

MAURICE J. BACH





## THE DESIGN OF THE UNIX® OPERATING SYSTEM

Maurice J. Bach

SEMINAR LIBRARY

Bepartment of Computer Science

UNIVERSITY OF KARAGHI

1-12-97

## CONTENTS

PREFACE	•		•	٠	X
CHAPTER 1 GENERAL OVERVIEW OF THE SYSTEM		-			1
1.1 HISTORY					1
1.2 SYSTEM STRUCTURE					4
1.3 USER PERSPECTIVE					
1.4 OPERATING SYSTEM SERVICES					
1.5 ASSUMPTIONS ABOUT HARDWARE					
16 SUMMARY					18

CHAPTER 2 INTRODUCTION TO THE KERNEL		
THE CITE OF THE CITE OF ENGLISH		
SYSTEM · · · ·		10
2.2 INTRODUCTION TO SYSTEM CONCEPTS		19
2.3 KERNEL DATA STRUCTURES		
2.4 SYSTEM ADMINISTRATION		
2.4 SYSTEM ADMINISTRATION		34
	•	36
2.6 EXERCISES		37
CHAPTER 3 THE BUFFER CACHE	٠.	38
3.1 BUFFER HEADERS		39
3.2 STRUCTURE OF THE BUFFER POOL		40
3.3 SCENARIOS FOR RETRIEVAL OF A BUFFER		42
3.4 READING AND WRITING DISK BLOCKS		53
3.5 ADVANTAGES AND DISADVANTAGES OF THE BUFFE	R	
CACHE		
3.6 SUMMARY		
3.7 EXERCISES	٠	58
CHAPTER 4 INTERNAL REPRESENTATION OF FILES		
4.1 INODES		
4.2 STRUCTURE OF A REGULAR FILE		
4.3 DIRECTORIES		
4.4 CONVERSION OF A PATH NAME TO AN INODE		74
4.5 SUPER BLOCK		
4.6 INODE ASSIGNMENT TO A NEW FILE		77
4.6 INODE ASSIGNMENT TO A NEW FILE		77 84
4.6 INODE ASSIGNMENT TO A NEW FILE		77 84

CHAPTER 5 SYSTEM CALLS FOR THE FILE SYSTEM					91
5.1 OPEN					
5.2 READ	,				96
5.3 WRITE					101
5.4 FILE AND RECORD LOCKING					103
5.5 ADJUSTING THE POSITION OF FILE 1/O -	LS	EE	K		103
5.6 CLOSE					103
5.7 FILE CREATION					
5.8 CREATION OF SPECIAL FILES					107
5.9 CHANGE DIRECTORY AND CHANGE ROOT					109
5.10 CHANGE OWNER AND CHANGE MODE .					110
5.11 STAT AND FSTAT					110
5.12 PIPES					111
5.13 DUP					117
5.14 MOUNTING AND UNMOUNTING FILE SYST	TE	MS			119
5.15 LINK	4				128
5.16 UNLINK	*				132
5.17 FILE SYSTEM ABSTRACTIONS					138
5.18 FILE SYSTEM MAINTENANCE					139
5.19 SUMMARY					140
5.20 EXERCISES					
CHAPTER 6 THE STRUCTURE OF PROCESSES					146
6.1 PROCESS STATES AND TRANSITIONS					147
6.2 LAYOUT OF SYSTEM MEMORY					
6.3 THE CONTEXT OF A PROCESS					
6.4 SAVING THE CONTEXT OF A PROCESS .					
6.5 MANIPULATION OF THE PROCESS ADDRES					
SPACE					171
AA CI DED					182

HE SE

6.7 SUMMARY 6.8 EXERCISES  CHAPTER 7 PROCESS CONTROL 7.1 PROCESS CREATION 7.2 SIGNALS 7.3 PROCESS TERMINATION 7.4 AWAITING PROCESS TERMINATION 7.5 INVOKING OTHER PROGRAMS 7.6 THE USER ID OF A PROCESS 7.7 CHANGING THE SIZE OF A PROCESS 7.8 THE SHELL 7.9 SYSTEM BOOT AND THE INIT PROCESS 7.10 SUMMARY 7.11 EXERCISES  CHAPTER 8 PROCESS SCHEDULING AND TIME 8.1 PROCESS SCHEDULING 8.2 SYSTEM CALLS FOR TIME 8.3 CLOCK 8.4 SUMMARY 8.5 EXERCISES  CHAPTER 9 MEMORY MANAGEMENT POLICIES 9.1 SWAPPING 9.2 DEMAND PAGING 9.3 A HYBRID SYSTEM WITH SWAPPING AND DEMAND 9.4 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.4 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.7 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.7 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.7 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.7 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES 9.7 OF THE SWAPPING AND DEMAND 9.8 SUMMARY 9.5 EXERCISES														18
CHAPTER 7 PROCESS CONTROL	6.7	SUMMARY				*								18
CHAPTER 7 PROCESS CONTROL 19 7.1 PROCESS CREATION 20 7.2 SIGNALS 21 7.3 PROCESS TERMINATION 21 7.4 AWAITING PROCESS TERMINATION 21 7.5 INVOKING OTHER PROGRAMS 21 7.6 THE USER ID OF A PROCESS 22 7.7 CHANGING THE SIZE OF A PROCESS 22 7.8 THE SHELL 23 7.9 SYSTEM BOOT AND THE INIT PROCESS 23 7.10 SUMMARY 23 7.11 EXERCISES 239  CHAPTER 8 PROCESS SCHEDULING AND TIME 247 8.1 PROCESS SCHEDULING 248 8.2 SYSTEM CALLS FOR TIME 258 8.3 CLOCK 260 8.4 SUMMARY 268 8.5 EXERCISES 268  CHAPTER 9 MEMORY MANAGEMENT POLICIES 271 9.1 SWAPPING 272 9.1 SWAPPING 272 9.2 DEMAND PAGING 285 PAGING 307	6.8	EXERCISES .												
7.1       PROCESS CREAT       20         7.2       SIGNALS       21         7.3       PROCESS TERMINATION       21         7.4       AWAITING PROCESS TERMINATION       21         7.5       INVOKING OTHER PROGRAMS       21         7.5       INVOKING OTHER PROCESS       225         7.6       THE USER ID OF A PROCESS       229         7.7       CHANGING THE SIZE OF A PROCESS       235         7.8       THE SHELL       232         7.9       SYSTEM BOOT AND THE INIT PROCESS       235         7.10       SUMMARY       238         7.11       EXERCISES       239         CHAPTER 8 PROCESS SCHEDULING AND TIME       247         8.1       PROCESS SCHEDULING       248         8.2       SYSTEM CALLS FOR TIME       258         8.3       CLOCK       260         8.4       SUMMARY       268         8.5       EXERCISES       268         CHAPTER 9       MEMORY MANAGEMENT POLICIES       271         9.1       SWAPPING       272         9.2       DEMAND PAGING       285         9.4       SUMMARY       307		of the last of the	ONTROL		40									19
7.1       PROCESS CREAT       20         7.2       SIGNALS       21         7.3       PROCESS TERMINATION       21         7.4       AWAITING PROCESS TERMINATION       21         7.5       INVOKING OTHER PROGRAMS       21         7.5       INVOKING OTHER PROCESS       225         7.6       THE USER ID OF A PROCESS       229         7.7       CHANGING THE SIZE OF A PROCESS       235         7.8       THE SHELL       232         7.9       SYSTEM BOOT AND THE INIT PROCESS       235         7.10       SUMMARY       238         7.11       EXERCISES       239         CHAPTER 8 PROCESS SCHEDULING AND TIME       247         8.1       PROCESS SCHEDULING       248         8.2       SYSTEM CALLS FOR TIME       258         8.3       CLOCK       260         8.4       SUMMARY       268         8.5       EXERCISES       268         CHAPTER 9       MEMORY MANAGEMENT POLICIES       271         9.1       SWAPPING       272         9.2       DEMAND PAGING       285         9.4       SUMMARY       307	CHAPTER	7 PROCESS C	ONIKOL											10
7.2 SIGNALS 7.3 PROCESS TERMINATION 217 7.4 AWAITING PROCESS TERMINATION 217 7.5 INVOKING OTHER PROGRAMS 217 7.5 INVOKING OTHER PROCESS 227 7.6 THE USER ID OF A PROCESS 227 7.7 CHANGING THE SIZE OF A PROCESS 229 7.8 THE SHELL 232 7.9 SYSTEM BOOT AND THE INIT PROCESS 235 7.10 SUMMARY 238 7.11 EXERCISES 239  CHAPTER 8 PROCESS SCHEDULING AND TIME 247 8.1 PROCESS SCHEDULING 248 8.2 SYSTEM CALLS FOR TIME 258 8.3 CLOCK 260 8.4 SUMMARY 268 8.5 EXERCISES 268  CHAPTER 9 MEMORY MANAGEMENT POLICIES 271 9.1 SWAPPING 272 9.1 SWAPPING 272 9.2 DEMAND PAGING 285 PAGING 307	7.1	PROCESS CR	EATION	-										201
7.3 PROCESS TERMINATION  7.4 AWAITING PROCESS TERMINATION  7.5 INVOKING OTHER PROGRAMS  7.6 THE USER ID OF A PROCESS  7.7 CHANGING THE SIZE OF A PROCESS  7.8 THE SHELL  7.9 SYSTEM BOOT AND THE INIT PROCESS  7.10 SUMMARY  7.11 EXERCISES  239  CHAPTER 8 PROCESS SCHEDULING AND TIME  8.1 PROCESS SCHEDULING  8.2 SYSTEM CALLS FOR TIME  248  8.3 CLOCK  8.4 SUMMARY  8.5 EXERCISES  268  CHAPTER 9 MEMORY MANAGEMENT POLICIES  9.1 SWAPPING  9.2 DEMAND PAGING  9.3 A HYBRID SYSTEM WITH SWAPPING AND DEMAND  9.4 SUMMARY  307														401
7.5 INVOKING OTHER PROGRAMS 217 7.6 THE USER ID OF A PROCESS 227 7.7 CHANGING THE SIZE OF A PROCESS 229 7.8 THE SHELL 232 7.9 SYSTEM BOOT AND THE INIT PROCESS 235 7.10 SUMMARY 238 7.11 EXERCISES 239  CHAPTER 8 PROCESS SCHEDULING AND TIME 247 8.1 PROCESS SCHEDULING 248 8.2 SYSTEM CALLS FOR TIME 258 8.3 CLOCK 260 8.4 SUMMARY 268 8.5 EXERCISES 268  CHAPTER 9 MEMORY MANAGEMENT POLICIES 271 9.1 SWAPPING 272 9.2 DEMAND PAGING 285 PAGING 285 PAGING 307		- arec TE	PMINAT	ION	100						-			417
7.5 INVOKING OTHER PROGRAMS 217 7.6 THE USER ID OF A PROCESS 227 7.7 CHANGING THE SIZE OF A PROCESS 229 7.8 THE SHELL 232 7.9 SYSTEM BOOT AND THE INIT PROCESS 235 7.10 SUMMARY 238 7.11 EXERCISES 239  CHAPTER 8 PROCESS SCHEDULING AND TIME 247 8.1 PROCESS SCHEDULING 248 8.2 SYSTEM CALLS FOR TIME 258 8.3 CLOCK 260 8.4 SUMMARY 268 8.5 EXERCISES 268  CHAPTER 9 MEMORY MANAGEMENT POLICIES 271 9.1 SWAPPING 272 9.2 DEMAND PAGING 285 PAGING 285 PAGING 307	7.4	AWAITING P	ROCESS	TERM	INA	1014	*		•					213
7.6 THE USER ID OF A PROCESS	7.5	INVOKING O	THER PR	ROGRA	MS			*					*	217
7.8 THE SHELL	7.6	THE USER ID	OF A PE	ROCES	S.									227
7.9 SYSTEM BOOT AND THE INIT PROCESS	7.7	CHANGING 1	THE SIZE	OF A	PRO	CESS								229
7.10 SUMMARY														
7.11 EXERCISES	7.9	SYSTEM BOO	T AND T	THE IN	IT PF	ROCE	SS							235
CHAPTER 8 PROCESS SCHEDULING AND TIME	7.10	SUMMARY .	0.10		Sec.		4.		*			٠		238
8.1 PROCESS SCHEDULING	7.11	EXERCISES					1.0							239
8.1 PROCESS SCHEDULING														
8.2 SYSTEM CALLS FOR TIME	CHAPTER	8 PROCESS SO	CHEDUL	ING A	ND T	IME								247
8.3 CLOCK	8.1	PROCESS SCH	EDULING	G.			13							248
8.3 CLOCK	8.2 5	SYSTEM CALL	S FOR T	IME	YNA	MAN	13		2					258
8.4 SUMMARY	8.3	CLOCK			150	190		The same						260
CHAPTER 9 MEMORY MANAGEMENT POLICIES														
CHAPTER 9 MEMORY MANAGEMENT POLICIES	8.5 1	EXERCISES .	Ma line				1							268
9.2 DEMAND PAGING  9.3 A HYBRID SYSTEM WITH SWAPPING AND DEMAND PAGING  9.4 SUMMARY  9.5 272  272  285														
9.2 DEMAND PAGING  9.3 A HYBRID SYSTEM WITH SWAPPING AND DEMAND PAGING  9.4 SUMMARY  9.5 272  272  285	CHAPTER	MEMORY M	IANAGE	MENT	POLI	CIES								271
9.3 A HYBRID SYSTEM WITH SWAPPING AND DEMAND PAGING 9.4 SUMMARY 9.5 285		MAPPING .												
PAGING		PAG PAG	ING .	Man a										
9.4 SUMMARY	9.3	HYBRID SYS	TEM WI	THE		No.				•			•	263
	94 9	AGING			APPI	NG A	ANI	DI	DEN	MA	NE	)		307
		The second secon	2		3 10	419	THE							307

CHARLES BY BOTH STORY OF STREET

CHAPTER 10 THE I/O SUBSYSTEM	,	6				312
10.1 DRIVER INTERFACES			9.	,		313
10.2 DISK DRIVERS	,		9	,		325
10.3 TERMINAL DRIVERS			,			329
10.4 STREAMS		1				344
10.5 SUMMARY	,					351
10.6 EXERCISES						
CHAPTER II INTERPROCESS COMMUNICATION						355
II.I PROCESS TRACING						
11.2 SYSTEM V IPC						
11.3 NETWORK COMMUNICATIONS						
11.4 SOCKETS						
II.5 SUMMARY						
11.6 EXERCISES						
CHAPTER 12 MULTIPROCESSOR SYSTEMS	1					301
12.1 PROBLEM OF MULTIPROCESSOR SYSTEM	MS		Ì			201
12.2 SOLUTION WITH MASTER AND SLAVE						
PROCESSORS				35		393
12.3 SOLUTION WITH SEMAPHORES						205
12.4 THE TUNIS SYSTEM						410
12.5 PERFORMANCE LIMITATIONS						410
12.6 EXERCISES						410
CHAPTER 13 DISTRIBUTED UNIX SYSTEMS						412
13.1 SATELLITE PROCESSORS						414
13.2 THE NEWCASTLE CONNECTION						422
13.3 TRANSPARENT DISTRIBUTED FILE SYST	E	MS			*	426
13.4 A TRANSPARENT DISTRIBUTED MODEL						

Multi-threading

ONIK

ONIK

-- Single thread

process

The UNIX system was first described in a 1974 paper in the Communications of the ACM [Thompson 74] by Ken Thompson and Dennis Ritchie. Since that time, it has become increasingly widespread and popular throughout the computer industry where more and more vendors are offering support for it on their machines. It is especially popular in universities where it is frequently used for operating systems research and case studies.

Many books and papers have described parts of the system, among them, two special issues of the Bell System Technical Journal in 1978 [BSTJ 78] and 1984 [BLTJ 84]. Many books describe the user level interface, particularly how to use electronic mail, how to prepare documents, or how to use the command interpreter called the shell; some books such as The UNIX Programming Environment [Kernighan 84] and Advanced UNIX Programming [Rochkind 85] describe the programming interface. This book describes the internal algorithms and structures that form the basis of the operating system (called the kernel) and their relationship to the programmer interface. It is thus applicable to several environments. First, it can be used as a textbook for an operating systems course at either the advanced undergraduate or first-year graduate level. It is most beneficial to reference the system source code when using the book, but the book can be read independently, too. Second, system programmers can use the book as a reference to gain better understanding of how the kernel works and to compare algorithms used in the UNIX system to algorithms used in other operating systems.