经典著作(共3卷)的中译本。本书是其中的第2卷:《TCP/IP详解 卷2: 实现》。

大家知道，TCP/IP已成为计算机网络的事实上的标准。在关于TCP/IP的论著中，最有影响的就是两部著作。一部是Douglas E. Comer写的《用TCP/IP进行网际互连》，一套共3卷(中译本已由电子工业出版社于1998年出版)，而另一部就是Stevens写的这3卷书。这两套巨著都很有名，各有其特点。无论是从事计算机网络教学的教师还是进行科研的技术人员，这两套书都应当是必读的。

本书的特点是内容丰富，概念清楚且准确，讲解详细，例子很多。作者在书中举出的所有例子均在作者安装的计算机网络上通过实际验证。各章都留有一定数量的习题。在附录A作者对部分习题给出了解答。在本书的最后，作者给出了许多经典的参考文献，并一一写出了评论。

第2卷是第1卷的继续深入。读者在学习这一卷时，应当先具备第1卷所阐述的关于TCP/IP的基本知识。本卷的特点是使用大量的源代码来讲述TCP/IP协议族中的各协议是怎样实现的。这些内容对于编写TCP/IP网络应用程序的程序员和负责维护基于TCP/IP协议的计算机网络的系统管理员来说，应当是必读的。

参加本书翻译的有：谢钧(序言和第1章~第7章)，蒋慧(第8章~第14章，第22章~第23章)，吴礼发(第15~第17章)，端义峰(第18章~第19章)，胥光辉(第20章~第21章)和陆雪莹(第24章~第32章以及全部附录)。全书由谢希仁教授审校。

限于水平，翻译中不妥或错误之处在所难免，敬请广大读者批评指正。

译 者

于解放军理工大学，南京

2000年2月

译、校者介绍

谢希仁，中国人民解放军理工大学(南京)计算机系教授，全军网络技术研究中心主任，博士研究生导师，1952年毕业于清华大学电机系电信专业。所编写的《计算机网络》于1992年获全国优秀教材奖。1999年再版的《计算机网络》第2版为普通高等教育“九五”国家级重点教材。近来还主持翻译了Comer写的《TCP/IP网际互联》计算机网络经典教材一套三卷本(电子工业出版社1998年出版)，Harnedy写的《简单网络管理协议教程》(电子工业出版社1999年出版)。

陆雪莹，女，1973年1月出生。1994年7月毕业于南京通信工程学院无线通信专业，获工学学士学位。1997年2月于南京通信工程学院计算机软件专业毕业，并获硕士学位。1997年9月至今，任南京通信工程学院计算机教研室教员，同时于解放军理工大学攻读军事通信学博士学位，讲师职称，主要研究方向：智能化网络管理，计算机网络分布式处理。曾参加国家“863”项目，并参加编写专业著作2本，翻译专业著作3本，在各级学术刊物上发表论文5篇。

蒋慧，女，1973年2月出生。1995年毕业于南京通信工程学院计算机系，获计算机应用专业工学学士学位。1998年于南京通信工程学院计算机软件专业毕业，并获硕士学位。1998年9月至今，于解放军理工大学攻读博士学位。自1995年以来，在国内外重要学术刊物和会议上发表8篇论文，其中2篇论文被IEEE国际会议录用。已出版3本有关网络的译作。目前从事软件需求工程、网络协议验证形式化方法以及函数式语言等方面的研究。

前　言

简介

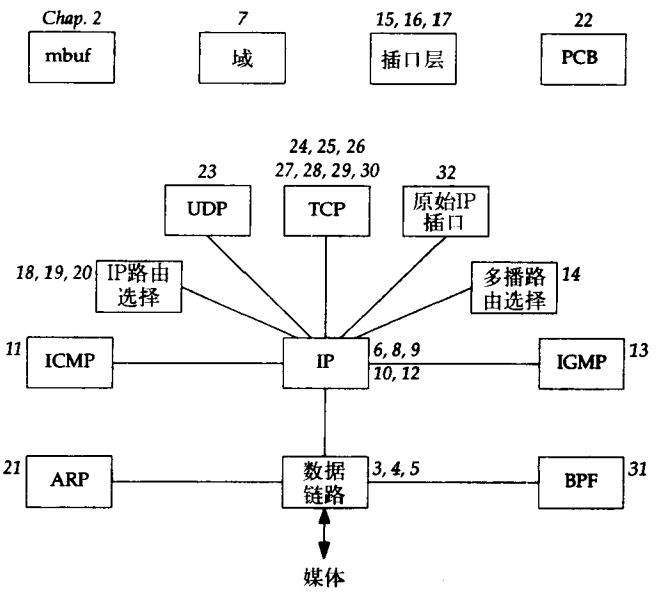
本书描述并给出了TCP/IP实现引用的源代码：加利福尼亚大学伯克利分校的计算机系统研究组(CSRG)的实现。历史上，它曾以4.x BSD系统(伯克利软件发行)发布。这个实现第一次发布是在1982年，经过了很多重大的改变和改进，并且有很多引入到其他Unix和非Unix系统中。这不是一个没有多大意义的实现，而是天天在世界上成千上万个系统上运行的TCP/IP实现的基础。这个实现还提供路由功能，显示在一个主机和一个路由器的TCP/IP实现间的区别。

我们描述这个实现并给出TCP/IP内核实现的完整源代码，大约15 000行C代码。在本文中描述的是4.4BSD-Lite版本。这个代码在1994年4月公开，包含很多增强的联网部分，它们被添加到1988年的4.3BSD Tahoe版、1990年的4.3BSD Reno版和1993年的4.4BSD版(附录B介绍了如何获得这些源代码)。4.4BSD版提供最新的TCP/IP特征，如多播和长肥管道支持(用于高宽带、长时延路径)。图1-1提供了伯克利联网代码的各种版本的其他细节。

本书适用于希望理解TCP/IP协议是如何实现的人：编写网络应用的程序员，负责利用TCP/IP维护计算机系统和网络的系统管理员，以及任何想理解大块的重要代码是如何满足一个真实操作系统的程序员。

本书的组织结构

下图显示的是所涉及的各种协议和子系统。每个方框旁的斜体数字指出方框中的论题在哪一章讨论。



我们采用自底向上的方法来讨论TCP/IP协议组，从数据链路层开始，然后是网络层(IP、ICMP、IGMP、IP路由选择和多播路由选择)，接下来是插口层，最后以运输层(UDP、TCP和原始IP)结束。

预期的读者

本书假设读者对TCP/IP是如何工作的有一个基本的理解。不熟悉TCP/IP的读者应该参考本套书中的第1卷，[Stevens 1994]，那本书对TCP/IP协议组进行了全面的描述。在本书中对第1卷的引用均为卷1。本书还假设读者对操作系统原理有一个基本的理解。

我们用一个数据结构方法来描述这个协议的实现。即，除了给出源代码外，每章还包括源代码使用和维护的数据结构的图和说明。我们显示了这些数据结构是如何适用于TCP/IP和内核使用的其他数据结构的。通篇使用大量的图表——超过250个图表。

这种数据结构方法允许读者采用各种方式使用本书。对所有实现细节感兴趣的读者可以从头到尾阅读全书，看完所有的源代码。可能只想理解协议是如何实现的其他读者，可通过理解所有数据结构并阅读所有文字可以达到目的，而不必看完所有的源代码。

我们预料很多读者会对书中的特定部分感兴趣并且想直接进入那一章。因此，通篇提供了很多向前或向后的引用，沿着完整的索引，允许单独学习某一章。在各章的结尾都提供了习题，并在附录A中给出大多数习题的答案作为自学的参考，使本书能发挥最大的作用。

源代码版权

本书中出现的所有代码，除了图1-2和图8-27，都是来自于4.4BSD-Lite发行版。这个软件是公开的，可从很多地方获得(附录B)。

这个源代码的所有部分都包含下列版权通告。

```
/*
 * Copyright (c) 1982, 1986, 1988, 1990, 1993, 1994
 *      The Regents of the University of California.  All rights reserved.
 *
 * Redistribution and use in source and binary forms, with or without
 * modification, are permitted provided that the following conditions
 * are met:
 * 1. Redistributions of source code must retain the above copyright
 *    notice, this list of conditions and the following disclaimer.
 * 2. Redistributions in binary form must reproduce the above copyright
 *    notice, this list of conditions and the following disclaimer in the
 *    documentation and/or other materials provided with the distribution.
 * 3. All advertising materials mentioning features or use of this software
 *    must display the following acknowledgement:
 *      This product includes software developed by the University of
 *      California, Berkeley and its contributors.
 * 4. Neither the name of the University nor the names of its contributors
 *    may be used to endorse or promote products derived from this software
 *    without specific prior written permission.
 *
 * THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS ``AS IS'' AND
 * ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE
 * IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE
 * ARE DISCLAIMED.  IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE
 * FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL
 * DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS
```

```
* OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION)
* HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT
* LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY
* OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF
* SUCH DAMAGE.
*/
```

Gary R.Wright

Middletown, Connecticut

W. Richard Stevens

Tucson, Arizona

1994年11月

目　录