

TCP/IP Illustrated, Volume 1

The Protocols

W. Richard Stevens



ADDISON-WESLEY

Boston • San Francisco • New York • Toronto • Montreal
London • Munich • Paris • Madrid
Canberra • Sydney • Tokyo • Singapore • Mexico City

Contents

Preface	xv
Chapter 1. Introduction	1
1.1 Introduction	1
1.2 Layering	1
1.3 TCP/IP Layering	6
1.4 Internet Addresses	7
1.5 The Domain Name System	9
1.6 Encapsulation	9
1.7 Demultiplexing	11
1.8 Client-Server Model	12
1.9 Port Numbers	12
1.10 Standardization Process	14
1.11 RFCs	14
1.12 Standard, Simple Services	15
1.13 The Internet	16
1.14 Implementations	16
1.15 Application Programming Interfaces	17
1.16 Test Network	18
1.17 Summary	19

Chapter 2.	Link Layer	21
2.1	Introduction	21
2.2	Ethernet and IEEE 802 Encapsulation	21
2.3	Trailer Encapsulation	23
2.4	SLIP: Serial Line IP	24
2.5	Compressed SLIP	25
2.6	PPP: Point-to-Point Protocol	26
2.7	Loopback Interface	28
2.8	MTU	29
2.9	Path MTU	30
2.10	Serial Line Throughput Calculations	30
2.11	Summary	31
Chapter 3.	IP: Internet Protocol	33
3.1	Introduction	33
3.2	IP Header	34
3.3	IP Routing	37
3.4	Subnet Addressing	42
3.5	Subnet Mask	43
3.6	Special Case IP Addresses	45
3.7	A Subnet Example	46
3.8	ifconfig Command	47
3.9	netstat Command	49
3.10	IP Futures	49
3.11	Summary	50
Chapter 4.	ARP: Address Resolution Protocol	53
4.1	Introduction	53
4.2	An Example	54
4.3	ARP Cache	56
4.4	ARP Packet Format	56
4.5	ARP Examples	57
4.6	Proxy ARP	60
4.7	Gratuitous ARP	62
4.8	arp Command	63
4.9	Summary	63
Chapter 5.	RARP: Reverse Address Resolution Protocol	65
5.1	Introduction	65
5.2	RARP Packet Format	65
5.3	RARP Examples	66
5.4	RARP Server Design	67
5.5	Summary	68

Chapter 6.	ICMP: Internet Control Message Protocol	69
6.1	Introduction	69
6.2	ICMP Message Types	70
6.3	ICMP Address Mask Request and Reply	72
6.4	ICMP Timestamp Request and Reply	74
6.5	ICMP Port Unreachable Error	77
6.6	4.4BSD Processing of ICMP Messages	81
6.7	Summary	83
Chapter 7.	Ping Program	85
7.1	Introduction	85
7.2	Ping Program	85
7.3	IP Record Route Option	91
7.4	IP Timestamp Option	95
7.5	Summary	96
Chapter 8.	Traceroute Program	97
8.1	Introduction	97
8.2	Traceroute Program Operation	97
8.3	LAN Output	99
8.4	WAN Output	102
8.5	IP Source Routing Option	104
8.6	Summary	109
Chapter 9.	IP Routing	111
9.1	Introduction	111
9.2	Routing Principles	112
9.3	ICMP Host and Network Unreachable Errors	117
9.4	To Forward or Not to Forward	119
9.5	ICMP Redirect Errors	119
9.6	ICMP Router Discovery Messages	123
9.7	Summary	125
Chapter 10.	Dynamic Routing Protocols	127
10.1	Introduction	127
10.2	Dynamic Routing	127
10.3	Unix Routing Daemons	128
10.4	RIP: Routing Information Protocol	129
10.5	RIP Version 2	136
10.6	OSPF: Open Shortest Path First	137
10.7	BGP: Border Gateway Protocol	138
10.8	CIDR: Classless Interdomain Routing	140
10.9	Summary	141

Chapter 11.	UDP: User Datagram Protocol	143
11.1	Introduction	143
11.2	UDP Header	144
11.3	UDP Checksum	144
11.4	A Simple Example	147
11.5	IP Fragmentation	148
11.6	ICMP Unreachable Error (Fragmentation Required)	151
11.7	Determining the Path MTU Using Traceroute	153
11.8	Path MTU Discovery with UDP	155
11.9	Interaction Between UDP and ARP	157
11.10	Maximum UDP Datagram Size	159
11.11	ICMP Source Quench Error	160
11.12	UDP Server Design	162
11.13	Summary	167
Chapter 12.	Broadcasting and Multicasting	169
12.1	Introduction	169
12.2	Broadcasting	171
12.3	Broadcasting Examples	172
12.4	Multicasting	175
12.5	Summary	178
Chapter 13.	IGMP: Internet Group Management Protocol	179
13.1	Introduction	179
13.2	IGMP Message	180
13.3	IGMP Protocol	180
13.4	An Example	183
13.5	Summary	186
Chapter 14.	DNS: The Domain Name System	187
14.1	Introduction	187
14.2	DNS Basics	188
14.3	DNS Message Format	191
14.4	A Simple Example	194
14.5	Pointer Queries	198
14.6	Resource Records	201
14.7	Caching	203
14.8	UDP or TCP	206
14.9	Another Example	206
14.10	Summary	208

Chapter 15.	TFTP: Trivial File Transfer Protocol	209
15.1	Introduction	209
15.2	Protocol	209
15.3	An Example	211
15.4	Security	213
15.5	Summary	213
Chapter 16.	BOOTP: Bootstrap Protocol	215
16.1	Introduction	215
16.2	BOOTP Packet Format	215
16.3	An Example	218
16.4	BOOTP Server Design	219
16.5	BOOTP Through a Router	220
16.6	Vendor-Specific Information	221
16.7	Summary	222
Chapter 17.	TCP: Transmission Control Protocol	223
17.1	Introduction	223
17.2	TCP Services	223
17.3	TCP Header	225
17.4	Summary	227
Chapter 18.	TCP Connection Establishment and Termination	229
18.1	Introduction	229
18.2	Connection Establishment and Termination	229
18.3	Timeout of Connection Establishment	235
18.4	Maximum Segment Size	236
18.5	TCP Half-Close	238
18.6	TCP State Transition Diagram	240
18.7	Reset Segments	245
18.8	Simultaneous Open	250
18.9	Simultaneous Close	252
18.10	TCP Options	253
18.11	TCP Server Design	254
18.12	Summary	260
Chapter 19.	TCP Interactive Data Flow	263
19.1	Introduction	263
19.2	Interactive Input	263
19.3	Delayed Acknowledgments	265
19.4	Nagle Algorithm	267
19.5	Window Size Advertisements	274
19.6	Summary	274

Chapter 20.	TCP Bulk Data Flow	275
20.1	Introduction	275
20.2	Normal Data Flow	275
20.3	Sliding Windows	280
20.4	Window Size	282
20.5	PUSH Flag	284
20.6	Slow Start	285
20.7	Bulk Data Throughput	286
20.8	Urgent Mode	292
20.9	Summary	296
Chapter 21.	TCP Timeout and Retransmission	297
21.1	Introduction	297
21.2	Simple Timeout and Retransmission Example	298
21.3	Round-Trip Time Measurement	299
21.4	An RTT Example	301
21.5	Congestion Example	306
21.6	Congestion Avoidance Algorithm	310
21.7	Fast Retransmit and Fast Recovery Algorithms	312
21.8	Congestion Example (Continued)	313
21.9	Per-Route Metrics	316
21.10	ICMP Errors	317
21.11	Repacketization	320
21.12	Summary	321
Chapter 22.	TCP Persist Timer	323
22.1	Introduction	323
22.2	An Example	323
22.3	Silly Window Syndrome	325
22.4	Summary	330
Chapter 23.	TCP Keepalive Timer	331
23.1	Introduction	331
23.2	Description	332
23.3	Keepalive Examples	333
23.4	Summary	337
Chapter 24.	TCP Futures and Performance	339
24.1	Introduction	339
24.2	Path MTU Discovery	340
24.3	Long Fat Pipes	344
24.4	Window Scale Option	347

24.5	Timestamp Option	349
24.6	PAWS: Protection Against Wrapped Sequence Numbers	351
24.7	T/TCP: A TCP Extension for Transactions	351
24.8	TCP Performance	354
24.9	Summary	356
Chapter 25.	SNMP: Simple Network Management Protocol	359
25.1	Introduction	359
25.2	Protocol	360
25.3	Structure of Management Information	363
25.4	Object Identifiers	364
25.5	Introduction to the Management Information Base	365
25.6	Instance Identification	367
25.7	Simple Examples	370
25.8	Management Information Base (Continued)	372
25.9	Additional Examples	382
25.10	Traps	385
25.11	ASN.1 and BER	386
25.12	SNMP Version 2	387
25.13	Summary	388
Chapter 26.	Telnet and Rlogin: Remote Login	389
26.1	Introduction	389
26.2	Rlogin Protocol	391
26.3	Rlogin Examples	396
26.4	Telnet Protocol	401
26.5	Telnet Examples	406
26.6	Summary	417
Chapter 27.	FTP: File Transfer Protocol	419
27.1	Introduction	419
27.2	FTP Protocol	419
27.3	FTP Examples	426
27.4	Summary	439
Chapter 28.	SMTP: Simple Mail Transfer Protocol	441
28.1	Introduction	441
28.2	SMTP Protocol	442
28.3	SMTP Examples	448
28.4	SMTP Futures	452
28.5	Summary	459

Chapter 29.	NFS: Network File System	461
29.1	Introduction	461
29.2	Sun Remote Procedure Call	461
29.3	XDR: External Data Representation	465
29.4	Port Mapper	465
29.5	NFS Protocol	467
29.6	NFS Examples	474
29.7	NFS Version 3	479
29.8	Summary	480
Chapter 30.	Other TCP/IP Applications	481
30.1	Introduction	481
30.2	Finger Protocol	481
30.3	Whois Protocol	483
30.4	Archie, WAIS, Gopher, Veronica, and WWW	484
30.5	X Window System	486
30.6	Summary	490
Appendix A.	The <code>tcpdump</code> Program	491
A.1	BSD Packet Filter	491
A.2	SunOS Network Interface Tap	493
A.3	SVR4 Data Link Provider Interface	494
A.4	<code>tcpdump</code> Output	495
A.5	Security Considerations	496
A.6	Socket Debug Option	496
Appendix B.	Computer Clocks	499
Appendix C.	The <code>sock</code> Program	503
Appendix D.	Solutions to Selected Exercises	507
Appendix E.	Configurable Options	525
E.1	BSD/386 Version 1.0	526
E.2	SunOS 4.1.3	527
E.3	System V Release 4	529
E.4	Solaris 2.2	529
E.5	AIX 3.2.2	536
E.6	4.4BSD	537
Appendix F.	Source Code Availability	539
Bibliography		543
Index		555