

TCP/IP Illustrated, Volume 3

TCP for Transactions, HTTP, NNTP, and the UNIX® Domain Protocols

W. Richard Stevens

Technische Universität Darmstadt FACHBEREICH INFORMATIK	
B I B L I O T H E K	
Inventar-Nr.:	<u>M10-00253</u>
Sachgebiete:	_____
Standort:	<u>C.2/Stev</u>



ADDISON-WESLEY PUBLISHING COMPANY
Reading, Massachusetts Menlo Park, California New York
Don Mills, Ontario Wokingham, England Amsterdam Bonn
Sydney Singapore Tokyo Madrid San Juan
Seoul Milan Mexico City Taipei

Contents

Preface		xv
Part 1. TCP for Transactions		1
<hr/>		
Chapter 1. T/TCP Introduction		3
1.1	Introduction	3
1.2	UDP Client–Server	3
1.3	TCP Client–Server	9
1.4	T/TCP Client–Server	17
1.5	Test Network	20
1.6	Timing Example	21
1.7	Applications	22
1.8	History	24
1.9	Implementations	26
1.10	Summary	28
Chapter 2. T/TCP Protocol		29
2.1	Introduction	29
2.2	New TCP Options for T/TCP	30
2.3	T/TCP Implementation Variables	33
2.4	State Transition Diagram	34
2.5	T/TCP Extended States	36
2.6	Summary	38

Chapter 3.	T/TCP Examples	39
3.1	Introduction	39
3.2	Client Reboot	40
3.3	Normal T/TCP Transaction	42
3.4	Server Receives Old Duplicate SYN	43
3.5	Server Reboot	44
3.6	Request or Reply Exceeds MSS	45
3.7	Backward Compatibility	49
3.8	Summary	51
Chapter 4.	T/TCP Protocol (Continued)	53
4.1	Introduction	53
4.2	Client Port Numbers and TIME_WAIT State	53
4.3	Purpose of the TIME_WAIT State	56
4.4	TIME_WAIT State Truncation	59
4.5	Avoiding the Three-Way Handshake with TAO	62
4.6	Summary	68
Chapter 5.	T/TCP Implementation: Socket Layer	69
5.1	Introduction	69
5.2	Constants	70
5.3	sosend Function	70
5.4	Summary	72
Chapter 6.	T/TCP Implementation: Routing Table	73
6.1	Introduction	73
6.2	Code Introduction	74
6.3	radix_node_head Structure	75
6.4	rtentry Structure	75
6.5	rt_metrics Structure	76
6.6	in_inithead Function	76
6.7	in_addroute Function	77
6.8	in_matroute Function	78
6.9	in_clsroute Function	78
6.10	in_rtqtime Function	79
6.11	in_rtqkill Function	82
6.12	Summary	85
Chapter 7.	T/TCP Implementation: Protocol Control Blocks	87
7.1	Introduction	87
7.2	in_pcbladdr Function	88
7.3	in_pcbconnect Function	89
7.4	Summary	90
Chapter 8.	T/TCP Implementation: TCP Overview	91
8.1	Introduction	91
8.2	Code Introduction	91

8.3	TCP protosw Structure	92	
8.4	TCP Control Block	93	
8.5	tcp_init Function	94	
8.6	tcp_slowtimo Function	94	
8.7	Summary	95	
Chapter 9.	T/TCP Implementation: TCP Output		97
9.1	Introduction	97	
9.2	tcp_output Function	97	
9.3	Summary	104	
Chapter 10.	T/TCP Implementation: TCP Functions		105
10.1	Introduction	105	
10.2	tcp_newtcpcb Function	105	
10.3	tcp_rtlookup Function	106	
10.4	tcp_gettaocache Function	108	
10.5	Retransmission Timeout Calculations	108	
10.6	tcp_close Function	112	
10.7	tcp_msssend Function	113	
10.8	tcp_mssrcvd Function	114	
10.9	tcp_dooptions Function	121	
10.10	tcp_reass Function	122	
10.11	Summary	124	
Chapter 11.	T/TCP Implementation: TCP Input		125
11.1	Introduction	125	
11.2	Preliminary Processing	125	
11.3	Header Prediction	129	
11.4	Initiation of Passive Open	130	
11.5	Initiation of Active Open	134	
11.6	PAWS: Protection Against Wrapped Sequence Numbers	141	
11.7	ACK Processing	142	
11.8	Completion of Passive Opens and Simultaneous Opens	142	
11.9	ACK Processing (Continued)	143	
11.10	FIN Processing	145	
11.11	Summary	147	
Chapter 12.	T/TCP Implementation: TCP User Requests		149
12.1	Introduction	149	
12.2	PRU_CONNECT Request	149	
12.3	tcp_connect Function	150	
12.4	PRU_SEND and PRU_SEND_EOF Requests	154	
12.5	tcp_usrclosed Function	155	
12.6	tcp_sysctl Function	155	
12.7	T/TCP Futures	156	
12.8	Summary	158	

Part 2. Additional TCP Applications **159**

Chapter 13. HTTP: Hypertext Transfer Protocol **161**

- 13.1 Introduction 161
- 13.2 Introduction to HTTP and HTML 162
- 13.3 HTTP Protocol 165
- 13.4 An Example 170
- 13.5 HTTP Statistics 172
- 13.6 Performance Problems 173
- 13.7 Summary 175

Chapter 14. Packets Found on an HTTP Server **177**

- 14.1 Introduction 177
- 14.2 Multiple HTTP Servers 180
- 14.3 Client SYN Interarrival Time 181
- 14.4 RTT Measurements 185
- 14.5 `listen` Backlog Queue 187
- 14.6 Client SYN Options 192
- 14.7 Client SYN Retransmissions 195
- 14.8 Domain Names 196
- 14.9 Timing Out Persist Probes 196
- 14.10 Simulation of T/TCP Routing Table Size 200
- 14.11 Mbuf Interaction 202
- 14.12 TCP PCB Cache and Header Prediction 203
- 14.13 Summary 205

Chapter 15. NNTP: Network News Transfer Protocol **207**

- 15.1 Introduction 207
- 15.2 NNTP Protocol 209
- 15.3 A Simple News Client 212
- 15.4 A More Sophisticated News Client 214
- 15.5 NNTP Statistics 215
- 15.6 Summary 216

Part 3. The Unix Domain Protocols **219**

Chapter 16. Unix Domain Protocols: Introduction **221**

- 16.1 Introduction 221
- 16.2 Usage 222
- 16.3 Performance 223
- 16.4 Coding Examples 224
- 16.5 Summary 225

Chapter 17.	Unix Domain Protocols: Implementation	227
17.1	Introduction	227
17.2	Code Introduction	227
17.3	Unix domain and protocols Structures	228
17.4	Unix Domain Socket Address Structures	230
17.5	Unix Domain Protocol Control Blocks	231
17.6	uipc_usrreq Function	233
17.7	PRU_ATTACH Request and unp_attach Function	233
17.8	PRU_DETACH Request and unp_detach Function	236
17.9	PRU_BIND Request and unp_bind Function	237
17.10	PRU_CONNECT Request and unp_connect Function	240
17.11	PRU_CONNECT2 Request and unp_connect2 Function	245
17.12	socketpair System Call	249
17.13	pipe System Call	253
17.14	PRU_ACCEPT Request	253
17.15	PRU_DISCONNECT Request and unp_disconnect Function	255
17.16	PRU_SHUTDOWN Request and unp_shutdown Function	257
17.17	PRU_ABORT Request and unp_drop Function	258
17.18	Miscellaneous Requests	259
17.19	Summary	261
Chapter 18.	Unix Domain Protocols: I/O and Descriptor Passing	263
18.1	Introduction	263
18.2	PRU_SEND and PRU_RCVD Requests	263
18.3	Descriptor Passing	269
18.4	unp_internalize Function	274
18.5	unp_externalize Function	276
18.6	unp_discard Function	277
18.7	unp_dispose Function	278
18.8	unp_scan Function	278
18.9	unp_gc Function	280
18.10	unp_mark Function	288
18.11	Performance (Revisited)	288
18.12	Summary	289
Appendix A.	Measuring Network Times	291
A.1	RTT Measurements Using Ping	292
A.2	Protocol Stack Measurements	294
A.3	Latency and Bandwidth	300
Appendix B.	Coding Applications for T/TCP	303
Bibliography		309
Index		315