

经 典 原 版 书 库

TCP/IP详解

卷1 协议

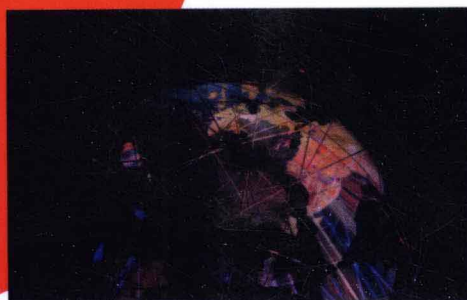
(美) Kevin R. Fall W. Richard Stevens 著

(英文版·第2版)

TCP/IP Illustrated, Volume 1 Second Edition

The Protocols

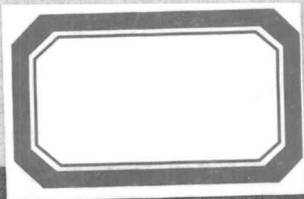
Kevin R. Fall
W. Richard Stevens



Foreword by Vint Cerf, *Internet pioneer*



机械工业出版社
China Machine Press



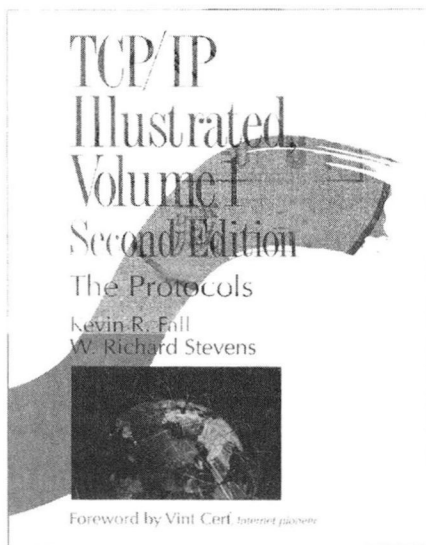
原 版 书

TCP/IP详解

卷1 协议

(英文版·第2版)

TCP/IP Illustrated, Volume 1
The Protocols (Second Edition)



(美) Kevin R. Fall W. Richard Stevens 著



机械工业出版社
China Machine Press

Original edition, entitled TCP/IP ILLUSTRATED, VOLUME 1: THE PROTOCOLS, 2E, 9780321336316 by FALL, KEVIN R.; STEVENS, W. RICHARD, published by Pearson Education, Inc, publishing as Addison-Wesley Professional, Copyright © 2012.

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 Pearson Education, Inc.

English reprint edition published by PEARSON EDUCATION ASIA LTD., and CHINA MACHINE PRESS, Copyright © 2012.

This edition is manufactured in the People's Republic of China, and is authorized for sale and distribution in the People's Republic of China exclusively (except Taiwan, Hong Kong SAR and Macau SAR).

本书英文影印版由 Pearson Education Asia Ltd. 授权机械工业出版社独家出版。未经出版者书面许可, 不得以任何方式复制或抄袭本书内容。

仅限于中华人民共和国境内(不包括中国香港、澳门特别行政区和中国台湾地区)销售发行。

本书封面贴有 Pearson Education (培生教育出版集团) 激光防伪标签, 无标签者不得销售。

封底无防伪标均为盗版

版权所有, 侵权必究

本书法律顾问 北京市展达律师事务所

本书版权登记号: 图字: 01-2012-2100

图书在版编目(CIP)数据

TCP/IP 详解 卷1: 协议(英文版·第2版)/(美)福尔(Fall, K. R.), (美)史蒂文斯(Stevens, W. R.) 著. —北京: 机械工业出版社, 2012.5

(经典原版书库)

书名原文: TCP/IP Illustrated, Volume 1: The Protocols, Second Edition

ISBN 978-7-111-38228-7

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

中国版本图书馆CIP数据核字(2012)第079994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: 迟振春

北京京师印务有限公司印刷

2012年5月第1版第1次印刷

186mm×240mm·65.75印张

标准书号: ISBN 978-7-111-38228-7

定价: 129.00元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991; 88361066

购书热线: (010) 68326294; 88379649; 68995259

投稿热线: (010) 88379604

读者信箱: hzjsj@hzbook.com

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域中取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自1998年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与Pearson, McGraw-Hill, Elsevier, MIT, John Wiley & Sons, Cengage等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出Andrew S. Tanenbaum, Bjarne Stroustrup, Brain W. Kernighan, Dennis Ritchie, Jim Gray, Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Abraham Silberschatz, William Stallings, Donald E. Knuth, John L. Hennessy, Larry L. Peterson等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及珍藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专程为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方式如下：

华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：(010) 88379604

联系地址：北京市西城区百万庄南街1号

邮政编码：100037



华章科技图书出版中心

Praise for the First Edition of *TCP/IP Illustrated, Volume 1: The Protocols*

"This is sure to be the bible for TCP/IP developers and users. Within minutes of picking up the text, I encountered several scenarios that had tripped up both my colleagues and myself in the past. Stevens reveals many of the mysteries once held tightly by the ever-elusive networking gurus. Having been involved in the implementation of TCP/IP for some years now, I consider this by far the finest text to date."

—Robert A. Ciampa, network engineer, Synneretics, division of 3COM

"While all of Stevens' books are readable and technically excellent, this new opus is awesome. Although many books describe the TCP/IP protocols, Stevens provides a level of depth and real-world detail lacking from the competition. He puts the reader inside TCP/IP using a visual approach and shows the protocols in action."

—Steven Baker, networking columnist, *Unix Review*

"*TCP/IP Illustrated, Volume 1*, is an excellent reference for developers, network administrators, or anyone who needs to understand TCP/IP technology. *TCP/IP Illustrated* is comprehensive in its coverage of TCP/IP topics, providing enough details to satisfy the experts while giving enough background and commentary for the novice."

—Bob Williams, vice president, Marketing, NetManage, Inc.

"... [T]he difference is that Stevens wants to show as well as tell about the protocols. His principal teaching tools are straightforward explanations, exercises at the ends of chapters, byte-by-byte diagrams of headers and the like, and listings of actual traffic as examples."

—Walter Zintz, *UnixWorld*

"Much better than theory only. . . . W. Richard Stevens takes a multihost-based configuration and uses it as a travelogue of TCP/IP examples with illustrations. *TCP/IP Illustrated, Volume 1*, is based on practical examples that reinforce the theory—distinguishing this book from others on the subject, and making it both readable and informative."

—Peter M. Haverlock, consultant, IBM TCP/IP Development

"The diagrams he uses are excellent and his writing style is clear and readable. In sum, Stevens has made a complex topic easy to understand. This book merits everyone's attention. Please read it and keep it on your bookshelf."

—Elizabeth Zinkann, sys admin

"W. Richard Stevens has produced a fine text and reference work. It is well organized and very clearly written with, as the title suggests, many excellent illustrations exposing the intimate details of the logic and operation of IP, TCP, and the supporting cast of protocols and applications."

—Scott Bradner, consultant, Harvard University OIT/NSD

Foreword

Rarely does one find a book on a well-known topic that is both historically and technically comprehensive and remarkably accurate. One of the things I admire about this work is the “warts and all” approach that gives it such credibility. The TCP/IP architecture is a product of the time in which it was conceived. That it has been able to adapt to growing requirements in many dimensions by factors of a million or more, to say nothing of a plethora of applications, is quite remarkable. Understanding the scope and limitations of the architecture and its protocols is a sound basis from which to think about future evolution and even revolution.

During the early formulation of the Internet architecture, the notion of “enterprise” was not really recognized. In consequence, most networks had their own IP address space and “announced” their addresses in the routing system directly. After the introduction of commercial service, Internet Service Providers emerged as intermediaries who “announced” Internet address blocks on behalf of their customers. Thus, most of the address space was assigned in a “provider dependent” fashion. “Provider independent” addressing was unusual. The net result (no pun intended) led to route aggregation and containment of the size of the global routing table. While this tactic had benefits, it also created the “multi-homing” problem since users of provider-dependent addresses did not have their own entries in the global routing table. The IP address “crunch” also led to Network Address Translation, which also did not solve provider dependence and multi-homing problems.

Reading through this book evokes a sense of wonder at the complexity that has evolved from a set of relatively simple concepts that worked with a small number of networks and application circumstances. As the chapters unfold, one can see the level of complexity that has evolved to accommodate an increasing number of requirements, dictated in part by new deployment conditions and challenges, to say nothing of sheer growth in the scale of the system.

The issues associated with securing “enterprise” users of the Internet also led to firewalls that are intended to supply perimeter security. While useful, it has become clear that attacks against local Internet infrastructure can come through

internal compromises (e.g., an infected computer is put onto an internal network or an infected thumb-drive is used to infect an internal computer through its USB port).

It has become apparent that, in addition to a need to expand the Internet address space through the introduction of IP version 6, with its 340 trillion trillion addresses, there is also a strong need to introduce various security-enhancing mechanisms such as the Domain Name System Security Extension (DNSSEC) among many others.

What makes this book unique, in my estimation, is the level of detail and attention to history. It provides background and a sense for the ways in which solutions to networking problems have evolved. It is relentless in its effort to achieve precision and to expose remaining problem areas. For an engineer determined to refine and secure Internet operation or to explore alternative solutions to persistent problems, the insights provided by this book will be invaluable. The authors deserve credit for a thorough rendering of the technology of today's Internet.

*Woodhurst
June 2011*

Vint Cerf

Preface to the Second Edition

Welcome to the second edition of *TCP/IP Illustrated, Volume 1*. This book aims to provide a detailed, current look at the TCP/IP protocol suite. Instead of just describing how the protocols operate, we show the protocols in operation using a variety of analysis tools. This helps you better understand the design decisions behind the protocols and how they interact with each other, and it simultaneously exposes you to implementation details without your having to read through the implementation's software source code or set up an experimental laboratory. Of course, reading source code or setting up a laboratory will only help to increase your understanding.

Networking has changed dramatically in the past three decades. Originally a research project and object of curiosity, the Internet has become a global communication fabric upon which governments, businesses, and individuals depend. The TCP/IP suite defines the underlying methods used to exchange information by every device on the Internet. After more than a decade of delay, the Internet and TCP/IP itself are now undergoing an evolution, to incorporate IPv6. Throughout the text we will discuss both IPv6 and the current IPv4 together, but we highlight the differences where they are important. Unfortunately, they do not directly interoperate, so some care and attention are required to appreciate the impact of the evolution.

The book is intended for anyone wishing to better understand the current set of TCP/IP protocols and how they operate: network operators and administrators, network software developers, students, and users who deal with TCP/IP. We have included material that should be of interest to both new readers as well as those familiar with the material from the first edition. We hope you will find the coverage of the new and older material useful and interesting.

Comments on the First Edition

Nearly two decades have passed since the publication of the first edition of *TCP/IP Illustrated, Volume 1*. It continues to be a valuable resource for both students and professionals in understanding the TCP/IP protocols at a level of detail difficult to

obtain in competing texts. Today it remains among the best references for detailed information regarding the operation of the TCP/IP protocols. However, even the best books concerned with information and communications technology become dated after a time, and the TCP/IP Illustrated series is no exception. In this edition, I hope to thoroughly update the pioneering work of Dr. Stevens with coverage of new material while maintaining the exceptionally high standard of presentation and detail common to his numerous books.

The first edition covers a broad set of protocols and their operation, ranging from the link layer all the way to applications and network management. Today, covering this breadth of material comprehensively in a single volume would produce a very lengthy text indeed. For this reason, the second edition focuses specifically on the *core* protocols: those relatively low-level protocols used most frequently in providing the basic services of configuration, naming, data delivery, and security for the Internet. Detailed discussions of applications, routing, Web services, and other important topics are postponed to subsequent volumes.

Considerable progress has been made in improving the robustness and compliance of TCP/IP implementations to their corresponding specifications since the publication of the first edition. While many of the examples in the first edition highlight implementation bugs or noncompliant behaviors, these problems have largely been addressed in currently available systems, at least for IPv4. This fact is not terribly surprising, given the greatly expanded use of the TCP/IP protocols in the last 18 years. Misbehaving implementations are a comparative rarity, which attests to a certain maturity of the protocol suite as a whole. The problems encountered in the operation of the core protocols nowadays often relate to intentional exploitation of infrequently used protocol features, a form of security concern that was not a primary focus in the first edition but one that we spend considerable effort to address in the second edition.

The Internet Milieu of the Twenty-first Century

The usage patterns and importance of the Internet have changed considerably since the publication of the first edition. The most obvious watershed event was the creation and subsequent intense commercialization of the World Wide Web starting in the early 1990s. This event greatly accelerated the availability of the Internet to large numbers of people with various (sometimes conflicting) motivations. As such, the protocols and systems originally implemented in a small-scale environment of academic cooperation have been stressed by limited availability of addresses and an increase of security concerns.

In response to the security threats, network and security administrators have introduced special control elements into the network. It is now common practice to place a *firewall* at the point of attachment to the Internet, for both large enterprises as well as small businesses and homes. As the demand for IP addresses and security has increased over the last decade, *Network Address Translation* (NAT) is now supported in virtually all current-generation routers and is in widespread use. It

has eased the pressure on Internet address availability by allowing sites to obtain a comparatively small number of routable Internet addresses from their service providers (one for each simultaneously online user), yet assign a very large number of addresses to local computers without further coordination. A consequence of NAT deployment has been a slowing of the migration to IPv6 (which provides for an almost incomprehensibly large number of addresses) and interoperability problems with some older protocols.

As the users of personal computers began to demand Internet connectivity by the mid-1990s, the largest supplier of PC software, Microsoft, abandoned its original policy of offering only proprietary alternatives to the Internet and instead undertook an effort to embrace TCP/IP compatibility in most of its products. Since then, personal computers running their Windows operating system have come to dominate the mix of PCs presently connected to the Internet. Over time, a significant rise in the number of Linux-based systems means that such systems now threaten to displace Microsoft as the frontrunner. Other operating systems, including Oracle Solaris and Berkeley's BSD-based systems, which once represented the majority of Internet-connected systems, are now a comparatively small component of the mix. Apple's OS X (Mach-based) operating system has risen as a new contender and is gaining in popularity, especially among portable computer users. In 2003, portable computer (laptop) sales exceeded desktop sales as the majority of personal computer types sold, and their proliferation has sparked a demand for widely deployed, high-speed Internet access supported by wireless infrastructure. It is projected that the most common method for accessing the Internet from 2012 and beyond will be smartphones. Tablet computers also represent an important growing contender.

Wireless networks are now available at a large number of locations such as restaurants, airports, coffeehouses, and other public places. They typically provide short-range free or pay-for-use (flat-rate) high-speed wireless Internet connections using hardware compatible with commonly used office or home local area network installations. A set of alternative "wireless broadband" technologies based on cellular telephone standards (e.g., LTE, HSPA, UMTS, EV-DO) are becoming widely available in developed regions of the world (and some developing regions of the world that are "leapfrogging" to newer wireless technology), offering longer-range operation, often at somewhat reduced bandwidths and with volume-based pricing. Both types of infrastructure address the desire of users to be mobile while accessing the Internet, using either portable computers or smaller devices. In either case, mobile end users accessing the Internet over wireless networks pose two significant technical challenges to the TCP/IP protocol architecture. First, mobility affects the Internet's routing and addressing structure by breaking the assumption that hosts have addresses assigned to them based upon the identity of their nearby router. Second, wireless links may experience outages and therefore cause data to be lost for reasons other than those typical of wired links (which generally do not lose data unless too much traffic is being injected into the network).

Finally, the Internet has fostered the rise of so-called peer-to-peer applications forming “overlay” networks. Peer-to-peer applications do not rely on a central server to accomplish a task but instead determine a set of *peer* computers with which they can communicate and interact to accomplish a task. The peer computers are operated by other end users and may come and go rapidly compared to a fixed server infrastructure. The “overlay” concept captures the fact that such interacting peers themselves form a network, overlaid atop the conventional TCP/IP-based network (which, one may observe, is itself an overlay above the underlying physical links). The development of peer-to-peer applications, while of intense interest to those who study traffic flows and electronic commerce, has not had a profound impact on the core protocols described in *Volume 1* per se, but the concept of overlay networks has become an important consideration for networking technology more generally.

Content Changes for the Second Edition

Regarding content in the text, the most important changes from the first edition are a restructuring of the scope of the overall text and the addition of significant material on security. Instead of attempting to cover nearly all common protocols in use at every layer in the Internet, the present text focuses in detail first on the non-security core protocols in widespread use, or that are expected to be in widespread use in the near future: Ethernet (802.3), Wi-Fi (802.11), PPP, ARP, IPv4, IPv6, UDP, TCP, DHCP, and DNS. These protocols are likely to be encountered by system administrators and users alike.

In the second edition, security is covered in two ways. First, in each appropriate chapter, a section devoted to describing known attacks and their countermeasures relating to the protocol described in the chapter is included. These descriptions are not presented as a recipe for constructing attacks but rather as a practical indication of the kinds of problems that may arise when protocol implementations (or specifications, in some cases) are insufficiently robust. In today’s Internet, incomplete specification or lax implementation practice can lead to mission-critical systems being compromised by even relatively unsophisticated attacks.

The second important discussion of security occurs in Chapter 18, where security and cryptography are studied in some detail, including protocols such as IPsec, TLS, DNSSEC, and DKIM. These protocols are now understood to be important for implementing any service or application expected to maintain integrity or secure operation. As the Internet has increased in commercial importance, the need for security (and the number of threats to it) has grown proportionally.

Although IPv6 was not included in the first edition, there is now reason to believe that the use of IPv6 may increase significantly with the exhaustion of unallocated IPv4 address groups in February 2011. IPv6 was conceived largely to address the problems of IPv4 address depletion and, while not nearly as common as IPv4 today, is becoming more important as a growing number of small devices (such as cellular telephones, household devices, and environmental

sensors) become attached to the Internet. Events such as the World IPv6 Day (June 8, 2011) helped to demonstrate that the Internet can continue to work even as the underlying protocols are modified and augmented in a significant way.

A second consideration for the structure of the second edition is a deemphasis of the protocols that are no longer commonly used and an update of the descriptions of those that have been revised substantially since the publication of the first edition. The chapters covering RARP, BOOTP, NFS, SMTP, and SNMP have been removed from the book, and the discussion of the SLIP protocol has been abandoned in favor of expanded coverage of DHCP and PPP (including PPPoE). The function of IP forwarding (described in Chapter 9 in the first edition) has been integrated with the overall description of the IPv4 and IPv6 protocols in Chapter 5 of this edition. The discussion of dynamic routing protocols (RIP, OSPF, and BGP) has been removed, as the latter two protocols alone could each conceivably merit a book-long discussion. Starting with ICMP, and continuing through IP, TCP, and UDP, the impact of operation using IPv4 versus IPv6 is discussed in any cases where the difference in operation is significant. There is no specific chapter devoted solely to IPv6; instead, its impact relative to each existing core protocol is described where appropriate. Chapters 15 and 25–30 of the first edition, which are devoted to Internet applications and their supporting protocols, have been largely removed; what remains only illustrates the operation of the underlying core protocols where necessary.

Several chapters covering new material have been added. The first chapter begins with a general introduction to networking issues and architecture, followed by a more Internet-specific orientation. The Internet's addressing architecture is covered in Chapter 2. A new chapter on host configuration and how a system "gets on" the network appears as Chapter 6. Chapter 7 describes firewalls and Network Address Translation (NAT), including how NATs are used in partitioning address space between routable and nonroutable portions. The set of tools used in the first edition has been expanded to include Wireshark (a free network traffic monitor application with a graphical user interface).

The target readership for the second edition remains identical to that of the first edition. No prior knowledge of networking concepts is required for approaching it, although the advanced reader should benefit from the level of detail and references. A rich collection of references is included in each chapter for the interested reader to pursue.

Editorial Changes for the Second Edition

The general flow of material in the second edition remains similar to that of the first edition. After the introductory material (Chapters 1 and 2), the protocols are presented in a bottom-up fashion to illustrate how the goal of network communication presented in the introduction is realized in the Internet architecture. As in the first edition, actual packet traces are used to illustrate the operational details of the protocols, where appropriate. Since the publication of the first edition, freely

available packet capture and analysis tools with graphical interfaces have become available, extending the capabilities of the `tcpdump` program used in the first edition. In the present text, `tcpdump` is used when the points to be illustrated are easily conveyed by examining the output of a text-based packet capture tool. In most other cases, however, screen shots of the Wireshark tool are used. Please be aware that some output listings, including snapshots of `tcpdump` output, are wrapped or simplified for clarity.

The packet traces shown typically illustrate the behavior of one or more parts of the network depicted on the inside of the front book cover. It represents a broad-band-connected “home” environment (typically used for client access or peer-to-peer networking), a “public” environment (e.g., coffee shop), and an enterprise environment. The operating systems used for examples include Linux, Windows, FreeBSD, and Mac OS X. Various versions are used, as many different OS versions are in use on the Internet today.

The structure of each chapter has been slightly modified from the first edition. Each chapter begins with an introduction to the chapter topic, followed in some cases by historical notes, the details of the chapter, a summary, and a set of references. A section near the end of most chapters describes security concerns and attacks. The per-chapter references represent a change for the second edition. They should make each chapter more self-contained and require the reader to perform fewer “long-distance page jumps” to find a reference. Some of the references are now enhanced with WWW URLs for easier access online. In addition, the reference format for papers and books has been changed to a somewhat more compact form that includes the first initial of each author’s last name followed by the last two digits of the year (e.g., the former [Cerf and Kahn 1974] is now shortened to [CK74]). For the numerous RFC references used, the RFC number is used instead of the author names. This follows typical RFC conventions and has the side benefit of grouping all the RFC references together in the reference lists.

On a final note, the typographical conventions of the TCP/IP Illustrated series have been maintained faithfully. However, the present author elected to use an editor and typesetting package other than the Troff system used by Dr. Stevens and some other authors of the Addison-Wesley Professional Computing Series collection. Thus, the particular task of final copyediting could take advantage of the significant expertise of Barbara Wood, the copy editor generously made available to me by the publisher. We hope you will be pleased with the results.

*Berkeley, California
September 2011*

Kevin R. Fall

Adapted Preface to the First Edition

Introduction

This book describes the TCP/IP protocol suite, but from a different perspective than other texts on TCP/IP. Instead of just describing the protocols and what they do, we'll use a popular diagnostic tool to watch the protocols in action. Seeing how the protocols operate in varying circumstances provides a greater understanding of how they work and why certain design decisions were made. It also provides a look into the implementation of the protocols, without having to wade through thousands of lines of source code.

When networking protocols were being developed in the 1960s through the 1980s, expensive, dedicated hardware was required to see the packets going "across the wire." Extreme familiarity with the protocols was also required to comprehend the packets displayed by the hardware. Functionality of the hardware analyzers was limited to that built in by the hardware designers.

Today this has changed dramatically with the ability of the ubiquitous workstation to monitor a local area network [Mogul 1990]. Just attach a workstation to your network, run some publicly available software, and watch what goes by on the wire. While many people consider this a tool to be used for *diagnosing* network problems, it is also a powerful tool for *understanding* how the network protocols operate, which is the goal of this book.

This book is intended for anyone wishing to understand how the TCP/IP protocols operate: programmers writing network applications, system administrators responsible for maintaining computer systems and networks utilizing TCP/IP, and users who deal with TCP/IP applications on a daily basis.

Typographical Conventions

When we display interactive input and output we'll show our typed input in a **bold font**, and the computer output like *this. Comments are added in italics.*

```
bsdi % telnet svr4 discard  
Trying 140.252.13.34...  
Connected to svr4.
```

*connect to the discard server
this line and next output by Telnet client*

Also, we always include the name of the system as part of the shell prompt (`bsdi` in this example) to show on which host the command was run.

Note

Throughout the text we'll use indented, parenthetical notes such as this to describe historical points or implementation details.

We sometimes refer to the complete description of a command on the Unix manual as in `ifconfig(8)`. This notation, the name of the command followed by a number in parentheses, is the normal way of referring to Unix commands. The number in parentheses is the section number in the Unix manual of the “manual page” for the command, where additional information can be located. Unfortunately not all Unix systems organize their manuals the same, with regard to the section numbers used for various groupings of commands. We'll use the BSD-style section numbers (which is the same for BSD-derived systems such as SunOS 4.1.3), but your manuals may be organized differently.

Acknowledgments

Although the author's name is the only one to appear on the cover, the combined effort of many people is required to produce a quality text book. First and foremost is the author's family, who put up with the long and weird hours that go into writing a book. Thank you once again, Sally, Bill, Ellen, and David.

The consulting editor, Brian Kernighan, is undoubtedly the best in the business. He was the first one to read various drafts of the manuscript and mark it up with his infinite supply of red pens. His attention to detail, his continual prodding for readable prose, and his thorough reviews of the manuscript are an immense resource to a writer.

Technical reviewers provide a different point of view and keep the author honest by catching technical mistakes. Their comments, suggestions, and (most importantly) criticisms add greatly to the final product. My thanks to Steve Bellovin, Jon Crowcroft, Pete Haverlock, and Doug Schmidt for comments on the entire manuscript. Equally valuable comments were provided on portions of the manuscript by Dave Borman for his thorough review of all the TCP chapters, and to Bob Gilligan who should be listed as a coauthor for Appendix E.

An author cannot work in isolation, so I would like to thank the following persons for lots of small favors, especially by answering my numerous e-mail questions: Joe Godsil, Jim Hogue, Mike Karels, Paul Lucchina, Craig Partridge, Thomas Skibo, and Jerry Toporek.

This book is the result of my being asked lots of questions on TCP/IP for which I could find no quick, immediate answer. It was then that I realized that the easiest way to obtain the answers was to run small tests, forcing certain conditions to occur, and just watch what happens. I thank Peter Haverlock for asking the probing questions and Van Jacobson for providing so much of the publicly available software that is used in this book to answer the questions.

A book on networking needs a real network to work with along with access to the Internet. My thanks to the National Optical Astronomy Observatories (NOAO), especially Sidney Wolff, Richard Wolff, and Steve Grandi, for providing access to their networks and hosts. A special thanks to Steve Grandi for answering lots of questions and providing accounts on various hosts. My thanks also to Keith Bostic and Kirk McKusick at the U.C. Berkeley CSRG for access to the latest 4.4BSD system.

Finally, it is the publisher that pulls everything together and does whatever is required to deliver the final product to the readers. This all revolves around the editor, and John Wait is simply the best there is. Working with John and the rest of the professionals at Addison-Wesley is a pleasure. Their professionalism and attention to detail show in the end result.

Camera-ready copy of the book was produced by the author, a Troff die-hard, using the Groff package written by James Clark.

Tucson, Arizona
October 1993

W. Richard Stevens

Contents

Foreword	v
Preface to the Second Edition	vii
Adapted Preface to the First Edition	xiii
Chapter 1 Introduction	1
1.1 Architectural Principles	2
1.1.1 Packets, Connections, and Datagrams	3
1.1.2 The End-to-End Argument and Fate Sharing	6
1.1.3 Error Control and Flow Control	7
1.2 Design and Implementation	8
1.2.1 Layering	8
1.2.2 Multiplexing, Demultiplexing, and Encapsulation in Layered Implementations	10
1.3 The Architecture and Protocols of the TCP/IP Suite	13
1.3.1 The ARPANET Reference Model	13
1.3.2 Multiplexing, Demultiplexing, and Encapsulation in TCP/IP	16
1.3.3 Port Numbers	17
1.3.4 Names, Addresses, and the DNS	19
1.4 Internets, Intranets, and Extranets	19
1.5 Designing Applications	20
1.5.1 Client/Server	20
1.5.2 Peer-to-Peer	21
1.5.3 Application Programming Interfaces (APIs)	22