Addison Wesley

TURING

图灵原版计算机科学系列

TCP/IP详解

卷2：实现

英文版

TCP/IP Illustrated

Volume 2: The Implementation

[美] Gary R. Wright
W. Richard Stevens 著

人民邮电出版社
POSTS & TELECOM PRESS

TCP/IP详解

卷2：实现

英文版

TCP/IP Illustrated
Volume 2: The Implementation

[美] Gary R. Wright
W. Richard Stevens 著

人民邮电出版社
北京

图书在版编目（CIP）数据

TCP/IP详解. 卷2，实现＝TCP/IP Illustrated, Volume 2: The Implementation, First Edition：英文 /（美）赖特（Wright, G. R .），（美）史蒂文斯（Stevens, W. R. ）著. —北京：人民邮电出版社，2010.4

（图灵原版计算机科学系列）

ISBN 978-7-115-22248-0

I. ①T … II. ①赖… ②史… III. ①计算机网络－通信协议－英文 IV. ①TN915.04

中国版本图书馆CIP数据核字（2010）第014975号

内 容 提 要

本书是TCP/IP领域的经典之作！书中完整而详细地介绍了TCP/IP协议是如何实现的。本书介绍了一个实际的TCP/IP实现，并给出了这一实现的完整源代码，帮助读者全面掌握TCP/IP的实现。本书内容详尽且具权威性，几乎每章都提供精选的习题，并在附录中提供了部分习题的答案。

本书适合任何希望了解TCP/IP协议如何实现的读者阅读，更是TCP/IP领域研究人员和开发人员的权威参考书。

图灵原版计算机科学系列

TCP/IP详解 卷2：实现（英文版）

- 著 [美] Gary R. Wright W. Richard Stevens

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

北京艺辉印刷有限公司印刷

- 开本：800×1000 1/16

印张：74.75

字数：1440千字 2010年4月第1版

印数：1 - 2 000册 2010年4月北京第1次印刷

著作权合同登记号 图字：01-2010-0310号

ISBN 978-7-115-22248-0

定价：139.00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

献给我的父母和姐姐，
感谢他们的爱与支持。
——Gary R. Wright

献给我的父母，
感谢他们让我接受了良好的教育，
并示我以优良的职业习惯。
——W. Richard Stevens

前　言

概述

本书介绍并给出了TCP/IP的常见参考实现的源代码，即由加州大学伯克利分校计算机系统研究组（CSRG）研发的实现。历史上该实现是随4.x BSD（Berkeley Software Distribution）系统一起发布的。最早的发布版本出现于1982年，随后经过了多次大改、多次微调，并实现了映射到其他Unix及非Unix系统的众多端口。这个实现不可小觑，它是世界范围内无数系统中日常运行的TCP/IP实现的基础。该实现还提供了路由器功能，使我们可以展示出TCP/IP的主机实现与路由器的区别。

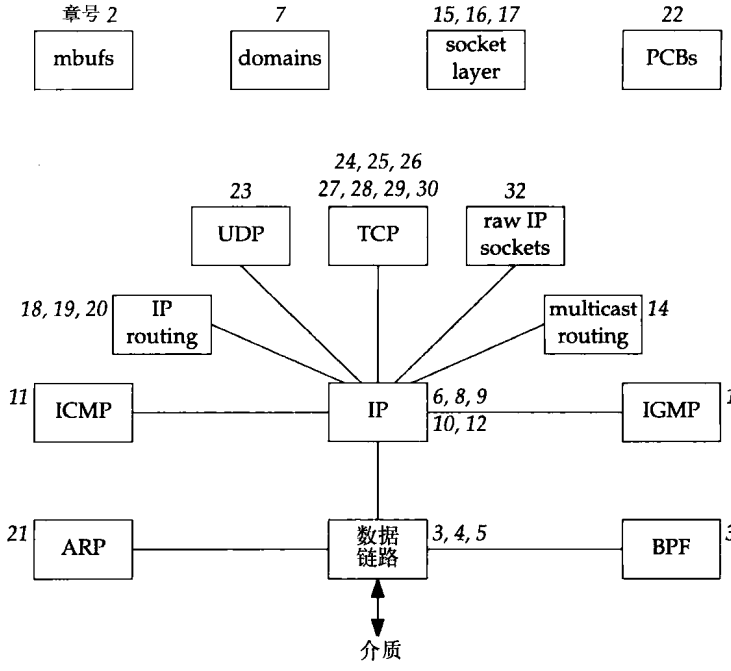
本书将描述该实现并给出TCP/IP核心实现的全部源代码，大约有15 000行C代码。书中介绍的Berkeley代码的版本是4.4BSD-Lite发布版本。这个代码于1994年4月公布，它在1988年的4.3BSD Tahoe发布版本、1990年的4.3BSD Reno发布版本以及1993年的4.4BSD发布版本基础上增加了大量的网络增强技术（附录B讲述了如何获取这些源代码）。4.4BSD发布版本提供了最新的TCP/IP特性，例如对多播（multicast）和长肥管道（long fat pipe）的支持（用于高带宽、长延迟的通道）。图1-1（第4页）提供了各版本的Berkeley网络代码的更多细节。

本书面向所有希望了解TCP/IP协议实现原理的读者：编写网络应用的程序员、利用TCP/IP维护计算机系统与网络的系统管理员以及希望了解实际操作系统中大型代码的集成

原理的程序员。

本书的结构

下面的图给出了本书涉及的各种协议和子系统，方框上的斜体数字指明了相应内容在哪一章讨论。



我们采用一种自底向上的方式来介绍TCP/IP协议族，首先是数据链路层，然后是网络层（IP、ICMP、IGMP、IP路由、多播路由）以及随后的套接字层，最后是传输层（UDP、TCP和原始IP）。

致读者

本书的读者应该对TCP/IP协议的工作原理有基本的了解。对TCP/IP协议不是很熟悉的读者应首先参考本套书的第1卷[Stevens, 1994]，该书对TCP/IP协议族有比较透彻的描述。本书将之前的第1卷称为卷1（Volume 1）。本书同时还需要读者对操作系统原理有初步的了解。

我们采用数据结构方法来描述协议的实现。每章不仅给出源代码，还用图片和文字描述源代码使用或者维护的数据结构。我们还展示了如何把这些数据结构集成到TCP/IP与内核所用的数据结构中。全书使用了大量的图表——超过250个。

这种数据结构方法使得读者可以用多种方式使用本书。那些对所有的实现细节都感兴趣的读者可以结合源代码，把本书从头一直读到最后。其他读者则可以通过理解数据

结构和阅读文字内容来理解协议的工作原理，而不一定非要参考源代码。

估计会有许多读者只是对本书中特定的章节感兴趣，并希望直接阅读这些章。全书贯穿了许多交叉引用，并提供了完整的索引，因此读者可以独立阅读各章的内容。索引后面还按照字母表顺序给出了书中所有函数和宏的交叉引用，以及相关详细信息的起始页码。每章的末尾都布置了一些习题，附录A给出了大部分习题的答案，这样做是为了本书更适合作为自学参考书。

源代码版权

本书中除了图1-2和图8-27之外的所有源代码都来自4.4BSD-Lite发布版。这个软件是公用程序，可以通过很多方式获得（详见附录B）。

书中的所有源代码版权说明如下：

```
/*
 * 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.
 */
```

致谢

感谢百忙之中拨冗阅读本书书稿并给出重要反馈的技术审稿人：Ragnvald Blindheim、

Jon Crowcroft、Sally Floyd、Glen Glater、John Gulbenkian、Don Hering、Mukesh Kacker、Berry Kercheval、Brian W. Kernighan、Ulf Kieber、Mark Laubach、Steven McCanne、Craig Partridge、Vern Paxson、Steve Rago、Chakravardhi Ravi、Peter Salus、Doug Schmidt、Keith Sklower、Ian Lance Taylor和 G. N. Ananda Vardhana。特别感谢顾问编辑Brian Kernighan，在完成本书的过程中，他提出了很多及时、透彻、很有帮助的评审意见，并始终鼓励和支持着我。

我们（再次）感谢美国国家光学天文台，尤其是授权我们接入其网络和主机的Sidney Wolff、Richard Wolff和Steve Grandi。我们还要感谢加州大学伯克利分校计算机系统研究组的Keith Bostic和Kirk McKusick授权我们使用最新的4.4BSD系统，Keith Sklower修改了4.4BSD-Lite软件使其能运行于BSD/386 V1.1系统。

Gary R. Wright感谢John Wait多年以来的好心督促、Dave Schaller的鼓励以及Jim Hogue在本书写作及出版过程中的支持。

W. Richard Stevens再次感谢他的家人，他们又一次忍受了他的“小型”著书项目。感谢你们：Sally、Bill、Ellen和David。

Addison-Wesley团队的辛勤工作、专业水准和鼎力支持使作者的工作轻松了很多。我们特别感谢John Wait的指导和Kim Dawley的创新思维。

作者制作了本书的最终电子版。描述工业级强度的软件系统就需要使用工业级强度的文本处理系统。本书的一位作者选择使用James Clark的Groff包，另一位作者只能勉强同意。

欢迎读者以电子邮件的方式反馈意见、提出建议或订正错误。两位作者都很乐意指出对方残存的错误。

Gary R. Wright
于康涅狄格州米德尔敦市

W. Richard Stevens
于亚利桑那州图森市
1994年11月

Contents