

Volume

1

INTERNETWORKING with

# TCP/IP

PRINCIPLES, PROTOCOLS, AND ARCHITECTURES

FOURTH EDITION



DOUGLAS E. COMER

---

# **Internetworking With TCP/IP**

## **Vol I:**

### **Principles, Protocols, and Architecture**

#### **Fourth Edition**

**DOUGLAS E. COMER**

*Department of Computer Sciences  
Purdue University  
West Lafayette, IN 47907*

**PRENTICE HALL**  
**Upper Saddle River, New Jersey 07458**

## Library of Congress Cataloging-in-Publication Data

Comer, Douglas

Internetworking with TCP/IP / Douglas E. Comer. -- 4th ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-13-018380-6

1. Principles, protocols, and architecture. 2. Client/server computing. 3. Internetworking (Telecommunications) I. Title

Publisher: *Alan Apt*

Project Manager: *Ana Arias Terry*

Editorial Assistant: *Toni Holm*

Vice-president and director of production and manufacturing, ESM: *David W. Riccardi*

Vice-president and editorial director of ECS: *Marcia Horton*

Executive Managing Editor: *Vince O'Brien*

Managing Editor: *David A. George*

Editorial/production supervision: *Irwin Zucker*

Art Director: *Heather Scott*

Assistant to Art Director: *John Christiana*

Manufacturing Buyer: *Pat Brown*

Marketing manager: *Danny Hoyt*



©2000, 1995 Prentice Hall

Prentice-Hall, Inc.

Upper Saddle River, New Jersey 07458

Prentice Hall books are widely used by corporations and government agencies for training, marketing, and resale. The publisher offers discounts on this book when ordered in bulk quantities.

For more information, contact Corporate Sales Department, Phone: 800-382-3419;

Fax: 201-236-7141; E-mail: [corpsales@prenhall.com](mailto:corpsales@prenhall.com)

Or write: Prentice Hall PTR, Corp. Sales Dept., One Lake Street, Upper Saddle River, NJ 07458.

UNIX is a registered trademark of UNIX System Laboratories, Incorporated

proNET-10 is a trademark of Proteon Corporation

LSI 11 is a trademark of Digital Equipment Corporation

Microsoft Windows is a trademark of Microsoft Corporation

EUI-64 is trademark of the Institute of Electrical and Electronics Engineers (IEEE)

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

ISBN 0-13-018380-6

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*

Pearson Education Asia Pte Ltd

Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

## Library of Congress Cataloging-in-Publication Data

Comer, Douglas  
Internetworking with TCP/IP / Douglas E. Comer. -- 4th ed.  
p. cm.  
Includes bibliographical references and index.  
ISBN 0-13-018380-6  
1. Principles, protocols, and architecture. 2. Client/server  
computing. 3. Internetworking (Telecommunications) I. Title

Publisher: *Alan Apt*  
Project Manager: *Ana Arias Terry*  
Editorial Assistant: *Toni Holm*  
Vice-president and director of production and manufacturing, ESM: *David W. Riccardi*  
Vice-president and editorial director of ECS: *Marcia Horton*  
Executive Managing Editor: *Vince O'Brien*  
Managing Editor: *David A. George*  
Editorial/production supervision: *Irwin Zucker*  
Art Director: *Heather Scott*  
Assistant to Art Director: *John Christiana*  
Manufacturing Buyer: *Pat Brown*  
Marketing manager: *Danny Hoyt*



©2000, 1995 Prentice Hall  
Prentice-Hall, Inc.  
Upper Saddle River, New Jersey 07458

Prentice Hall books are widely used by corporations and government agencies for training, marketing, and resale. The publisher offers discounts on this book when ordered in bulk quantities. For more information, contact Corporate Sales Department, Phone: 800-382-3419; Fax: 201-236-7141; E-mail: [corpsales@prenhall.com](mailto:corpsales@prenhall.com) Or write: Prentice Hall PTR, Corp. Sales Dept., One Lake Street, Upper Saddle River, NJ 07458.

UNIX is a registered trademark of UNIX System Laboratories, Incorporated  
proNET-10 is a trademark of Proteon Corporation  
LSI 11 is a trademark of Digital Equipment Corporation  
Microsoft Windows is a trademark of Microsoft Corporation  
EUI-64 is trademark of the Institute of Electrical and Electronics Engineers (IEEE)

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

ISBN 0-13-018380-6

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*  
Pearson Education Asia Pte Ltd  
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

## About The Author

Dr. Douglas Comer is an internationally recognized expert on TCP/IP protocols and the Internet. One of the researchers who contributed to the Internet as it was being formed in the late 1970s and 1980s, he was a member of the Internet Architecture Board, the group responsible for guiding the Internet's development. He was also chairman of the CSNET technical committee and a member of the CSNET executive committee.

Comer consults for companies on the design and implementation of networks, and gives professional seminars on TCP/IP and internetworking to both technical and nontechnical audiences around the world. His operating system, Xinu, and implementation of TCP/IP protocols are documented in his books, and used in commercial products.

Comer is a professor of computer science at Purdue University, where he teaches courses and does research on computer networking, internetworking, and operating systems. In addition to writing a series of best-selling technical books, he serves as the North American editor of the journal *Software — Practice and Experience*. Comer is a Fellow of the ACM.

Additional information can be found at:

[www.cs.purdue.edu/people/comer](http://www.cs.purdue.edu/people/comer)

# Contents

|                 |              |
|-----------------|--------------|
| <b>Foreword</b> | <b>xxiii</b> |
|-----------------|--------------|

|                |              |
|----------------|--------------|
| <b>Preface</b> | <b>xxvii</b> |
|----------------|--------------|

|                                            |          |
|--------------------------------------------|----------|
| <b>Chapter 1 Introduction And Overview</b> | <b>1</b> |
|--------------------------------------------|----------|

|      |                                               |    |
|------|-----------------------------------------------|----|
| 1.1  | <i>The Motivation For Internetworking</i>     | 1  |
| 1.2  | <i>The TCP/IP Internet</i>                    | 2  |
| 1.3  | <i>Internet Services</i>                      | 3  |
| 1.4  | <i>History And Scope Of The Internet</i>      | 6  |
| 1.5  | <i>The Internet Architecture Board</i>        | 8  |
| 1.6  | <i>The IAB Reorganization</i>                 | 9  |
| 1.7  | <i>The Internet Society</i>                   | 11 |
| 1.8  | <i>Internet Request For Comments</i>          | 11 |
| 1.9  | <i>Internet Protocols And Standardization</i> | 12 |
| 1.10 | <i>Future Growth And Technology</i>           | 12 |
| 1.11 | <i>Organization Of The Text</i>               | 13 |
| 1.12 | <i>Summary</i>                                | 14 |

|                                                            |           |
|------------------------------------------------------------|-----------|
| <b>Chapter 2 Review Of Underlying Network Technologies</b> | <b>17</b> |
|------------------------------------------------------------|-----------|

|     |                                                   |    |
|-----|---------------------------------------------------|----|
| 2.1 | <i>Introduction</i>                               | 17 |
| 2.2 | <i>Two Approaches To Network Communication</i>    | 18 |
| 2.3 | <i>Wide Area And Local Area Networks</i>          | 19 |
| 2.4 | <i>Ethernet Technology</i>                        | 20 |
| 2.5 | <i>Fiber Distributed Data Interconnect (FDDI)</i> | 33 |
| 2.6 | <i>Asynchronous Transfer Mode</i>                 | 37 |
| 2.7 | <i>WAN Technologies: ARPANET</i>                  | 38 |
| 2.8 | <i>National Science Foundation Networking</i>     | 40 |

- 2.9 *ANSNET* 44
- 2.10 *A Very High Speed Backbone (vBNS)* 45
- 2.11 *Other Technologies Over Which TCP/IP Has Been Used* 46
- 2.12 *Summary And Conclusion* 50

### **Chapter 3 Internetworking Concept And Architectural Model 53**

- 3.1 *Introduction* 53
- 3.2 *Application-Level Interconnection* 53
- 3.3 *Network-Level Interconnection* 54
- 3.4 *Properties Of The Internet* 55
- 3.5 *Internet Architecture* 56
- 3.6 *Interconnection Through IP Routers* 56
- 3.7 *The User's View* 58
- 3.8 *All Networks Are Equal* 58
- 3.9 *The Unanswered Questions* 59
- 3.10 *Summary* 60

### **Chapter 4 Classful Internet Addresses 63**

- 4.1 *Introduction* 63
- 4.2 *Universal Identifiers* 63
- 4.3 *The Original Classful Addressing Scheme* 64
- 4.4 *Addresses Specify Network Connections* 65
- 4.5 *Network And Directed Broadcast Addresses* 65
- 4.6 *Limited Broadcast* 66
- 4.7 *Interpreting Zero To Mean "This"* 67
- 4.8 *Subnet And Supernet Extensions* 67
- 4.9 *IP Multicast Addresses* 68
- 4.10 *Weaknesses In Internet Addressing* 68
- 4.11 *Dotted Decimal Notation* 69
- 4.12 *Loopback Address* 70
- 4.13 *Summary Of Special Address Conventions* 70
- 4.14 *Internet Addressing Authority* 71
- 4.15 *Reserved Address Prefixes* 72
- 4.16 *An Example* 72
- 4.17 *Network Byte Order* 74
- 4.18 *Summary* 75

**Chapter 5 Mapping Internet Addresses To Physical Addresses (ARP) 77**

- 5.1 Introduction 77
- 5.2 The Address Resolution Problem 77
- 5.3 Two Types Of Physical Addresses 78
- 5.4 Resolution Through Direct Mapping 78
- 5.5 Resolution Through Dynamic Binding 79
- 5.6 The Address Resolution Cache 80
- 5.7 ARP Cache Timeout 81
- 5.8 ARP Refinements 82
- 5.9 Relationship Of ARP To Other Protocols 82
- 5.10 ARP Implementation 82
- 5.11 ARP Encapsulation And Identification 84
- 5.12 ARP Protocol Format 84
- 5.13 Summary 86

**Chapter 6 Determining An Internet Address At Startup (RARP) 89**

- 6.1 Introduction 89
- 6.2 Reverse Address Resolution Protocol (RARP) 90
- 6.3 Timing RARP Transactions 92
- 6.4 Primary And Backup RARP Servers 92
- 6.5 Summary 93

**Chapter 7 Internet Protocol: Connectionless Datagram Delivery 95**

- 7.1 Introduction 95
- 7.2 A Virtual Network 95
- 7.3 Internet Architecture And Philosophy 96
- 7.4 The Conceptual Service Organization 96
- 7.5 Connectionless Delivery System 97
- 7.6 Purpose Of The Internet Protocol 97
- 7.7 The Internet Datagram 97
- 7.8 Internet Datagram Options 107
- 7.9 Summary 113

**Chapter 8 Internet Protocol: Routing IP Datagrams 115**

- 8.1 Introduction 115
- 8.2 Routing In An Internet 115
- 8.3 Direct And Indirect Delivery 117

|      |                                    |     |
|------|------------------------------------|-----|
| 8.4  | <i>Table-Driven IP Routing</i>     | 119 |
| 8.5  | <i>Next-Hop Routing</i>            | 119 |
| 8.6  | <i>Default Routes</i>              | 121 |
| 8.7  | <i>Host-Specific Routes</i>        | 121 |
| 8.8  | <i>The IP Routing Algorithm</i>    | 121 |
| 8.9  | <i>Routing With IP Addresses</i>   | 122 |
| 8.10 | <i>Handling Incoming Datagrams</i> | 124 |
| 8.11 | <i>Establishing Routing Tables</i> | 125 |
| 8.12 | <i>Summary</i>                     | 125 |

## **Chapter 9 Internet Protocol: Error And Control Messages (ICMP) 129**

|      |                                                           |     |
|------|-----------------------------------------------------------|-----|
| 9.1  | <i>Introduction</i>                                       | 129 |
| 9.2  | <i>The Internet Control Message Protocol</i>              | 129 |
| 9.3  | <i>Error Reporting vs. Error Correction</i>               | 130 |
| 9.4  | <i>ICMP Message Delivery</i>                              | 131 |
| 9.5  | <i>ICMP Message Format</i>                                | 132 |
| 9.6  | <i>Testing Destination Reachability And Status (Ping)</i> | 133 |
| 9.7  | <i>Echo Request And Reply Message Format</i>              | 134 |
| 9.8  | <i>Reports Of Unreachable Destinations</i>                | 134 |
| 9.9  | <i>Congestion And Datagram Flow Control</i>               | 136 |
| 9.10 | <i>Source Quench Format</i>                               | 136 |
| 9.11 | <i>Route Change Requests From Routers</i>                 | 137 |
| 9.12 | <i>Detecting Circular Or Excessively Long Routes</i>      | 139 |
| 9.13 | <i>Reporting Other Problems</i>                           | 140 |
| 9.14 | <i>Clock Synchronization And Transit Time Estimation</i>  | 140 |
| 9.15 | <i>Information Request And Reply Messages</i>             | 142 |
| 9.16 | <i>Obtaining A Subnet Mask</i>                            | 142 |
| 9.17 | <i>Router Discovery</i>                                   | 143 |
| 9.18 | <i>Router Solicitation</i>                                | 144 |
| 9.19 | <i>Summary</i>                                            | 145 |

## **Chapter 10 Classless And Subnet Address Extensions (CIDR) 147**

|      |                                                 |     |
|------|-------------------------------------------------|-----|
| 10.1 | <i>Introduction</i>                             | 147 |
| 10.2 | <i>Review Of Relevant Facts</i>                 | 147 |
| 10.3 | <i>Minimizing Network Numbers</i>               | 148 |
| 10.4 | <i>Transparent Routers</i>                      | 149 |
| 10.5 | <i>Proxy ARP</i>                                | 150 |
| 10.6 | <i>Subnet Addressing</i>                        | 152 |
| 10.7 | <i>Flexibility In Subnet Address Assignment</i> | 154 |
| 10.8 | <i>Variable-Length Subnets</i>                  | 155 |

|       |                                                            |     |
|-------|------------------------------------------------------------|-----|
| 10.9  | <i>Implementation Of Subnets With Masks</i>                | 156 |
| 10.10 | <i>Subnet Mask Representation</i>                          | 157 |
| 10.11 | <i>Routing In The Presence Of Subnets</i>                  | 158 |
| 10.12 | <i>The Subnet Routing Algorithm</i>                        | 159 |
| 10.13 | <i>A Unified Routing Algorithm</i>                         | 160 |
| 10.14 | <i>Maintenance Of Subnet Masks</i>                         | 161 |
| 10.15 | <i>Broadcasting To Subnets</i>                             | 161 |
| 10.16 | <i>Anonymous Point-To-Point Networks</i>                   | 162 |
| 10.17 | <i>Classless Addressing (Supernetting)</i>                 | 164 |
| 10.18 | <i>The Effect Of Supernetting On Routing</i>               | 165 |
| 10.19 | <i>CIDR Address Blocks And Bit Masks</i>                   | 165 |
| 10.20 | <i>Address Blocks And CIDR Notation</i>                    | 166 |
| 10.21 | <i>A Classless Addressing Example</i>                      | 167 |
| 10.22 | <i>Data Structures And Algorithms For Classless Lookup</i> | 167 |
| 10.23 | <i>Longest-Match Routing And Mixtures Of Route Types</i>   | 170 |
| 10.24 | <i>CIDR Blocks Reserved For Private Networks</i>           | 172 |
| 10.25 | <i>Summary</i>                                             | 173 |

## **Chapter 11 Protocol Layering**

**177**

|       |                                                              |     |
|-------|--------------------------------------------------------------|-----|
| 11.1  | <i>Introduction</i>                                          | 177 |
| 11.2  | <i>The Need For Multiple Protocols</i>                       | 177 |
| 11.3  | <i>The Conceptual Layers Of Protocol Software</i>            | 178 |
| 11.4  | <i>Functionality Of The Layers</i>                           | 181 |
| 11.5  | <i>X.25 And Its Relation To The ISO Model</i>                | 182 |
| 11.6  | <i>Differences Between ISO And Internet Layering</i>         | 185 |
| 11.7  | <i>The Protocol Layering Principle</i>                       | 187 |
| 11.8  | <i>Layering In The Presence Of Network Substructure</i>      | 189 |
| 11.9  | <i>Two Important Boundaries In The TCP/IP Model</i>          | 191 |
| 11.10 | <i>The Disadvantage Of Layering</i>                          | 192 |
| 11.11 | <i>The Basic Idea Behind Multiplexing And Demultiplexing</i> | 192 |
| 11.12 | <i>Summary</i>                                               | 194 |

## **Chapter 12 User Datagram Protocol (UDP)**

**197**

|      |                                                  |     |
|------|--------------------------------------------------|-----|
| 12.1 | <i>Introduction</i>                              | 197 |
| 12.2 | <i>Identifying The Ultimate Destination</i>      | 197 |
| 12.3 | <i>The User Datagram Protocol</i>                | 198 |
| 12.4 | <i>Format Of UDP Messages</i>                    | 199 |
| 12.5 | <i>UDP Pseudo-Header</i>                         | 200 |
| 12.6 | <i>UDP Encapsulation And Protocol Layering</i>   | 201 |
| 12.7 | <i>Layering And The UDP Checksum Computation</i> | 203 |

|       |                                                    |     |
|-------|----------------------------------------------------|-----|
| 12.8  | <i>UDP Multiplexing, Demultiplexing, And Ports</i> | 203 |
| 12.9  | <i>Reserved And Available UDP Port Numbers</i>     | 204 |
| 12.10 | <i>Summary</i>                                     | 206 |

## **Chapter 13 Reliable Stream Transport Service (TCP)**

**209**

|       |                                                    |     |
|-------|----------------------------------------------------|-----|
| 13.1  | <i>Introduction</i>                                | 209 |
| 13.2  | <i>The Need For Stream Delivery</i>                | 209 |
| 13.3  | <i>Properties Of The Reliable Delivery Service</i> | 210 |
| 13.4  | <i>Providing Reliability</i>                       | 211 |
| 13.5  | <i>The Idea Behind Sliding Windows</i>             | 213 |
| 13.6  | <i>The Transmission Control Protocol</i>           | 215 |
| 13.7  | <i>Ports, Connections, And Endpoints</i>           | 216 |
| 13.8  | <i>Passive And Active Opens</i>                    | 218 |
| 13.9  | <i>Segments, Streams, And Sequence Numbers</i>     | 219 |
| 13.10 | <i>Variable Window Size And Flow Control</i>       | 220 |
| 13.11 | <i>TCP Segment Format</i>                          | 221 |
| 13.12 | <i>Out Of Band Data</i>                            | 222 |
| 13.13 | <i>Maximum Segment Size Option</i>                 | 223 |
| 13.14 | <i>TCP Checksum Computation</i>                    | 224 |
| 13.15 | <i>Acknowledgements And Retransmission</i>         | 225 |
| 13.16 | <i>Timeout And Retransmission</i>                  | 226 |
| 13.17 | <i>Accurate Measurement Of Round Trip Samples</i>  | 228 |
| 13.18 | <i>Karn's Algorithm And Timer Backoff</i>          | 229 |
| 13.19 | <i>Responding To High Variance In Delay</i>        | 230 |
| 13.20 | <i>Response To Congestion</i>                      | 232 |
| 13.21 | <i>Congestion, Tail Drop, And TCP</i>              | 234 |
| 13.22 | <i>Random Early Discard (RED)</i>                  | 235 |
| 13.23 | <i>Establishing A TCP Connection</i>               | 237 |
| 13.24 | <i>Initial Sequence Numbers</i>                    | 239 |
| 13.25 | <i>Closing a TCP Connection</i>                    | 239 |
| 13.26 | <i>TCP Connection Reset</i>                        | 241 |
| 13.27 | <i>TCP State Machine</i>                           | 241 |
| 13.28 | <i>Forcing Data Delivery</i>                       | 243 |
| 13.29 | <i>Reserved TCP Port Numbers</i>                   | 243 |
| 13.30 | <i>TCP Performance</i>                             | 243 |
| 13.31 | <i>Silly Window Syndrome And Small Packets</i>     | 245 |
| 13.32 | <i>Avoiding Silly Window Syndrome</i>              | 246 |
| 13.33 | <i>Summary</i>                                     | 249 |

**Chapter 14 Routing: Cores, Peers, And Algorithms 253**

- 14.1 Introduction* 253
- 14.2 The Origin Of Routing Tables* 254
- 14.3 Routing With Partial Information* 255
- 14.4 Original Internet Architecture And Cores* 256
- 14.5 Core Routers* 257
- 14.6 Beyond The Core Architecture To Peer Backbones* 260
- 14.7 Automatic Route Propagation* 262
- 14.8 Distance Vector (Bellman-Ford) Routing* 262
- 14.9 Gateway-To-Gateway Protocol (GGP)* 264
- 14.10 Distance Factoring* 265
- 14.11 Reliability And Routing Protocols* 265
- 14.12 Link-State (SPF) Routing* 266
- 14.13 Summary* 267

**Chapter 15 Routing: Exterior Gateway Protocols And Autonomous Systems (BGP) 269**

- 15.1 Introduction* 269
- 15.2 Adding Complexity To The Architectural Model* 269
- 15.3 Determining A Practical Limit On Group Size* 270
- 15.4 A Fundamental Idea: Extra Hops* 271
- 15.5 Hidden Networks* 273
- 15.6 Autonomous System Concept* 274
- 15.7 From A Core To Independent Autonomous Systems* 275
- 15.8 An Exterior Gateway Protocol* 276
- 15.9 BGP Characteristics* 277
- 15.10 BGP Functionality And Message Types* 278
- 15.11 BGP Message Header* 278
- 15.12 BGP OPEN Message* 279
- 15.13 BGP UPDATE Message* 280
- 15.14 Compressed Mask-Address Pairs* 281
- 15.15 BGP Path Attributes* 282
- 15.16 BGP KEEPALIVE Message* 283
- 15.17 Information From The Receiver's Perspective* 284
- 15.18 The Key Restriction Of Exterior Gateway Protocols* 285
- 15.19 The Internet Routing Arbiter System* 287
- 15.20 BGP NOTIFICATION Message* 288
- 15.21 Decentralization Of Internet Architecture* 289
- 15.22 Summary* 290

**Chapter 16 Routing: In An Autonomous System (RIP, OSPF, HELLO) 293**

- 16.1 Introduction 293
- 16.2 Static Vs. Dynamic Interior Routes 293
- 16.3 Routing Information Protocol (RIP) 296
- 16.4 The Hello Protocol 305
- 16.5 Delay Metrics And Oscillation 305
- 16.6 Combining RIP, Hello, And BGP 307
- 16.7 Inter-Autonomous System Routing 307
- 16.8 Gated: Inter-Autonomous System Communication 308
- 16.9 The Open SPF Protocol (OSPF) 308
- 16.10 Routing With Partial Information 315
- 16.11 Summary 315

**Chapter 17 Internet Multicasting 319**

- 17.1 Introduction 319
- 17.2 Hardware Broadcast 319
- 17.3 Hardware Origins Of Multicast 320
- 17.4 Ethernet Multicast 321
- 17.5 IP Multicast 321
- 17.6 The Conceptual Pieces 322
- 17.7 IP Multicast Addresses 323
- 17.8 Multicast Address Semantics 325
- 17.9 Mapping IP Multicast To Ethernet Multicast 325
- 17.10 Hosts And Multicast Delivery 326
- 17.11 Multicast Scope 326
- 17.12 Extending Host Software To Handle Multicasting 327
- 17.13 Internet Group Management Protocol 328
- 17.14 IGMP Implementation 328
- 17.15 Group Membership State Transitions 329
- 17.16 IGMP Message Format 331
- 17.17 Multicast Forwarding And Routing Information 332
- 17.18 Basic Multicast Routing Paradigms 334
- 17.19 Consequences Of TRPF 335
- 17.20 Multicast Trees 337
- 17.21 The Essence Of Multicast Routing 338
- 17.22 Reverse Path Multicasting 338
- 17.23 Distance Vector Multicast Routing Protocol 339
- 17.24 The Mrouted Program 340
- 17.25 Alternative Protocols 343
- 17.26 Core Based Trees (CBT) 343
- 17.27 Protocol Independent Multicast (PIM) 344

- 17.28 *Multicast Extensions To OSPF (MOSPF)* 347
- 17.29 *Reliable Multicast And ACK Implosions* 347
- 17.30 *Summary* 349

## **Chapter 18 TCP/IP Over ATM Networks**

**353**

- 18.1 *Introduction* 353
- 18.2 *ATM Hardware* 354
- 18.3 *Large ATM Networks* 354
- 18.4 *The Logical View Of An ATM Network* 355
- 18.5 *The Two ATM Connection Paradigms* 356
- 18.6 *Paths, Circuits, And Identifiers* 357
- 18.7 *ATM Cell Transport* 358
- 18.8 *ATM Adaptation Layers* 358
- 18.9 *ATM Adaptation Layer 5* 360
- 18.10 *AAL5 Convergence, Segmentation, And Reassembly* 361
- 18.11 *Datagram Encapsulation And IP MTU Size* 361
- 18.12 *Packet Type And Multiplexing* 362
- 18.13 *IP Address Binding In An ATM Network* 363
- 18.14 *Logical IP Subnet Concept* 364
- 18.15 *Connection Management* 365
- 18.16 *Address Binding Within An LIS* 366
- 18.17 *ATMARP Packet Format* 366
- 18.18 *Using ATMARP Packets To Determine An Address* 369
- 18.19 *Obtaining Entries For A Server Database* 370
- 18.20 *Timing Out ATMARP Information In A Server* 370
- 18.21 *Timing Out ATMARP Information In A Host Or Router* 371
- 18.22 *IP Switching Technologies* 371
- 18.23 *Switch Operation* 372
- 18.24 *Optimized IP Forwarding* 372
- 18.25 *Classification, Flows, And Higher Layer Switching* 373
- 18.26 *Applicability Of Switching Technology* 374
- 18.27 *Summary* 374

## **Chapter 19 Mobile IP**

**377**

- 19.1 *Introduction* 377
- 19.2 *Mobility, Routing, and Addressing* 377
- 19.3 *Mobile IP Characteristics* 378
- 19.4 *Overview Of Mobile IP Operation* 378
- 19.5 *Mobile Addressing Details* 379
- 19.6 *Foreign Agent Discovery* 380

|                                                                  |                                                         |            |
|------------------------------------------------------------------|---------------------------------------------------------|------------|
| 19.7                                                             | <i>Agent Registration</i>                               | 381        |
| 19.8                                                             | <i>Registration Message Format</i>                      | 381        |
| 19.9                                                             | <i>Communication With A Foreign Agent</i>               | 383        |
| 19.10                                                            | <i>Datagram Transmission And Reception</i>              | 383        |
| 19.11                                                            | <i>The Two-Crossing Problem</i>                         | 384        |
| 19.12                                                            | <i>Communication With Computers On the Home Network</i> | 385        |
| 19.13                                                            | <i>Summary</i>                                          | 386        |
| <br><b>Chapter 20 Private Network Interconnection (NAT, VPN)</b> |                                                         | <b>389</b> |
| 20.1                                                             | <i>Introduction</i>                                     | 389        |
| 20.2                                                             | <i>Private And Hybrid Networks</i>                      | 389        |
| 20.3                                                             | <i>A Virtual Private Network (VPN)</i>                  | 390        |
| 20.4                                                             | <i>VPN Addressing And Routing</i>                       | 392        |
| 20.5                                                             | <i>A VPN With Private Addresses</i>                     | 393        |
| 20.6                                                             | <i>Network Address Translation (NAT)</i>                | 394        |
| 20.7                                                             | <i>NAT Translation Table Creation</i>                   | 395        |
| 20.8                                                             | <i>Multi-Address NAT</i>                                | 396        |
| 20.9                                                             | <i>Port-Mapped NAT</i>                                  | 396        |
| 20.10                                                            | <i>Interaction Between NAT And ICMP</i>                 | 398        |
| 20.11                                                            | <i>Interaction Between NAT And Applications</i>         | 398        |
| 20.12                                                            | <i>Conceptual Address Domains</i>                       | 399        |
| 20.13                                                            | <i>Slirp And Masquerade</i>                             | 399        |
| 20.14                                                            | <i>Summary</i>                                          | 400        |
| <br><b>Chapter 21 Client-Server Model Of Interaction</b>         |                                                         | <b>403</b> |
| 21.1                                                             | <i>Introduction</i>                                     | 403        |
| 21.2                                                             | <i>The Client-Server Model</i>                          | 403        |
| 21.3                                                             | <i>A Simple Example: UDP Echo Server</i>                | 404        |
| 21.4                                                             | <i>Time And Date Service</i>                            | 406        |
| 21.5                                                             | <i>The Complexity of Servers</i>                        | 407        |
| 21.6                                                             | <i>RARP Server</i>                                      | 408        |
| 21.7                                                             | <i>Alternatives To The Client-Server Model</i>          | 409        |
| 21.8                                                             | <i>Summary</i>                                          | 410        |
| <br><b>Chapter 22 The Socket Interface</b>                       |                                                         | <b>413</b> |
| 22.1                                                             | <i>Introduction</i>                                     | 413        |
| 22.2                                                             | <i>The UNIX I/O Paradigm And Network I/O</i>            | 414        |
| 22.3                                                             | <i>Adding Network I/O to UNIX</i>                       | 414        |

|       |                                                       |     |
|-------|-------------------------------------------------------|-----|
| 22.4  | <i>The Socket Abstraction</i>                         | 415 |
| 22.5  | <i>Creating A Socket</i>                              | 415 |
| 22.6  | <i>Socket Inheritance And Termination</i>             | 416 |
| 22.7  | <i>Specifying A Local Address</i>                     | 417 |
| 22.8  | <i>Connecting Sockets To Destination Addresses</i>    | 418 |
| 22.9  | <i>Sending Data Through A Socket</i>                  | 419 |
| 22.10 | <i>Receiving Data Through A Socket</i>                | 421 |
| 22.11 | <i>Obtaining Local And Remote Socket Addresses</i>    | 422 |
| 22.12 | <i>Obtaining And Setting Socket Options</i>           | 423 |
| 22.13 | <i>Specifying A Queue Length For A Server</i>         | 424 |
| 22.14 | <i>How A Server Accepts Connections</i>               | 424 |
| 22.15 | <i>Servers That Handle Multiple Services</i>          | 425 |
| 22.16 | <i>Obtaining And Setting Host Names</i>               | 426 |
| 22.17 | <i>Obtaining And Setting The Internal Host Domain</i> | 427 |
| 22.18 | <i>Socket Library Calls</i>                           | 427 |
| 22.19 | <i>Network Byte Order Conversion Routines</i>         | 428 |
| 22.20 | <i>IP Address Manipulation Routines</i>               | 429 |
| 22.21 | <i>Accessing The Domain Name System</i>               | 431 |
| 22.22 | <i>Obtaining Information About Hosts</i>              | 432 |
| 22.23 | <i>Obtaining Information About Networks</i>           | 433 |
| 22.24 | <i>Obtaining Information About Protocols</i>          | 434 |
| 22.25 | <i>Obtaining Information About Network Services</i>   | 434 |
| 22.26 | <i>An Example Client</i>                              | 435 |
| 22.27 | <i>An Example Server</i>                              | 437 |
| 22.28 | <i>Summary</i>                                        | 440 |

## **Chapter 23 Bootstrap And Autoconfiguration (BOOTP, DHCP)**

**443**

|       |                                            |     |
|-------|--------------------------------------------|-----|
| 23.1  | <i>Introduction</i>                        | 443 |
| 23.2  | <i>The Need For An Alternative To RARP</i> | 444 |
| 23.3  | <i>Using IP To Determine An IP Address</i> | 444 |
| 23.4  | <i>The BOOTP Retransmission Policy</i>     | 445 |
| 23.5  | <i>The BOOTP Message Format</i>            | 446 |
| 23.6  | <i>The Two-Step Bootstrap Procedure</i>    | 447 |
| 23.7  | <i>Vendor-Specific Field</i>               | 448 |
| 23.8  | <i>The Need For Dynamic Configuration</i>  | 448 |
| 23.9  | <i>Dynamic Host Configuration</i>          | 450 |
| 23.10 | <i>Dynamic IP Address Assignment</i>       | 450 |
| 23.11 | <i>Obtaining Multiple Addresses</i>        | 451 |
| 23.12 | <i>Address Acquisition States</i>          | 452 |
| 23.13 | <i>Early Lease Termination</i>             | 452 |
| 23.14 | <i>Lease Renewal States</i>                | 454 |
| 23.15 | <i>DHCP Message Format</i>                 | 455 |

- 23.16 *DHCP Options And Message Type* 456
- 23.17 *Option Overload* 457
- 23.18 *DHCP And Domain Names* 457
- 23.19 *Summary* 458

## **Chapter 24 The Domain Name System (DNS)**

461

- 24.1 *Introduction* 461
- 24.2 *Names For Machines* 462
- 24.3 *Flat Namespace* 462
- 24.4 *Hierarchical Names* 463
- 24.5 *Delegation Of Authority For Names* 464
- 24.6 *Subset Authority* 464
- 24.7 *Internet Domain Names* 465
- 24.8 *Official And Unofficial Internet Domain Names* 466
- 24.9 *Named Items And Syntax Of Names* 468
- 24.10 *Mapping Domain Names To Addresses* 469
- 24.11 *Domain Name Resolution* 471
- 24.12 *Efficient Translation* 472
- 24.13 *Caching: The Key To Efficiency* 473
- 24.14 *Domain Server Message Format* 474
- 24.15 *Compressed Name Format* 477
- 24.16 *Abbreviation Of Domain Names* 477
- 24.17 *Inverse Mappings* 478
- 24.18 *Pointer Queries* 479
- 24.19 *Object Types And Resource Record Contents* 479
- 24.20 *Obtaining Authority For A Subdomain* 480
- 24.21 *Summary* 481

## **Chapter 25 Applications: Remote Login (TELNET, Rlogin)**

485

- 25.1 *Introduction* 485
- 25.2 *Remote Interactive Computing* 485
- 25.3 *TELNET Protocol* 486
- 25.4 *Accommodating Heterogeneity* 488
- 25.5 *Passing Commands That Control The Remote Side* 490
- 25.6 *Forcing The Server To Read A Control Function* 492
- 25.7 *TELNET Options* 492
- 25.8 *TELNET Option Negotiation* 493
- 25.9 *Rlogin (BSD UNIX)* 494
- 25.10 *Summary* 495

**Chapter 26 Applications: File Transfer And Access (FTP, TFTP, NFS) 497**

- 26.1 *Introduction* 497
- 26.2 *File Access And Transfer* 497
- 26.3 *On-line Shared Access* 498
- 26.4 *Sharing By File Transfer* 499
- 26.5 *FTP: The Major TCP/IP File Transfer Protocol* 499
- 26.6 *FTP Features* 500
- 26.7 *FTP Process Model* 500
- 26.8 *TCP Port Number Assignment* 502
- 26.9 *The User's View Of FTP* 502
- 26.10 *An Example Anonymous FTP Session* 504
- 26.11 *TFTP* 505
- 26.12 *NFS* 507
- 26.13 *NFS Implementation* 507
- 26.14 *Remote Procedure Call (RPC)* 508
- 26.15 *Summary* 509

**Chapter 27 Applications: Electronic Mail (SMTP, POP, IMAP, MIME) 511**

- 27.1 *Introduction* 511
- 27.2 *Electronic Mail* 511
- 27.3 *Mailbox Names And Aliases* 513
- 27.4 *Alias Expansion And Mail Forwarding* 513
- 27.5 *The Relationship Of Internetworking And Mail* 514
- 27.6 *TCP/IP Standards For Electronic Mail Service* 516
- 27.7 *Electronic Mail Addresses* 516
- 27.8 *Pseudo Domain Addresses* 518
- 27.9 *Simple Mail Transfer Protocol (SMTP)* 518
- 27.10 *Mail Retrieval And Mailbox Manipulation Protocols* 521
- 27.11 *The MIME Extension For Non-ASCII Data* 522
- 27.12 *MIME Multipart Messages* 523
- 27.13 *Summary* 524

**Chapter 28 Applications: World Wide Web (HTTP) 527**

- 28.1 *Introduction* 527
- 28.2 *Importance Of The Web* 527
- 28.3 *Architectural Components* 528
- 28.4 *Uniform Resource Locators* 528
- 28.5 *An Example Document* 529
- 28.6 *Hypertext Transfer Protocol* 530

|       |                                           |     |
|-------|-------------------------------------------|-----|
| 28.7  | <i>HTTP GET Request</i>                   | 530 |
| 28.8  | <i>Error Messages</i>                     | 531 |
| 28.9  | <i>Persistent Connections And Lengths</i> | 532 |
| 28.10 | <i>Data Length And Program Output</i>     | 532 |
| 28.11 | <i>Length Encoding And Headers</i>        | 533 |
| 28.12 | <i>Negotiation</i>                        | 534 |
| 28.13 | <i>Conditional Requests</i>               | 535 |
| 28.14 | <i>Support For Proxy Servers</i>          | 535 |
| 28.15 | <i>Caching</i>                            | 536 |
| 28.16 | <i>Summary</i>                            | 537 |

## **Chapter 29 Applications: Voice And Video Over IP (RTP) 539**

|       |                                                      |     |
|-------|------------------------------------------------------|-----|
| 29.1  | <i>Introduction</i>                                  | 539 |
| 29.2  | <i>Audio Clips And Encoding Standards</i>            | 539 |
| 29.3  | <i>Audio And Video Transmission And Reproduction</i> | 540 |
| 29.4  | <i>Jitter And Playback Delay</i>                     | 541 |
| 29.5  | <i>Real-Time Transport Protocol (RTP)</i>            | 542 |
| 29.6  | <i>Streams, Mixing, And Multicasting</i>             | 543 |
| 29.7  | <i>RTP Encapsulation</i>                             | 544 |
| 29.8  | <i>RTP Control Protocol (RTCP)</i>                   | 544 |
| 29.9  | <i>RTCP Operation</i>                                | 545 |
| 29.10 | <i>IP Telephony And Signaling</i>                    | 546 |
| 29.11 | <i>Resource Reservation And Quality Of Service</i>   | 548 |
| 29.12 | <i>QoS, Utilization, And Capacity</i>                | 549 |
| 29.13 | <i>RSVP</i>                                          | 549 |
| 29.14 | <i>COPS</i>                                          | 550 |
| 29.15 | <i>Summary</i>                                       | 551 |

## **Chapter 30 Applications: Internet Management (SNMP) 553**

|       |                                                         |     |
|-------|---------------------------------------------------------|-----|
| 30.1  | <i>Introduction</i>                                     | 553 |
| 30.2  | <i>The Level Of Management Protocols</i>                | 553 |
| 30.3  | <i>Architectural Model</i>                              | 554 |
| 30.4  | <i>Protocol Framework</i>                               | 556 |
| 30.5  | <i>Examples of MIB Variables</i>                        | 557 |
| 30.6  | <i>The Structure Of Management Information</i>          | 558 |
| 30.7  | <i>Formal Definitions Using ASN.1</i>                   | 559 |
| 30.8  | <i>Structure And Representation Of MIB Object Names</i> | 559 |
| 30.9  | <i>Simple Network Management Protocol</i>               | 564 |
| 30.10 | <i>SNMP Message Format</i>                              | 566 |
| 30.11 | <i>Example Encoded SNMP Message</i>                     | 569 |

- 30.12 *New Features In SNMPv3* 572
- 30.13 *Summary* 572

## **Chapter 31 Summary Of Protocol Dependencies 575**

- 31.1 *Introduction* 575
- 31.2 *Protocol Dependencies* 575
- 31.3 *The Hourglass Model* 577
- 31.4 *Application Program Access* 578
- 31.5 *Summary* 579

## **Chapter 32 Internet Security And Firewall Design (IPsec) 581**

- 32.1 *Introduction* 581
- 32.2 *Protecting Resources* 582
- 32.3 *Information Policy* 583
- 32.4 *Internet Security* 583
- 32.5 *IP Security (IPsec)* 584
- 32.6 *IPsec Authentication Header* 584
- 32.7 *Security Association* 585
- 32.8 *IPsec Encapsulating Security Payload* 586
- 32.9 *Authentication And Mutable Header Fields* 587
- 32.10 *IPsec Tunneling* 588
- 32.11 *Required Security Algorithms* 588
- 32.12 *Secure Sockets* 589
- 32.13 *Firewalls And Internet Access* 589
- 32.14 *Multiple Connections And Weakest Links* 589
- 32.15 *Firewall Implementation* 590
- 32.16 *Packet-Level Filters* 590
- 32.17 *Security And Packet Filter Specification* 591
- 32.18 *The Consequence Of Restricted Access For Clients* 592
- 32.19 *Proxy Access Through A Firewall* 592
- 32.20 *The Details Of Firewall Architecture* 593
- 32.21 *Stub Network* 594
- 32.22 *An Alternative Firewall Implementation* 595
- 32.23 *Monitoring And Logging* 596
- 32.24 *Summary* 596

|                                                                           |                |
|---------------------------------------------------------------------------|----------------|
| <b>Chapter 33 The Future Of TCP/IP (IPv6)</b>                             | <b>599</b>     |
| 33.1 Introduction                                                         | 599            |
| 33.2 Why Change?                                                          | 600            |
| 33.3 New Policies                                                         | 600            |
| 33.4 Motivation For Changing IPv4                                         | 600            |
| 33.5 The Road To A New Version Of IP                                      | 601            |
| 33.6 The Name Of The Next IP                                              | 602            |
| 33.7 Features Of IPv6                                                     | 602            |
| 33.8 General Form Of An IPv6 Datagram                                     | 603            |
| 33.9 IPv6 Base Header Format                                              | 603            |
| 33.10 IPv6 Extension Headers                                              | 605            |
| 33.11 Parsing An IPv6 Datagram                                            | 606            |
| 33.12 IPv6 Fragmentation And Reassembly                                   | 607            |
| 33.13 The Consequence Of End-To-End Fragmentation                         | 607            |
| 33.14 IPv6 Source Routing                                                 | 608            |
| 33.15 IPv6 Options                                                        | 609            |
| 33.16 Size Of The IPv6 Address Space                                      | 610            |
| 33.17 IPv6 Colon Hexadecimal Notation                                     | 610            |
| 33.18 Three Basic IPv6 Address Types                                      | 612            |
| 33.19 The Duality Of Broadcast And Multicast                              | 612            |
| 33.20 An Engineering Choice And Simulated Broadcast                       | 613            |
| 33.21 Proposed IPv6 Address Space Assignment                              | 613            |
| 33.22 Embedded IPv4 Addresses And Transition                              | 614            |
| 33.23 Unspecified And Loopback Addresses                                  | 616            |
| 33.24 Unicast Address Hierarchy                                           | 616            |
| 33.25 Aggregatable Global Unicast Address Structure                       | 617            |
| 33.26 Interface Identifiers                                               | 618            |
| 33.27 Additional Hierarchy                                                | 619            |
| 33.28 Local Addresses                                                     | 619            |
| 33.29 Autoconfiguration And Renumbering                                   | 620            |
| 33.30 Summary                                                             | 620            |
| <br><b>Appendix 1 A Guide To RFCs</b>                                     | <br><b>623</b> |
| <br><b>Appendix 2 Glossary Of Internetworking Terms And Abbreviations</b> | <br><b>673</b> |
| <br><b>Bibliography</b>                                                   | <br><b>721</b> |
| <br><b>Index</b>                                                          | <br><b>729</b> |

## *Applications: World Wide Web (HTTP)*

### **28.1 Introduction**

This chapter continues the discussion of applications that use TCP/IP technology by focusing on the application that has had the most impact: the *World Wide Web* (WWW). After a brief overview of concepts, the chapter examines the primary protocol used to transfer a Web page from a server to a Web browser. The discussion covers caching as well as the basic transfer mechanism.

### **28.2 Importance Of The Web**

During the early history of the Internet, FTP data transfers accounted for approximately one third of Internet traffic, more than any other application. From its inception in the early 1990s, however, the Web had a much higher growth rate. By 1995, Web traffic overtook FTP to become the largest consumer of Internet backbone bandwidth, and has remained the leader ever since. By 2000, Web traffic completely overshadowed other applications.

Although traffic is easy to measure and cite, the impact of the Web cannot be understood from such statistics. More people know about and use the Web than any other Internet application. Most companies have Web sites and on-line catalogs; references to the Web appear in advertising. In fact, for many users, the Internet and the Web are indistinguishable.

## 28.3 Architectural Components

Conceptually, the Web consists of a large set of documents, called *Web pages*, that are accessible over the Internet. Each Web page is classified as a *hypermedia* document. The suffix *media* is used to indicate that a document can contain items other than text (e.g., graphics images); the prefix *hyper* is used because a document can contain *selectable links* that refer to other, related documents.

Two main building blocks are used to implement the Web on top of the global Internet. A *Web browser* consists of an application program that a user invokes to access and display a Web page. The browser becomes a client that contacts the appropriate *Web server* to obtain a copy of the specified page. Because a given server can manage more than one Web page, a browser must specify the exact page when making a request.

The data representation standard used for a Web page depends on its contents. For example, standard graphics representations such as *Graphics Interchange Format (GIF)* or *Joint Picture Encoding Group (JPEG)* can be used for a page that contains a single graphics image. Pages that contain a mixture of text and other items are represented using *HyperText Markup Language (HTML)*. An HTML document consists of a file that contains text along with embedded commands, called *tags*, that give guidelines for display. A tag is enclosed in less-than and greater-than symbols; some tags come in pairs that apply to all items between the pair. For example, the two commands `<CENTER>` and `</CENTER>` cause items between them to be centered in the browser's window.

## 28.4 Uniform Resource Locators

Each Web page is assigned a unique name that is used to identify it. The name, which is called a *Uniform Resource Locator (URL)*<sup>†</sup>, begins with a specification of the *scheme* used to access the item. In effect, the scheme specifies the transfer protocol; the format of the remainder of the URL depends on the scheme. For example, a URL that follows the *http scheme* has the following form<sup>‡</sup>:

`http: // hostname [: port] / path [; parameters] [? query]`

where brackets denote an optional item. For now, it is sufficient to understand that the *hostname* string specifies the domain name or IP address of the computer on which the server for the item operates, *:port* is an optional protocol port number needed only in cases where the server does not use the well-known port (80), *path* is a string that identifies one particular document on the server, *;parameters* is an optional string that specifies additional parameters supplied by the client, and *?query* is an optional string used when the browser sends a question. A user is unlikely to ever see or use the optional parts directly. Instead, URLs that a user enters contain only a *hostname* and *path*. For example, the URL:

<sup>†</sup>A URL is a specific type of the more general *Uniform Resource Identifier (URI)*.

<sup>‡</sup>Some of the literature refers to the initial string, *http:*, as a *pragma*.

`http://www.cs.purdue.edu/people/comer/`

specifies the author's Web page. The server operates on computer *www.cs.purdue.edu*, and the document is named */people/comer/*.

The protocol standards distinguish between the *absolute* form of a URL illustrated above, and a *relative* form. A relative URL, which is seldom seen by a user, is only meaningful when the server has already been determined. Relative URLs are useful once communication has been established with a specific server. For example, when communicating with server *www.cs.purdue.edu*, only the string */people/comer/* is needed to specify the document named by the absolute URL above. We can summarize.

*Each Web page is assigned a unique identifier known as a Uniform Resource Locator (URL). The absolute form of a URL contains a full specification; a relative form that omits the address of the server is only useful when the server is implicitly known.*

## 28.5 An Example Document

In principle, Web access is straightforward. All access originates with a URL — a user either enters a URL via the keyboard or selects an item which provides the browser with a URL. The browser parses the URL, extracts the information, and uses it to obtain a copy of the requested page. Because the format of the URL depends on the scheme, the browser begins by extracting the scheme specification, and then uses the scheme to determine how to parse the rest of the URL.

An example will illustrate how a URL is produced from a *selectable link* in a document. In fact, a document contains a pair of values for each link: an item to be displayed on the screen and a URL to follow if the user selects the item. In HTML, the pair of tags `<A>` and `</A>` are known as an *anchor*. The anchor defines a link; a URL is added to the first tag, and items to be displayed are placed between the two tags. The browser stores the URL internally, and follows it when the user selects the link. For example, the following HTML document contains a selectable link:

```
<HTML>
  The author of this text is
  <A HREF="http://www.cs.purdue.edu/people/comer">
  Douglas Comer.</A>
</HTML>
```

When the document is displayed, a single line of text appears on the screen:

The author of this text is Douglas Comer.

The browser underlines the phrase *Douglas Comer* to indicate that it corresponds to a selectable link. Internally, of course, the browser stores the URL from the <A> tag, which it follows when the user selects the link.

## 28.6 Hypertext Transfer Protocol

The protocol used for communication between a browser and a Web server or between intermediate machines and Web servers is known as the *HyperText Transfer Protocol* (*HTTP*). HTTP has the following set of characteristics:

*Application Level.* HTTP operates at the application level. It assumes a reliable, connection-oriented transport protocol such as TCP, but does not provide reliability or retransmission itself.

*Request/Response.* Once a transport session has been established, one side (usually a browser) must send an HTTP request to which the other side responds.

*Stateless.* Each HTTP request is self-contained; the server does not keep a history of previous requests or previous sessions.

*Bi-Directional Transfer.* In most cases, a browser requests a Web page, and the server transfers a copy to the browser. HTTP also allows transfer from a browser to a server (e.g., when a user submits a so-called “form”).

*Capability Negotiation.* HTTP allows browsers and servers to negotiate details such as the character set to be used during transfers. A sender can specify the capabilities it offers and a receiver can specify the capabilities it accepts.

*Support For Caching.* To improve response time, a browser caches a copy of each Web page it retrieves. If a user requests a page again, HTTP allows the browser to interrogate the server to determine whether the contents of the page has changed since the copy was cached.

*Support For Intermediaries.* HTTP allows a machine along the path between a browser and a server to act as a *proxy server* that caches Web pages and answers a browser’s request from its cache.

## 28.7 HTTP GET Request

In the simplest case, a browser contacts a Web server directly to obtain a page. The browser begins with a URL, extracts the hostname section, uses DNS to map the name into an equivalent IP address, and uses the IP address to form a TCP connection

Volume  
1

COMER

Fourth  
Edition

INTERNETWORKING WITH

**TCP/IP**

Principles, Protocols,  
and Architectures

PRENTICE  
HALL

Fourth Edition

INTERNETWORKING with

# TCP/IP

VOLUME I

PRINCIPLES, PROTOCOLS, AND ARCHITECTURES

DOUGLAS E. COMER

"Comer continues to prove himself the Baedeker of the Internet Protocols with this fine 4th edition."

-Vint Cerf, Senior Vice President, MCI WorldCom, "Father of the Internet"

"Doug's deep insight into the Internet, coupled with his talent for teaching, yields an updated classic that once again leads the pack." -Paul Mockapetris, Inventor of the Domain Name System

"There are many TCP/IP books on the shelves today, but Doug Comer's 'Internetworking with TCP/IP' is the one that comes off the shelf for accessible and authoritative answers to questions about Internet technology."

-Lyman Chapin, Chief Scientist, BBN Technologies

The TCP/IP "bible"—now completely updated to reflect all the latest technologies! Leading Internet expert Douglas Comer has updated his classic, superb introduction to TCP/IP-based internetworking that is renowned for its clarity and accessibility in explaining internetworking and routing. Discover how the basic TCP/IP technology has survived and evolved over two decades of exponential growth, and understand the TCP/IP protocols and technical advances. This edition explains emerging technologies such as Mobile IP, Virtual Private Networks, resource reservation with RSVP, and IPv6. Comer reveals how to master TCP/IP and how the Internet works. The reader is required to have a modest background in the fundamentals of computer systems, but does not need sophisticated mathematics. As with previous editions, this edition provides an introduction to physical networks and then shows how they are combined to form an internet. It states design principles clearly, and discusses motivations and consequences.

## THIS NEW EDITION OF VOLUME 1:

- Explains how voice and video are sent over IP internets and how IP Telephony operates.
- Describes Mobile IP (a technology that allows a computer to move from one network to another without changing its IP address).
- Discusses IP security and the security standard, IPsec.
- Revises the discussion of IPv6 to incorporate the latest changes.
- Shows how to interconnect private intranets and the global Internet using Virtual Private Network (VPN) and Network Address Translation (NAT) technologies.
- Expands the description of IP multicasting to cover multicast routing protocols.
- Presents the Differentiated Services (DiffServe) scheme for classes of services as well as Path MTU discovery and routing for anonymous serial networks.
- Explains Random Early Discard (RED), which is now recommended for routers.
- Updates the coverage of all protocols to the latest versions, including RIP, IGMP, and SNMP.

**Douglas E. Comer** is a Professor of Computer Science at Purdue University and a Fellow of the ACM. He is the author of many best-selling books: the three-volume series, *Internetworking with TCP/IP*, *Computer Networks and Internets*, *Second Edition*, and *The Internet Book*. Comer served as chairman of the DARPA Distributed Systems Architecture Board and the CSNET Technical Committees, and is a former member of the Internet Architecture Board (IAB).

Prentice Hall  
Upper Saddle River, NJ 07458  
<http://www.prenhall.com>

ISBN 0-13-018380-6



9 0000

