

Contents

Preface.....	xv
---------------------	-----------

Notational Conventions	xix
Constants	xix
Word Size	xx
Byte Order	xx
Bit Order	xx
Bit Ranges	xxi
Memory Size Units	xxi

Introduction.....	xxiii
--------------------------	--------------

Chapter 1: Review of UNIX Kernel Internals	1
1.1 Introduction.....	2
1.2 Processes, Programs, and Threads.....	3
1.3 The Process Address Space	5
1.3.1 Address Space Mapping	7
1.4 Context Switch	8
1.5 Memory and Process Management System Calls	9
1.5.1 The Fork System Call.....	9
1.5.2 The Exec System Call	12
1.5.3 The Exit System Call	13
1.5.4 The Sbrk and Brk System Calls.....	13
1.5.5 Shared Memory	14
1.5.6 I/O Operations	14
1.5.7 Mapped Files	15
1.6 Summary	15
1.7 Exercises	15
1.8 Further Reading	17

Part I: Cache Memory Systems.....21
Chapter 2: Introduction to Cache Memory Systems.....23

2.1	Memory Hierarchies.....	24
2.2	Cache Fundamentals	26
2.2.1	How a Cache Is Accessed.....	26
2.2.2	Virtual or Physical Addresses?	29
2.2.3	Searching the Cache	29
2.2.4	Replacement Policies	30
2.2.5	Write Policies	30
2.3	Direct Mapped Caches	33
2.3.1	Hashing Algorithms for Direct Mapped Caches.....	34
2.3.2	Direct Mapped Cache Example	37
2.3.3	Miss Processing and Replacement Policy for Direct Mapped Caches.....	40
2.3.4	Summary of Direct Mapped Caching	42
2.4	Two-Way Set Associative Caches.....	42
2.4.1	Summary of Two-Way Set Associative Caches	44
2.5	<i>n</i> -Way Set Associative Caches.....	44
2.6	Fully Associative Caches.....	45
2.7	Summary of <i>n</i> -Way Set Associative Caches	45
2.8	Cache Flushing.....	46
2.9	Uncached Operation.....	47
2.10	Separate Instruction and Data Caches.....	48
2.11	Cache Performance.....	49
2.12	How Cache Architectures Differ.....	51
2.13	Exercises	52
2.14	Further Reading.....	55

Chapter 3: Virtual Caches59

3.1	Virtual Cache Operation	60
3.2	Problems with Virtual Caches.....	62
3.2.1	Ambiguities	62
3.2.2	Aliases	64
3.3	Managing a Virtual Cache	67
3.3.1	Context Switch.....	67
3.3.2	Fork.....	68
3.3.3	Exec.....	71
3.3.4	Exit.....	71
3.3.5	Brk and Sbrk.....	72

3.3.6	Shared Memory and Mapped Files	72
3.3.7	I/O	73
3.3.8	User–Kernel Data Ambiguities	77
3.4	Summary	78
3.5	Exercises	78
3.6	Further Reading	81

Chapter 4: Virtual Caches with Keys 83

4.1	The Operation of a Virtual Cache with Keys	84
4.2	Managing a Virtual Cache with Keys.....	85
4.2.1	Context Switch.....	85
4.2.2	Fork	87
4.2.3	Exec	89
4.2.4	Exit	90
4.2.5	Brk and Sbrk.....	90
4.2.6	Shared Memory and Mapped Files	90
4.2.7	I/O.....	94
4.2.8	User–Kernel Data Ambiguities	94
4.3	Virtual Cache Usage in MMUs	94
4.4	Summary	95
4.5	Exercises	96
4.6	Further Reading	98

Chapter 5: Virtual Caches with Physical Address Tags 99

5.1	The Organization of a Virtual Cache with Physical Tags	100
5.2	Managing a Virtual Cache with Physical Tags	103
5.2.1	Context Switch.....	104
5.2.2	Fork	104
5.2.3	Exec	105
5.2.4	Exit	105
5.2.5	Brk and Sbrk.....	106
5.2.6	Shared Memory and Mapped Files	106
5.2.7	I/O.....	107
5.2.8	User–Kernel Data Ambiguities	107
5.3	Summary	107
5.4	Exercises	108
5.5	Further Reading	109

Chapter 6: Physical Caches 111

6.1	The Organization of a Physical Cache	112
6.2	Managing a Physical Cache	114
6.2.1	Context Switch.....	114
6.2.2	Fork.....	114
6.2.3	Exec, Exit, Brk, and Sbrk	114
6.2.4	Shared Memory and Mapped Files.....	115
6.2.5	User–Kernel Data Ambiguities.....	115
6.2.6	I/O and Bus Watching.....	115
6.3	Multilevel Caches.....	121
6.3.1	Primary Virtual Cache with Secondary Physical Cache	122
6.3.2	Primary Virtual Cache with Physical Tags and a Secondary Physical Cache	124
6.4	Summary	126
6.5	Exercises	127
6.6	Further Reading.....	128

Chapter 7: Efficient Cache Management Techniques 131

7.1	Introduction	132
7.2	Address Space Layout	132
7.2.1	Virtually Indexed Caches	132
7.2.2	Dynamic Address Binding.....	136
7.2.3	Physically Indexed Caches.....	137
7.3	Cache Size Bounded Flushing	139
7.4	Delayed Cache Invalidations	139
7.4.1	Virtual Caches with Keys	141
7.4.2	Physically Tagged Caches without Bus Watching.....	141
7.5	Cache-Aligning Data Structures	142
7.6	Summary	144
7.7	Exercises	145
7.8	Further Reading.....	146

Part II: Multiprocessor Systems 147
Chapter 8: Introduction to Multiprocessor Systems 149

8.1	Introduction	150
8.1.1	MP Operating Systems.....	151

8.2	The Tightly Coupled, Shared Memory, Symmetric Multiprocessor	152
8.3	The MP Memory Model	154
8.3.1	The Sequential Memory Model.....	155
8.3.2	Atomic Reads and Writes	155
8.3.3	Atomic Read-Modify-Write Operations	158
8.4	Mutual Exclusion	160
8.5	Review of Mutual Exclusion on Uniprocessor	
	UNIX Systems	162
8.5.1	Short-Term Mutual Exclusion	163
8.5.2	Mutual Exclusion with Interrupt Handlers	163
8.5.3	Long-Term Mutual Exclusion.....	164
8.6	Problems Using UP Mutual Exclusion Policies on MPs	167
8.7	Summary	168
8.8	Exercises	169
8.9	Further Reading	172

Chapter 9: Master–Slave Kernels..... 175

9.1	Introduction.....	176
9.2	Spin Locks	177
9.3	Deadlocks.....	179
9.4	Master–Slave Kernel Implementation	181
9.4.1	Run Queue Implementation.....	181
9.4.2	Process Selection for Slaves.....	184
9.4.3	Process Selection for the Master.....	186
9.4.4	Clock Interrupt Handling	186
9.5	Performance Considerations.....	187
9.5.1	Master–Slave Improvements.....	188
9.6	Summary	189
9.7	Exercises	190
9.8	Further Reading	192

Chapter 10: Spin-Locked Kernels 195

10.1	Introduction.....	196
10.2	Giant Locking	196
10.3	Multithreading Cases Requiring No Locks	199
10.4	Coarse-Grained Locking	200
10.5	Fine-Grained Locking	203
10.5.1	Short-Term Mutual Exclusion	203
10.5.2	Long-Term Mutual Exclusion.....	204
10.5.3	Mutual Exclusion with Interrupt Handlers	206

10.5.4	Lock Granularity.....	207
10.5.5	Performance	208
10.5.6	Kernel Preemption	209
10.6	Effects of Sleep and Wakeup on Multiprocessors	209
10.7	Summary	211
10.8	Exercises	212
10.9	Further Reading.....	215

Chapter 11: Semaphored Kernels217

11.1	Introduction	218
11.1.1	Mutual Exclusion with Semaphores.....	219
11.1.2	Synchronization with Semaphores	220
11.1.3	Resource Allocation with Semaphores	220
11.2	Deadlocks	221
11.3	Implementing Semaphores.....	223
11.4	Coarse-Grained Semaphore Implementations	226
11.5	Multithreading with Semaphores	227
11.5.1	Long-Term Mutual Exclusion	227
11.5.2	Short-Term Mutual Exclusion	228
11.5.3	Synchronization	228
11.6	Performance Considerations	230
11.6.1	Measuring Lock Contention	230
11.6.2	Convoys.....	231
11.6.3	Multireader Locks	234
11.7	Summary	237
11.8	Exercises	239
11.9	Further Reading.....	240

Chapter 12: Other MP Primitives245

12.1	Introduction	246
12.2	Monitors.....	246
12.3	Eventcounts and Sequencers.....	248
12.4	The MP Primitives of SVR4.2 MP.....	251
12.4.1	Spin Locks	251
12.4.2	Sleep Locks	253
12.4.3	Synchronization Variables	255
12.4.4	Multireader Locks	258
12.5	Comparison of MP Synchronization Primitives.....	258
12.6	Summary	261
12.7	Exercises	262

12.8	Further Reading	263
------	-----------------------	-----

Chapter 13: Other Memory Models 267

13.1	Introduction.....	268
13.2	Dekker's Algorithm	268
13.3	Other Memory Models	271
13.4	Total Store Ordering.....	273
13.5	Partial Store Ordering.....	278
13.6	The Store Buffer as Part of the Memory Hierarchy	280
13.7	Summary	281
13.8	Exercises.....	281
13.9	Further Reading	282

Part III: Multiprocessor Systems with Caches..... 285

Chapter 14: Introduction to MP Cache Consistency..... 287

14.1	Introduction.....	288
14.2	The Cache Consistency Problem.....	290
14.3	Software Cache Consistency	293
	14.3.1 Uncached Shared Data	294
	14.3.2 Selective Cache Flushing.....	296
	14.3.3 Handling Other Memory Models.....	300
14.4	Summary	301
14.5	Exercises.....	302
14.6	Further Reading	303

Chapter 15: Hardware Cache Consistency 307

15.1	Introduction.....	308
15.2	Write-Invalidate Protocols	310
	15.2.1 The Write-Through Invalidate Protocol.....	310
	15.2.2 The Write-Once Protocol.....	311
	15.2.3 MESI Protocols.....	314
15.3	Write-Update Protocols	315
	15.3.1 The Firefly Protocol	315
	15.3.2 The MIPS R4000 Update Protocol	316
15.4	Consistency of Read-Modify-Write Operations	317
15.5	Hardware Consistency for Multilevel Caches.....	318

15.6	Other Main Memory Architectures	319
15.6.1	The Cross-bar Interconnect	320
15.6.2	Directory-based Hardware Cache Consistency	322
15.7	Effects on the Software	324
15.8	Hardware Consistency for Nonsequential Memory Models	327
15.9	Performance Considerations for Software	327
15.9.1	Cache-Aligning Data Structures	327
15.9.2	Reducing Cache Line Contention When Acquiring a Spin Lock	329
15.9.3	Matching Consistency Protocols to Data Usage	330
15.10	Summary	332
15.11	Exercises	333
15.12	Further Reading	335
 Appendix A: Architecture Summary		341
 Appendix B: Answers to Selected Exercises		349
 Index		387