

CONTENTS

PREFACE

xv

1 INTRODUCTION

1

- 1.1 WHAT IS AN OPERATING SYSTEM? 3
 - 1.1.1 The Operating System as an Extended Machine 3
 - 1.1.2 The Operating System as a Resource Manager 4
- 1.2 HISTORY OF OPERATING SYSTEMS 5
 - 1.2.1 The First Generation (1945–55) Vacuum Tubes and Plugboards 6
 - 1.2.2 The Second Generation (1955–65) Transistors and Batch Systems 6
 - 1.2.3 The Third Generation (1965–1980): ICs and Multiprogramming 8
 - 1.2.4 The Fourth Generation (1980–Present): Personal Computers 12
 - 1.2.5 History of MINIX 13
- 1.3 OPERATING SYSTEM CONCEPTS 15
 - 1.3.1 Processes 15
 - 1.3.2 Files 17
 - 1.3.3 The Shell 20
- 1.4 SYSTEM CALLS 21
 - 1.4.1 System Calls for Process Management 22
 - 1.4.2 System Calls for Signaling 26
 - 1.4.3 System Calls for File Management 28
 - 1.4.4 System Calls for Directory Management 33
 - 1.4.5 System Calls for Protection 35
 - 1.4.6 System Calls for Time Management 36

- 1.5 OPERATING SYSTEM STRUCTURE 37
 - 1.5.1 Monolithic Systems 37
 - 1.5.2 Layered Systems 39
 - 1.5.3 Virtual Machines 40
 - 1.5.4 Client-Server Model 42
- 1.6 OUTLINE OF THE REST OF THIS BOOK 44
- 1.7 SUMMARY 44

2 PROCESSES

47

- 2.1 INTRODUCTION TO PROCESSES 47
 - 2.1.1 The Process Model 48
 - 2.1.2 Implementation of Processes 52
 - 2.1.3 Threads 53
- 2.2 INTERPROCESS COMMUNICATION 57
 - 2.2.1 Race Conditions 57
 - 2.2.2 Critical Sections 58
 - 2.2.3 Mutual Exclusion with Busy Waiting 59
 - 2.2.4 Sleep and Wakeup 63
 - 2.2.5 Semaphores 66
 - 2.2.6 Monitors 68
 - 2.2.7 Message Passing 72
- 2.3 CLASSICAL IPC PROBLEMS 75
 - 2.3.1 The Dining Philosophers Problem 75
 - 2.3.2 The Readers and Writers Problem 77
 - 2.3.3 The Sleeping Barber Problem 80
- 2.4 PROCESS SCHEDULING 82
 - 2.4.1 Round Robin Scheduling 84
 - 2.4.2 Priority Scheduling 85
 - 2.4.3 Multiple Queues 86
 - 2.4.4 Shortest Job First 87
 - 2.4.5 Guaranteed Scheduling 89
 - 2.4.6 Lottery Scheduling 89
 - 2.4.7 Real-Time Scheduling 90
 - 2.4.8 Two-level Scheduling 92
 - 2.4.9 Policy versus Mechanism 93

- 2.5 OVERVIEW OF PROCESSES IN MINIX 93
 - 2.5.1 The Internal Structure of MINIX 93
 - 2.5.2 Process Management in MINIX 95
 - 2.5.3 Interprocess Communication in MINIX 97
 - 2.5.4 Process Scheduling in MINIX 98
- 2.6 IMPLEMENTATION OF PROCESSES IN MINIX 98
 - 2.6.1 Organization of the MINIX Source Code 99
 - 2.6.2 The Common Header Files 102
 - 2.6.3 The MINIX Header Files 107
 - 2.6.4 Process Data Structures and Header Files 112
 - 2.6.5 Bootstrapping MINIX 120
 - 2.6.6 System Initialization 122
 - 2.6.7 Interrupt Handling in MINIX 128
 - 2.6.8 Interprocess Communication in MINIX 137
 - 2.6.9 Scheduling in MINIX 140
 - 2.6.10 Hardware-Dependent Kernel Support 142
 - 2.6.11 Utilities and the Kernel Library 145
- 2.7 SUMMARY 147

3 INPUT/OUTPUT

153

- 3.1 PRINCIPLES OF I/O HARDWARE 154
 - 3.1.1 I/O Devices 154
 - 3.1.2 Device Controllers 155
 - 3.1.3 Direct Memory Access (DMA) 157
- 3.2 PRINCIPLES OF I/O SOFTWARE 159
 - 3.2.1 Goals of the I/O Software 159
 - 3.2.2 Interrupt Handlers 161
 - 3.2.3 Device Drivers 161
 - 3.2.4 Device-Independent I/O Software 162
 - 3.2.5 User-Space I/O Software 164
- 3.3 DEADLOCKS 166
 - 3.3.1 Resources 167
 - 3.3.2 Principles of Deadlock 168
 - 3.3.3 The Ostrich Algorithm 170
 - 3.3.4 Detection and Recovery 172
 - 3.3.5 Deadlock Prevention 173
 - 3.3.6 Deadlock Avoidance 175

- 3.4 OVERVIEW OF I/O IN MINIX 179
 - 3.4.1 Interrupt Handlers in MINIX 180
 - 3.4.2 Device Drivers in MINIX 181
 - 3.4.3 Device-Independent I/O Software in MINIX 185
 - 3.4.4 User-level I/O Software in MINIX 185
 - 3.4.5 Deadlock Handling in MINIX 186

- 3.5 BLOCK DEVICES IN MINIX 187
 - 3.5.1 Overview of Block Device Drivers in MINIX 187
 - 3.5.2 Common Block Device Driver Software 190
 - 3.5.3 The Driver Library 193

- 3.6 RAM DISKS 195
 - 3.6.1 RAM Disk Hardware and Software 196
 - 3.6.2 Overview of the RAM Disk Driver in MINIX 197
 - 3.6.3 Implementation of the RAM Disk Driver in MINIX 198

- 3.7 DISKS 200
 - 3.7.1 Disk Hardware 200
 - 3.7.2 Disk Software 202
 - 3.7.3 Overview of the Hard Disk Driver in MINIX 208
 - 3.7.4 Implementation of the Hard Disk Driver in MINIX 211
 - 3.7.5 Floppy Disk Handling 220

- 3.8 CLOCKS 222
 - 3.8.1 Clock Hardware 223
 - 3.8.2 Clock Software 224
 - 3.8.3 Overview of the Clock Driver in MINIX 227
 - 3.8.4 Implementation of the Clock Driver in MINIX 230

- 3.9 TERMINALS 235
 - 3.9.1 Terminal Hardware 235
 - 3.9.2 Terminal Software 240
 - 3.9.3 Overview of the Terminal Driver in MINIX 249
 - 3.9.4 Implementation of the Device-Independent Terminal Driver 264
 - 3.9.5 Implementation of the Keyboard Driver 282
 - 3.9.6 Implementation of the Display Driver 288

- 3.10 THE SYSTEM TASK IN MINIX 296

- 3.11 SUMMARY 304

- 4.1 BASIC MEMORY MANAGEMENT 310
 - 4.1.1 Monoprogramming without Swapping or Paging 310
 - 4.1.2 Multiprogramming with Fixed Partitions 311
- 4.2 SWAPPING 313
 - 4.2.1 Memory Management with Bit Maps 316
 - 4.2.2 Memory Management with Linked Lists 317
- 4.3 VIRTUAL MEMORY 319
 - 4.3.1 Paging 319
 - 4.3.2 Page Tables 322
 - 4.3.3 TLBs—Translation Lookaside Buffers 327
 - 4.3.4 Inverted Page Tables 330
- 4.4 PAGE REPLACEMENT ALGORITHMS 331
 - 4.4.1 The Optimal Page Replacement Algorithm 331
 - 4.4.2 The Not-Recently-Used Page Replacement Algorithm 332
 - 4.4.3 The First-In, First-Out (FIFO) Page Replacement Algorithm 333
 - 4.4.4 The Second Chance Page Replacement Algorithm 333
 - 4.4.5 The Clock Page Replacement Algorithm 334
 - 4.4.6 The Least Recently Used (LRU) Page Replacement Algorithm 334
 - 4.4.7 Simulating LRU in Software 336
- 4.5 DESIGN ISSUES FOR PAGING SYSTEMS 338
 - 4.5.1 The Working Set Model 338
 - 4.5.2 Local versus Global Allocation Policies 339
 - 4.5.3 Page Size 341
 - 4.5.4 Virtual Memory Interface 343
- 4.6 SEGMENTATION 343
 - 4.6.1 Implementation of Pure Segmentation 347
 - 4.6.2 Segmentation with Paging: MULTICS 348
 - 4.6.3 Segmentation with Paging: The Intel Pentium 352
- 4.7 OVERVIEW OF MEMORY MANAGEMENT IN MINIX 356
 - 4.7.1 Memory Layout 358
 - 4.7.2 Message Handling 361
 - 4.7.3 Memory Manager Data Structures and Algorithms 363
 - 4.7.4 The FORK, EXIT, and WAIT System Calls 367
 - 4.7.5 The EXEC System Call 368
 - 4.7.6 The BRK System Call 371
 - 4.7.7 Signal Handling 372
 - 4.7.8 Other System Calls 378

- 4.8 IMPLEMENTATION OF MEMORY MANAGEMENT IN MINIX 379
 - 4.8.1 The Header Files and Data Structures 379
 - 4.8.2 The Main Program 382
 - 4.8.3 Implementation of FORK, EXIT, and WAIT 382
 - 4.8.4 Implementation of EXEC 385
 - 4.8.5 Implementation of BRK 386
 - 4.8.6 Implementation of Signal Handling 387
 - 4.8.7 Implementation of the Other System Calls 393
 - 4.8.8 Memory Manager Utilities 394
- 4.9 SUMMARY 396

5 FILE SYSTEMS

401

- 5.1 FILES 402
 - 5.1.1 File Naming 402
 - 5.1.2 File Structure 404
 - 5.1.3 File Types 405
 - 5.1.4 File Access 407
 - 5.1.5 File Attributes 408
 - 5.1.6 File Operations 409
- 5.2 DIRECTORIES 410
 - 5.2.1 Hierarchical Directory Systems 411
 - 5.2.2 Path Names 412
 - 5.2.3 Directory Operations 414
- 5.3 FILE SYSTEM IMPLEMENTATION 415
 - 5.3.1 Implementing Files 415
 - 5.3.2 Implementing Directories 419
 - 5.3.3 Disk Space Management 422
 - 5.3.4 File System Reliability 424
 - 5.3.5 File System Performance 429
 - 5.3.6 Log-Structured File Systems 432
- 5.4 SECURITY 434
 - 5.4.1 The Security Environment 434
 - 5.4.2 Famous Security Flaws 436
 - 5.4.3 Generic Security Attacks 439
 - 5.4.4 Design Principles for Security 441
 - 5.4.5 User Authentication 442

- 5.5 PROTECTION MECHANISMS 446
 - 5.5.1 Protection Domains 446
 - 5.5.2 Access Control Lists 448
 - 5.5.3 Capabilities 450
 - 5.5.4 Covert Channels 451
- 5.6 OVERVIEW OF THE MINIX FILE SYSTEM 453
 - 5.6.1 Messages 454
 - 5.6.2 File System Layout 454
 - 5.6.3 Bit Maps 458
 - 5.6.4 I-nodes 460
 - 5.6.5 The Block Cache 461
 - 5.6.6 Directories and Paths 463
 - 5.6.7 File Descriptors 465
 - 5.6.8 File Locking 467
 - 5.6.9 Pipes and Special Files 467
 - 5.6.10 An Example: The READ System Call 469
- 5.7 IMPLEMENTATION OF THE MINIX FILE SYSTEM 470
 - 5.7.1 Header Files and Global Data Structures 470
 - 5.7.2 Table Management 474
 - 5.7.3 The Main Program 482
 - 5.7.4 Operations on Individual Files 485
 - 5.7.5 Directories and Paths 493
 - 5.7.6 Other System Calls 498
 - 5.7.7 The I/O Device Interface 501
 - 5.7.8 General Utilities 503
- 5.8 SUMMARY 503

6 READING LIST AND BIBLIOGRAPHY

507

- 6.1 SUGGESTIONS FOR FURTHER READING 507
 - 6.1.1 Introduction and General Works 507
 - 6.1.2 Processes 509
 - 6.1.3 Input/Output 510
 - 6.1.4 Memory Management 511
 - 6.1.5 File Systems 511
- 6.2 ALPHABETICAL BIBLIOGRAPHY 512

APPENDICES

A	MINIX SOURCE CODE LISTING	521
B	INDEX TO FILES	905
C	INDEX TO SYMBOLS	909
	INDEX	925