

UNIX Network Programming Volume 1

Second Edition

Networking APIs: Sockets and XTI

by W. Richard Stevens

To join a Prentice Hall PTR Internet mailing list, point to
<http://www.prehall.com/mail-lists/>



Prentice Hall PTR
Upper Saddle River, NJ 07458

ISBN 0 13 460122-X



0134601224



Contents

Preface	xv
Part 1. Introduction and TCP/IP	1
Chapter 1. Introduction	3
1.1 Introduction	3
1.2 A Simple Daytime Client	6
1.3 Protocol Independence	9
1.4 Error Handling: Wrapper Functions	11
1.5 A Simple Daytime Server	13
1.6 Road Map to Client-Server Examples in the Text	16
1.7 OSI Model	18
1.8 BSD Networking History	19
1.9 Test Networks and Hosts	20
1.10 Unix Standards	24
1.11 64-bit Architectures	27
1.12 Summary	28
Chapter 2. The Transport Layer: TCP and UDP	29
2.1 Introduction	29
2.2 The Big Picture	30
2.3 UDP: User Datagram Protocol	32
2.4 TCP: Transmission Control Protocol	32
2.5 TCP Connection Establishment and Termination	34

2.6	TIME_WAIT State	40
2.7	Port Numbers	41
2.8	TCP Port Numbers and Concurrent Servers	44
2.9	Buffer Sizes and Limitations	46
2.10	Standard Internet Services	50
2.11	Protocol Usage by Common Internet Applications	52
2.12	Summary	52

Part 2. Elementary Sockets

55

Chapter 3. Sockets Introduction

57

3.1	Introduction	57
3.2	Socket Address Structures	57
3.3	Value-Result Arguments	63
3.4	Byte Ordering Functions	66
3.5	Byte Manipulation Functions	69
3.6	inet_aton, inet_addr, and inet_ntoa Functions	70
3.7	inet_pton and inet_ntop Functions	72
3.8	sock_ntop and Related Functions	75
3.9	readn, writen, and readline Functions	77
3.10	isfdtype Function	81
3.11	Summary	82

Chapter 4. Elementary TCP Sockets

85

4.1	Introduction	85
4.2	socket Function	85
4.3	connect Function	89
4.4	bind Function	91
4.5	listen Function	93
4.6	accept Function	99
4.7	fork and exec Functions	102
4.8	Concurrent Servers	104
4.9	close Function	107
4.10	getsockname and getpeername Functions	107
4.11	Summary	110

Chapter 5. TCP Client-Server Example

111

5.1	Introduction	111
5.2	TCP Echo Server: main Function	112
5.3	TCP Echo Server: str_echo Function	113
5.4	TCP Echo Client: main Function	113
5.5	TCP Echo Client: str_cli Function	115
5.6	Normal Startup	115
5.7	Normal Termination	117
5.8	Posix Signal Handling	119
5.9	Handling SIGCHLD Signals	122
5.10	wait and waitpid Functions	124

5.11	Connection Abort before accept Returns	129
5.12	Termination of Server Process	130
5.13	SIGPIPE Signal	132
5.14	Crashing of Server Host	133
5.15	Crashing and Rebooting of Server Host	134
5.16	Shutdown of Server Host	135
5.17	Summary of TCP Example	135
5.18	Data Format	137
5.19	Summary	140
Chapter 6.	I/O Multiplexing: The select and poll Functions	143
6.1	Introduction	143
6.2	I/O Models	144
6.3	select Function	150
6.4	str_cli Function (Revisited)	155
6.5	Batch Input	157
6.6	shutdown Function	160
6.7	str_cli Function (Revisited Again)	161
6.8	TCP Echo Server (Revisited)	162
6.9	pselect Function	168
6.10	poll Function	169
6.11	TCP Echo Server (Revisited Again)	172
6.12	Summary	175
Chapter 7.	Socket Options	177
7.1	Introduction	177
7.2	getsockopt and setsockopt Functions	178
7.3	Checking If an Option Is Supported and Obtaining the Default	178
7.4	Socket States	183
7.5	Generic Socket Options	183
7.6	IPv4 Socket Options	197
7.7	ICMPv6 Socket Option	199
7.8	IPv6 Socket Options	199
7.9	TCP Socket Options	201
7.10	fcntl Function	205
7.11	Summary	207
Chapter 8.	Elementary UDP Sockets	211
8.1	Introduction	211
8.2	recvfrom and sendto Functions	212
8.3	UDP Echo Server: main Function	213
8.4	UDP Echo Server: dg_echo Function	214
8.5	UDP Echo Client: main Function	216
8.6	UDP Echo Client: dg_cli Function	217
8.7	Lost Datagrams	217
8.8	Verifying Received Response	218
8.9	Server Not Running	220
8.10	Summary of UDP example	221

16.4	File Operations	427	
16.5	Interface Configuration	428	
16.6	get_ifi_info Function	429	
16.7	Interface Operations	439	
16.8	ARP Cache Operations	440	
16.9	Routing Table Operations	442	
16.10	Summary	443	
Chapter 17.	Routing Sockets		445
17.1	Introduction	445	
17.2	Datalink Socket Address Structure	446	
17.3	Reading and Writing	447	
17.4	sysctl Operations	454	
17.5	get_ifi_info Function	459	
17.6	Interface Name and Index Functions	463	
17.7	Summary	467	
Chapter 18.	Broadcasting		469
18.1	Introduction	469	
18.2	Broadcast Addresses	470	
18.3	Unicast versus Broadcast	472	
18.4	dg_cli Function Using Broadcasting	475	
18.5	Race Conditions	478	
18.6	Summary	486	
Chapter 19.	Multicasting		487
19.1	Introduction	487	
19.2	Multicast Addresses	487	
19.3	Multicasting versus Broadcasting on a LAN	490	
19.4	Multicasting on a WAN	493	
19.5	Multicast Socket Options	495	
19.6	mcast_join and Related Functions	499	
19.7	dg_cli Function Using Multicasting	502	
19.8	Receiving Mbone Session Announcements	504	
19.9	Sending and Receiving	507	
19.10	SNTP: Simple Network Time Protocol	510	
19.11	SNTP (Continued)	515	
19.12	Summary	528	
Chapter 20.	Advanced UDP Sockets		531
20.1	Introduction	531	
20.2	Receiving Flags, Destination IP Address, and Interface Index	532	
20.3	Datagram Truncation	539	
20.4	When to Use UDP Instead Of TCP	539	
20.5	Adding Reliability to a UDP Application	542	
20.6	Binding Interface Addresses	553	
20.7	Concurrent UDP Servers	557	
20.8	IPv6 Packet Information	560	
20.9	Summary	562	

Chapter 21.	Out-of-Band Data	565
21.1	Introduction	565
21.2	TCP Out-of-Band Data	565
21.3	socketpair Function	572
21.4	TCP Out-of-Band Data Summary	580
21.5	Client-Server Heartbeat Functions	581
21.6	Summary	586
Chapter 22.	Signal-Driven I/O	589
22.1	Introduction	589
22.2	Signal-Driven I/O for Sockets	590
22.3	UDP Echo Server Using SIGIO	592
22.4	Summary	598
Chapter 23.	Threads	601
23.1	Introduction	601
23.2	Basic Thread Functions: Creation and Termination	602
23.3	str_cil Function Using Threads	605
23.4	TCP Echo Server Using Threads	607
23.5	Thread-Specific Data	611
23.6	Web Client and Simultaneous Connections (Continued)	620
23.7	Mutexes: Mutual Exclusion	622
23.8	Condition Variables	627
23.9	Web Client and Simultaneous Connections (Continued)	631
23.10	Summary	633
Chapter 24.	IP Options	635
24.1	Introduction	635
24.2	IPv4 Options	635
24.3	IPv4 Source Route Options	637
24.4	IPv6 Extension Headers	645
24.5	IPv6 Hop-by-Hop Options and Destination Options	645
24.6	IPv6 Routing Header	649
24.7	IPv6 Sticky Options	653
24.8	Summary	654
Chapter 25.	Raw Sockets	655
25.1	Introduction	655
25.2	Raw Socket Creation	656
25.3	Raw Socket Output	657
25.4	Raw Socket Input	659
25.5	Ping Program	661
25.6	Traceroute Program	672
25.7	An ICMP Message Daemon	685
25.8	Summary	702

Chapter 26.	Datalink Access	703
26.1	Introduction	703
26.2	BPF: BSD Packet Filter	704
26.3	DLPI: Data Link Provider Interface	706
26.4	Linux: SOCK_PACKET	707
26.5	libpcap: Packet Capture Library	707
26.6	Examining the UDP Checksum Field	708
26.7	Summary	725
Chapter 27.	Client-Server Design Alternatives	727
27.1	Introduction	727
27.2	TCP Client Alternatives	730
27.3	TCP Test Client	730
27.4	TCP Iterative Server	732
27.5	TCP Concurrent Server, One Child per Client	732
27.6	TCP Preforked Server, No Locking around accept	736
27.7	TCP Preforked Server, File Locking around accept	742
27.8	TCP Preforked Server, Thread Locking around accept	745
27.9	TCP Preforked Server, Descriptor Passing	746
27.10	TCP Concurrent Server, One Thread per Client	752
27.11	TCP Prethreaded Server, per-Thread accept	754
27.12	TCP Prethreaded Server, Main Thread accept	756
27.13	Summary	759
Part 4.	XTI: X/Open Transport Interface	761
Chapter 28.	XTI: TCP Clients	763
28.1	Introduction	763
28.2	t_open Function	764
28.3	t_error and t_strerror Functions	767
28.4	netbuf Structures and XTI Structures	769
28.5	t_bind Function	770
28.6	t_connect Function	772
28.7	t_rcv and t_snd Functions	773
28.8	t_look Function	774
28.9	t_andrel and t_rcvrel Functions	775
28.10	t_anddis and t_rcvdis Functions	777
28.11	XTI TCP Daytime Client	778
28.12	xtl_rdrw Function	781
28.13	Summary	782
Chapter 29.	XTI: Name and Address Functions	783
29.1	Introduction	783
29.2	/etc/netconfig File and netconfig Functions	784
29.3	NETPATH Variable and netpath Functions	785
29.4	netdir Functions	786

29.5	<code>t_malloc</code> and <code>t_free</code> Functions	788
29.6	<code>t_getprotaddr</code> Functions	790
29.7	<code>xti_ntop</code> Function	791
29.8	<code>tcp_connect</code> Function	792
29.9	Summary	796
Chapter 30.	XTI: TCP Servers	797
30.1	Introduction	797
30.2	<code>t_listen</code> Function	799
30.3	<code>tcp_listen</code> Function	800
30.4	<code>t_accept</code> Function	802
30.5	<code>xti_accept</code> Function	803
30.6	Simple Daytime Server	804
30.7	Multiple Pending Connections	806
30.8	<code>xti_accept</code> Function (Revisited)	808
30.9	Summary	816
Chapter 31.	XTI: UDP Clients and Servers	819
31.1	Introduction	819
31.2	<code>t_rcvudata</code> and <code>t_sndudata</code> Functions	819
31.3	<code>udp_client</code> Function	820
31.4	<code>t_rcvuderr</code> Function: Asynchronous Errors	824
31.5	<code>udp_server</code> Function	826
31.6	Reading a Datagram in Pieces	829
31.7	Summary	831
Chapter 32.	XTI Options	833
32.1	Introduction	833
32.2	<code>t_optbdr</code> Structure	835
32.3	XTI Options	837
32.4	<code>t_optmgmt</code> Function	840
32.5	Checking if an Option Is Supported and Obtaining the Default	841
32.6	Getting and Setting XTI Options	844
32.7	Summary	848
Chapter 33.	Streams	849
33.1	Introduction	849
33.2	Overview	850
33.3	<code>getmsg</code> and <code>putmsg</code> Functions	854
33.4	<code>getpsg</code> and <code>putpsg</code> Functions	855
33.5	<code>ioctl</code> Function	855
33.6	TPI: Transport Provider Interface	856
33.7	Summary	856
Chapter 34.	XTI: Additional Functions	867
34.1	Introduction	867
34.2	Nonblocking I/O	867
34.3	<code>t_reconnect</code> Function	868

34.4	<code>t_getinfo</code> Function	869
34.5	<code>t_getstate</code> Function	869
34.6	<code>t_sync</code> Function	870
34.7	<code>t_unbind</code> Function	872
34.8	<code>t_rcv</code> and <code>t_rcvdata</code> Functions	872
34.9	<code>t_snd</code> and <code>t_sndvdata</code> Functions	873
34.10	<code>t_rcvreldata</code> and <code>t_sndreldata</code> Functions	874
34.11	Signal-Driven I/O	874
34.12	Out-of-Band Data	875
34.13	Loopback Transport Providers	880
34.14	Summary	881
Appendix A.	IPv4, IPv6, ICMPv4, and ICMPv6	883
A.1	Introduction	883
A.2	IPv4 Header	883
A.3	IPv6 Header	885
A.4	IPv4 Addresses	887
A.5	IPv6 Addresses	892
A.6	ICMPv4 and ICMPv6: Internet Control Message Protocol	896
Appendix B.	Virtual Networks	899
B.1	Introduction	899
B.2	The Mbone	899
B.3	The Gbone	901
Appendix C.	Debugging Techniques	903
C.1	System Call Tracing	903
C.2	Standard Internet Services	906
C.3	<code>sock</code> Program	908
C.4	Small Test Programs	911
C.5	<code>tcpdump</code> Program	913
C.6	<code>netstat</code> Program	914
C.7	<code>lsof</code> Program	914
Appendix D.	Miscellaneous Source Code	915
D.1	<code>unp.h</code> Header	915
D.2	<code>config.h</code> Header	919
D.3	<code>unpxti.h</code> Header	920
D.4	Standard Error Functions	922
Appendix E.	Solutions to Selected Exercises	925
Bibliography		963
Index		971