

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.prenhall.com/mail_lists/

ISBN 0-13-490012-X



Prentice Hall PTR
Upper Saddle River, NJ 07458

Contents

Preface	xv
Part 1. Introduction and TCP/IP	1
<hr/>	
Chapter 1. Introduction	
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 <code>accept</code> Returns	129
5.12	Termination of Server Process	130
5.13	<code>SIGPIPE</code> 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 <code>select</code> and <code>poll</code> Functions	143
6.1	Introduction	143
6.2	I/O Models	144
6.3	<code>select</code> Function	150
6.4	<code>str_cli</code> Function (Revisited)	155
6.5	Batch Input	157
6.6	<code>shutdown</code> Function	160
6.7	<code>str_cli</code> Function (Revisited Again)	161
6.8	TCP Echo Server (Revisited)	162
6.9	<code>pselect</code> Function	168
6.10	<code>poll</code> 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	<code>getsockopt</code> and <code>setsockopt</code> 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	<code>fcntl</code> Function	205
7.11	Summary	207
Chapter 8.	Elementary UDP Sockets	211
8.1	Introduction	211
8.2	<code>recvfrom</code> and <code>sendto</code> Functions	212
8.3	UDP Echo Server: <code>main</code> Function	213
8.4	UDP Echo Server: <code>dg_echo</code> Function	214
8.5	UDP Echo Client: <code>main</code> Function	216
8.6	UDP Echo Client: <code>dg_cli</code> 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

8.11	connect Function with UDP	224	
8.12	dg_cli Function (Revisited)	227	
8.13	Lack of Flow Control with UDP	228	
8.14	Determining Outgoing Interface with UDP	231	
8.15	TCP and UDP Echo Server Using <code>select</code>	233	
8.16	Summary	235	
Chapter 9.	Elementary Name and Address Conversions		237
9.1	Introduction	237	
9.2	Domain Name System	237	
9.3	gethostbyname Function	240	
9.4	RES_USE_INET6 Resolver Option	245	
9.5	gethostbyname2 Function and IPv6 Support	246	
9.6	gethostbyaddr Function	248	
9.7	uname Function	249	
9.8	gethostname Function	250	
9.9	getservbyname and getservbyport Functions	251	
9.10	Other Networking Information	255	
9.11	Summary	256	
Part 3.	Advanced Sockets		259
Chapter 10.	IPv4 and IPv6 Interoperability		261
10.1	Introduction	261	
10.2	IPv4 Client, IPv6 Server	262	
10.3	IPv6 Client, IPv4 Server	265	
10.4	IPv6 Address Testing Macros	267	
10.5	IPV6_ADDRFORM Socket Option	268	
10.6	Source Code Portability	270	
10.7	Summary	271	
Chapter 11.	Advanced Name and Address Conversions		273
11.1	Introduction	273	
11.2	getaddrinfo Function	273	
11.3	gai_strerror Function	278	
11.4	freeaddrinfo Function	279	
11.5	getaddrinfo Function: IPv6 and Unix Domain	279	
11.6	getaddrinfo Function: Examples	282	
11.7	host_serv Function	284	
11.8	tcp_connect Function	285	
11.9	tcp_listen Function	288	
11.10	udp_client Function	293	
11.11	udp_connect Function	295	
11.12	udp_server Function	296	
11.13	getnameinfo Function	298	
11.14	Reentrant Functions	300	
11.15	gethostbyname-r and gethostbyaddr-r Functions	303	

11.16	Implementation of <code>getaddrinfo</code> and <code>getnameinfo</code> Functions	305
11.17	Summary	328
Chapter 12.	Daemon Processes and <code>inetd</code> Superserver	331
12.1	Introduction	331
12.2	<code>syslogd</code> Daemon	332
12.3	<code>syslog</code> Function	333
12.4	<code>daemon-init</code> Function	335
12.5	<code>inetd</code> Daemon	339
12.6	<code>daemon-inetd</code> Function	344
12.7	Summary	346
Chapter 13.	Advanced I/O Functions	349
13.1	Introduction	349
13.2	Socket Timeouts	349
13.3	<code>recv</code> and <code>send</code> Functions	354
13.4	<code>readv</code> and <code>writev</code> Functions	357
13.5	<code>recvmsg</code> and <code>sendmsg</code> Functions	358
13.6	Ancillary Data	362
13.7	How Much Data Is Queued?	365
13.8	Sockets and Standard I/O	366
13.9	T/TCP: TCP for Transactions	369
13.10	Summary	371
Chapter 14.	Unix Domain Protocols	373
14.1	Introduction	373
14.2	Unix Domain Socket Address Structure	374
14.3	<code>socketpair</code> Function	376
14.4	Socket Functions	377
14.5	Unix Domain Stream Client-Server	378
14.6	Unix Domain Datagram Client-Server	379
14.7	Passing Descriptors	381
14.8	Receiving Sender Credentials	390
14.9	Summary	394
Chapter 15.	Nonblocking I/O	397
15.1	Introduction	397
15.2	Nonblocking Reads and Writes: <code>str_cli</code> Function (Revisited)	399
15.3	Nonblocking <code>connect</code>	409
15.4	Nonblocking <code>connect</code> : Daytime Client	410
15.5	Nonblocking <code>connect</code> : Web Client	413
15.6	Nonblocking <code>accept</code>	422
15.7	Summary	424
Chapter 16.	<code>ioctl</code> Operations	425
16.1	Introduction	425
16.2	<code>ioctl</code> Function	426
16.3	Socket Operations	426

16.4	File Operations	427	
16.5	Interface Configuration	428	
16.6	<code>get—ifi_info</code> 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	<code>sysctl</code> Operations	454	
17.5	<code>get—ifi_info</code> 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	<code>dg_cli</code> 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	<code>mcast_join</code> and Related Functions	499	
19.7	<code>dg_cli</code> 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	<code>socketmark</code> 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 <code>SIGIO</code>	592
22.4	Summary	598
Chapter 23.	Threads	601
23.1	Introduction	601
23.2	Basic Thread Functions: Creation and Termination	602
23.3	<code>str_cli</code> 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_sndrel and t_rcvrel Functions	775
28.10	t_snddis and t_rcvdis Functions	777
28.11	XTI TCP Daytime Client	778
28.12	xti_rdwrr 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	t_alloc and t_free Functions	788	
29.6	t_getprotaddr Functions	790	
29.7	xti_ntop Function	791	
29.8	tcp_connect Function	792	
29.9	Summary	796	
Chapter 30.	XTI: TCP Servers		797
30.1	Introduction	797	
30.2	t_listen Function	799	
30.3	tcp_listen Function	800	
30.4	t_accept Function	802	
30.5	xti_accept Function	803	
30.6	Simple Daytime Server	804	
30.7	Multiple Pending Connections	806	
30.8	xti_accept Function (Revisited)	808	
30.9	Summary	816	
Chapter 31.	XTI: UDP Clients and Servers		819
31.1	Introduction	819	
31.2	t_rcvudata and t_sndudata Functions	819	
31.3	udp_client Function	820	
31.4	t_rcvuderr Function: Asynchronous Errors	824	
31.5	udp_server 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	t_opthdr Structure	835	
32.3	XTI Options	837	
32.4	t_optmgmt 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	getmsg and putmsg Functions	854	
33.4	getpmsg and putpmsg Functions	855	
33.5	ioctl1 Function	855	
33.6	TPI: Transport Provider Interface	856	
33.7	Summary	866	
Chapter 34.	XTI: Additional Functions		867
34.1	Introduction	867	
34.2	Nonblocking I/O	867	
34.3	t_rcvconnect Function	868	

34.4	t_getinfo Function	869	
34.5	t_getstate Function	869	
34.6	t_sync Function	870	
34.7	t_unbind Function	872	
34.8	t_rcvv and t_rcvvudata Functions	872	
34.9	t_sndv and t_sndvudata Functions	873	
34.10	t_rcvreldata and t_sndreldata 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 6bone	901	
Appendix C. Debugging Techniques			903
C.1	System Call Tracing	903	
C.2	Standard Internet Services	908	
C.3	sock Program	908	
C.4	Small Test Programs	911	
C.5	tcpdump Program	913	
C.6	netstat Program	914	
C.7	lsolf Program	914	
Appendix D. Miscellaneous Source Code			915
D.1	unp.h Header	915	
D.2	config.h Header	919	
D.3	unpxti.h Header	920	
D.4	Standard Error Functions	922	
Appendix E. Solutions to Selected Exercises			925
Bibliography			963
Index			971