

重点大学计算机教材



华章教育

操作系统 原理与设计

曹先彬 陈香兰 编著
中国科学技术大学



机械工业出版社
China Machine Press

重点大学计算机教材



华章教育

操作系统 原理与设计

曹先彬 陈香兰
中国科学技术大学



机械工业出版社
China Machine Press

本书的内容尽量全面覆盖操作系统课程的所有基础知识,主要包括操作系统概论、进程管理、内存管理、输入输出管理、文件管理、操作系统示例分析等。全书在强调基本原理与主要知识点的介绍的同时,力求原理、设计与实现并重;使学生能在掌握基本原理之后,通过对设计与具体实现的学习进一步加深对操作系统基础知识的理解。

本书既可作为高等学校计算机专业本、专科教材,也可作为非计算机专业的操作系统课程教材,同时也适合广大学生自学和考研复习使用。

版权所有,侵权必究。

本书法律顾问 北京市展达律师事务所

图书在版编目(CIP)数据

操作系统原理与设计/曹先彬,陈香兰编著. —北京:机械工业出版社,2009.9
(重点大学计算机教材)

ISBN 978-7-111-27377-6

I. 操… II. ①曹… ②陈… III. 操作系统—研究生—入学考试—自学参考资料 IV. TP316

中国版本图书馆CIP数据核字(2009)第090148号

机械工业出版社(北京市西城区百万庄大街22号 邮政编码 100037)

责任编辑:王璐

北京市荣盛彩色印刷有限公司印刷

2009年9月第1版第1次印刷

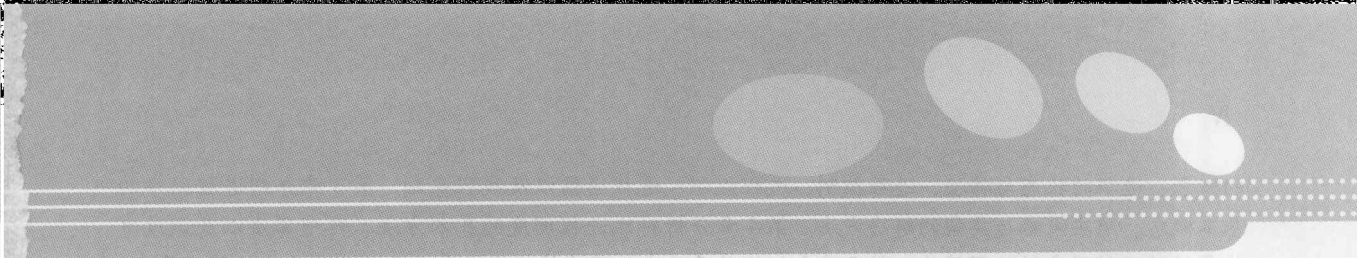
184mm×260mm·22.25印张

标准书号:ISBN 978-7-111-27377-6

定 价:36.00元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换

本社购书热线:(010) 68326294



的概念、进程的状态与转换、进程控制、进程组织、进程通信以及线程基础。第3章介绍进程同步的基本概念、解决策略与典型的解决方法。第4章分析对比进程调度的基本方法、典型算法以及调度可能导致的死锁现象的处置策略。第5章介绍操作系统中的用户接口方式和用于批处理系统的作业管理。

第三部分是内存管理。本部分包括第6、7章。其中，第6章介绍存储管理的基础知识，主要包含三个部分：存储管理的概念、连续分配管理和非连续存储管理。第7章介绍虚拟存储管理知识，主要包括虚拟存储的基本概念和虚拟存储的实现技术。

第四部分是输入输出管理。本部分包括第8章。首先介绍I/O过程、包含的功能部件、I/O控制方式等基础知识；然后介绍I/O核心子系统，其中分析指出I/O管理需要解决的关键技术，并给出相应的实现方法。

第五部分是文件管理。本部分包括第9、10章。其中，第9章介绍文件管理的一些基础知识。主要包括：文件系统的基本概念、文件存储介质、单个文件的逻辑组织和多个文件的目录组织等基础知识。最后简要地介绍了实现文件共享与保护的基本方法。第10章侧重介绍文件系统实现时需要的相关技术。包括外存空间管理技术、文件系统实现时需要的结构信息、文件在外存空间存储可以采用的存储方法与存取方式、支持文件存储的磁盘空间优化调度与提高文件系统可靠性等性能的方法。

第六部分是操作系统示例分析。本部分包括第11章。本章选用Linux操作系统作为示例平台，从源码级出发，详细介绍了操作系统设计的具体思路与实现方法。本章内容顺序是完全按照前面章节对操作系统原理的介绍思路展开的，这样可以将示例操作系统的知识与前面的基础知识对照起来进行学习与理解。

第七部分是操作系统的扩展内容。本部分包括第12、13、14章。在本部分，我们选择介绍了操作系统的安全性、分布式操作系统以及多核操作系统。其中，第12章介绍了操作系统本身提供的一些经典的安全措施，便于读者对操作系统安全有一个最基本了解。第13章阐释了与分布式操作系统相关的基础知识，重点介绍分布式操作系统设计涉及的分布式通信、命名系统、分布式资源管理与分布式处理机管理等内容，以期对分布式操作系统有一宏观上的了解。第14章对多核技术和多核操作系统做简单介绍，其中重点指出多核操作系统实现的关键技术，目的是使读者对多核操作系统有概念性的了解。

本教材全面系统地介绍操作系统的经典内容和若干最新发展。本书既可作为高等学校计算机专业本、专科教材，也可作为非计算机专业的操作系统课程教材，同时也适合自学和考试复习使用。

□ 致谢

本教材是在学习、借鉴已有多个操作系统教材的基础上，整理我们十几年的教学讲义而成的。在此，我们对 Andrew S. Tanenbaum、Gary Nutt、Abraham Silberschatz、William Stallings 等国外著名专家以及我国一线的操作系统主讲老师陈向群、何炎祥、罗宇、陈渝、张亮、王丹、蒲晓蓉、齐勇、郑扣根、陆丽娜、汤小丹、张尧学、孟静、陈莉君等表示衷心的感谢。同时，在本教材的形成过程中，许言午、杜文博、李彤、杨浩、俞盛朋等在文字、绘图等方面给予了很大帮助，在此也表示谢意。特别感谢多年来中国科技大学计算机专业的本科生，他们对本书内容的组织、教学的方法等提出了许多有益的建议。

虽然我们尽可能地避免教材的错误，但限于水平、时间等种种原因，书中难免还存在错误与不妥之处，恳请读者批评指正。

编者

2009年6月

CONTENTS 目录

前 言

第一部分 操作系统概论

第1章 操作系统概述	2
1.1 操作系统的概念	2
1.1.1 操作系统层次模型	2
1.1.2 操作系统的含义	4
1.1.3 操作系统的目的	5
1.1.4 操作系统的功能	5
1.2 操作系统的特点	8
1.2.1 操作系统的研究观点	8
1.2.2 操作系统的特征	9
1.3 操作系统的发展与类型	10
1.3.1 操作系统的发展动力	10
1.3.2 操作系统的发展历程	10
1.3.3 操作系统的类型	15
1.4 操作系统的运行环境	16
1.4.1 操作系统运行的硬件 环境组成	16
1.4.2 操作系统运行涉及的 硬件支持	17
1.5 操作系统的结构设计方法	18
1.5.1 模块接口法	18
1.5.2 有序分层法	19
1.5.3 微核法	19
1.6 本章小结	20
习题	20

第二部分 进程管理

第2章 进程与线程	24
-----------------	----

2.1 多道程序设计与进程概念	24
2.1.1 多道程序设计技术	24
2.1.2 程序的顺序执行与 并发执行	25
2.1.3 进程的概念	29
2.1.4 进程特征	30
2.2 进程的状态与转换	31
2.2.1 三状态模型及其转换	31
2.2.2 五状态模型及其转换	32
2.2.3 七状态模型及其转换	33
2.3 进程控制	35
2.3.1 操作系统内核	35
2.3.2 进程控制手段	36
2.3.3 主要进程控制原语	36
2.4 进程组织	39
2.4.1 进程控制块	39
2.4.2 进程的组织方式	41
2.4.3 进程组织示例	42
2.5 进程通信	42
2.5.1 进程通信概述	43
2.5.2 消息传递系统的实现	44
2.6 线程基础	47
2.6.1 线程引入	47
2.6.2 线程的定义与特征	48
2.6.3 线程的实现	49
2.7 本章小结	51
习题	51

第3章 进程同步	53
----------------	----

3.1 进程同步的基本概念	53
---------------------	----

3.1.1 进程之间的关系	53	4.5.1 死锁的基本概念	107
3.1.2 临界资源与临界区	54	4.5.2 死锁的分析模型	109
3.1.3 进程同步机制的准则	55	4.5.3 死锁的处置策略	111
3.2 实现进程互斥的基本方法	56	4.6 本章小结	117
3.2.1 软件方法	56	习题	118
3.2.2 硬件方法	60	第5章 用户接口与作业管理	120
3.3 信号量机制	61	5.1 用户与操作系统的接口方式	120
3.3.1 单信号量机制	62	5.1.1 作业级接口	120
3.3.2 多信号量机制	66	5.1.2 程序级接口	121
3.4 管程机制	68	5.2 批处理系统的作业管理	122
3.4.1 管程的引入	68	5.2.1 作业的基本概念	122
3.4.2 管程的使用	69	5.2.2 批处理系统的作业组织	123
3.5 典型进程同步机制问题	75	5.2.3 批处理作业的状态 转换与控制	124
3.5.1 生产者—消费者问题	75	5.2.4 批处理系统的作业调度	126
3.5.2 读者—写者问题	80	5.3 交互式系统的作业管理	130
3.5.3 哲学家就餐问题	84	5.3.1 命令式接口	130
3.6 本章小结	87	5.3.2 图形用户界面式接口	131
习题	88	5.4 本章小结	132
第4章 进程调度与死锁	92	习题	132
4.1 处理器调度的基本概念	92	第三部分 内存管理	
4.1.1 处理器调度的类型	92	第6章 存储管理基础	136
4.1.2 调度模型	94	6.1 存储管理的概念	136
4.2 进程调度准则	95	6.1.1 存储管理的目的	136
4.3 进程调度基础	97	6.1.2 存储管理的基本过程	137
4.4 经典进程调度算法	99	6.1.3 覆盖与交换技术	142
4.4.1 先来先服务调度算法	99	6.2 连续分配方式与管理	144
4.4.2 最短优先调度算法	100	6.2.1 单一连续分配方式	144
4.4.3 最短剩余时间调度 算法	101	6.2.2 分区分配方式	145
4.4.4 时间片轮转调度算法	101	6.3 非连续分配方式与管理	152
4.4.5 优先权调度算法	102	6.3.1 分页管理方式	152
4.4.6 最高响应比优先调度 算法	103	6.3.2 分段管理方式	159
4.4.7 多级队列反馈调度算法	104	6.3.3 段页式管理方式	164
4.4.8 进程调度算法示例	105	6.4 本章小结	167
4.5 进程死锁与处置	107	习题	168

第7章 虚拟内存管理 170

- 7.1 虚拟存储基本概念 170
 - 7.1.1 虚拟存储的引入 170
 - 7.1.2 虚拟存储实现技术 172
- 7.2 请求分页管理方式 173
 - 7.2.1 请求分页分配基本思想 ... 173
 - 7.2.2 请求分页分配管理 175
 - 7.2.3 页面分配与调入 177
 - 7.2.4 页面置换 180
 - 7.2.5 抖动处理 187
- 7.3 请求分段管理方式 190
 - 7.3.1 请求分段分配基本思想 ... 190
 - 7.3.2 请求分段分配管理 191
- 7.4 请求段页式管理方式 192
 - 7.4.1 请求段页式分配基本思想 192
 - 7.4.2 请求段页式分配管理 193
- 7.5 本章小结 194
- 习题 194

第四部分 输入输出管理**第8章 I/O 管理** 198

- 8.1 I/O 管理概述 198
 - 8.1.1 I/O 系统与过程 198
 - 8.1.2 I/O 功能部件 199
 - 8.1.3 I/O 控制方式 203
 - 8.1.4 I/O 管理目标与功能 206
- 8.2 输入输出核心子系统 207
 - 8.2.1 I/O 调度 207
 - 8.2.2 缓冲区管理 208
 - 8.2.3 设备分配与回收 210
 - 8.2.4 实现虚拟设备的
SPOOLing 技术 214
 - 8.2.5 出错处理 216
- 8.3 本章小结 216
- 习题 216

第五部分 文件管理**第9章 文件管理基础** 220

- 9.1 概述 220
 - 9.1.1 文件与文件系统 220
 - 9.1.2 文件的类型 222
 - 9.1.3 文件的存储介质 222
- 9.2 文件的逻辑组织 224
 - 9.2.1 堆文件 224
 - 9.2.2 顺序文件 224
 - 9.2.3 索引顺序文件 225
 - 9.2.4 索引文件 226
- 9.3 文件目录 226
 - 9.3.1 基本概念 226
 - 9.3.2 文件目录结构 227
 - 9.3.3 目录查询 231
- 9.4 文件共享 232
- 9.5 文件保护 235
- 9.6 本章小结 236
- 习题 236

第10章 文件系统实现 238

- 10.1 文件系统层次结构 238
 - 10.1.1 文件系统管理对象 238
 - 10.1.2 系统管理软件集合 238
 - 10.1.3 用户接口 239
- 10.2 外存空间管理 239
 - 10.2.1 列表 239
 - 10.2.2 位图 240
 - 10.2.3 链表 240
 - 10.2.4 索引 240
- 10.3 文件系统实现需要的结构
信息 241
 - 10.3.1 磁盘结构信息 241
 - 10.3.2 内存结构信息 241
 - 10.3.3 目录实现的数据结构 ... 242
- 10.4 文件的存储组织与存取方式 243

10.4.1	文件的存储组织	243
10.4.2	文件的存取方式	247
10.5	文件系统的性能问题	248
10.5.1	磁盘优化调度	248
10.5.2	磁盘高速缓冲	250
10.5.3	文件系统的可靠性	250
10.6	文件管理示例	251
10.6.1	文件存储空间的管理 ...	252
10.6.2	文件的存储组织	252
10.6.3	支持文件存取的用户 文件描述符表和 文件表	252
10.6.4	目录管理	254
10.7	本章小结	255
	习题	255

第六部分 操作系统示例分析

第 11 章 Linux 示例分析

11.1	Linux 操作系统中的进程 和线程	258
11.1.1	进程描述符	258
11.1.2	Linux 进程的内核栈 ...	261
11.1.3	Linux 进程的状态	262
11.1.4	Linux 的 0 号进程 init_task	262
11.1.5	Linux 系统中进程的 组织	262
11.1.6	Linux 系统中进程的 控制	263
11.1.7	Linux 中的进程通信 机制	265
11.1.8	Linux 中的线程及其 控制	266
11.2	Linux 中的信号量	266
11.2.1	Linux 内核中的 信号量	266

11.2.2	Linux 用户态的信号 量集接口	267
11.3	Linux 中的调度	268
11.3.1	Linux 中的调度策略 ...	269
11.3.2	Linux 进程的时间片 ...	269
11.3.3	Linux 进程调度的 时机	270
11.3.4	Linux 中的调度算法 ...	271
11.4	Linux 中的物理内存管理	273
11.4.1	Linux 中物理页框的 描述	274
11.4.2	Linux 中的动态可分配 物理内存	275
11.4.3	Linux 中物理内存的 zone 划分	275
11.4.4	Linux 中动态可分配 物理内存的使用	276
11.4.5	伙伴算法	277
11.5	Linux 进程的地址空间及其 维护	278
11.5.1	Linux 进程的用户态 地址空间和内核态 地址空间	278
11.5.2	Linux 的进程地址空间 描述符	278
11.5.3	线性区及其组织	279
11.5.4	线性区的属性标志	280
11.5.5	线性区相关操作	280
11.5.6	页表	281
11.6	Linux 的 I/O 子系统	281
11.6.1	Linux 内核中的 I/O 低级操作	281
11.6.2	Linux 的设备文件和 I/O 高级操作	282
11.6.3	设备驱动	285
11.7	Linux 的文件系统	286

11.7.1 VFS 中的目录组织	286
11.7.2 VFS 提供的系统 调用	288
11.7.3 VFS 的实现	292
11.7.4 EXT2 文件系统	299
11.8 本章小结	302
习题	302

第七部分 操作系统的扩展内容

第 12 章 操作系统的安全

12.1 操作系统安全的重要性	306
12.2 操作系统的安全等级	306
12.2.1 TCSEC	306
12.2.2 国际通用准则 CC	307
12.2.3 我国的安全评估标准	307
12.3 操作系统的安全服务	308
12.3.1 用户管理的安全性	308
12.3.2 访问控制安全服务	308
12.4 主流操作系统的安全措施	309
12.4.1 Windows NT 的 安全性	309
12.4.2 UNIX/Linux 的 安全性	312
12.5 本章小结	313
习题	313

第 13 章 分布式操作系统基础

13.1 概述	314
13.1.1 操作系统的概念	314
13.1.2 分布式计算技术	316
13.1.3 分布式操作系统 概述	318
13.2 经典的分布式操作系统	322
13.2.1 Mach	322
13.2.2 Amoeba	323
13.2.3 Chorus	324
13.2.4 Spring	324

13.3 通信	325
13.3.1 远程过程调用 RPC	325
13.3.2 Mach 中的端口及其 通信	326
13.3.3 Spring 中的 Door	326
13.3.4 通信中的代理	327
13.3.5 代理的建立	327
13.4 命名系统	328
13.4.1 Mach 中端口的命名 和端口名字空间	328
13.4.2 统一资源定位符和 统一资源命名	329
13.4.3 域名系统 DNS 的 层次组织	329
13.4.4 目录服务 X.500 属性 名空间的组织	330
13.5 资源管理	330
13.5.1 管理策略	330
13.5.2 集中式算法	331
13.5.3 分布式的时间戳方法	331
13.5.4 令牌环网算法	332
13.6 处理器管理	332
13.6.1 基于图论的分配	332
13.6.2 发送者或接收者发起的 分布式启发式算法	333
13.7 本章小结	333
习题	334

第 14 章 多核操作系统

14.1 多核技术简介	335
14.2 多核操作系统的类型	336
14.3 多核操作系统关键技术	337
14.3.1 多核中的同步与互斥	338
14.3.2 多核中的调度	339
14.4 本章小结	342
习题	342

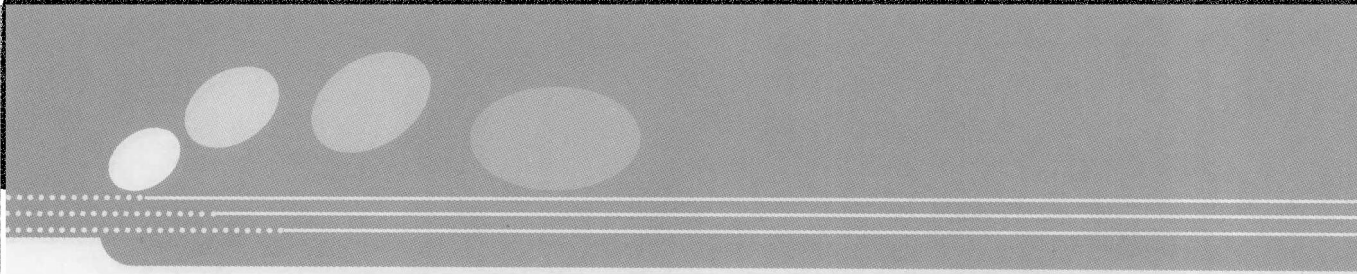
参考文献

第一部分

操作系统概论

操作系统是计算机系统的一个重要组成部分，对其学习与掌握是非常重要的。由于操作系统的发展很快，包含的内容也很多，因此在学习时应该有一个总体思路，由浅入深地把握学习的重点和具体知识点。

本部分即本书的第1章。在第1章中，我们首先从操作系统的层次模型出发，延伸出操作系统的定义、目的、功能等基础知识，使我们对操作系统有一个基本了解。然后从不同角度认识操作系统，并归纳了现代操作系统呈现出的特征。接着介绍操作系统的发展历程，并对操作系统进行了归类，其中，多道批处理操作系统、分时操作系统、实时操作系统这3种典型操作系统是我们关注的重点。还简单介绍操作系统硬件对操作系统的支持。最后对操作系统的结构设计方法进行了介绍。



备。常用的输出设备有：显示器、打印机、绘图仪等。

计算机软件系统的目的是管理、支持计算机运行和完成指定任务。一般可分为系统软件和应用软件两部分。其中：

- 系统软件根据功能又可分为：操作系统、各种语言处理程序和面向应用的数据库管理系统等。
- 应用软件是用户编写的完成具体任务的程序。

值得指出，计算机硬件系统和软件系统是相辅相成的。硬件系统是计算机系统的物理基础，提供了计算机完成各种任务需要的硬件资源。而软件系统是计算机系统的灵魂，正是有了软件系统，计算机系统才能协调各种硬件部件运行，完成用户指定的特定任务。

2. 操作系统模型

在计算机软件系统中，操作系统负责承担的是管理计算机各种软硬件资源并协调计算机工作的流程。因此，我们有必要在计算机系统的整体背景下，更深入地认识操作系统这一最重要的系统软件。

在操作系统设计中，一般将运算器与控制器统一称为**处理器（Processor）**或**中央处理器（Central Processing Unit, CPU）**；由于内存是直接支持计算机运行的，操作系统将单独考虑对它的管理，称为**存储器（Memory）**；输入/输出（I/O）设备统一作为**设备（Device）**进行管理。另外，由于信息在计算机中主要是以文件的形式存储在外存中的，因此一般将文件和外存统一进行管理。

图 1-1 为操作系统的层次模型，主要由计算机硬件和软件组成。

从图 1-1 可以看出：作为最重要的系统软件，操作系统是叠加在计算机硬件之上的第一层软件，是其他软件和硬件之间的接口。具体地说，操作系统起着承上启下的桥梁作用：向上，操作系统支撑着上层的语言处理程序、应用软件等软件部件。一个应用软件必须在操作系统的支撑下，才能调动相关的计算机硬件和信息资源完成特定的任务。为此，操作系统需要设计专门的**外壳（Shell）**程序和接口，以便向上层应用提供功能接口并保证用户应用的方便性；向下，操作系统需要具有对所有计算机硬件和信息资源的管理功能，包括处理器管理、存储器管理、I/O 管理等 3 个硬件资源的管理，以及一个信息资源的管理（文件管理），并应该尽可能将这些资源高效地管理好。

图 1-1 除描述了操作系统对外所起的作用外，同时也指明操作系统在内部组成方面也是一个分层的结构。简单来看，操作系统由位于底层的内核和高层的外壳组成。其中，内核包括处理器、存储器、设备等硬件资源管理模块和一个信息资源管理模块（文件管理模块），它是操作系统的**关键和根本**。因此，操作系统内核既是我们学习操作系统的主要内容，也是操作系统设计与实现时需要重点关注的内容。

总之，图 1-1 反映了操作系统在整个计算机系统中的地位 and 组成，也示意了操作系统的内部组织结构，可以看作是操作系统相关知识的一个总领。后面介绍的操作系统的含义、目的、功能及特点等都可从图 1-1 出发展开与分析得到。

分析现有的典型操作系统，一个具体的操作系统在设计时，也可以说是按照图 1-1 所示的操

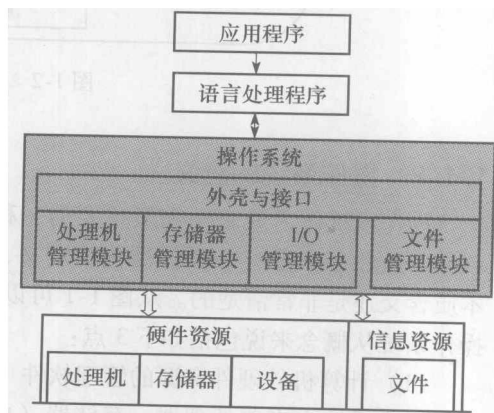


图 1-1 操作系统的层次模型

作系统层次模型来进行的。例如,图 1-2 为 Windows NT 的总体结构,图 1-3 为 UNIX 的总体结构。可以看出,它们就是对图 1-1 的细化。

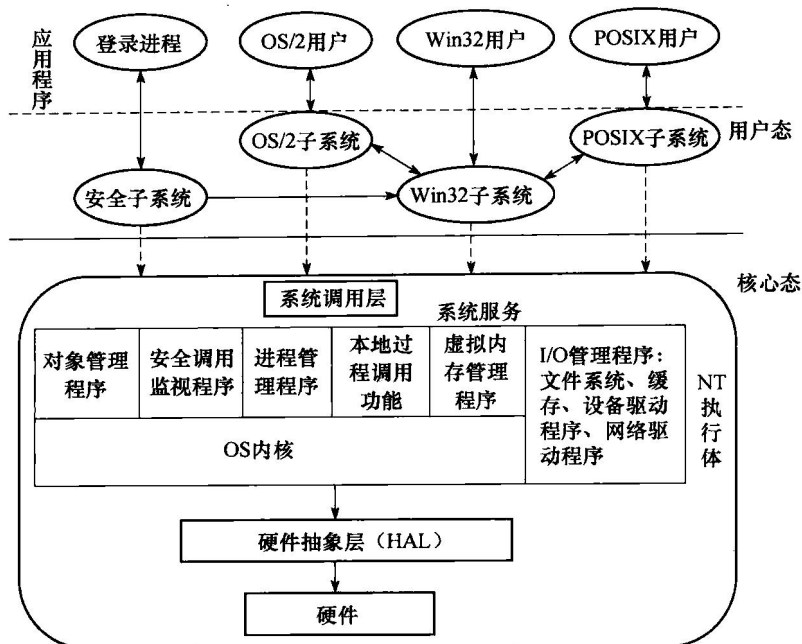


图 1-2 Windows NT 的总体结构

1.1.2 操作系统的含义

操作系统作为最基本和最重要的一种系统软件,虽然至今没有一个标准化的定义,但其本质含义还是非常清楚的。从图 1-1 可以看出,操作系统从概念来说包括如下 3 点:

(1) 计算机软硬件资源的管理软件

操作系统充当着处理器、存储器(内存)、设备等硬件资源以及信息软件资源的管理者,使它们能够正常地运行。相应地,操作系统必须包括处理器管理模块、存储器管理模块、I/O 管理模块和文件管理模块。

值得指出,操作系统中的资源概念是一个泛化的概念。它既包括物理资源,如 CPU、内存、输入/输出设备、网络设备等硬件资源;也包括逻辑资源,如作业、进程、文件、共享的程序和数据等软件资源。

(2) 计算机工作流程的组织管理软件

操作系统充当着从数据输入、计算处理直到结果输出等全过程的协调者,通过合理地组织计算机的工作流程,使计算机能够合理、高效地运行。为此,操作系统必须考虑如何合理地安排计算机的工作流程,如处理器的调度策略、存储器的分配策略、设备的管理策略等都要精心设计。

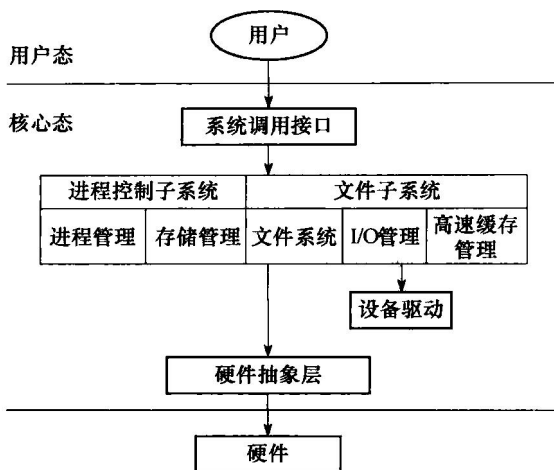


图 1-3 UNIX 的总体结构

(3) 支持用户使用的服务软件

操作系统充当着服务与接口的提供者,并具有响应用户需求的能力,使用户能够方便地通过这些服务和接口来使用操作系统。为此,操作系统必须为应用程序和终端用户提供有用、方便和高效的软件环境。

1.1.3 操作系统的目的

操作系统的目的也可以从图 1-1 分析得出。对应图 1-1,操作系统的基本目的在于:一方面管理好整个计算机系统的软硬件资源和工作流程;另一方面为用户提供有效的服务;另外,它还要能适应不断发展的计算机组成部件与体系结构的变化。因此,虽然目前存在各种类型的操作系统,且不同类型的操作系统其目标各有侧重,但一般而言,操作系统的目的可以归纳如下:

(1) 迎合用户使用需求的方便性目的

方便性目的要求操作系统能让用户便利地使用操作系统完成自己的任务。为此,每种操作系统都提供了易于用户使用的用户界面和编程接口。

(2) 迎合系统管理需求的合理性与有效性目的

系统管理时的**合理性**要求操作系统能公平地安排计算机的工作流程,防止系统运行出现异常现象(如多个程序在内存的执行顺序因不合理而产生“死锁”或“饥饿”)。进一步,在保证合理性的基础上,系统管理时的**有效性目的**要求操作系统能通过各种不同角度的管理策略的优化设计,使计算机系统能高效地运行,特别是尽可能地提高系统吞吐量和系统资源利用率(如 CPU 利用率等)。

(3) 适应计算机硬件和体系结构变化的可扩展性目的

操作系统作为计算机软件系统的一部分,其发展与计算机硬件系统及体系结构的发展是相辅相成的。随着计算机硬件和体系结构的迅猛发展,操作系统也必须具有良好的**可扩展性**,才能适应自身迅速发展的要求。为此,好的操作系统应采用某种模块化结构,以便能方便地对之增加新的功能模块和修改旧的功能模块。正因如此,基于**客户/服务器**(Client/Server, C/S)结构的操作系统在这点上优于集中式的操作系统。

(4) 适应操作系统运行环境动态变化的开放性目的

在操作系统的运行环境中,由于存在着计算机与设备可能来源于不同厂商、多个计算机可能通过网络加以集成以便协同工作、应用软件需要提高可移植性和互操作性等因素,要求操作系统必须具有统一的开放环境。例如,在操作系统设计时,为适应不同硬件的特点,需要专门考虑如何读写存储器的物理地址、设备接口寄存器、设备接口缓冲区等。这样,由于操作系统承担了上述工作,具有了适应运行环境动态变化的**开放性**,所以在其之上的各类程序就没有必要直接同硬件打交道了。当硬件改变时,操作系统相应变化即可,而其他程序无须作出改变。现代操作系统如 Windows NT 中的**硬件抽象层**(Hardware Abstract Level, HAL)的作用主要就在于此。

值得说明的是:在操作系统的上述目的中,它们的权重是变化的。例如,在过去很长一段时间里,由于计算机系统非常昂贵,操作系统的有效性目的比方便性目的更为重要。但随着计算机的普及,价格已不再是首要的考虑因素,而是更加重视操作系统的方便性。方便性和有效性是操作系统最基础性的两个目标。

1.1.4 操作系统的功能

由于操作系统起着全面管理计算机软硬件资源的作用,因此它需要采取各种措施,合理、有

效地管理好计算机的处理器、存储器、I/O 设备、文件管理以及人机接口等功能部件,使它们能相互配合、协同工作,实现对计算机系统的资源管理,并控制程序的执行。

1. 处理器管理

处理器管理的工作是:对处理器资源进行合理的分配使用,以提高处理器利用率,并使各用户公平地得到处理器资源。处理器管理(如处理器调度)的单位可为进程或线程。为此,处理器管理又包括如下一些功能。

(1) 进程描述与组织

基于对进程逻辑特性的分析,操作系统提供了相应的手段来描述和存储组织内存中的进程。

(2) 进程控制

操作系统需要对进程在内存中的活动进行全生命期的控制,包括创建、撤销、挂起、改变运行优先级等,以主动改变进程的状态。其中,由于多个进程可能都提出了使用处理器资源的申请,还需要重点考虑进程的调度机制。进程调度机制,即根据一定的准则,设计合适的进程调度算法,从多个申请占有处理器的进程中选择出一个进程并将处理器分配给它;同时,由于进程调度不当时可能导致不好的后果(如死锁和饥饿),还需要考虑死锁和饥饿等不良后果的处置方法。进程控制的最终目的是充分利用处理器资源和提高系统性能。

(3) 进程同步与通信

由于内存中多个进程之间存在相互协作和通信,操作系统必须考虑如何保证进程之间正确的互斥与同步,以及如何高效地传送数据。为此,操作系统提供了相应的进程同步机制和进程通信手段。

另外,处理器管理还包括对作业的管理功能,以保证作业周转的快速性以及系统处理用户作业的高效性等。

2. 存储器管理

存储器管理的工作是:为系统和用户提供一个独立于实际内存的存储空间,使信息可以方便、准确、安全地存放,并使存储访问可以合理、高效地进行。存储器管理的目标可总结为:提高存储资源的利用率、方便用户使用、方便进程的并发运行。具体地,存储器管理包括下列一些功能。

(1) 虚拟存储

虚拟存储技术可构建一个远大于实际内存的虚拟存储空间。提供地址映射(变换)手段,实现进程逻辑地址到内存物理地址的映射。

(2) 内存访问管理

重点管理进程对内存的访问全过程,包括:内存申请处理、存储分配管理、内存使用管理与内存回收管理等。

(3) 存储保护

设置存储保护功能,使每个进程运行在其独立的虚拟地址空间中,保证进程间互不干扰。

3. I/O 管理

I/O 管理的工作是:向用户提供方便、统一的外部设备接口,按照用户的要求和设备的类型等属性,控制设备工作并完成用户的 I/O 请求;利用中断技术、通道技术和缓冲技术,提高 CPU 与设备、设备与设备之间的并行工作能力,以充分利用设备资源,提高外部设备的使用效率;保证在多道程序环境下,当多个进程竞争使用设备时,能按照一定的策略分配和管理设备,使系统能有条不紊地工作。I/O 管理的目标为:方便设备使用、提高 CPU 与 I/O 设备的利用率。I/O 管理的具体功能包括:

(1) 设备驱动

利用设备驱动程序和设备中断处理,实现物理 I/O 操作。

(2) 设备独立性

提供统一的 I/O 设备接口,使应用程序独立于物理设备,提高系统的可适应性。

(3) 虚拟设备及其实现

采用某种技术(如 SPOOLing 技术),将本质上独占性的物理设备转换为逻辑上可供多个申请进程共享的虚拟设备。

(4) 设备分配和回收

当多个进程竞争使用设备时,按照一定的策略进行设备分配;并在进程释放设备后,回收设备到系统中,使系统能有条不紊地工作。

(5) 缓冲区管理

提供缓冲存储空间,以缓解 CPU、外部设备等在处理速度上的不匹配。同时,设计高效的缓冲区管理机制,提高设备的利用率。

4. 文件管理

文件管理的工作是:管理信息的存储介质(磁盘空间);按照一定的结构组织多个文件并采取合适的方式存储文件;提供文件使用的操作;提高文件系统的性能。文件管理的具体功能有:

(1) 文件存储空间管理

对信息的存储介质进行管理,解决在何处存放信息的问题,以提高空间利用率和读写性能。

(2) 信息组织与存储

以文件、目录等形式来组织与存储信息,重点实现信息的高效检索。

(3) 文件操作

提供文件的读、写等操作,方便用户使用。

(4) 文件系统的性能管理

提供一定的安全性、可靠性的保障措施,提高文件系统的性能。

5. 人机接口

人机接口的工作是:为用户提供一个友好的访问操作系统的接口,以方便用户使用操作系统。操作系统提供了三种主要的接口方式:

(1) 系统命令方式

供用户在系统状态下直接组织和控制自己的作业运行。如:命令行、菜单。

(2) 编程接口方式

提供一系列的系统调用和高级语言库函数,供用户程序和系统程序内部调用操作系统功能。

(3) 图形接口方式

采用图形用户界面(Graphical User Interface, GUI),用非常容易识别的图标(icon)表示系统的各种功能、应用程序以及文件。用户可以通过鼠标、菜单和对话框来完成对应用程序和文件的操作。此时,用户不必记住命令名及格式,从而把用户从繁琐且枯燥的操作中解放出来。它实际上是图形化的系统命令方式。

上面的第(1)、(3)两种接口又可以统称为作业级接口;这时,第(2)种接口对应称为程序级接口。