

Advanced Programming in the UNIX[®] Environment

Third Edition

W. Richard Stevens
Stephen A. Rago

◆ Addison-Wesley

Upper Saddle River, NJ • Boston • Indianapolis • San Francisco
New York • Toronto • Montreal • London • Munich • Paris • Madrid
Capetown • Sydney • Tokyo • Singapore • Mexico City

Contents

Foreword to the Second Edition	xix
Preface	xxi
Preface to the Second Edition	xxv
Preface to the First Edition	xxix
Chapter 1. UNIX System Overview	1
1.1 Introduction	1
1.2 UNIX Architecture	1
1.3 Logging In	2
1.4 Files and Directories	4
1.5 Input and Output	8
1.6 Programs and Processes	10
1.7 Error Handling	14
1.8 User Identification	16
1.9 Signals	18
1.10 Time Values	20
1.11 System Calls and Library Functions	21
1.12 Summary	23
Chapter 2. UNIX Standardization and Implementations	25
2.1 introduction	25

2.2	UNIX Standardization	25	
2.2.1	ISO C	25	
2.2.2	IEEE POSIX	26	
2.2.3	The Single UNIX Specification	30	
2.2.4	FIPS	32	
2.3	UNIX System Implementations	33	
2.3.1	UNIX System V Release 4	33	
2.3.2	4.4BSD	34	
2.3.3	FreeBSD	34	
2.3.4	Linux	35	
2.3.5	Mac OS X	35	
2.3.6	Solaris	35	
2.3.7	Other UNIX Systems	35	
2.4	Relationship of Standards and Implementations	36	
2.5	Limits	36	
2.5.1	ISO C Limits	37	
2.5.2	POSIX Limits	38	
2.5.3	XSI Limits	41	
2.5.4	sysconf, pathconf, and fpathconf Functions	42	
2.5.5	Indeterminate Runtime Limits	49	
2.6	Options	53	
2.7	Feature Test Macros	57	
2.8	Primitive System Data Types	58	
2.9	Differences Between Standards	58	
2.10	Summary	60	
Chapter 3.	File I/O		61
3.1	Introduction	61	
3.2	File Descriptors	61	
3.3	open and openat Functions	62	
3.4	creat Function	66	
3.5	close Function	66	
3.6	lseek Function	66	
3.7	read Function	71	
3.8	write Function	72	
3.9	I/O Efficiency	72	
3.10	File Sharing	74	
3.11	Atomic Operations	77	
3.12	dup and dup2 Functions	79	
3.13	sync, fsync, and fdatasync Functions	81	
3.14	fcntl Function	82	

3.15	ioctl Function	87	
3.16	/dev/fd	88	
3.17	Summary	90	
Chapter 4.	Files and Directories		93
4.1	Introduction	93	
4.2	stat, fstat, fstatat, and lstat Functions		93
4.3	File Types	95	
4.4	Set-User-ID and Set-Group-ID	98	
4.5	File Access Permissions	99	
4.6	Ownership of New Files and Directories		101
4.7	access and faccessat Functions		102
4.8	umask Function	104	
4.9	chmod, fchmod, and fchmodat Functions		106
4.10	Sticky Bit	108	
4.11	chown, fchown, fchownat, and lchown Functions		109
4.12	File Size	111	
4.13	File Truncation	112	
4.14	File Systems	113	
4.15	link, linkat, unlink, unlinkat, and remove Functions		116
4.16	rename and renameat Functions		119
4.17	Symbolic Links	120	
4.18	Creating and Reading Symbolic Links		123
4.19	File Times	124	
4.20	futimens, utimensat, and utimes Functions		126
4.21	mkdir, mkdirat, and rmdir Functions		129
4.22	Reading Directories	130	
4.23	chdir, fchdir, and getcwd Functions		135
4.24	Device Special Files	137	
4.25	Summary of File Access Permission Bits		140
4.26	Summary	140	
Chapter 5.	Standard I/O Library		143
5.1	Introduction	143	
5.2	Streams and FILE Objects	143	
5.3	Standard Input, Standard Output, and Standard Error		145
5.4	Buffering	145	
5.5	Opening a Stream	148	

5.6	Reading and Writing a Stream	150
5.7	Line-at-a-Time I/O	152
5.8	Standard I/O Efficiency	153
5.9	Binary I/O	156
5.10	Positioning a Stream	157
5.11	Formatted I/O	159
5.12	Implementation Details	164
5.13	Temporary Files	167
5.14	Memory Streams	171
5.15	Alternatives to Standard I/O	174
5.16	Summary	175
Chapter 6.	System Data Files and Information	177
6.1	Introduction	177
6.2	Password File	177
6.3	Shadow Passwords	181
6.4	Group File	182
6.5	Supplementary Group IDs	183
6.6	Implementation Differences	184
6.7	Other Data Files	185
6.8	Login Accounting	186
6.9	System Identification	187
6.10	Time and Date Routines	189
6.11	Summary	196
Chapter 7.	Process Environment	197
7.1	Introduction	197
7.2	main Function	197
7.3	Process Termination	198
7.4	Command-Line Arguments	203
7.5	Environment List	203
7.6	Memory Layout of a C Program	204
7.7	Shared Libraries	206
7.8	Memory Allocation	207
7.9	Environment Variables	210
7.10	setjmp and longjmp Functions	213
7.11	getrlimit and setrlimit Functions	220
7.12	Summary	225
Chapter 8.	Process Control	227
8.1	Introduction	227

8.2	Process Identifiers	227	
8.3	fork Function	229	
8.4	vfork Function	234	
8.5	exit Functions	236	
8.6	wait and waitpid Functions	238	
8.7	waitid Function	244	
8.8	wait3 and wait4 Functions	245	
8.9	Race Conditions	245	
8.10	exec Functions	249	
8.11	Changing User IDs and Group IDs	255	
8.12	Interpreter Files	260	
8.13	system Function	264	
8.14	Process Accounting	269	
8.15	User Identification	275	
8.16	Process Scheduling	276	
8.17	Process Times	280	
8.18	Summary	282	
Chapter 9.	Process Relationships		285
9.1	Introduction	285	
9.2	Terminal Logins	285	
9.3	Network Logins	290	
9.4	Process Groups	293	
9.5	Sessions	295	
9.6	Controlling Terminal	296	
9.7	tcgetpgrp, tcsetpgrp, and tcgetsid Functions	298	
9.8	Job Control	299	
9.9	Shell Execution of Programs	303	
9.10	Orphaned Process Groups	307	
9.11	FreeBSD Implementation	310	
9.12	Summary	312	
Chapter 10.	Signals		313
10.1	Introduction	313	
10.2	Signal Concepts	313	
10.3	signal Function	323	
10.4	Unreliable Signals	326	
10.5	Interrupted System Calls	327	
10.6	Reentrant Functions	330	
10.7	SIGCLD Semantics	332	

10.8	Reliable-Signal Terminology and Semantics	335
10.9	<i>kill</i> and <i>raise</i> Functions	336
10.10	<i>alarm</i> and <i>pause</i> Functions	338
10.11	Signal Sets	344
10.12	<i>sigprocmask</i> Function	346
10.13	<i>sigpending</i> Function	347
10.14	<i>sigaction</i> Function	349
10.15	<i>sigsetjmp</i> and <i>siglongjmp</i> Functions	355
10.16	<i>sigsuspend</i> Function	359
10.17	<i>abort</i> Function	365
10.18	<i>system</i> Function	367
10.19	<i>sleep</i> , <i>nanosleep</i> , and <i>clock_nanosleep</i> Functions	373
10.20	<i>sigqueue</i> Function	376
10.21	Job-Control Signals	377
10.22	Signal Names and Numbers	379
10.23	Summary	381
Chapter 11.	Threads	383
11.1	Introduction	383
11.2	Thread Concepts	383
11.3	Thread Identification	384
11.4	Thread Creation	385
11.5	Thread Termination	388
11.6	Thread Synchronization	397
11.6.1	Mutexes	399
11.6.2	Deadlock Avoidance	402
11.6.3	<i>pthread_mutex_timedlock</i> Function	407
11.6.4	Reader–Writer Locks	409
11.6.5	Reader–Writer Locking with Timeouts	413
11.6.6	Condition Variables	413
11.6.7	Spin Locks	417
11.6.8	Barriers	418
11.7	Summary	422
Chapter 12.	Thread Control	425
12.1	Introduction	425
12.2	Thread Limits	425
12.3	Thread Attributes	426
12.4	Synchronization Attributes	430
12.4.1	Mutex Attributes	430

12.4.2	Reader–Writer Lock Attributes	439
12.4.3	Condition Variable Attributes	440
12.4.4	Barrier Attributes	441
12.5	Reentrancy	442
12.6	Thread-Specific Data	446
12.7	Cancel Options	451
12.8	Threads and Signals	453
12.9	Threads and fork	457
12.10	Threads and I/O	461
12.11	Summary	462
Chapter 13.	Daemon Processes	463
13.1	Introduction	463
13.2	Daemon Characteristics	463
13.3	Coding Rules	466
13.4	Error Logging	469
13.5	Single-Instance Daemons	473
13.6	Daemon Conventions	474
13.7	Client–Server Model	479
13.8	Summary	480
Chapter 14.	Advanced I/O	481
14.1	Introduction	481
14.2	Nonblocking I/O	481
14.3	Record Locking	485
14.4	I/O Multiplexing	500
14.4.1	select and pselect Functions	502
14.4.2	poll Function	506
14.5	Asynchronous I/O	509
14.5.1	System V Asynchronous I/O	510
14.5.2	BSD Asynchronous I/O	510
14.5.3	POSIX Asynchronous I/O	511
14.6	readv and writev Functions	521
14.7	readn and writen Functions	523
14.8	Memory-Mapped I/O	525
14.9	Summary	531
Chapter 15.	Interprocess Communication	533
15.1	Introduction	533
15.2	Pipes	534
15.3	popen and pclose Functions	541

15.4	Coprocesses	548	
15.5	FIFOs	552	
15.6	XSI IPC	556	
15.6.1	Identifiers and Keys	556	
15.6.2	Permission Structure	558	
15.6.3	Configuration Limits	559	
15.6.4	Advantages and Disadvantages	559	
15.7	Message Queues	561	
15.8	Semaphores	565	
15.9	Shared Memory	571	
15.10	POSIX Semaphores	579	
15.11	Client–Server Properties	585	
15.12	Summary	587	
Chapter 16.	Network IPC: Sockets		589
16.1	Introduction	589	
16.2	Socket Descriptors	590	
16.3	Addressing	593	
16.3.1	Byte Ordering	593	
16.3.2	Address Formats	595	
16.3.3	Address Lookup	597	
16.3.4	Associating Addresses with Sockets	604	
16.4	Connection Establishment	605	
16.5	Data Transfer	610	
16.6	Socket Options	623	
16.7	Out-of-Band Data	626	
16.8	Nonblocking and Asynchronous I/O	627	
16.9	Summary	628	
Chapter 17.	Advanced IPC		629
17.1	Introduction	629	
17.2	UNIX Domain Sockets	629	
17.2.1	Naming UNIX Domain Sockets	634	
17.3	Unique Connections	635	
17.4	Passing File Descriptors	642	
17.5	An Open Server, Version 1	653	
17.6	An Open Server, Version 2	659	
17.7	Summary	669	
Chapter 18.	Terminal I/O		671
18.1	Introduction	671	

18.2	Overview	671	
18.3	Special Input Characters	678	
18.4	Getting and Setting Terminal Attributes	683	
18.5	Terminal Option Flags	683	
18.6	<code>stty</code> Command	691	
18.7	Baud Rate Functions	692	
18.8	Line Control Functions	693	
18.9	Terminal Identification	694	
18.10	Canonical Mode	700	
18.11	Noncanonical Mode	703	
18.12	Terminal Window Size	710	
18.13	<code>termcap</code> , <code>terminfo</code> , and <code>curses</code>	712	
18.14	Summary	713	
Chapter 19.	Pseudo Terminals		715
19.1	Introduction	715	
19.2	Overview	715	
19.3	Opening Pseudo-Terminal Devices	722	
19.4	<code>pty_fork</code> Function	726	
19.5	<code>pty</code> Program	729	
19.6	Using the <code>pty</code> Program	733	
19.7	Advanced Features	740	
19.8	Summary	741	
Chapter 20.	A Database Library		743
20.1	Introduction	743	
20.2	History	743	
20.3	The Library	744	
20.4	Implementation Overview	746	
20.5	Centralized or Decentralized?	750	
20.6	Concurrency	752	
20.7	Building the Library	753	
20.8	Source Code	753	
20.9	Performance	781	
20.10	Summary	786	
Chapter 21.	Communicating with a Network Printer		789
21.1	Introduction	789	
21.2	The Internet Printing Protocol	789	
21.3	The Hypertext Transfer Protocol	792	
21.4	Printer Spooling	793	

21.5	Source Code	795	
21.6	Summary	843	
Appendix A.	Function Prototypes		845
Appendix B.	Miscellaneous Source Code		895
B.1	Our Header File	895	
B.2	Standard Error Routines	898	
Appendix C.	Solutions to Selected Exercises		905
	Bibliography		947
	Index		955