

Contents

Foreword	xxi
Preface	xxiii
Chapter 1 Introduction And Overview	1
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
Chapter 2 Review Of Underlying Network Technologies	17
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>	32
2.6 <i>Asynchronous Transfer Mode</i>	36
2.7 <i>ARPANET Technology</i>	37
2.8 <i>National Science Foundation Networking</i>	39
2.9 <i>ANSNET</i>	44

2.10	<i>A Planned Wide Area Backbone</i>	44
2.11	<i>Other Technologies Over Which TCP/IP Has Been Used</i>	44
2.12	<i>Summary And Conclusion</i>	47
Chapter 3 Internetworking Concept And Architectural Model		49
3.1	<i>Introduction</i>	49
3.2	<i>Application-Level Interconnection</i>	49
3.3	<i>Network-Level Interconnection</i>	50
3.4	<i>Properties Of The Internet</i>	51
3.5	<i>Internet Architecture</i>	52
3.6	<i>Interconnection Through IP Routers</i>	52
3.7	<i>The User's View</i>	54
3.8	<i>All Networks Are Equal</i>	54
3.9	<i>The Unanswered Questions</i>	55
3.10	<i>Summary</i>	56
Chapter 4 Internet Addresses		59
4.1	<i>Introduction</i>	59
4.2	<i>Universal Identifiers</i>	59
4.3	<i>Three Primary Classes Of IP Addresses</i>	60
4.4	<i>Addresses Specify Network Connections</i>	61
4.5	<i>Network And Broadcast Addresses</i>	61
4.6	<i>Limited Broadcast</i>	62
4.7	<i>Interpreting Zero To Mean "This"</i>	62
4.8	<i>Weaknesses In Internet Addressing</i>	63
4.9	<i>Dotted Decimal Notation</i>	65
4.10	<i>Loopback Address</i>	65
4.11	<i>Summary Of Special Address Conventions</i>	66
4.12	<i>Internet Addressing Authority</i>	66
4.13	<i>An Example</i>	67
4.14	<i>Network Byte Order</i>	69
4.15	<i>Summary</i>	70
Chapter 5 Mapping Internet Addresses To Physical Addresses (ARP)		73
5.1	<i>Introduction</i>	73
5.2	<i>The Address Resolution Problem</i>	73
5.3	<i>Two Types Of Physical Addresses</i>	74
5.4	<i>Resolution Through Direct Mapping</i>	74

- 5.5 *Resolution Through Dynamic Binding* 75
- 5.6 *The Address Resolution Cache* 76
- 5.7 *ARP Refinements* 77
- 5.8 *Relationship Of ARP To Other Protocols* 77
- 5.9 *ARP Implementation* 77
- 5.10 *ARP Encapsulation And Identification* 79
- 5.11 *ARP Protocol Format* 79
- 5.12 *Summary* 81

Chapter 6 Determining An Internet Address At Startup (RARP) 83

- 6.1 *Introduction* 83
- 6.2 *Reverse Address Resolution Protocol (RARP)* 84
- 6.3 *Timing RARP Transactions* 86
- 6.4 *Primary And Backup RARP Servers* 86
- 6.5 *Summary* 87

Chapter 7 Internet Protocol: Connectionless Datagram Delivery 89

- 7.1 *Introduction* 89
- 7.2 *A Virtual Network* 89
- 7.3 *Internet Architecture And Philosophy* 90
- 7.4 *The Concept Of Unreliable Delivery* 90
- 7.5 *Connectionless Delivery System* 91
- 7.6 *Purpose Of The Internet Protocol* 91
- 7.7 *The Internet Datagram* 91
- 7.8 *Internet Datagram Options* 100
- 7.9 *Summary* 106

Chapter 8 Internet Protocol: Routing IP Datagrams 109

- 8.1 *Introduction* 109
- 8.2 *Routing In An Internet* 109
- 8.3 *Direct And Indirect Delivery* 111
- 8.4 *Table-Driven IP Routing* 113
- 8.5 *Next-Hop Routing* 113
- 8.6 *Default Routes* 115
- 8.7 *Host-Specific Routes* 115
- 8.8 *The IP Routing Algorithm* 116
- 8.9 *Routing With IP Addresses* 116
- 8.10 *Handling Incoming Datagrams* 118

- 8.11 *Establishing Routing Tables* 119
- 8.12 *Summary* 119

Chapter 9 Internet Protocol: Error And Control Messages (ICMP) 123

- 9.1 *Introduction* 123
- 9.2 *The Internet Control Message Protocol* 123
- 9.3 *Error Reporting vs. Error Correction* 124
- 9.4 *ICMP Message Delivery* 125
- 9.5 *ICMP Message Format* 126
- 9.6 *Testing Destination Reachability And Status (Ping)* 127
- 9.7 *Echo Request And Reply Message Format* 128
- 9.8 *Reports Of Unreachable Destinations* 128
- 9.9 *Congestion And Datagram Flow Control* 130
- 9.10 *Source Quench Format* 130
- 9.11 *Route Change Requests From Routers* 131
- 9.12 *Detecting Circular Or Excessively Long Routes* 133
- 9.13 *Reporting Other Problems* 134
- 9.14 *Clock Synchronization And Transit Time Estimation* 134
- 9.15 *Information Request And Reply Messages* 136
- 9.16 *Obtaining A Subnet Mask* 136
- 9.17 *Summary* 137

Chapter 10 Subnet And Supernet Address Extensions 139

- 10.1 *Introduction* 139
- 10.2 *Review Of Relevant Facts* 139
- 10.3 *Minimizing Network Numbers* 140
- 10.4 *Transparent Routers* 141
- 10.5 *Proxy ARP* 142
- 10.6 *Subnet Addressing* 143
- 10.7 *Flexibility In Subnet Address Assignment* 146
- 10.8 *Implementation Of Subnets With Masks* 147
- 10.9 *Subnet Mask Representation* 148
- 10.10 *Routing In The Presence Of Subnets* 149
- 10.11 *The Subnet Routing Algorithm* 150
- 10.12 *A Unified Routing Algorithm* 151
- 10.13 *Maintenance Of Subnet Masks* 152
- 10.14 *Broadcasting To Subnets* 152
- 10.15 *Supernet Addressing* 153
- 10.16 *The Effect Of Supernetting On Routing* 154
- 10.17 *Summary* 155

Chapter 11 Protocol Layering 159

- 11.1 *Introduction* 159
- 11.2 *The Need For Multiple Protocols* 159
- 11.3 *The Conceptual Layers Of Protocol Software* 160
- 11.4 *Functionality Of The Layers* 163
- 11.5 *X.25 And Its Relation To The ISO Model* 164
- 11.6 *Differences Between X.25 And Internet Layering* 167
- 11.7 *The Protocol Layering Principle* 169
- 11.8 *Layering In The Presence Of Network Substructure* 171
- 11.9 *Two Important Boundaries In The TCP/IP Model* 173
- 11.10 *The Disadvantage Of Layering* 174
- 11.11 *The Basic Idea Behind Multiplexing And Demultiplexing* 174
- 11.12 *Summary* 176

Chapter 12 User Datagram Protocol (UDP) 179

- 12.1 *Introduction* 179
- 12.2 *Identifying The Ultimate Destination* 179
- 12.3 *The User Datagram Protocol* 180
- 12.4 *Format Of UDP Messages* 181
- 12.5 *UDP Pseudo-Header* 182
- 12.6 *UDP Encapsulation And Protocol Layering* 183
- 12.7 *Layering And The UDP Checksum Computation* 185
- 12.8 *UDP Multiplexing, Demultiplexing, And Ports* 185
- 12.9 *Reserved And Available UDP Port Numbers* 186
- 12.10 *Summary* 188

Chapter 13 Reliable Stream Transport Service (TCP) 191

- 13.1 *Introduction* 191
- 13.2 *The Need For Stream Delivery* 191
- 13.3 *Properties Of The Reliable Delivery Service* 192
- 13.4 *Providing Reliability* 193
- 13.5 *The Idea Behind Sliding Windows* 195
- 13.6 *The Transmission Control Protocol* 198
- 13.7 *Ports, Connections, And Endpoints* 199
- 13.8 *Passive And Active Opens* 201
- 13.9 *Segments, Streams, And Sequence Numbers* 201
- 13.10 *Variable Window Size And Flow Control* 202
- 13.11 *TCP Segment Format* 203

13.12	<i>Out Of Band Data</i>	205	
13.13	<i>Maximum Segment Size Option</i>	206	
13.14	<i>TCP Checksum Computation</i>	207	
13.15	<i>Acknowledgements And Retransmission</i>	208	
13.16	<i>Timeout And Retransmission</i>	209	
13.17	<i>Accurate Measurement Of Round Trip Samples</i>	211	
13.18	<i>Karn's Algorithm And Timer Backoff</i>	212	
13.19	<i>Responding To High Variance In Delay</i>	213	
13.20	<i>Response To Congestion</i>	214	
13.21	<i>Establishing A TCP Connection</i>	216	
13.22	<i>Initial Sequence Numbers</i>	217	
13.23	<i>Closing a TCP Connection</i>	217	
13.24	<i>TCP Connection Reset</i>	219	
13.25	<i>TCP State Machine</i>	219	
13.26	<i>Forcing Data Delivery</i>	221	
13.27	<i>Reserved TCP Port Numbers</i>	221	
13.28	<i>TCP Performance</i>	221	
13.29	<i>Silly Window Syndrome And Small Packets</i>	223	
13.30	<i>Avoiding Silly Window Syndrome</i>	224	
13.31	<i>Summary</i>	227	
 Chapter 14 Routing: Cores, Peers, And Algorithms (GGP)			231
14.1	<i>Introduction</i>	231	
14.2	<i>The Origin Of Routing Tables</i>	232	
14.3	<i>Routing With Partial Information</i>	233	
14.4	<i>Original Internet Architecture And Cores</i>	234	
14.5	<i>Core Routers</i>	235	
14.6	<i>Beyond The Core Architecture To Peer Backbones</i>	238	
14.7	<i>Automatic Route Propagation</i>	240	
14.8	<i>Vector Distance (Bellman-Ford) Routing</i>	240	
14.9	<i>Gateway-To-Gateway Protocol (GGP)</i>	242	
14.10	<i>GGP Message Formats</i>	243	
14.11	<i>Link-State (SPF) Routing</i>	245	
14.12	<i>SPF Protocols</i>	246	
14.13	<i>Summary</i>	246	
 Chapter 15 Routing: Autonomous Systems (EGP)			249
15.1	<i>Introduction</i>	249	
15.2	<i>Adding Complexity To The Architectural Model</i>	249	
15.3	<i>A Fundamental Idea: Extra Hops</i>	250	

15.4	<i>Autonomous System Concept</i>	252
15.5	<i>Exterior Gateway Protocol (EGP)</i>	254
15.6	<i>EGP Message Header</i>	255
15.7	<i>EGP Neighbor Acquisition Messages</i>	256
15.8	<i>EGP Neighbor Reachability Messages</i>	257
15.9	<i>EGP Poll Request Messages</i>	258
15.10	<i>EGP Routing Update Messages</i>	259
15.11	<i>Measuring From The Receiver's Perspective</i>	261
15.12	<i>The Key Restriction Of EGP</i>	262
15.13	<i>Technical Problems</i>	264
15.14	<i>Decentralization Of Internet Architecture</i>	264
15.15	<i>Beyond Autonomous Systems</i>	264
15.16	<i>Summary</i>	265

Chapter 16 Routing: In An Autonomous System (RIP, OSPF, HELLO) 267

16.1	<i>Introduction</i>	267
16.2	<i>Static Vs. Dynamic Interior Routes</i>	267
16.3	<i>Routing Information Protocol (RIP)</i>	270
16.4	<i>The Hello Protocol</i>	276
16.5	<i>Combining RIP, Hello, And EGP</i>	278
16.6	<i>The Open SPF Protocol (OSPF)</i>	279
16.7	<i>Routing With Partial Information</i>	286
16.8	<i>Summary</i>	286

Chapter 17 Internet Multicasting (IGMP)

289

17.1	<i>Introduction</i>	289
17.2	<i>Hardware Broadcast</i>	289
17.3	<i>Hardware Multicast</i>	290
17.4	<i>IP Multicast</i>	291
17.5	<i>IP Multicast Addresses</i>	291
17.6	<i>Mapping IP Multicast To Ethernet Multicast</i>	292
17.7	<i>Extending IP To Handle Multicasting</i>	293
17.8	<i>Internet Group Management Protocol</i>	294
17.9	<i>IGMP Implementation</i>	294
17.10	<i>Group Membership State Transitions</i>	295
17.11	<i>IGMP Message Format</i>	296
17.12	<i>Multicast Address Assignment</i>	297
17.13	<i>Propagating Routing Information</i>	297
17.14	<i>The MROUTED Program</i>	298
17.15	<i>Summary</i>	300

Chapter 18 TCP/IP Over ATM Networks	303
18.1 Introduction	303
18.2 ATM Hardware	304
18.3 Large ATM Networks	304
18.4 The Logical View Of An ATM Network	305
18.5 The Two ATM Connection Paradigms	306
18.6 Paths, Circuits, And Identifiers	307
18.7 ATM Cell Transport	308
18.8 ATM Adaptation Layers	308
18.9 AAL5 Convergence, Segmentation, And Reassembly	311
18.10 Datagram Encapsulation And IP MTU Size	311
18.11 Packet Type And Multiplexing	312
18.12 IP Address Binding In An ATM Network	313
18.13 Logical IP Subnet Concept	314
18.14 Connection Management	315
18.15 Address Binding Within An LIS	316
18.16 ATMARP Packet Format	316
18.17 Using ATMARP Packets To Determine An Address	318
18.18 Obtaining Entries For A Server Database	320
18.19 Timing Out ATMARP Information In A Server	320
18.20 Timing Out ATMARP Information In A Host Or Router	320
18.21 Summary	321
Chapter 19 Client-Server Model Of Interaction	325
19.1 Introduction	325
19.2 The Client-Server Model	325
19.3 A Simple Example: UDP Echo Server	326
19.4 Time And Date Service	328
19.5 The Complexity of Servers	329
19.6 RARP Server	330
19.7 Alternatives To The Client-Server Model	331
19.8 Summary	332
Chapter 20 The Socket Interface	335
20.1 Introduction	335
20.2 The UNIX I/O Paradigm And Network I/O	336
20.3 Adding Network I/O to UNIX	336
20.4 The Socket Abstraction	337

- 20.5 *Creating A Socket* 337
- 20.6 *Socket Inheritance And Termination* 338
- 20.7 *Specifying A Local Address* 339
- 20.8 *Connecting Sockets To Destination Addresses* 340
- 20.9 *Sending Data Through A Socket* 341
- 20.10 *Receiving Data Through A Socket* 343
- 20.11 *Obtaining Local And Remote Socket Addresses* 344
- 20.12 *Obtaining And Setting Socket Options* 345
- 20.13 *Specifying A Queue Length For A Server* 346
- 20.14 *How A Server Accepts Connections* 346
- 20.15 *Servers That Handle Multiple Services* 347
- 20.16 *Obtaining And Setting Host Names* 348
- 20.17 *Obtaining And Setting The Internal Host Domain* 349
- 20.18 *BSD UNIX Network Library Calls* 349
- 20.19 *Network Byte Order Conversion Routines* 350
- 20.20 *IP Address Manipulation Routines* 351
- 20.21 *Accessing The Domain Name System* 352
- 20.22 *Obtaining Information About Hosts* 354
- 20.23 *Obtaining Information About Networks* 355
- 20.24 *Obtaining Information About Protocols* 355
- 20.25 *Obtaining Information About Network Services* 356
- 20.26 *An Example Client* 357
- 20.27 *An Example Server* 359
- 20.28 *Summary* 362

Chapter 21 Bootstrap And Autoconfiguration (BOOTP, DHCP)

365

- 21.1 *Introduction* 365
- 21.2 *The Need For An Alternative To RARP* 366
- 21.3 *Using IP To Determine An IP Address* 366
- 21.4 *The BOOTP Retransmission Policy* 367
- 21.5 *The BOOTP Message Format* 368
- 21.6 *The Two-Step Bootstrap Procedure* 369
- 21.7 *Vendor-Specific Field* 370
- 21.8 *The Need For Dynamic Configuration* 370
- 21.9 *Dynamic Host Configuration* 372
- 21.10 *Dynamic IP Address Assignment* 372
- 21.11 *Obtaining Multiple Addresses* 373
- 21.12 *Address Acquisition States* 374
- 21.13 *Early Lease Termination* 374
- 21.14 *Lease Renewal States* 376
- 21.15 *DHCP Message Format* 377
- 21.16 *DHCP Options And Message Type* 378

- 21.17 *Option Overload* 379
- 21.18 *DHCP And Domain Names* 379
- 21.19 *Summary* 380

Chapter 22 The Domain Name System (DNS) 383

- 22.1 *Introduction* 383
- 22.2 *Names For Machines* 384
- 22.3 *Flat Namespace* 384
- 22.4 *Hierarchical Names* 385
- 22.5 *Delegation Of Authority For Names* 386
- 22.6 *Subset Authority* 386
- 22.7 *TCP/IP Internet Domain Names* 387
- 22.8 *Official And Unofficial Internet Domain Names* 388
- 22.9 *Items Named And Syntax Of Names* 390
- 22.10 *Mapping Domain Names To Addresses* 391
- 22.11 *Domain Name Resolution* 393
- 22.12 *Efficient Translation* 394
- 22.13 *Caching: The Key To Efficiency* 395
- 22.14 *Domain Server Message Format* 396
- 22.15 *Compressed Name Format* 399
- 22.16 *Abbreviation Of Domain Names* 399
- 22.17 *Inverse Mappings* 400
- 22.18 *Pointer Queries* 401
- 22.19 *Object Types And Resource Record Contents* 401
- 22.20 *Obtaining Authority For A Subdomain* 402
- 22.21 *Summary* 403

Chapter 23 Applications: Remote Login (TELNET, Rlogin) 407

- 23.1 *Introduction* 407
- 23.2 *Remote Interactive Computing* 407
- 23.3 *TELNET Protocol* 408
- 23.4 *Accommodating Heterogeneity* 410
- 23.5 *Passing Commands That Control The Remote Side* 412
- 23.6 *Forcing The Server To Read A Control Function* 414
- 23.7 *TELNET Options* 414
- 23.8 *TELNET Option Negotiation* 415
- 23.9 *Rlogin (BSD UNIX)* 416
- 23.10 *Summary* 417

Chapter 24 Applications: File Transfer And Access (FTP, TFTP, NFS) 419

- 24.1 *Introduction* 419
- 24.2 *File Access And Transfer* 419
- 24.3 *On-line Shared Access* 420
- 24.4 *Sharing By File Transfer* 421
- 24.5 *FTP: The Major TCP/IP File Transfer Protocol* 421
- 24.6 *FTP Features* 422
- 24.7 *FTP Process Model* 422
- 24.8 *TCP Port Number Assignment* 424
- 24.9 *The User's View Of FTP* 424
- 24.10 *An Example Anonymous FTP Session* 426
- 24.11 *TFTP* 427
- 24.12 *NFS* 429
- 24.13 *NFS Implementation* 429
- 24.14 *Remote Procedure Call (RPC)* 430
- 24.15 *Summary* 431

Chapter 25 Applications: Electronic Mail (822, SMTP, MIME) 433

- 25.1 *Introduction* 433
- 25.2 *Electronic Mail* 433
- 25.3 *Mailbox Names And Aliases* 435
- 25.4 *Alias Expansion And Mail Forwarding* 435
- 25.5 *The Relationship Of Internetworking And Mail* 436
- 25.6 *TCP/IP Standards For Electronic Mail Service* 438
- 25.7 *Electronic Mail Addresses* 438
- 25.8 *Pseudo Domain Addresses* 440
- 25.9 *Simple Mail Transfer Protocol (SMTP)* 440
- 25.10 *The MIME Extension For Non-ASCII Data* 443
- 25.11 *MIME Multipart Messages* 444
- 25.12 *Summary* 445

Chapter 26 Applications: Internet Management (SNMP, SNMPv2) 447

- 26.1 *Introduction* 447
- 26.2 *The Level Of Management Protocols* 447
- 26.3 *Architectural Model* 448
- 26.4 *Protocol Architecture* 450
- 26.5 *Examples of MIB Variables* 451
- 26.6 *The Structure Of Management Information* 452

26.7	<i>Formal Definitions Using ASN.1</i>	453
26.8	<i>Structure And Representation Of MIB Object Names</i>	453
26.9	<i>Simple Network Management Protocol</i>	458
26.10	<i>SNMP Message Format</i>	460
26.11	<i>Example Encoded SNMP Message</i>	462
26.12	<i>Summary</i>	463
Chapter 27	Summary Of Protocol Dependencies	465
27.1	<i>Introduction</i>	465
27.2	<i>Protocol Dependencies</i>	465
27.3	<i>Application Program Access</i>	467
27.4	<i>Summary</i>	468
Chapter 28	Internet Security And Firewall Design	471
28.1	<i>Introduction</i>	471
28.2	<i>Protecting Resources</i>	472
28.3	<i>The Need For An Information Policy</i>	472
28.4	<i>Communication, Cooperation, And Mutual Mistrust</i>	474
28.5	<i>Mechanisms For Internet Security</i>	475
28.6	<i>Firewalls And Internet Access</i>	476
28.7	<i>Multiple Connections And Weakest Links</i>	477
28.8	<i>Firewall Implementation And High-Speed Hardware</i>	478
28.9	<i>Packet-Level Filters</i>	479
28.10	<i>Security And Packet Filter Specification</i>	480
28.11	<i>The Consequence Of Restricted Access For Clients</i>	481
28.12	<i>Accessing Services Through A Firewall</i>	481
28.13	<i>The Details Of Firewall Architecture</i>	483
28.14	<i>Stub Network</i>	484
28.15	<i>An Alternative Firewall Implementation</i>	484
28.16	<i>Monitoring And Logging</i>	485
28.17	<i>Summary</i>	486
Chapter 29	The Future Of TCP/IP (IPng, IPv6)	489
29.1	<i>Introduction</i>	489
29.2	<i>Why Change TCP/IP And The Internet?</i>	490
29.3	<i>Motivation For Changing IPv4</i>	491
29.4	<i>The Road To A New Version Of IP</i>	492
29.5	<i>The Name Of The Next IP</i>	492

29.6	<i>Features Of IPv6</i>	493
29.7	<i>General Form Of An IPv6 Datagram</i>	494
29.8	<i>IPv6 Base Header Format</i>	494
29.9	<i>IPv6 Extension Headers</i>	496
29.10	<i>Parsing An IPv6 Datagram</i>	497
29.11	<i>IPv6 Fragmentation And Reassembly</i>	498
29.12	<i>The Consequence Of End-To-End Fragmentation</i>	498
29.13	<i>IPv6 Source Routing</i>	500
29.14	<i>IPv6 Options</i>	500
29.15	<i>Size Of The IPv6 Address Space</i>	502
29.16	<i>IPv6 Colon Hexadecimal Notation</i>	502
29.17	<i>Three Basic IPv6 Address Types</i>	503
29.18	<i>The Duality Of Broadcast And Multicast</i>	504
29.19	<i>An Engineering Choice And Simulated Broadcast</i>	504
29.20	<i>Proposed IPv6 Address Space Assignment</i>	504
29.21	<i>IPv4 Address Encoding And Transition</i>	506
29.22	<i>Providers, Subscribers, And Address Hierarchy</i>	506
29.23	<i>Additional Hierarchy</i>	507
29.24	<i>Summary</i>	508
Appendix 1	A Guide To RFCs	511
Appendix 2	Glossary Of Internetworking Terms And Abbreviations	557
Bibliography		591
Index		599

Contents

Foreword	xvii
Preface	xix
Chapter 1 Introduction And Overview	1
1.1 TCP/IP Protocols	1
1.2 The Need To Understand Details	1
1.3 Complexity Of Interactions Among Protocols	2
1.4 The Approach In This Text	2
1.5 The Importance Of Studying Code	3
1.6 The Xinu Operating System	3
1.7 Organization Of The Remainder Of The Book	4
1.8 Summary	4
Chapter 2 The Structure Of TCP/IP Software In An Operating System	7
2.1 Introduction	7
2.2 The Process Concept	8
2.3 Process Priority	9
2.4 Communicating Processes	9
2.5 Interprocess Communication	12
2.6 Device Drivers, Input, And Output	14
2.7 Network Input and Interrupts	14
2.8 Passing Packets To Higher Level Protocols	16
2.9 Passing Datagrams From IP To Transport Protocols	16
2.10 Delivery To Application Programs	18
2.11 Information Flow On Output	19
2.12 From TCP Through IP To Network Output	20

- 2.13 *UDP Output* 21
- 2.14 *Summary* 21

Chapter 3 Network Interface Layer **27**

- 3.1 *Introduction* 27
- 3.2 *The Network Interface Abstraction* 28
- 3.3 *Logical State Of An Interface* 31
- 3.4 *Local Host Interface* 31
- 3.5 *Buffer Management* 32
- 3.6 *Demultiplexing Incoming Packets* 35
- 3.7 *Summary* 36

Chapter 4 Address Discovery And Binding (ARP) **39**

- 4.1 *Introduction* 39
- 4.2 *Conceptual Organization Of ARP Software* 40
- 4.3 *Example ARP Design* 40
- 4.4 *Data Structures For The ARP Cache* 41
- 4.5 *ARP Output Processing* 44
- 4.6 *ARP Input Processing* 49
- 4.7 *ARP Cache Management* 53
- 4.8 *ARP Initialization* 58
- 4.9 *ARP Configuration Parameters* 59
- 4.10 *Summary* 59

Chapter 5 IP: Global Software Organization **61**

- 5.1 *Introduction* 61
- 5.2 *The Central Switch* 61
- 5.3 *IP Software Design* 62
- 5.4 *IP Software Organization And Datagram Flow* 63
- 5.5 *Byte-Ordering In The IP Header* 76
- 5.6 *Sending A Datagram To IP* 77
- 5.7 *Table Maintenance* 80
- 5.8 *Summary* 82

Chapter 6 IP: Routing Table And Routing Algorithm 85

- 6.1 *Introduction* 85
- 6.2 *Route Maintenance And Lookup* 85
- 6.3 *Routing Table Organization* 86
- 6.4 *Routing Table Data Structures* 87
- 6.5 *Origin Of Routes And Persistence* 89
- 6.6 *Routing A Datagram* 89
- 6.7 *Periodic Route Table Maintenance* 96
- 6.8 *IP Options Processing* 104
- 6.9 *Summary* 105

Chapter 7 IP: Fragmentation And Reassembly 107

- 7.1 *Introduction* 107
- 7.2 *Fragmenting Datagrams* 107
- 7.3 *Implementation Of Fragmentation* 108
- 7.4 *Datagram Reassembly* 113
- 7.5 *Maintenance Of Fragment Lists* 122
- 7.6 *Initialization* 124
- 7.7 *Summary* 124

Chapter 8 IP: Error Processing (ICMP) 127

- 8.1 *Introduction* 127
- 8.2 *ICMP Message Formats* 127
- 8.3 *Implementation Of ICMP Messages* 127
- 8.4 *Handling Incoming ICMP Messages* 130
- 8.5 *Handling An ICMP Redirect Message* 132
- 8.6 *Setting A Subnet Mask* 133
- 8.7 *Choosing A Source Address For An ICMP Packet* 135
- 8.8 *Generating ICMP Error Messages* 136
- 8.9 *Avoiding Errors About Errors* 139
- 8.10 *Allocating A Buffer For ICMP* 140
- 8.11 *The Data Portion Of An ICMP Message* 142
- 8.12 *Generating An ICMP Redirect Message* 144
- 8.13 *Summary* 145

Chapter 9	IP: Multicast Processing (IGMP)	147
9.1	<i>Introduction</i>	147
9.2	<i>Maintaining Multicast Group Membership Information</i>	147
9.3	<i>A Host Group Table</i>	148
9.4	<i>Searching For A Host Group</i>	150
9.5	<i>Adding A Host Group Entry To The Table</i>	151
9.6	<i>Configuring The Network Interface For A Multicast Address</i>	152
9.7	<i>Translation Between IP and Hardware Multicast Addresses</i>	154
9.8	<i>Removing A Multicast Address From The Host Group Table</i>	156
9.9	<i>Joining A Host Group</i>	157
9.10	<i>Maintaining Contact With A Multicast Router</i>	158
9.11	<i>Implementing IGMP Membership Reports</i>	160
9.12	<i>Computing A Random Delay</i>	161
9.13	<i>A Process To Send IGMP Reports</i>	163
9.14	<i>Handling Incoming IGMP Messages</i>	164
9.15	<i>Leaving A Host Group</i>	165
9.16	<i>Initialization Of IGMP Data Structures</i>	167
9.17	<i>Summary</i>	168
Chapter 10	UDP: User Datagrams	171
10.1	<i>Introduction</i>	171
10.2	<i>UDP Ports And Demultiplexing</i>	171
10.3	<i>UDP</i>	175
10.4	<i>UDP Output Processing</i>	185
10.5	<i>Summary</i>	188
Chapter 11	TCP: Data Structures And Input Processing	191
11.1	<i>Introduction</i>	191
11.2	<i>Overview Of TCP Software</i>	192
11.3	<i>Transmission Control Blocks</i>	192
11.4	<i>TCP Segment Format</i>	196
11.5	<i>Sequence Space Comparison</i>	198
11.6	<i>TCP Finite State Machine</i>	199
11.7	<i>Example State Transition</i>	200
11.8	<i>Declaration Of The Finite State Machine</i>	200
11.9	<i>TCB Allocation And Initialization</i>	202
11.10	<i>Implementation Of The Finite State Machine</i>	204
11.11	<i>Handling An Input Segment</i>	205

11.12 *Summary* 214

Chapter 12 TCP: Finite State Machine Implementation

217

- 12.1 *Introduction* 217
- 12.2 *CLOSED State Processing* 217
- 12.3 *Graceful Shutdown* 218
- 12.4 *Timed Delay After Closing* 218
- 12.5 *TIME-WAIT State Processing* 219
- 12.6 *CLOSING State Processing* 221
- 12.7 *FIN-WAIT-2 State Processing* 222
- 12.8 *FIN-WAIT-1 State Processing* 223
- 12.9 *CLOSE-WAIT State Processing* 225
- 12.10 *LAST-ACK State Processing* 227
- 12.11 *ESTABLISHED State Processing* 228
- 12.12 *Processing Urgent Data In A Segment* 229
- 12.13 *Processing Other Data In A Segment* 231
- 12.14 *Keeping Track Of Received Octets* 233
- 12.15 *Aborting A TCP Connection* 236
- 12.16 *Establishing A TCP Connection* 237
- 12.17 *Initializing A TCB* 237
- 12.18 *SYN-SENT State Processing* 239
- 12.19 *SYN-RECEIVED State Processing* 240
- 12.20 *LISTEN State Processing* 243
- 12.21 *Initializing Window Variables For A New TCB* 244
- 12.22 *Summary* 246

Chapter 13 TCP: Output Processing

247

- 13.1 *Introduction* 247
- 13.2 *Controlling TCP Output Complexity* 247
- 13.3 *The Four TCP Output States* 248
- 13.4 *TCP Output As A Process* 248
- 13.5 *TCP Output Messages* 249
- 13.6 *Encoding Output States And TCB Numbers* 250
- 13.7 *Implementation Of The TCP Output Process* 250
- 13.8 *Mutual Exclusion* 251
- 13.9 *Implementation Of The IDLE State* 252
- 13.10 *Implementation Of The PERSIST State* 252
- 13.11 *Implementation Of The TRANSMIT State* 253
- 13.12 *Implementation Of The RETRANSMIT State* 255
- 13.13 *Sending A Segment* 255

- 13.14 *Computing The TCP Data Length* 259
- 13.15 *Computing Sequence Counts* 260
- 13.16 *Other TCP Procedures* 261
- 13.17 *Summary* 267

Chapter 14 TCP: Timer Management

269

- 14.1 *Introduction* 269
- 14.2 *A General Data Structure For Timed Events* 269
- 14.3 *A Data Structure For TCP Events* 270
- 14.4 *Timers, Events, And Messages* 271
- 14.5 *The TCP Timer Process* 272
- 14.6 *Deleting A TCP Timer Event* 274
- 14.7 *Deleting All Events For A TCB* 275
- 14.8 *Determining The Time Remaining For An Event* 276
- 14.9 *Inserting A TCP Timer Event* 277
- 14.10 *Starting TCP Output Without Delay* 279
- 14.11 *Summary* 280

Chapter 15 TCP: Flow Control And Adaptive Retransmission

283

- 15.1 *Introduction* 283
- 15.2 *The Difficulties With Adaptive Retransmission* 284
- 15.3 *Tuning Adaptive Retransmission* 284
- 15.4 *Retransmission Timer And Backoff* 284
- 15.5 *Window-Based Flow Control* 287
- 15.6 *Maximum Segment Size Computation* 291
- 15.7 *Congestion Avoidance And Control* 295
- 15.8 *Slow-Start And Congestion Avoidance* 296
- 15.9 *Round Trip Estimation And Timeout* 299
- 15.10 *Miscellaneous Notes And Techniques* 305
- 15.11 *Summary* 306

Chapter 16 TCP: Urgent Data Processing And The Push Function

309

- 16.1 *Introduction* 309
- 16.2 *Out-Of-Band Signaling* 309
- 16.3 *Urgent Data* 310
- 16.4 *Interpreting The Standard* 310
- 16.5 *Configuration For Berkeley Urgent Pointer Interpretation* 313
- 16.6 *Informing An Application* 313

- 16.7 *Reading Data From TCP* 314
- 16.8 *Sending Urgent Data* 316
- 16.9 *TCP Push Function* 317
- 16.10 *Interpreting Push With Out-Of-Order Delivery* 318
- 16.11 *Implementation Of Push On Input* 319
- 16.12 *Summary* 320

Chapter 17 Socket-Level Interface 323

- 17.1 *Introduction* 323
- 17.2 *Interfacing Through A Device* 323
- 17.3 *TCP Connections As Devices* 325
- 17.4 *An Example TCP Client Program* 326
- 17.5 *An Example TCP Server Program* 327
- 17.6 *Implementation Of The TCP Master Device* 329
- 17.7 *Implementation Of A TCP Slave Device* 337
- 17.8 *Initialization Of A Slave Device* 351
- 17.9 *Summary* 352

Chapter 18 RIP: Active Route Propagation And Passive Acquisition 355

- 18.1 *Introduction* 355
- 18.2 *Active And Passive Mode Participants* 356
- 18.3 *Basic RIP Algorithm And Cost Metric* 356
- 18.4 *Instabilities And Solutions* 357
- 18.5 *Message Types* 361
- 18.6 *Protocol Characterization* 361
- 18.7 *Implementation Of RIP* 362
- 18.8 *The Principle RIP Process* 365
- 18.9 *Responding To An Incoming Request* 370
- 18.10 *Generating Update Messages* 372
- 18.11 *Initializing Copies Of An Update Message* 373
- 18.12 *Generating Periodic RIP Output* 378
- 18.13 *Limitations Of RIP* 379
- 18.14 *Summary* 379

Chapter 19 OSPF: Route Propagation With An SPF Algorithm 381

- 19.1 *Introduction* 381
- 19.2 *OSPF Configuration And Options* 382
- 19.3 *OSPF's Graph-Theoretic Model* 382

19.4	<i>OSPF Declarations</i>	386
19.5	<i>Adjacency And Link State Propagation</i>	391
19.6	<i>Discovering Neighboring Gateways With Hello</i>	392
19.7	<i>Sending Hello Packets</i>	394
19.8	<i>Designated Router Concept</i>	399
19.9	<i>Electing A Designated Router</i>	400
19.10	<i>Reforming Adjacencies After A Change</i>	404
19.11	<i>Handling Arriving Hello Packets</i>	406
19.12	<i>Adding A Gateway To The Neighbor List</i>	408
19.13	<i>Neighbor State Transitions</i>	410
19.14	<i>OSPF Timer Events And Retransmissions</i>	412
19.15	<i>Determining Whether Adjacency Is Permitted</i>	414
19.16	<i>Handling OSPF input</i>	415
19.17	<i>Declarations And Procedures For Link State Processing</i>	418
19.18	<i>Generating Database Description Packets</i>	421
19.19	<i>Creating A Template</i>	422
19.20	<i>Transmitting A Database Description Packet</i>	424
19.21	<i>Handling An Arriving Database Description Packet</i>	426
19.22	<i>Handling Link State Request Packets</i>	432
19.23	<i>Building A Link State Summary</i>	434
19.24	<i>OSPF Utility Procedures</i>	435
19.25	<i>Summary</i>	439

Chapter 20 SNMP: MIB Variables, Representations, And Bindings 441

20.1	<i>Introduction</i>	441
20.2	<i>Server Organization And Name Mapping</i>	442
20.3	<i>MIB Variables</i>	443
20.4	<i>MIB Variable Names</i>	444
20.5	<i>Lexicographic Ordering Among Names</i>	445
20.6	<i>Prefix Removal</i>	445
20.7	<i>Operations Applied To MIB Variables</i>	446
20.8	<i>Names For Tables</i>	446
20.9	<i>Conceptual Threading Of The Name Hierarchy</i>	447
20.10	<i>Data Structure For MIB Variables</i>	448
20.11	<i>A Data Structure For Fast Lookup</i>	450
20.12	<i>Implementation Of The Hash Table</i>	452
20.13	<i>Specification Of MIB Bindings</i>	452
20.14	<i>Internal Variables Used In Bindings</i>	457
20.15	<i>Hash Table Lookup</i>	458
20.16	<i>SNMP Structures And Constants</i>	461
20.17	<i>ASN.1 Representation Manipulation</i>	464
20.18	<i>Summary</i>	474

8.81
 9.81
 01.81
 11.81
 21.81
 31.81

PRINTED BY TAYLOR & FRANCIS

Chapter 21 SNMP: Client And Server 477

- 21.1 *Introduction* 477
- 21.2 *Data Representation In The Server* 477
- 21.3 *Server Implementation* 478
- 21.4 *Parsing An SNMP Message* 480
- 21.5 *Converting ASN.1 Names In The Binding List* 484
- 21.6 *Resolving A Query* 485
- 21.7 *Interpreting The Get-Next Operation* 487
- 21.8 *Indirect Application Of Operations* 487
- 21.9 *Indirection For Tables* 490
- 21.10 *Generating A Reply Message Backward* 491
- 21.11 *Converting From Internal Form to ASN.1* 494
- 21.12 *Utility Functions Used By The Server* 495
- 21.13 *Implementation Of An SNMP Client* 496
- 21.14 *Initialization Of Variables* 498
- 21.15 *Summary* 500

Chapter 22 SNMP: Table Access Functions 503

- 22.1 *Introduction* 503
- 22.2 *Table Access* 504
- 22.3 *Object Identifiers For Tables* 504
- 22.4 *Address Entry Table Functions* 504
- 22.5 *Address Translation Table Functions* 511
- 22.6 *Network Interface Table Functions* 521
- 22.7 *Routing Table Functions* 529
- 22.8 *TCP Connection Table Functions* 538
- 22.9 *Summary* 545

Chapter 23 Implementation In Retrospect 547

- 23.1 *Introduction* 547
- 23.2 *Statistical Analysis Of The Code* 547
- 23.3 *Lines Of Code For Each Protocol* 548
- 23.4 *Functions And Procedures For Each Protocol* 550
- 23.5 *Summary* 551

Appendix 1 Cross Reference Of Procedure Calls	553
Appendix 2 Xinu Functions And Constants Used In The Code	573
Bibliography	589
Index	597

1.15
 1.16
 1.17
 Chapter 2: XINU
 2.1
 2.2
 2.3
 2.4
 2.5
 2.6
 2.7
 2.8
 2.9
 2.10
 2.11
 2.12
 2.13
 2.14
 2.15
 2.16
 2.17
 2.18
 2.19
 2.20
 2.21
 2.22
 2.23
 2.24
 2.25
 2.26
 2.27
 2.28
 2.29
 2.30
 2.31
 2.32
 2.33
 2.34
 2.35
 2.36
 2.37
 2.38
 2.39
 2.40
 2.41
 2.42
 2.43
 2.44
 2.45
 2.46
 2.47
 2.48
 2.49
 2.50
 2.51
 2.52
 2.53
 2.54
 2.55
 2.56
 2.57
 2.58
 2.59
 2.60
 2.61
 2.62
 2.63
 2.64
 2.65
 2.66
 2.67
 2.68
 2.69
 2.70
 2.71
 2.72
 2.73
 2.74
 2.75
 2.76
 2.77
 2.78
 2.79
 2.80
 2.81
 2.82
 2.83
 2.84
 2.85
 2.86
 2.87
 2.88
 2.89
 2.90
 2.91
 2.92
 2.93
 2.94
 2.95
 2.96
 2.97
 2.98
 2.99
 2.100

Contents

Foreword	xxi
Preface	xxiii
Chapter 1 Introduction And Overview	1
1.1 <i>Use Of TCP/IP</i>	1
1.2 <i>Designing Applications For A Distributed Environment</i>	2
1.3 <i>Standard And Nonstandard Application Protocols</i>	2
1.4 <i>An Example Of Standard Application Protocol Use</i>	2
1.5 <i>An Example Connection</i>	3
1.6 <i>Using TELNET To Access An Alternative Service</i>	4
1.7 <i>Application Protocols And Software Flexibility</i>	5
1.8 <i>Viewing Services From The Provider's Perspective</i>	6
1.9 <i>The Remainder Of This Text</i>	7
1.10 <i>Summary</i>	7
Chapter 2 The Client Server Model And Software Design	9
2.1 <i>Introduction</i>	9
2.2 <i>Motivation</i>	10
2.3 <i>Terminology And Concepts</i>	10
2.4 <i>Summary</i>	18
Chapter 3 Concurrent Processing In Client-Server Software	21
3.1 <i>Introduction</i>	21
3.2 <i>Concurrency In Networks</i>	21

- 3.3 *Concurrency In Servers* 23
- 3.4 *Terminology And Concepts* 24
- 3.5 *An Example Of Concurrent Process Creation* 26
- 3.6 *Executing New Code* 31
- 3.7 *Context Switching And Protocol Software Design* 31
- 3.8 *Concurrency And Asynchronous I/O* 32
- 3.9 *Summary* 32

Chapter 4 Program Interface To Protocols

35

- 4.1 *Introduction* 35
- 4.2 *Loosely Specified Protocol Software Interface* 35
- 4.3 *Interface Functionality* 36
- 4.4 *Conceptual Interface Specification* 37
- 4.5 *System Calls* 37
- 4.6 *Two Basic Approaches To Network Communication* 38
- 4.7 *The Basic I/O Functions Available In UNIX* 39
- 4.8 *Using UNIX I/O With TCP/IP* 40
- 4.9 *Summary* 40

Chapter 5 The Socket Interface

43

- 5.1 *Introduction* 43
- 5.2 *Berkeley Sockets* 43
- 5.3 *Specifying A Protocol Interface* 44
- 5.4 *The Socket Abstraction* 45
- 5.5 *Specifying An Endpoint Address* 47
- 5.6 *A Generic Address Structure* 48
- 5.7 *Major System Calls Used With Sockets* 49
- 5.8 *Utility Routines For Integer Conversion* 52
- 5.9 *Using Socket Calls In A Program* 53
- 5.10 *Symbolic Constants For Socket Call Parameters* 54
- 5.11 *Summary* 54

Chapter 6 Algorithms And Issues In Client Software Design

57

- 6.1 *Introduction* 57
- 6.2 *Learning Algorithms Instead Of Details* 57
- 6.3 *Client Architecture* 58
- 6.4 *Identifying The Location Of A Server* 58
- 6.5 *Parsing An Address Argument* 60

6.6	<i>Looking Up A Domain Name</i>	61
6.7	<i>Looking Up A Well-Known Port By Name</i>	62
6.8	<i>Port Numbers And Network Byte Order</i>	62
6.9	<i>Looking Up A Protocol By Name</i>	63
6.10	<i>The TCP Client Algorithm</i>	63
6.11	<i>Allocating A Socket</i>	64
6.12	<i>Choosing A Local Protocol Port Number</i>	65
6.13	<i>A Fundamental Problem In Choosing A Local IP Address</i>	65
6.14	<i>Connecting A TCP Socket To A Server</i>	66
6.15	<i>Communicating With The Server Using TCP</i>	66
6.16	<i>Reading A Response From A TCP Connection</i>	67
6.17	<i>Closing A TCP Connection</i>	68
6.18	<i>Programming A UDP Client</i>	69
6.19	<i>Connected And Unconnected UDP Sockets</i>	69
6.20	<i>Using Connect With UDP</i>	70
6.21	<i>Communicating With A Server Using UDP</i>	70
6.22	<i>Closing A Socket That Uses UDP</i>	70
6.23	<i>Partial Close For UDP</i>	71
6.24	<i>A Warning About UDP Unreliability</i>	71
6.25	<i>Summary</i>	71

Chapter 7 Example Client Software

75

7.1	<i>Introduction</i>	75
7.2	<i>The Importance Of Small Examples</i>	75
7.3	<i>Hiding Details</i>	76
7.4	<i>An Example Procedure Library For Client Programs</i>	76
• 7.5	<i>Implementation Of ConnectTCP</i>	77
7.6	<i>Implementation Of ConnectUDP</i>	78
7.7	<i>A Procedure That Forms Connections</i>	78
7.8	<i>Using The Example Library</i>	81
7.9	<i>The DAYTIME Service</i>	81
7.10	<i>Implementation Of A TCP Client For DAYTIME</i>	82
7.11	<i>Reading From A TCP Connection</i>	83
7.12	<i>The TIME Service</i>	84
7.13	<i>Accessing The TIME Service</i>	84
7.14	<i>Accurate Times And Network Delays</i>	85
7.15	<i>A UDP Client For The TIME Service</i>	85
7.16	<i>The ECHO Service</i>	87
7.17	<i>A TCP Client For The ECHO Service</i>	88
7.18	<i>A UDP Client For The ECHO Service</i>	89
7.19	<i>Summary</i>	91

Chapter 8 Algorithms And Issues In Server Software Design**95**

- 8.1 *Introduction* 95
- 8.2 *The Conceptual Server Algorithm* 95
- 8.3 *Concurrent Vs. Iterative Servers* 96
- 8.4 *Connection-Oriented Vs. Connectionless Access* 96
- 8.5 *Connection-Oriented Servers* 97
- 8.6 *Connectionless Servers* 97
- 8.7 *Failure, Reliability, And Statelessness* 98
- 8.8 *Optimizing Stateless Servers* 99
- 8.9 *Four Basic Types Of Servers* 101
- 8.10 *Request Processing Time* 102
- 8.11 *Iterative Server Algorithms* 102
- 8.12 *An Iterative, Connection-Oriented Server Algorithm* 103
- 8.13 *Binding To A Well-Known Address Using INADDR_ANY* 103
- 8.14 *Placing The Socket In Passive Mode* 104
- 8.15 *Accepting Connections And Using Them* 104
- 8.16 *An Iterative, Connectionless Server Algorithm* 104
- 8.17 *Forming A Reply Address In A Connectionless Server* 105
- 8.18 *Concurrent Server Algorithms* 106
- 8.19 *Master And Slave Processes* 106
- 8.20 *A Concurrent, Connectionless Server Algorithm* 107
- 8.21 *A Concurrent, Connection-Oriented Server Algorithm* 107
- 8.22 *Using Separate Programs As Slaves* 108
- 8.23 *Apparent Concurrency Using A Single Process* 109
- 8.24 *When To Use Each Server Type* 110
- 8.25 *A Summary of Server Types* 111
- 8.26 *The Important Problem Of Server Deadlock* 112
- 8.27 *Alternative Implementations* 112
- 8.28 *Summary* 113

Chapter 9 Iterative, Connectionless Servers (UDP)**115**

- 9.1 *Introduction* 115
- 9.2 *Creating A Passive Socket* 115
- 9.3 *Process Structure* 118
- 9.4 *An Example TIME Server* 119
- 9.5 *Summary* 121

Chapter 10	Iterative, Connection-Oriented Servers (TCP)	123
10.1	<i>Introduction</i>	123
10.2	<i>Allocating A Passive TCP Socket</i>	123
10.3	<i>A Server For The DAYTIME Service</i>	124
10.4	<i>Process Structure</i>	124
10.5	<i>An Example DAYTIME Server</i>	125
10.6	<i>Closing Connections</i>	128
10.7	<i>Connection Termination And Server Vulnerability</i>	128
10.8	<i>Summary</i>	129
Chapter 11	Concurrent, Connection-Oriented Servers (TCP)	131
11.1	<i>Introduction</i>	131
11.2	<i>Concurrent ECHO</i>	131
11.3	<i>Iterative Vs. Concurrent Implementations</i>	132
11.4	<i>Process Structure</i>	132
11.5	<i>An Example Concurrent ECHO Server</i>	133
11.6	<i>Cleaning Up Errant Processes</i>	137
11.7	<i>Summary</i>	138
Chapter 12	Single-Process, Concurrent Servers (TCP)	139
12.1	<i>Introduction</i>	139
12.2	<i>Data-driven Processing In A Server</i>	139
12.3	<i>Data-Driven Processing With A Single Process</i>	140
12.4	<i>Process Structure Of A Single-Process Server</i>	141
12.5	<i>An Example Single-Process ECHO Server</i>	142
12.6	<i>Summary</i>	144
Chapter 13	Multiprotocol Servers (TCP, UDP)	147
13.1	<i>Introduction</i>	147
13.2	<i>The Motivation For Reducing The Number Of Servers</i>	147
13.3	<i>Multiprotocol Server Design</i>	148
13.4	<i>Process Structure</i>	148
13.5	<i>An Example Multiprotocol DAYTIME Server</i>	149
13.6	<i>The Concept Of Shared Code</i>	153
13.7	<i>Concurrent Multiprotocol Servers</i>	153
13.8	<i>Summary</i>	153

Chapter 14	Multiservice Servers (TCP, UDP)	155
14.1	<i>Introduction</i>	155
14.2	<i>Consolidating Servers</i>	155
14.3	<i>A Connectionless, Multiservice Server Design</i>	156
14.4	<i>A Connection-Oriented, Multiservice Server Design</i>	157
14.5	<i>A Concurrent, Connection-Oriented, Multiservice Server</i>	158
14.6	<i>A Single-Process, Multiservice Server Implementation</i>	158
14.7	<i>Invoking Separate Programs From A Multiservice Server</i>	159
14.8	<i>Multiservice, Multiprotocol Designs</i>	160
14.9	<i>An Example Multiservice Server</i>	161
14.10	<i>The BSD UNIX Super Server, Inetd</i>	168
14.11	<i>Summary</i>	168
Chapter 15	Uniform, Efficient Management Of Server Concurrency	171
15.1	<i>Introduction</i>	171
15.2	<i>Choosing Between An Iterative And A Concurrent Design</i>	171
15.3	<i>Level Of Concurrency</i>	172
15.4	<i>Demand-Driven Concurrency</i>	173
15.5	<i>The Cost Of Concurrency</i>	173
15.6	<i>Overhead And Delay</i>	173
15.7	<i>Small Delays Can Matter</i>	174
15.8	<i>Process Preallocation</i>	175
15.9	<i>Delayed Process Allocation</i>	179
15.10	<i>The Uniform Basis For Both Techniques</i>	180
15.11	<i>Combining Techniques</i>	181
15.12	<i>Summary</i>	181
Chapter 16	Concurrency In Clients	183
16.1	<i>Introduction</i>	183
16.2	<i>The Advantages Of Concurrency</i>	183
16.3	<i>The Motivation For Exercising Control</i>	184
16.4	<i>Concurrent Contact With Multiple Servers</i>	185
16.5	<i>Implementing Concurrent Clients</i>	185
16.6	<i>Single-Process Implementations</i>	187
16.7	<i>An Example Concurrent Client That Uses ECHO</i>	188
16.8	<i>Execution Of The Concurrent Client</i>	192
16.9	<i>Concurrency In The Example Code</i>	193
16.10	<i>Summary</i>	194

Chapter 17 Tunneling At The Transport And Application Levels 195

- 17.1 *Introduction* 195
- 17.2 *Multiprotocol Environments* 195
- 17.3 *Mixing Network Technologies* 197
- 17.4 *Dynamic Circuit Allocation* 198
- 17.5 *Encapsulation And Tunneling* 199
- 17.6 *Tunneling Through An IP Internet* 199
- 17.7 *Application-Level Tunneling Between Clients And Servers* 200
- 17.8 *Summary* 201

Chapter 18 Application Level Gateways 203

- 18.1 *Introduction* 203
- 18.2 *Clients And Servers In Constrained Environments* 203
- 18.3 *Using Application Gateways* 205
- 18.4 *Interoperability Through A Mail Gateway* 206
- 18.5 *Implementation Of A Mail Gateway* 207
- 18.6 *A Comparison Of Application Gateways And Tunneling* 207
- 18.7 *Application Gateways And Limited Functionality Systems* 209
- 18.8 *Application Gateways Used For Security* 210
- 18.9 *Application Gateways And The Extra Hop Problem* 211
- 18.10 *An Example Application Gateway* 213
- 18.11 *Implementation Of An Application Gateway* 214
- 18.12 *Code For The Application Gateway* 216
- 18.13 *An Example Gateway Exchange* 217
- 18.14 *Using Rfcd With UNIX's forward* 218
- 18.15 *Summary* 218

Chapter 19 External Data Representation (XDR) 221

- 19.1 *Introduction* 221
- 19.2 *Representations For Data In Computers* 221
- 19.3 *The N-Squared Conversion Problem* 222
- 19.4 *Network Standard Byte Order* 223
- 19.5 *A De Facto Standard External Data Representation* 224
- 19.6 *XDR Data Types* 225
- 19.7 *Implicit Types* 226
- 19.8 *Software Support For Using XDR* 226
- 19.9 *XDR Library Routines* 226
- 19.10 *Building A Message One Piece At A Time* 226

- 19.11 *Conversion Routines In The XDR Library* 228
- 19.12 *XDR Streams, I/O, and TCP* 230
- 19.13 *Records, Record Boundaries, And Datagram I/O* 231
- 19.14 *Summary* 231

Chapter 20 Remote Procedure Call Concept (RPC)

233

- 20.1 *Introduction* 233
- 20.2 *Remote Procedure Call Model* 233
- 20.3 *Two Paradigms For Building Distributed Programs* 234
- 20.4 *A Conceptual Model For Conventional Procedure Calls* 235
- 20.5 *An Extension Of the Procedural Model* 235
- 20.6 *Execution Of Conventional Procedure Call And Return* 236
- 20.7 *The Procedural Model In Distributed Systems* 237
- 20.8 *Analogy Between Client-Server And RPC* 238
- 20.9 *Distributed Computation As A Program* 239
- 20.10 *Sun Microsystems' Remote Procedure Call Definition* 240
- 20.11 *Remote Programs And Procedures* 240
- 20.12 *Reducing The Number Of Arguments* 241
- 20.13 *Identifying Remote Programs And Procedures* 241
- 20.14 *Accommodating Multiple Versions Of A Remote Program* 242
- 20.15 *Mutual Exclusion For Procedures In A Remote Program* 243
- 20.16 *Communication Semantics* 244
- 20.17 *At Least Once Semantics* 244
- 20.18 *RPC Retransmission* 245
- 20.19 *Mapping A Remote Program To A Protocol Port* 245
- 20.20 *Dynamic Port Mapping* 246
- 20.21 *RPC Port Mapper Algorithm* 247
- 20.22 *Sun RPC Message Format* 249
- 20.23 *Marshaling Arguments For A Remote Procedure* 250
- 20.24 *Authentication* 250
- 20.25 *An Example Of RPC Message Representation* 251
- 20.26 *An Example Of The UNIX Authentication Field* 252
- 20.27 *Summary* 253

Chapter 21 Distributed Program Generation (Rpcgen Concept)

255

- 21.1 *Introduction* 255
- 21.2 *Using Remote Procedure Calls* 256
- 21.3 *Programming Mechanisms To Support RPC* 257
- 21.4 *Dividing A Program Into Local And Remote Procedures* 258
- 21.5 *Adding Code For RPC* 259

- 21.6 *Stub Procedures* 259
- 21.7 *Multiple Remote Procedures And Dispatching* 260
- 21.8 *Name Of The Client-Side Stub Procedure* 261
- 21.9 *Using Rpcgen To Generate Distributed Programs* 262
- 21.10 *Rpcgen Output And Interface Procedures* 262
- 21.11 *Rpcgen Input And Output* 263
- 21.12 *Using Rpcgen To Build A Client And Server* 264
- 21.13 *Summary* 264

Chapter 22 Distributed Program Generation (Rpcgen Example) 267

- 22.1 *Introduction* 267
- 22.2 *An Example To Illustrate Rpcgen* 268
- 22.3 *Dictionary Lookup* 268
- 22.4 *Eight Steps To A Distributed Application* 269
- 22.5 *Step 1: Build A Conventional Application Program* 270
- 22.6 *Step 2: Divide The Program Into Two Parts* 274
- 22.7 *Step 3: Create An Rpcgen Specification* 280
- 22.8 *Step 4: Run Rpcgen* 282
- 22.9 *The .h File Produced By Rpcgen* 282
- 22.10 *The XDR Conversion File Produced By Rpcgen* 283
- 22.11 *The Client Code Produced By Rpcgen* 284
- 22.12 *The Server Code Produced By Rpcgen* 286
- 22.13 *Step 5: Write Stub Interface Procedures* 289
- 22.14 *Step 6: Compile And Link The Client Program* 293
- 22.15 *Step 7: Compile And Link The Server Program* 297
- 22.16 *Step 8: Start The Server And Execute The Client* 299
- 22.17 *Using The UNIX Make Utility* 299
- 22.18 *Summary* 302

Chapter 23 Network File System Concepts (NFS) 305

- 23.1 *Introduction* 305
- 23.2 *Remote File Access Vs. Transfer* 305
- 23.3 *Operations On Remote Files* 306
- 23.4 *File Access Among Heterogeneous Computers* 306
- 23.5 *Stateless Servers* 307
- 23.6 *NFS And UNIX File Semantics* 307
- 23.7 *Review Of The UNIX File System* 307
- 23.8 *Files Under NFS* 321
- 23.9 *NFS File Types* 322
- 23.10 *NFS File Modes* 322

23.11	<i>NFS File Attributes</i>	323
23.12	<i>NFS Client And Server</i>	324
23.13	<i>NFS Client Operation</i>	325
23.14	<i>NFS Client And UNIX</i>	326
23.15	<i>NFS Mounts</i>	327
23.16	<i>File Handle</i>	328
23.17	<i>Handles Replace Path Names</i>	328
23.18	<i>An NFS Client In UNIX</i>	330
23.19	<i>File Positioning With A Stateless Server</i>	330
23.20	<i>Operations On Directories</i>	331
23.21	<i>Reading A Directory Statelessly</i>	331
23.22	<i>Multiple Hierarchies In An NFS Server</i>	332
23.23	<i>The Mount Protocol</i>	332
23.24	<i>Summary</i>	333

Chapter 24 Network File System Protocol (NFS, Mount)

335

24.1	<i>Introduction</i>	335
24.2	<i>Using RPC To Define A Protocol</i>	335
24.3	<i>Defining A Protocol With Data Structures And Procedures</i>	336
24.4	<i>NFS Constant, Type, And Data Declarations</i>	337
24.5	<i>NFS Procedures</i>	340
24.6	<i>Semantics Of NFS Operations</i>	341
24.7	<i>The Mount Protocol</i>	345
24.8	<i>Procedures In The Mount Protocol</i>	347
24.9	<i>Semantics of Mount Operations</i>	347
24.10	<i>NFS And Mount Authentication</i>	348
24.11	<i>Summary</i>	350

Chapter 25 A TELNET Client (Program Structure)

353

25.1	<i>Introduction</i>	353
25.2	<i>Overview</i>	354
25.3	<i>A TELNET Client Algorithm</i>	356
25.4	<i>Terminal I/O In UNIX</i>	357
25.5	<i>Establishing Terminal Modes</i>	361
25.6	<i>Global Variable Used For Stored State</i>	362
25.7	<i>Restoring Terminal Modes Before Exit</i>	363
25.8	<i>Client Suspension And Resumption</i>	364
25.9	<i>Finite State Machine Specification</i>	365
25.10	<i>Embedding Commands In A TELNET Data Stream</i>	366
25.11	<i>Option Negotiation</i>	367

25.12	<i>Request/Offer Symmetry</i>	367
25.13	<i>TELNET Character Definitions</i>	367
25.14	<i>A Finite State Machine For Data From The Server</i>	369
25.15	<i>Transitions Among States</i>	370
25.16	<i>A Finite State Machine Implementation</i>	372
25.17	<i>A Compact FSM Representation</i>	372
25.18	<i>Keeping The Compact Representation At Run-Time</i>	374
25.19	<i>Implementation Of A Compact Representation</i>	374
25.20	<i>Building An FSM Transition Matrix</i>	376
25.21	<i>The Socket Output Finite State Machine</i>	378
25.22	<i>Definitions For The Socket Output FSM</i>	380
25.23	<i>The Option Subnegotiation Finite State Machine</i>	381
25.24	<i>Definitions For The Option Subnegotiation FSM</i>	382
25.25	<i>FSM Initialization</i>	383
25.26	<i>Arguments For The TELNET Client</i>	383
25.27	<i>The Heart Of The TELNET Client</i>	385
25.28	<i>Implementation Of The Main FSM</i>	388
25.29	<i>Summary</i>	390

Chapter 26 A TELNET Client (Implementation Details)

393

26.1	<i>Introduction</i>	393
26.2	<i>The FSM Action Procedures</i>	393
26.3	<i>Recording The Type Of An Option Request</i>	394
26.4	<i>Performing No Operation</i>	395
26.5	<i>Responding To WILL/WONT For The Echo Option</i>	395
26.6	<i>Responding To WILL/WONT For Unsupported Options</i>	397
26.7	<i>Responding To WILL/WONT For The No Go-Ahead Option</i>	397
26.8	<i>Generating DO/DONT For Binary Transmission</i>	399
26.9	<i>Responding To DO/DONT For Unsupported Options</i>	400
26.10	<i>Responding To DO/DONT For The Transmit Binary Option</i>	401
26.11	<i>Responding To DO/DONT For The Terminal Type Option</i>	402
26.12	<i>Option Subnegotiation</i>	404
26.13	<i>Sending Terminal Type Information</i>	405
26.14	<i>Terminating Subnegotiation</i>	406
26.15	<i>Sending A Character To The Server</i>	407
26.16	<i>Displaying Incoming Data On The User's Terminal</i>	408
26.17	<i>Using Curses To Control The User's Terminal</i>	412
26.18	<i>Writing A Block Of Data To The Server</i>	413
26.19	<i>Interacting With The Client Process</i>	414
26.20	<i>Responding To Illegal Commands</i>	415
26.21	<i>Scripting To A File</i>	415
26.22	<i>Implementation Of Scripting</i>	416

26.23	<i>Initialization Of Scripting</i>	416
26.24	<i>Collecting Characters Of The Script File Name</i>	417
26.25	<i>Opening A Script File</i>	418
26.26	<i>Terminating Scripting</i>	420
26.27	<i>Printing Status Information</i>	421
26.28	<i>Summary</i>	422
Chapter 27	Practical Hints And Techniques For UNIX Servers	425
27.1	<i>Introduction</i>	425
27.2	<i>Operating In Background</i>	425
27.3	<i>Programming A Server To Operate In Background</i>	426
27.4	<i>Open Descriptors And Inheritance</i>	427
27.5	<i>Programming A Server To Close Inherited Descriptors</i>	428
27.6	<i>Signals From The Controlling TTY</i>	428
27.7	<i>Programming A Server To Change Its Controlling TTY</i>	428
27.8	<i>Moving To A Safe And Known Directory</i>	429
27.9	<i>Programming A Server To Change Directories</i>	429
27.10	<i>The UNIX Umask</i>	430
27.11	<i>Programming A Server To Set Its Umask</i>	430
27.12	<i>Process Groups</i>	430
27.13	<i>Programming A Server To Set Its Process Group</i>	431
27.14	<i>Descriptors For Standard I/O</i>	431
27.15	<i>Programming A Server To Open Standard Descriptors</i>	431
27.16	<i>Mutual Exclusion For The Server</i>	432
27.17	<i>Programming A Server To Avoid Multiple Copies</i>	432
27.18	<i>Recording A Server's Process ID</i>	433
27.19	<i>Programming A Server To Record Its Process ID</i>	433
27.20	<i>Waiting For A Child Process To Exit</i>	434
27.21	<i>Programming A Server To Wait For Each Child To Exit</i>	434
27.22	<i>Extraneous Signals</i>	434
27.23	<i>Programming A Server To Ignore Extraneous Signals</i>	435
27.24	<i>Using A System Log Facility</i>	435
27.25	<i>Summary</i>	440
Appendix 1	System Calls And Library Routines Used With Sockets	443
Appendix 2	Manipulation Of UNIX File And Socket Descriptors	473

Bibliography

477

Index

485