



Internetworking with TCP/IP, Volume 1

Principles, Protocols, and Architecture (Fifth Edition)

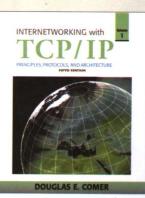
[美] Douglas E. Comer 著

TCP/IP

网络互连卷I原理、协议和体系结构(第5版)

(英文版)









典藏原版书苑

TCP/IP 网络互连 卷 I

原理、协议和体系结构

(第5版)(英文版)

Internetworkh

Principles, Protocols and Architectura (Firm Edition)

and Arabbect + CEITHE Edition

枫 力 早

Douglas E. Comer

人民邮电出版社

图书在版编目(CIP)数据

TCP/IP 网络互连. 第 1 卷, 原理、协议和体系结构: 第 5 版 / (美) 库默 (Comer, D. E.) 著. 一北京: 人民邮电出版社, 2006.11

国外著名高等院校信息科学与技术优秀教材

ISBN 7-115-15272-1

I.T... II. 库... III. 计算机网络—通信协议—高等学校—教材—英文 IV. TN915.04 中国版本图书馆 CIP 数据核字(2006)第 107545 号

版权声明

Original edition, entitled Internetworking With TCP/IP, Volume I——Principles, Protocols, and Architecture, 5th Edition, 0131876716 by Douglas E. Comer, published by Pearson Education, Inc, publishing as Prentice Hall, Copyright © 2006 Pearson Education, Inc.

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.

China edition published by PEARSON EDUCATION ASIA LTD., and POSTS & TELECOMMUNICATIONS PRESS Copyright \bigcirc 2006.

This edition is manufactured in the People's Republic of China, and is authorized for sale only in People's Republic of China excluding Hong Kong, Macau and Taiwan.

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。

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。

典藏原版书苑

TCP/IP 网络互连 卷 I 原理、协议和体系结构 (第 5 版) (英文版)

- ◆ 著 [美] Douglas E. Comer 责任编辑 王 琳
- ◆ 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 北京顺义振华印刷厂印刷 新华书店总店北京发行所经销

◆ 开本: 800×1000 1/16

印张: 42.5

字数: 904 千字

2006年11月第1版

印数: 1-4000 册

2006年11月北京第1次印刷

著作权合同登记号 图字: 01-2006-5378 号

ISBN 7-115-15272-1/TP • 5694

定价: 69.00 元

读者服务热线: (010)67132705 印装质量热线: (010)67129223

内容提要

本书是《TCP/IP 网络互连 卷 I 原理、协议和体系结构》这一经典计算机网络教科书的第 5 版,全面介绍了 TCP/IP 技术和因特网的体系结构。在阅读本书之后,你将了解多个物理网络是如何相互连接成一个协调的系统、网际协议在这个环境中如何运作以及应用程序如何使用产生的系统。

全书分为 4 个部分: 第 1 章~第 2 章对网络技术进行概述和讨论,深入讨论物理网络硬件; 第 3 章~第 12 章从单一主机的角度描述 TCP/IP 因特网,解释主机包含的协议及其操作方式,讨论了因特网寻址和路由的基本知识和协议层的概念; 第 13 章~第 19 章和第 31 章从全局的视角描述因特网的体系结构,讨论路由体系结构和路由器用来交换路由信息的协议; 最后,第 20 章~第 30 章讨论因特网上可用的应用层服务。

和以前的版本相比,本书各章都进行了更新以反映协议的最新版本和最新的技术,并且删减了陈旧技术的相关内容,讨论的重点转移到当前因特网上使用的协议和技术上。

本书可供计算机和通信专业的研究生、高年级本科生学习参考,也可供从事科研和技术开发的人员阅读。它是任何一个想要了解网络互连技术的人所必不可少的参考书。



为了回应世界各地的读者们要求推出新的版本的建议,尤其许多读者提出了需要强调的主题,我将讨论的重点转移到当前因特网上使用的协议和技术上,相应减少或删除了许多陈旧的内容。其中,讨论 IP 切换和 MPLS 的新的一章替换了讨论 ATM 的一章; 删减了关于因特网上已经不再使用的 RARP 和 BOOTP 等的内容; 删除了对 GGP 和现在已经废弃的路由仲裁器系统的讨论; 有状态寻址机制只作为历史进行讨论,对因特网完全使用无状态寻址进行了表述。

各章都进行了更新,以反映协议的最新版本和最新的技术。例如,讨论 TCP 的一章现在讨论选择性确认(SACK)、显式拥塞通知(ECN)和快速重传机制。另外,关于流量管理的新章节讨论了队列和调度机制,重写后的安全性一章讨论了有状态防火墙[†]。

与以前的版本一样,整本书主要关注网络互连的一般性概念,以及 TCP/IP 互连技术。网络互连是一个强大的抽象概念,使我们能够处理多种底层通信技术的复杂性。它隐藏了网络硬件的细节,提供了一个高层的通信环境。本书回顾了网络互连的体系结构和协议背后的原理,帮助读者了解这些体系结构和协议如何使相互连接的网络成为统一的单一通信系统。本书还解释了相互连网的通信系统如何有助于分布式计算。

在阅读本书之后,你将了解多个物理网络如何相互连接成一个协调的系统、网际协议在这个环境中如何运作以及应用程序如何使用该系统。作为一个具体示例,你将了解全球 TCP/IP 因特网的细节,包括它的路由器系统的体系结构和它支持的应用协议。另外,还会了解网络互连方式的一些局限性。

本书可以作为大学本科高年级或研究生的教材,也可以作为专业人员的参考书。对于专业人员,本书全面介绍了 TCP/IP 技术和因特网的体系结构。本书目的不在于替代协议标准文档,但它是学习网络互连的好起点,因为它对原理做了通用的概述。因此,它为读者提供了高层的"透视图",而这是很难从各个协议文档中获得的。

[†]除了其他改变之外,作者不情愿地将术语"数据报路由"改成了"数据包转发"。

如果在学校中使用,本书提供的内容超过了本科生或研究生网络课程一个学期的学习量。如果加上编程项目和阅读文献,课程可以扩展到两个学期。对于本科生课程,许多细节是不必要的。学生应该掌握本书描述的基本概念,他们应该能够描述或使用这些概念。对于研究生课程,学生应该以本书中的资料作为基础进行进一步的探索。他们应该详细理解细节,从而能够回答习题或解决需要探索扩展的细节问题。许多习题对细节之处做了提示,解决这些问题常常要求学生阅读协议标准并且应用创造力来理解结果。

无论是在任何级别的课程中,动手练习都有助于学生对概念形成直观认识。因此,我希望教师设计出实践项目,让学生在项目中使用因特网服务和协议。我在普度大学的研究生网络互连课程的学期项目要求学生构建一个 IP 路由器。我们提供硬件和操作系统的源代码,包括网络接口的设备驱动程序;学生构建一个可以运行的路由器,它能够将具有不同 MTU 的三个网络相互连接起来。这个课程非常有难度,学生分成小组工作,他们的成果让人印象深刻(许多行业新人都经过了这个课程的考验)。尽管这样的试验在与生产性计算设施隔离的实验室网络中进行是最安全的,但是我们发现,如果能够访问因特网,学生会表现出最大的热情并且获得最大的收益。

本书组织成 4 个主要部分: 第 1 章~第 2 章为简介部分,对网络技术进行概述和讨论,特别是对物理网络硬件进行了讨论,其目的是提供对网络硬件的基本认识,而不纠缠于硬件的细节;第 3 章~第 12 章从单一主机的角度描述 TCP/IP 因特网,解释主机包含的协议及其操作方式,讨论因特网寻址和路由的基本知识和协议层的概念;第 13 章~第 19 章和第 31 章从全局的视角描述因特网的体系结构,讨论路由体系结构和路由器用来交换路由信息的协议;最后,第 20 章~第 30 章讨论因特网上可用的应用层服务;讨论客户机-服务器交互模型,并且给出了客户机和服务器软件的案例。

本书内容的组织方式是自底向上的。首先概述硬件,然后在硬件上构建新功能。这种方式对于开发过因特网软件的人是很自然的,因为在实现软件时就是采用这种模式。在第 10 章之前,没有出现分层的概念。对分层的讨论主要强调功能的概念性层与分层的协议软件(其中每层上出现多个对象)之间的区别。

理解本书需要一定的背景知识。读者应该基本了解计算机系统,熟悉堆栈、队列和树等数据结构。读者需要基本了解计算机软件在支持并发编程的操作系统中的组织方式,以及执行计算所调用的应用程序。读者不需要掌握高级的数学知识,也不需要了解信息论或数据通信原理:本书把物理网络描述成黑盒,可以在它们上面进行网络互连。这样可以清楚地阐述设计原理,并且讨论意图和结果。

我要感谢为本书的各个版本做出过贡献的所有人。Dan Ardelean 和 Max Martynov 为这个版本提供了许多帮助,包括阅读 RFC 并且提出更新建议。Dave Roberts 和 Fernando Lichtschein 阅读了早期的书稿。特别感谢我的妻子和伙伴 Chris,她仔细地编辑了全书并且提出了许多改进意见。

Douglas E. Comer 2005年5月

Foreword

Here, better than ever, is the fifth edition of a landmark book, the book that signaled the coming of age of the Internet.

Development of the protocols for the Internet started around 1974. The protocols had been in limited but real use starting in the early 80's, but as of 1987, there was still no good introduction to how they worked or how to code them. The standards documents for TCP, IP and the other protocols existed, of course, but the true truth — the collection of knowledge and wisdom necessary to implement a protocol stack and actually expect it to work — was a mystery, known only to a small band of the initiated. This lack of knowledge was not a good thing, and the initiated knew it. But it takes a lot of effort to pull all the right stuff together and write it down. We waited, knowing that a good book explaining TCP/IP would be an important step towards the broad acceptance of our protocols.

And then Doug wrote the book. The Internet took a big step on its way from a small guild to a global community.

Of course, knowing that this was a landmark book back then is not enough to make you buy it now. Collectors might want to find the first edition, but 18 years ago is a long time in Internet years, and a lot has changed since then. We have learned a lot more, the field has grown up, whole new protocols have emerged, and Doug has rewritten the book four times. That is a measure both of how much and how fast the field changes, and how much work must go into keeping this book current. This book has all the new stuff, and our best current knowledge about all the old stuff. It is much more than simply TCP and IP. The book will give you an introduction to network technology like Ethernet, the design principles of the Internet, addressing and routing, programming over TCP, and examples of applications like email and the world wide web. This book has been updated with such things as an introduction to IP switching and MPLS, an updated discussion of mobile IP, private network interconnection, and secure alternatives to older protocols. Lots of sections have been revised to make sure that the reader gets the very latest information, as well as the basic understanding that has always been the hallmark of this book.

Other things have changed since the first edition. Not only has the Internet grown up, but some of our heroes have grown old, and some have died. The foreword to the first edition was written by Jon Postel, one of the true Internet pioneers, who died in the fall of 1998. Below, we have reprinted the foreword he wrote for the first edition. Much is the same, but much has changed. In 1987, Jon wrote "Computer communication systems and networks are currently separated and fragmented. The goal of inter-

2 Foreword

connection and internetworking, to have a single powerful computer communication network, is fundamental to the design of TCP/IP." Only 18 years ago, networks were fragmented; today, the Internet unites the world. And TCP/IP is still the glue at the core of the Internet that makes all this work. And this is still the book to read to learn about it.

David Clark Massachusetts Institute of Technology

May, 2005

Foreword To The First Edition By The Late Jon Postel

In this book, Professor Douglas Comer has provided a long sought overview and introduction to TCP/IP. There have been many requests for "the" article, report, or book to read to get started on understanding the TCP/IP protocols. At last, this book satisfies those requests. Writing an introduction to TCP/IP for the uninitiated is a very difficult task. While combining the explanation of the general principles of computer communication with the specific examples from the TCP/IP protocol suite, Doug Comer has provided a very readable book.

While this book is specifically about the TCP/IP protocol suite, it is a good book for learning about computer communications protocols in general. The principles of architecture, layering, multiplexing, encapsulation, addressing and address mapping, routing, and naming are quite similar in any protocol suite, though, of course, different in detail (See Chapters 3, 10, 17, and 18)†. Computer communication protocols do not do anything themselves. Like operating systems, they are in the service of applications processes. Processes are the active elements that request communication and are the ultimate senders and receivers of the data transmitted. The various layers of protocols are like the various layers in a computer operating system, especially the file system. Understanding protocol architecture is like understanding operating system architecture. In this book, Doug Comer has taken the "bottom up" approach — starting with the physical networks and moving up in levels of abstraction to the applications.

Since application processes are the active elements using the communication supported by the protocols, TCP/IP is an "interprocess communication" (IPC) mechanism. While there are several experiments in progress with operating system style message passing and procedure call types of IPC based on IP, the focus in this book is on more traditional applications that use the UDP datagram or TCP logical connection forms of IPC (See Chapters 11, 12, 17, 18, and 19).

One of the key ideas inherent in TCP/IP and in the title of this book is "internet-working." The power of a communication system is directly related to the number of entities in that system. The telephone network is very useful because (nearly) all of the telephones are in (as it appears to the users) one network. Computer communication systems and networks are currently separated and fragmented. The goal of interconnection and internetworking, to have a single powerful computer communication network, is fundamental to the design of TCP/IP. Essential to internetworking is addressing (See

[†]Editor's note: chapter numbers have changed since the first edition.

Chapters 4, 5, and 6), and a universal protocol — the Internet Protocol (See Chapters 7, 8, and 9).

To have an internetwork the individual networks must be connected. The connecting devices are called gateways. Further, these gateways must have some procedures for forwarding data from one network to the next. The data is in the form of IP datagrams and the destination is specified by an IP address, but the gateway must make a routing decision based on the IP address and what it knows about the connectivity of the networks making up the Internet. The procedures for distributing the current connectivity information to the gateways are called routing algorithms, and these are currently the subject of much study and development (See Chapters 13, 14, 15, and 16).

Like all communication systems, the TCP/IP protocol suite is an unfinished system. It is evolving to meet changing requirements and new opportunities. Thus, this book is, in a sense, a snapshot of TCP/IP circa 1987. And, as Doug Comer points out, there are many loose ends (See Chapter 20).

Most chapters end with a few pointers to material "for further study." Many of these refer to memos of the RFC series of notes. This series of notes is the result of a policy of making the working ideas and the protocol specifications developed by the TCP/IP research and development community widely available. This availability of the basic and detailed information about these protocols, and the availability of the early implementations of them, has had much to do with their current widespread use. This commitment to public documentation at this level of detail is unusual for a research effort, and has had significant benefits for the development of computer communication (See Appendix 3).

This book brings together information about the various parts of the TCP/IP architecture and protocols and makes it accessible. Its publication is a very significant milestone in the evolution of computer communications.

Jon Postel, Internet Protocol Designer and Deputy Internet Architect

December, 1987

What Others Have Said About The Fifth Edition Of Internetworking With TCP/IP

"This is the book I go to for clear explanations of the basic principles and latest developments in TCP/IP technologies. It's a 'must have' reference for networking professionals."

Dr. Ralph Droms Cisco Systems Chair of the DHCP working group

"When the Nobel committee turns its attention to the Internet, Doug gets the prize for literature. The new fifth edition of this classic is the best way to master Internet technology."

Dr. Paul V. Mockapetris
Inventor of the Domain Name System

"The best-written TCP/IP book I have ever read. Dr. Comer explains complex ideas clearly, with excellent diagrams and explanations. With this edition, Dr. Comer makes this classic textbook contemporary."

Dr. John Lin, Bell Laboratories

"The Internet continues to evolve and so does Comer's comprehensive coverage of the protocol architecture of this global communication system. As the convergence of voice and data networks continues apace, understanding the Internet is all the more important for the 21st century telecommunications engineer. The fifth edition of Comer's classic continues to serve as the premiere guide."

Dr. Vinton Cerf SVP Technology Strategy, MCI and co-inventor of TCP/IP "Internetworking with TCP/IP has always been the definitive reference for the Internet's key technologies. This new edition is an important update that confirms Doug Comer's reputation for clear and accurate presentation of essential information for every Internet professional."

Dr. Lyman Chapin, Interisle Consulting Group Former IAB Chair

"One of the greatest books I have read. True genius is when you are not only fluent in your field, but can get your point across simply. Thank you Dr. Comer for writing a great book!"

Marvin E. Miller CIO, The ACS Corporation

"... a true masterpiece."

Mr. Javier Sandino Systems Engineer

About The Author

Douglas Comer, Distinguished Professor of Computer Science at Purdue University and Visiting Faculty at Cisco Systems Inc., is an internationally recognized expert on computer networking, the TCP/IP protocols, and the Internet. The author of numerous refereed articles and technical books, he is a pioneer in the development of curriculum and laboratories for research and education.

A prolific author, Comer's popular books have been translated into 16 languages, and are used in industry as well as computer science, engineering, and business departments around the world. His landmark three-volume series *Internetworking With TCP/IP* revolutionized networking and network education. His textbooks and innovative laboratory manuals have and continue to shape graduate and undergraduate curricula.

The accuracy and insight of Dr. Comer's books reflect his extensive background in computer systems. His research spans both hardware and software. He has created a complete operating system, written device drivers, and implemented network protocol software for conventional computers as well as network processors. The resulting software has been used by industry in a variety of products.

Comer has created and teaches courses on network protocols and computer technologies for a variety of audiences, including courses for engineers as well as academic audiences. His innovative educational laboratories allow him and his students to design and implement working prototypes of large, complex systems, and measure the performance of the resulting prototypes. He continues to teach at industries, universities, and conferences around the world. In addition, Comer consults for industry on the design of computer networks and systems.

For over eighteen years, Professor Comer has served as editor-in-chief of the research journal *Software* — *Practice and Experience*. He is a Fellow of the ACM, a Fellow of the Purdue Teaching Academy, and a recipient of numerous awards, including a Usenix Lifetime Achievement award.

Additional information can be found at:

www.cs.purdue.edu/people/comer

and information about Comer's books can be found at:

www.comerbooks.com

Contents

Cnapte	r 1 Introduction And Overview	1
1.1	The Motivation For Internetworking 1	
1.2	The TCP/IP Internet 2	
1.3	Internet Services 2	
1.4	History And Scope Of The Internet 5	
1.5	The Internet Architecture Board 7	
1.6	The IAB Reorganization 7	
1.7	Internet Request For Comments 8	
1.8	Future Growth And Technology 9	
1.9	Organization Of The Text 10	
1.10	Summary 10	
Chapte	r 2 Review Of Underlying Network Technologies	13
2.1	Introduction 13	
2.2	Two Approaches To Network Communication 14	
2.3	Wide Area And Local Area Networks 15	
2.4	Ethernet Technology 16	
2.5	Switched Ethernet 24	
2.6	Asynchronous Transfer Mode 25	
2.7	Summary 27	
Chapte	r 3 Internetworking Concept And Architectural Model	31
		J 1
3.1	Introduction 31	
3.2	Application-Level Interconnection 31	
3.3	Network-Level Interconnection 32	

Co	

3.4 3.5 3.6 3.7 3.8 3.9 3.10	Properties Of The Internet 33 Internet Architecture 34 Interconnection Through IP Routers 35 The User's View 36 All Networks Are Equal 37 The Unanswered Questions 38 Summary 38	
Chapte	r 4 Classful Internet Addresses	41
4.1	Introduction 41	
4.2	Universal Identifiers 41	
4.3	The Original Classful Addressing Scheme 42	
4.4	Addresses Specify Network Connections 43	
4.5	Network And Directed Broadcast Addresses 43	
4.6	Limited Broadcast 44	
4.7	The All-0s Address 45	
4.8	Subnet And Classless Extensions 45	
4.9	IP Multicast Addresses 45	
4.10	Weaknesses In Internet Addressing 46	
4.11	Dotted Decimal Notation 47	
4.12	Loopback Address 48	
4.13	Summary Of Special Address Conventions 48	
4.14	Internet Addressing Authority 48	
4.15	Reserved Address Prefixes 50	
4.16	An Example 50	
4.17	Network Byte Order 52	
4.18	Summary 52	
Chapte	er 5 Mapping Internet Addresses To Physical Addresses (ARP)	57
5.1	Introduction 57	
5.2	The Address Resolution Problem 57	
5.3	Two Types Of Physical Addresses 58	
5.4	Resolution Through Direct Mapping 58	
5.5	Resolution Through Dynamic Binding 59	
5.6	The Address Resolution Cache 60	
5.7	ARP Cache Timeout 61	
5.8	ARP Refinements 62	
5.9	Relationship Of ARP To Other Protocols 62	
5.10	ARP Implementation 62	
5.11	ARP Encapsulation And Identification 64	

5.12	ARP Protocol Format 64	
<i>5.13</i>	Automatic ARP Cache Revalidation 66	
5.14	Reverse Address Resolution (RARP) 66	
5.15	Summary 67	
Chapte	r 6 Internet Protocol: Connectionless Datagram Delivery (IPv4)	71
6.1	Introduction 71	
6.2	A Virtual Network 71	
6.3	Internet Architecture And Philosophy 72	
6.4	The Conceptual Service Organization 72	
6.5	Connectionless Delivery System 73	
6.6	Purpose Of The Internet Protocol 73	
6.7	The IPv4 Datagram 74	
6.8	Internet Datagram Options 83	
6.9	Summary 89	
Chapte	r 7 Internet Protocol: Forwarding IP Datagrams	93
7.1	Introduction 93	
7.2	Forwarding In An Internet 93	
<i>7.3</i>	Direct And Indirect Delivery 95	
7.4	Table-Driven IP Forwarding 97	
7.5	Next-Hop Forwarding 97	
7.6	Default Routes 100	
7.7	Host-Specific Routes 100	
<i>7.8</i>	The IP Forwarding Algorithm 100	
7.9	Forwarding With IP Addresses 101	
7.10	Handling Incoming Datagrams 103	
7.11	Establishing Routing Tables 104	
7.12	Summary 104	
Chapte	r 8 Internet Protocol: Error And Control Messages (ICMP)	109
8.1	Introduction 109	
8.2	The Internet Control Message Protocol 109	
8.3	Error Reporting Vs. Error Correction 110	
<i>8.4</i>	ICMP Message Delivery 111	
8.5	ICMP Message Format 112	
8.6	Testing Destination Reachability And Status (Ping) 113	
<i>8.7</i>	Echo Request And Reply Message Format 114	

8.8	Reports Of Unreachable Destinations 115	
8.9	Congestion And Datagram Flow Control 116	
8.10	Source Quench Format 117	
8.11	Route Change Requests From Routers 118	
<i>8.12</i>	Detecting Circular Or Excessively Long Routes 119	
<i>8.13</i>	Reporting Other Problems 120	
8.14	Clock Synchronization And Transit Time Estimation 121	
<i>8.15</i>	Older ICMP Messages No Longer Needed 122	
8.16	Summary 123	
Chapte	er 9 Classiess And Subnet Address Extensions (CIDR)	127
9.1	Introduction 127	
9.2	Review Of Relevant Facts 127	
9.3	Minimizing Network Numbers 128	
9.4	Proxy ARP 128	
9.5	Subnet Addressing 130	
9.6	Flexibility In Subnet Address Assignment 132	
9.7	Variable-Length Subnets 134	
9.8	Implementation Of Subnets With Masks 135	
9.9	Subnet Mask Representation 136	
9.10	Forwarding In The Presence Of Subnets 137	
9.11	The Subnet Forwarding Algorithm 138	
9.12	A Unified Forwarding Algorithm 139	
9.13	Maintenance Of Subnet Masks 140	
9.14	Broadcasting To Subnets 140	
9.15	Anonymous Point-To-Point Networks 141	
9.16	Classless Addressing And Supernetting 142	
9.17	CIDR Address Blocks And Bit Masks 143	
9.18	Address Blocks And CIDR Notation 144	
9.19	A Classless Addressing Example 145	
9.20	Data Structures And Algorithms For Classless Lookup 145	
9.21	Longest-Match And Mixtures Of Route Types 148	
9.22	CIDR Blocks Reserved For Private Networks 149	
9.23	Summary 150	
Chapte	r 10 Protocol Layering	155
10.1	Introduction 155	
10.2	The Need For Multiple Protocols 155	
10.3	The Conceptual Layers Of Protocol Software 156	
10.4	Functionality Of The Layers 159	