INTERNETWORKING WITH TCP/IP

PRINCIPLES, PROTOCOLS, AND ARCHITECTURE



Prentice-Hall International Editions

DOUGLAS COMER

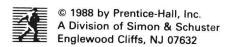
Internetworking With TCP/IP Principles, Protocols, and Architecture

DOUGLAS E. COMER

Department of Computer Sciences Purdue University West Lafayette, IN 47907



This edition may be sold only in those countries to which it is consigned by Prentice-Hall International. It is not to be re-exported and it is not for sale in the U.S.A., Mexico, or Canada.



UNIX is a registered trademark of AT&T Bell Laboratories. proNET-10 is a trademark of Proteon Corporation.

VAX, Microvax, and LSI 11 are trademarks of Digital Equipment Corporation

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

10 9 8 7 6 5 4 3 2

Prentice-Hall International (UK) Limited, London Prentice-Hall of Australia Pty. Limited, Sydney Prentice-Hall Canada Inc., Toronto Prentice-Hall Hispanoamericana, S.A., Mexico Prentice-Hall of India Private Limited, New Delhi Prentice-Hall of Japan, Inc., Tokyo Simon & Schuster Asia Pte. Ltd., Singapore Editora Prentice-Hall do Brasil, Ltda., Rio de Janeiro Prentice-Hall, Inc., Englewood Cliffs, New Jersey

Foreword

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 application 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 "internetworking." 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 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 Chapters 4, 5, and 6), and a universal protocol – the Internet Protocol (See Chapters 7, 8, and 9).

xiv Foreword

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

Preface

In the last century, railroads revolutionized the world by providing a transportation network that moved raw materials and manufactured products. They made an industrialized society possible. Digital communication networks have started a new revolution by providing the technology that transports the data needed by a society in which information plays a key role. Networking already permeates industry, education, and government. It has already begun to change the way we view the world by shrinking geographic distances and forming new communities of people who interact frequently. More important, network growth is explosive. The revolution is well underway.)

To understand networking and the selection of topics discussed in this book, it is important to realize that network research and development occurred in three stages. Before the 1960s, the main question was, "How can we transmit bits across a communication medium efficiently and reliably?" The results include the development of information theory, the sampling theorem, and other ideas commonly referred to as signal processing. Beginning around the mid 1960s, emphasis shifted to packet switching and the question became, "How can we transmit packets across a communication medium efficiently and reliably?" The results include the development of packet switching technologies, local area networks, and statistical analysis of network response to load. From approximately the mid 1970s to the present, emphasis has centered on network architecture and the question, "How can we provide communication services across a series of interconnected networks?" The results include the development of internetwork technologies, protocol layering models, datagram and stream transport services, and the client-server interaction paradigm.

Most textbooks and network courses concentrate on the first two stages of network research, presenting the well-known theories of data communications and queueing analysis. Although such information is important to engineers who design network technologies and hardware products, most network architects purchase commercially available network hardware. Instead of detailed knowledge about how bits or packets flow across communication media, they need to know how to interconnect such hardware and how to use the resulting system.

This text concentrates on the third stage of networking. It examines the architecture of interconnected networks and explains the principles and protocols that make such interconnected architectures function as a single unified communication system. More important, it shows how an interconnected architecture can be used for distributed computation.

xvi Preface

The entire text focuses on the concept of internetworking in general and the TCP/IP internet technology in particular. Internetworking is a powerful abstraction that allows us to deal with the complexity of multiple underlying communication technologies. It hides the details of network hardware and provides a high level communication environment. As the book shows, the ultimate goal of internetworking is maximal interoperability, that is, maximizing the ability of programs on diverse computer and network systems to communicate reliably and efficiently.

The text reviews both the architecture of network interconnections as well as internet communication services and the protocols needed to provide those services. By the end of the book, the reader will understand how it is possible to interconnect multiple physical networks into a coordinated system, how internet protocols operate in that environment, and how application programs use the resulting system. As a specific example, the reader learns the details of the DARPA (TCP/IP) Internet, including the architecture of the gateway system and the application protocols it supports. In addition, the book discusses some of the limitations of the internet approach.

Writing about internetworking is both exciting and challenging. It is challenging because, as in any rapidly changing research area, nothing is stable. It is exciting because the TCP/IP Internet is an active, rapidly expanding entity. Researchers working on it generate new ideas constantly and the possibilities seem endless. Looking back over TCP/IP and the DARPA Internet makes it clear that much has been accomplished. Knowing that the research has taken a little over a decade makes one realize how intense the effort has been.

Designed as both a college text and as a professional reference, the book is written at an advanced undergraduate or graduate level. For professionals, the book provides a comprehensive introduction to the TCP/IP technology and the architecture of the Internet. Although it is not intended to replace protocol standards, the book is a good starting point for learning about internetworking because it provides a uniform overview that emphasizes principles. Moreover, it gives the reader perspective that can be extremely difficult to obtain from individual protocol documents.

When used in the classroom, the text provides more than sufficient material for a single semester network course at either the undergraduate or graduate level. Such a course can be extended to a 2-semester sequence if accompanied by programming projects and readings from the literature. For undergraduate courses, it can be taken at face value. Students should be expected to grasp the basic concepts described in the text, and they should be able to describe or use them. At the graduate level, students should be expected to use the material here as a basis for further exploration of current research. They should understand it well enough to answer exercises or solve problems that require them to explore subtleties and consequences. Many of the exercises suggest such subtleties; solving them often requires students to read protocol standards and apply creative energy to comprehend consequences.

At all levels, hands-on experience sharpens the concepts and helps students gain intuition. Thus, I encourage instructors to invent projects that force students to use internetwork services and protocols. Although such experimentation is safest when the instructional laboratory network is isolated from production computing facilities, we have found that students exhibit the most enthusiasm, and benefit the most, when they have access to the "real" TCP/IP Internet.

The book is organized into four main parts. Chapters 1 and 2 form an introduction that provides an overview and discusses existing technologies. In particular, Chapter 2 reviews physical network hardware. The intention is to provide basic intuition about what is possible, not to spend inordinate time on hardware details. Chapters 3-12 describe the TCP/IP Internet from the viewpoint of a single host, showing the basic services available and the protocols a host uses to access them. They cover the basics of Internet addressing and routing as well as the notion of protocol layering. Chapters 13-16 describe the architecture of the Internet when viewed globally. They explore the core gateway system and the protocols gateways use to exchange routing information. Finally, Chapters 17-19 discuss application level services available in the Internet. They present the client-server model of interaction and give several examples of how one can organize client and server software. The last section discusses electronic mail and the domain name system, two topics that are extremely popular.

The chapters have been organized "bottom up." They begin with an overview of hardware and continue to build new functionality on top of it. This view will appeal to anyone who has developed Internet software because it follows the same pattern one uses in implementation. The concept of layering does not appear until Chapter 10. The discussion of layering emphasizes the distinction between conceptual layers of functionality and the reality of layered protocol software in which multiple objects appear at each layer.

Although it is difficult to omit any chapter completely, the instructor will find that students are often satisfied to know that something is possible without knowing the details. For example, one can skim through Chapters 5, 6, and 9 by covering only the functionality and not the details of the protocols. In addition, several chapters (especially 16) contain engineering techniques. While such techniques are crucial to efficient implementations, they can be skipped to save time.

A modest background is required to understand the material. The reader is expected to have programmed in a high level language and to be familiar with basic data structures like stacks, queues, and trees. Readers need basic intuition about the organization of computer software into an operating system that supports concurrent programming and application programs that users invoke to perform computation. Readers do not need sophisticated mathematics, nor do they need to know information theory or theorems from data communications; the book describes the physical network as a black box around which an internetwork can be built. It states design principles in English and discusses motivations and consequences.

Many people have contributed to this book. I thank Steve Chapin, Jim Griffioen, Chris Kent, Tim Korb, Dan Lynch, Thomas Narten, Shawn Ostermann, John Steele, and Raj Yavatkar, who all read drafts and made valuable comments. Craig Partridge supplied numerous suggestions, including a few exercises, and corrected several technical errors. He and Van Jacobson supplied the graph of Internet round trip delays in Chapter 12. Barry Shein graciously allowed me to use his example UNIX client and server code in Appendix 1. Charlotte Tubis provided valuable editing. Special thanks go to my wife, Chris, who has read the text more times than I can count and made extensive suggestions.

Contents

Foreword	xiii
Preface	xv
Chapter 1 Introduction and Overview	1
1.1 The Need For An Internet 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 Activities Board 7 1.6 Internet Research And Request For Comments 1.7 Internet Protocols and Standardization 8 1.8 Future Growth 8 1.9 Organization Of This Text 9 1.10 Summary 9	7
Chapter 2 Review of Underlying Network Tech	nologies 11
2.1 Introduction 11 2.2 Two Approaches To Network Communication 1 2.3 Long Haul And Local Area Networks 12 2.4 Ethernet Technology 13 2.5 ProNET Ring Technology 19 2.6 ARPANET Technology 22 2.7 National Science Foundation Networking 25 2.8 X25NET 27 2.9 Cypress Network 28 2.10 Summary And Conclusion 29	12

Chapter 3 Internetworking Concept and Architectural Model	31
3.1 Introduction 31 3.2 Application-Level Interconnection 31 3.3 Network-Level Interconnection 32 3.4 Properties Of The Internet 33 3.5 Internet Architecture 33 3.6 Interconnection Through Gateways 34 3.7 The User's View 35 3.8 The Unanswered Questions 36 3.9 Summary 37	
Chapter 4 Internet Addresses	39
4.1 Introduction 39 4.2 Universal Identifiers 39 4.3 Three Primary Classes Of Internet Addresses 40 4.4 Addresses Specify Network Connections 41 4.5 Network And Broadcast Addresses 41 4.6 Interpreting Zero To Mean "This" 42 4.7 Weaknesses In Internet Addressing 42 4.8 Dotted Decimal Notation 43 4.9 Internet Addressing Authority 44 4.10 An Example 44 4.11 Network Byte Order 46 4.12 Summary 47	
Chapter 5 Mapping Internet Addresses to Physical Addresses (ARP) 5.1 Introduction 49 5.2 The Address Resolution Problem 49 5.3 Two Types Of Physical Addresses 50 5.4 Resolution Through Direct Mapping 50 5.5 Resolution Through Dynamic Binding 51 5.6 The Address Resolution Cache 52 5.7 ARP Refinements 53 5.8 Relationship Of ARP To Other Protocols 53 5.9 ARP Implementation 53 5.10 ARP Encapsulation And Identification 54 5.11 ARP Protocol Format 55 5.12 Summary 56	49

Chapter 6 Determining an Internet Address at Startup (RARP)	59
 6.1 Introduction 59 6.2 Reverse Address Resolution Protocol (RARP) 60 6.3 Timing RARP Transactions 62 6.4 Primary And Backup RARP Servers 62 6.5 Summary 63 	
Chapter 7 Internet Protocol: Connectionless Datagram Delivery (IP)	65
 7.1 Introduction 65 7.2 A Virtual Network 65 7.3 Internet Architecture And Philosophy 66 7.4 The Concept Of Unreliable Delivery 66 7.5 Connectionless Delivery System 67 7.6 Purpose Of The Internet Protocol 67 7.7 The Internet Datagram 67 7.8 Internet Datagram Options 72 7.9 Summary 76 	
Chapter 8 Routing IP Datagrams	79
 8.1 Introduction 79 8.2 Routing In The Internet 79 8.3 Datagram Delivery Over A Single Network 80 8.4 Indirect Routing 81 8.5 Table-Driven IP Routing 81 8.6 Default Routes 83 8.7 Host-Specific Routes 83 8.8 The Final Algorithm 84 8.9 Handling Incoming Datagrams 84 8.10 Establishing Routing Tables 85 8.11 Summary 85 	
Chapter 9 Internet Protocol: Error and Control Messages (ICMP)	89
 9.1 Introduction 89 9.2 The Internet Control Message System 89 9.3 ICMP Message Format And Delivery 90 9.4 Testing Destination Reachability And Status 92 	

Contents

viii

 9.5 Reports Of Unreachable Destinations 92 9.6 Datagram Flow Control 93 9.7 Route Change Requests From Gateways 94 9.8 Detecting Circular Or Excessively Long Routes 95 9.9 Reporting Incorrect Datagram Headers 96 9.10 Clock Synchronization And Transit Time Estimation 96 9.11 Obtaining A Network Address 97 9.12 Obtaining A Subnet Mask 97 9.13 Summary 98 	
Chapter 10 Protocol Layering	10
10.1 Introduction 101 10.2 The Need For Multiple Protocols 101 10.3 The Conceptual Layers Of Protocol Software 102 10.4 Functionality Of The Layers 105 10.5 X.25 And Its Relation To The ISO Model 106 10.6 Differences Between ISO And Internet Layering 109 10.7 The Protocol Layering Principle 110 10.8 Layering In The Presence Of Network Substructure 113 10.9 The Disadvantage Of Layering 114 10.10 The Basic Idea Behind Multiplexing And Demultiplexing 115 10.11 ISO Connectionless Delivery Protocols 116 10.12 Summary 117	
Chapter 11 User Datagram Protocol 11.1 Introduction 119 11.2 Identifying The Ultimate Destination 119 11.3 The User Datagram Protocol 120 11.4 Format Of UDP Messages 121 11.5 UDP Encapsulation And Protocol Layering 122 11.6 Layering And The UDP Checksum Computation 124 11.7 UDP Demultiplexing 124 11.8 Reserved And Available UDP Port Numbers 125 11.9 Summary 126	119
Chapter 12 Reliable Stream Transport Service (TCP)	129
12.1 Introduction 129 12.2 The Need For Stream Delivery 129	

Contents

12.3 Application Interface To The Reliable Delivery Service 130	
12.4 Providing Reliability 131	
12.5 The Idea Behind Sliding Windows 133	
12.6 The Transmission Control Protocol 136	
12.7 TCP Ports And Connections 137	
12.8 Segments, Streams, And Sequence Numbers 138	
12.9 Variable Window Size And Flow Control 139	
12.10 TCP Segment Format 139	
12.11 TCP Checksum Computation 142	
12.12 Acknowledgements And Retransmission 142	
12.13 Timeout And Retransmission 143	
12.14 Response To Congestion 145	
12.15 Establishing A TCP Connection 145	
12.16 Closing a TCP Connection 147	
12.17 TCP Connection Reset 147	
12.18 Forcing Data Delivery 148	
12.19 Reserved TCP Port Numbers 148	
12.20 Summary 150	
Chapter 13 The Core Gateway System (GGP) 13.1 Introduction 153 13.2 The Origin Of Gateway Routing Tables 153 13.3 Routing With Partial Information 154 13.4 Core And Noncore Gateways 156 13.5 Gateway-To-Gateway Protocol (GGP) 157 13.6 GGP Message Formats 158 13.7 Alternatives To GGP 160 13.8 Summary 161	
Chapter 14 Autonomous Systems and Confederations (EGP)	163
14.1 Introduction 163	
14.2 Adding Complexity To The Architectural Model 163	
14.3 Autonomous System Concept 166	
14.4 Exterior Gateway Protocol 168	
14.5 EGP Message Header 168	
14.6 EGP Neighbor Acquisition Messages 170	
14.7 EGP Neighbor Reachability Messages 170	
14.8 EGP Poll Request Messages 172	
14.8 EGP Foll Request Messages 172 14.9 EGP Routing Update Messages 173 14.10 The Key Restriction Of EGP 175	

14.11 Problems Addressed In EGP2/EGP3 177 14.12 Decentralization Of The Internet Architecture 177 14.13 The Autonomous Confederation Concept 177 14.14 Summary 178	
Chapter 15 Interior Gateway Protocols (RIP, HELLO, GATED)	181
15.1 Introduction 181 15.2 Static Vs. Dynamic Interior Routes 181 15.3 Routing Information Protocol (RIP) 183 15.4 RIP Message Format 185 15.5 HELLO Protocol As An IGP 187 15.6 Combining RIP, Hello, And EGP 188 15.7 Routing With Partial Information 189 15.8 Summary 190	
Chapter 16 Transparent Gateways and Subnet Addressing	193
16.1 Introduction 193 16.2 Review Of Relevant Facts 193 16.3 Minimizing Network Numbers 194 16.4 Transparent Gateways 194 16.5 Proxy ARP 196 16.6 Subnet Addresses 197 16.7 Flexibility In Subnet Addressing 199 16.8 Implementation Of Subnets With Masks 200 16.9 Routing In The Presence Of Subnets 201 16.10 The Subnet Routing Algorithm 202 16.11 Unified Routing Algorithm 202 16.12 Maintenance Of Subnets Masks 203 16.13 Broadcasting To Subnets 203 16.14 Summary 203	
Chapter 17 Client-Server Model Of Interaction	207
 17.1 Introduction 207 17.2 The Client-Server Model 207 17.3 A Simple Example: UDP Echo Server 208 17.4 Time And Date Service 210 17.5 The Complexity of Servers 211 17.6 RARP Server 212 	

17.7 Alternatives To The Client-Server Model 212

17.8 Summary 213	
Chapter 18 The Domain Name System	215
 18.1 Introduction 215 18.2 Names For Machines 215 18.3 Flat Namespace 216 18.4 Hierarchical Names 217 18.5 Delegation Of Authority For Names 217 18.6 Subset Authority 218 18.7 Internet Domain Names 219 18.8 Mapping Domain Names To Addresses 221 	
18.9 Domain Name Resolution 222 18.10 Efficient Translation With Caching 223 18.11 Domain Server Message Format 224 18.12 Compressed Name Format 227 18.13 Abbreviation Of Domain Names 228 18.14 Inverse Mappings 228 18.15 Pointer Queries 229 18.16 Extensions To Other Types Of Names 229 18.17 Obtaining Authority For A Subdomain 230 18.18 Summary 230	
Chapter 19 Application Level Services	233
19.1 Introduction 233 19.2 Remote Interactive Computing 233 19.3 Internet TELNET Protocol 234 19.4 Rlogin (4.3 BSD UNIX) 236 19.5 File Transfer 237 19.6 Internet File Transfer Protocol (FTP) 237 19.7 Internet Trivial File Transfer Protocol 239 19.8 Electronic Mail 240 19.9 Mailbox Names And Aliases 240 19.10 The Relationship Of Internetworking And Mail 241 19.11 Internet Standard For Electronic Mail Service 243 19.12 Internet Simple Mail Transfer Protocol (SMTP) 244 19:13 Protocol Dependencies 246 19.14 Summary 247	

Chapter 20 Internet Research And Engineering Problems	249
 20.1 Introduction 249 20.2 The Forces Stimulating Evolution 250 20.3 Accommodating Expansion And Increased Load 251 20.4 New Applications 254 20.5 New Hardware And Communication Technologies 257 20.6 Changes To Accommodate New Groups 258 20.7 Summary 260 	
Appendix 1 4.3 BSD UNIX Interface to Internet Protocols	263
Appendix 2 Hints And Suggestions For Implementors	291
Appendix 3 A Guide To RFCs	299
Appendix 4 Glossary of Internet Terms and Abbreviations	327
Appendix 5 Official DARPA Internet Protocols	357
Bibliography	369
Index	375

Introduction and Overview

1.1 The Need For An Internet

Data communication has become a fundamental part of computing. World-wide networks gather data about such diverse subjects as atmospheric conditions, crop production, and airline traffic. Groups establish electronic mailing lists so they can share information of common interest. Hobbyists exchange programs for their home computers. In the scientific world, data networks are essential because they allow scientists to send programs and data to remote supercomputers for processing, to retrieve the results, and to exchange scientific information with colleagues.

Unfortunately, most networks are independent entities, established to serve the needs of a single group. The users choose a hardware technology appropriate to their communication problems. More important, it is impossible to build a universal network from a single hardware technology, because no single network suffices for all uses. Some users need a high-speed network to connect machines, but such networks cannot be expanded to span large distances. Others settle for a slower speed network that connects machines thousands of miles apart.

Recently, however, a new technology has emerged that makes it possible to interconnect many disparate physical networks and make them function as a coordinated unit. The new technology, called an *internetwork*, or *internet*, accommodates multiple, diverse underlying hardware technologies by adding both physical connections and a new set of conventions. The internet technology hides the details of network hardware and permits computers to communicate independent of their physical network connections.

To appreciate internet technology, think of how it affects research. Imagine for a minute the effects of interconnecting all the computers used by scientists. Any scientist would be able to exchange data resulting from an experiment with any other scientist. It would be possible to establish national data centers to collect data from natural phenome-