



普通高等教育“十一五”国家级规划教材

高等学校计算机系列

计算机操作系统

刘循 朱敏 文艺 编著



人民邮电出版社
POSTS & TELECOM PRESS

普通高等教育“十一五”国家级规划教材
高等学校计算机系列

计算机操作系统

刘 循 朱 敏 文 艺 编著

人民邮电出版社
北 京

图书在版编目 (C I P) 数据

计算机操作系统 / 刘循, 朱敏, 文艺编著. —北京: 人民邮电出版社, 2009.2 (2009.6 重印)
普通高等教育“十一五”国家级规划教材. 高等学校计算机系列
ISBN 978-7-115-19037-6

I. 计… II. ①刘…②朱…③文… III. 操作系统—高等学校—教材 IV. TP316

中国版本图书馆CIP数据核字 (2009) 第000052号

内 容 提 要

操作系统是现代计算机系统中必不可少的重要系统软件, 也是计算机专业的必修课程。本书深入浅出地对操作系统的基本原理进行了描述, 并以 UNIX 和 Windows 操作系统为例, 分析了操作系统的实现思想。

全书共分为 11 章。主要内容包括操作系统的基本概念和作用、不同操作系统的特点, 进程和线程的基本概念、处理器调度、进程的同步与通信、死锁, 传统存储器管理和虚拟存储器管理, 设备管理, 文件管理和磁盘管理, 用户接口, 操作系统结构等。

本书的内容符合计算机专业“计算机操作系统”教学大纲要求, 并涵盖 2009 年硕士研究生全国统一入学考试大纲的内容。

本书可作为计算机及相关专业的教材, 也可供初学者学习参考。

普通高等教育“十一五”国家级规划教材

高等学校计算机系列

计算机操作系统

-
- ◆ 编 著 刘 循 朱 敏 文 艺
责任编辑 邹文波
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京楠萍印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 21
字数: 507 千字 2009 年 2 月第 1 版
印数: 3 001—4 500 册 2009 年 6 月北京第 2 次印刷

ISBN 978-7-115-19037-6/TP

定价: 34.00 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223
反盗版热线: (010)67171154

前 言

操作系统是计算机系统的重要组成部分，是计算机及相关专业的必修课程。

随着我国高等教育改革的深入和计算机科学技术的发展，近几年来，计算机及相关专业本科教学发生了很大的变化，对传统的课程内容及其课时进行了精简，并设置了许多新的课程，如计算机网络、分布式处理系统、计算机及信息安全、UNIX 操作系统、Linux 操作系统和大型计算机操作系统等。

为了适应教学改革，操作系统课程的课时从过去的近 80 学时降到 68 学时，再降到现在的 48 学时，课程内容和教学大纲内容也进行了相应调整，并增设了 20 学时的操作系统课程设计。针对这些变化，现在操作系统课堂授课内容主要围绕操作系统的管理功能展开，除进程与线程管理、存储器管理、设备管理、文件管理和用户接口外，传统操作系统教程中的内容，如网络操作系统、分布式操作系统和操作系统安全等内容，基本不包括在操作系统课程教学中，而分别设置在计算机网络、分布式处理系统、计算机及信息安全、Linux 操作系统和 UNIX 操作系统课程中。

本书的内容组织和编写正是为了适应操作系统课程的变化。

本书的内容在完全围绕操作系统的处理器管理、存储器管理、设备管理、文件管理和用户接口功能的基础上，将操作系统结构作为独立的章节。这样，既体现了课程授课内容，又体现了操作系统课程设计的内容。

为了适应计算机新技术发展的特点，特别是多核处理器的出现，本书扩大了线程管理和多处理器调度等内容。

在操作系统课程学习中，学生普遍感到学习内容太抽象，许多原理难以掌握。本书在介绍重要的概念和原理之后，针对实际操作系统中的实现，进行了实例分析。

(1) 在第 1 章导论中有许多操作系统的基本概念（如操作系统概念、操作系统的作用和功能、批处理系统、单道系统和多道系统、联机输入\输出、脱机输入\输出等），为了帮助读者理解这些概念和了解这些概念出现的历史背景，作者以 IBM 计算机为体系，说明近几十年计算机和计算机操作系统的发展历程。

(2) 在进程描述和控制中，以 UNIX 操作系统为实例，分析了进程的描述和控制思想，并以 Solairis 操作系统为实例，分析了线程、线程和进程之间的关系。

(3) 在处理器调度中，以 Windows 操作系统为实例，分析了进程调度、线程调度和多处理器调度问题。

(4) 在进程同步和通信中，以 UNIX 操作系统和 Windows 操作系统为实例，分析了操作系统的进程同步和进程通信的实施。

(5) 在存储器管理中，以 Windows 操作系统和 MULTICS 操作系统为实例，分析了分页、分段和段页式存储器管理的原理。

(6) 在操作系统结构中，以 Windows 操作系统和 UNIX 操作系统结构为实例，分析操作系统的结构方式和特点。

除了从章节上力求做到重点突出、内容精选、结构合理外，本书对每章后的练习进行了精选，体现为原理和综合分析两部分。综合分析题力求具有代表性，能够突出重点内容。

本书的内容符合本科教学 48~51 学时（3 学分）。学时安排如下：

章	内 容	学 时 数
1	导论	3
2	进程的描述与控制	3
3	处理器调度	6
4	进程同步与进程通信	6
5	死锁	3
6	存储器管理	6
7	虚拟存储器管理	3
8	输入/输出设备管理	6
9	文件管理	6
10	用户接口	3
11	操作系统的性能指标和系统结构	6
合 计		51

在本书的编写过程中，刘萍、肖新、张静、刘小刚等参与了资料和练习的整理以及部分内容的校验工作，在此表示感谢。

限于编著者水平，书中难免有错误和不妥之处，恳请读者批评指正。

编著者

2008 年 12 月于四川大学

目 录

第 1 章 导论	1	2.2.1 进程创建	38
1.1 操作系统与计算机	1	2.2.2 进程的结束	39
1.1.1 计算机的组织结构	1	2.2.3 进程阻塞和唤醒	40
1.1.2 操作系统与计算机	2	2.2.4 进程的挂起和激活	40
1.1.3 操作系统的概念和使用操作系统的目的	2	2.2.5 执行的模式	41
1.2 操作系统的功能和特征	3	2.2.6 操作系统的运行	41
1.2.1 操作系统的功能	3	2.3 前趋图	42
1.2.2 操作系统的特征	6	2.4 UNIX 操作系统的进程描述和控制	43
1.3 操作系统的发展	7	2.4.1 UNIX 操作系统的进程状态	43
1.3.1 早期的系统	7	2.4.2 UNIX 操作系统进程的描述和控制	44
1.3.2 批处理系统	8	2.4.3 UNIX 操作系统的多级进程关系	45
1.3.3 多道程序系统	10	2.5 线程的描述与控制	46
1.3.4 分时操作系统	11	2.5.1 线程及其特征	46
1.3.5 实时操作系统	13	2.5.2 线程的状态与控制	48
1.3.6 个人计算机操作系统	15	2.5.3 线程库	48
1.3.7 网络操作系统	16	2.5.4 线程的实现	49
1.3.8 分布式操作系统	17	2.5.5 线程模型	51
1.3.9 嵌入式操作系统	19	2.6 Solaris 操作系统线程	52
1.4 IBM 计算机与操作系统	20	2.6.1 Solaris 系统中的进程与线程	52
1.4.1 早期的 IBM 计算机	20	2.6.2 Solaris 系统中的线程状态	54
1.4.2 具有批处理系统的计算机	22	2.7 本章小结	55
1.4.3 IBM 360 家族	23	练习 2	55
1.4.4 IBM AS400	25	第 3 章 处理器调度	57
1.4.5 IBM 大型主机	27	3.1 处理器调度的层次	57
1.5 本章小结	28	3.1.1 高级调度	57
练习 1	29	3.1.2 中级调度	60
第 2 章 进程的描述与控制	30	3.1.3 低级调度	60
2.1 进程描述	30	3.2 评价调度算法的准则	63
2.1.1 进程概念和特征	30	3.3 调度算法	65
2.1.2 进程的基本状态及转换	32	3.3.1 作业调度算法	65
2.1.3 具有挂起功能的进程状态和转换	33	3.3.2 进程调度算法	71
2.1.4 进程的描述	35	3.4 线程调度	75
2.2 进程控制	38		

3.5 实时调度.....76	4.4.3 哲学家就餐问题..... 124
3.5.1 实时调度需要满足的条件.....76	4.5 管程.....124
3.5.2 实时调度算法.....78	4.5.1 管程的定义.....125
3.6 多处理器调度.....79	4.5.2 Hoare 和 Hanse 观点.....127
3.6.1 多处理器中同步的粒度.....80	4.5.3 管程的应用.....133
3.6.2 多处理器调度的设计要点.....80	4.6 进程通信.....136
3.6.3 线程调度策略.....82	4.6.1 共享存储区通信机制.....136
3.7 Windows 2000/XP 系统的处理器 调度.....85	4.6.2 消息传递通信机制.....137
3.7.1 Windows 2000/XP 中的线程 调度.....85	4.6.3 管道通信机制.....140
3.7.2 Windows 2000/XP 中线程调度 的数据结构.....86	4.7 线程的同步和通信.....143
3.7.3 Windows 2000/XP 线程调度的 相关问题.....86	4.7.1 线程之间的同步.....143
3.7.4 对称多处理器系统上的线程 调度.....87	4.7.2 线程之间的通信.....145
3.8 本章小结.....89	4.7.3 Windows 2000/XP 中的同步和 通信.....145
练习 3.....90	4.8 本章小结.....147
第 4 章 进程同步与进程通信93	练习 4.....147
4.1 进程并发.....93	第 5 章 死锁150
4.1.1 程序的顺序执行.....93	5.1 死锁的产生.....150
4.1.2 进程的并发性.....94	5.1.1 死锁产生的原因.....150
4.1.3 进程间的竞争和协作.....96	5.1.2 死锁产生的条件.....151
4.1.4 进程同步.....97	5.1.3 处理死锁的基本方法.....152
4.2 临界区管理.....99	5.1.4 资源分配图.....152
4.2.1 临界资源和临界区.....99	5.2 死锁预防.....153
4.2.2 进程同步准则.....99	5.3 死锁避免.....156
4.2.3 早期的临界区管理方法.....100	5.3.1 系统的安全状态.....156
4.3 信号量机制.....108	5.3.2 银行家算法.....158
4.3.1 整型信号量.....108	5.4 死锁的检测和解除.....161
4.3.2 记录型信号量.....113	5.4.1 检测死锁.....161
4.3.3 AND 型信号量集.....114	5.4.2 死锁解除.....163
4.3.4 信号量集.....116	5.5 线程死锁.....165
4.3.5 一些信号量实现问题.....116	5.6 本章小结.....165
4.4 用信号量解决经典进程同步 问题.....118	练习 5.....166
4.4.1 生产者和消费者问题.....118	第 6 章 存储器管理167
4.4.2 读者和写者问题.....121	6.1 存储器管理概述.....167
	6.1.1 存储器的层次.....167
	6.1.2 程序准备执行.....168
	6.1.3 覆盖技术.....171
	6.1.4 紧凑技术.....172
	6.1.5 对换技术.....172

6.2 连续存储空间管理	173	7.3.4 时钟 (clock) 置换算法	203
6.2.1 单一连续分配	173	7.4 页面调度性能	204
6.2.2 固定分区分配方式	173	7.4.1 页面调度对系统性能的影响分析	204
6.2.3 可变分区分配	175	7.4.2 工作集模型	207
6.3 分页式存储管理	178	7.5 请求分段存储管理方式	207
6.3.1 分页存储管理的基本原理	179	7.5.1 请求分段的实现	208
6.3.2 页表	180	7.5.2 段页式虚拟存储器管理的实现	208
6.3.3 地址变换机构	181	7.5.3 具有分页的分段实例: MULTICS 操作系统	209
6.3.4 快表	181	7.6 Windows 2000/XP 系统存储器管理实例	209
6.3.5 多级页表	182	7.6.1 基于分页管理的 Windows 2000/XP	210
6.3.6 反置页表 (Inverted Page Table)	185	7.6.2 Windows 2000/XP 系统的虚拟存储管理实现	211
6.3.7 分页存储管理中的页面共享和保护	185	7.6.3 Windows 2000/XP 的内存空间分配	212
6.4 分段式存储管理	186	7.6.4 Windows 2000/XP 内存页面级保护机制	214
6.4.1 分段存储管理	186	7.7 本章小结	214
6.4.2 分段存储管理中的段表和地址变换机构	187	练习 7	214
6.4.3 分段式存储管理与分页存储管理的区别	188	第 8 章 输入/输出设备管理	216
6.4.4 具有分页的分段	188	8.1 输入/输出系统	216
6.4.5 段的共享和保护	189	8.1.1 输入/输出系统	217
6.5 本章小结	190	8.1.2 输入/输出设备	219
练习 6	191	8.1.3 设备管理的功能	220
第 7 章 虚拟存储器管理	192	8.2 设备控制器	220
7.1 虚拟存储器的基本概念	192	8.2.1 设备控制器及其主要功能	220
7.1.1 虚拟存储器的概念	192	8.2.2 设备控制器的组成	221
7.1.2 虚拟存储器的特征	194	8.3 输入/输出控制方式	222
7.2 请求分页虚拟存储管理	194	8.3.1 轮流测试 (polling)	222
7.2.1 请求分页的硬件支持	194	8.3.2 中断方式 (interrupts)	223
7.2.2 分页虚拟存储器管理实施中的策略问题	197	8.3.3 直接存储器访问 (DMA) 方式	224
7.3 页面置换算法	199	8.3.4 通道 (channel) 方式	225
7.3.1 先进先出 (FIFO) 页面置换算法	200	8.4 缓冲管理	229
7.3.2 最佳 (optimal) 页面置换算法	201	8.4.1 缓冲的引入	229
7.3.3 最近最久未使用 (LRU) 页面置换算法	202		

8.4.2 单缓冲区	229	9.5 磁盘存储器管理	267
8.4.3 双缓冲区	230	9.5.1 磁盘存储器的物理结构	268
8.4.4 循环缓冲区	231	9.5.2 对磁盘的访问	268
8.4.5 缓冲池 (buffer pool)	233	9.5.3 磁盘格式化	269
8.5 输入/输出软件	235	9.5.4 磁盘的访问时间	269
8.5.1 中断处理程序	235	9.5.5 磁盘调度算法	270
8.5.2 设备驱动程序	236	9.5.6 磁盘的分配和回收	274
8.5.3 与设备无关的操作 系统软件	237	9.5.7 磁盘可靠性	278
8.5.4 用户级输入/输出软件	238	9.5.8 独立磁盘冗余阵列 (RAID)	280
8.6 设备分配和回收	240	9.6 文件系统性能和可靠性	281
8.6.1 对设备的控制和描述信息	240	9.6.1 磁盘高速缓存	281
8.6.2 设备分配原则	241	9.6.2 缓冲区的提前读与延迟写	282
8.6.3 设备分配与回收过程	243	9.6.3 减少磁臂运动	282
8.7 本章小结	244	9.6.4 文件系统可靠性	283
练习 8	245	9.7 文件的共享和保护	283
第 9 章 文件管理	246	9.7.1 基于 i 节点的文件共享 方法	283
9.1 文件和文件属性	246	9.7.2 基于符号链方法的文件 共享	284
9.1.1 文件的概念	246	9.7.3 文件的存取权限及验证	284
9.1.2 文件的分类	247	9.8 文件系统实例	285
9.1.3 文件属性	248	9.8.1 CD-ROM 文件系统	286
9.2 文件结构和文件实现	249	9.8.2 MS-DOS 文件系统	287
9.2.1 文件的逻辑结构	249	9.8.3 Microsoft Windows 98 文件系统	288
9.2.2 文件的物理结构	252	9.8.4 UNIX 文件系统	289
9.2.3 文件的实现	256	9.9 本章小结	293
9.2.4 文件的操作	257	练习 9	293
9.3 文件系统	257	第 10 章 用户接口	295
9.3.1 文件系统及其功能	257	10.1 命令接口	295
9.3.2 文件系统模型	258	10.1.1 命令接口及命令分类	295
9.3.3 文件系统的实例	259	10.1.2 命令方式	296
9.4 目录管理	260	10.1.3 常用的 UNIX 命令	296
9.4.1 文件控制块	260	10.2 图形接口	297
9.4.2 文件索引节点	261	10.3 程序接口	298
9.4.3 单级目录结构	262	10.3.1 系统调用函数	299
9.4.4 两级目录结构	263	10.3.2 UNIX 系统中的系统调用	299
9.4.5 树形目录结构	263	10.4 本章小结	309
9.4.6 非循环图目录结构	264		
9.4.7 一般图目录结构	265		
9.4.8 目录实现	266		

练习 10.....	309	11.2.4 基于微内核结构形式.....	314
第 11 章 操作系统的性能指标和		11.2.5 虚拟机结构形式.....	315
系统结构	310	11.3 UNIX 和 Windows 的系统结构...	317
11.1 操作系统的性能指标.....	310	11.3.1 UNIX 的系统结构.....	317
11.2 操作系统结构.....	311	11.3.2 Windows 的系统结构.....	318
11.2.1 简单结构形式	311	练习 11	322
11.2.2 模块化结构形式	311	参考文献	323
11.2.3 基于层结构形式	312		

第 1 章 导 论

操作系统是计算机系统中非常重要的部分。对操作系统的研究与实现一直是计算机科学中挑战性和实用性完美结合的典范。不同的计算机环境和不同的应用环境，需要不同的操作系统。从个人计算机到大型计算机，从办公自动化到电子商务应用环境，从小型企业化管理到大型企业的工业自动化控制，都需要有操作系统的支持。操作系统的职责是管理计算机系统的硬件资源、软件资源，控制计算机的整个工作流程。一方面，计算机系统结构和硬件技术的发展推动了操作系统的发展，因此而出现了多种形式、能够满足多种应用、资源利用效率不断提高的操作系统；另一方面，操作系统技术的不断发展又使得计算机的处理器、存储器等硬件资源的利用率得到了很大提高，能够不断满足用户方便、高效、快捷应用计算机的需求。

本章讲述的主要内容如下：

- 操作系统与计算机；
- 操作系统的功能和特征；
- 操作系统的发展；
- IBM 计算机与操作系统。

1.1 操作系统与计算机

操作系统构建在计算机的硬件结构之上，计算机的组织结构是指计算机的各硬件部件之间的组织形式。在学习操作系统之前首先需要了解计算机的组织结构。

1.1.1 计算机的组织结构

计算机的组织结构决定了计算机中信息传送速度和计算机的性能。虽然不同规模的计算机系统其组织结构形式存在一些差别，但是，计算机的基本组织结构形式是相同的，即由处理器、内存储器、输入部件和输出部件组成。计算机的基本组织结构如图 1.1 所示。

处理器是计算机的核心，由控制器、运算器及高速存储器组成。控制器从计算机的内存储器中取出指令并进行分析，运算器实现指令所需要的操作，高速存储器由若干寄存器组成，存储一些暂时的结果和控制信息。内存储器存储处理器运行的程序和数据，处理器直接读写内存储器。输入部件和输出部件实现计算机的信息输入与输出。磁盘作为计算机的外存，能够长期保存大量的信息，是重要的输入与输出部件。

1.1.2 操作系统与计算机

计算机系统由硬件和软件两大部分组成。硬件是软件运行的平台，软件是硬件功能的体现和扩充，计算机硬件和软件相辅相成、互为促进、缺一不可。

软件包括应用软件和系统软件。应用软件用于解决各种具体的应用问题，如财务软件用于财务管理，办公软件用于处理办公事务，网络软件用于网络应用。系统软件为各种应用提供使用计算机环境和服务支持。

系统软件主要包括操作系统和数据库系统等。在系统软件中，操作系统是基础，操作系统为其他系统软件的运行提供支持与服务，其他的系统软件可以看成为运行在操作系统之上的应用软件或应用程序。操作系统、计算机硬件、应用程序、用户之间的关系如图 1.2 所示。

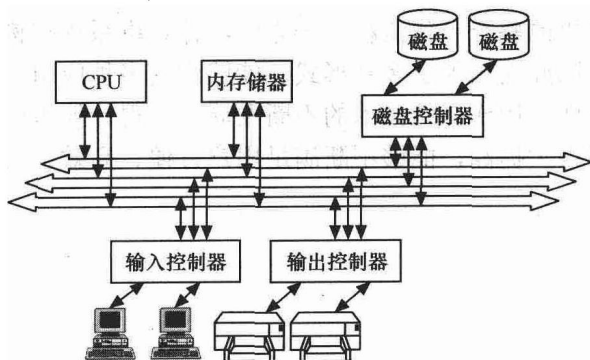


图 1.1 计算机的基本组织结构

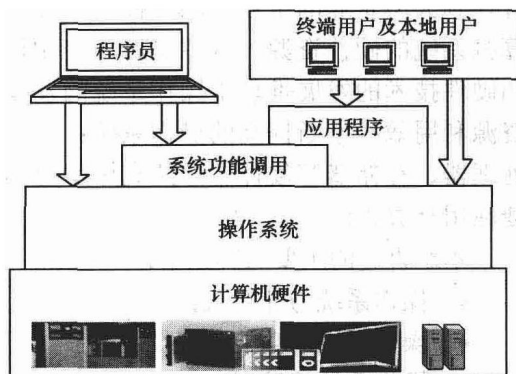


图 1.2 操作系统与计算机硬件之间的关系

计算机的价值通过计算机应用体现。程序员、应用程序和终端用户及本地用户应用通过操作系统应用计算机，操作系统为他们提供运行平台。操作系统的作用如下：

- 直接位于计算机硬件之上，为计算机的应用提供接口；
- 提供通用的计算机服务，与专用的应用领域无关；
- 实现资源管理策略，为不同的应用提供共享资源。

操作系统作为计算机资源的管理者，能够协调和指挥计算机的各个组件按照一定的计划协同工作，有序地控制计算机中的处理器、存储器和输入/输出设备的分配，在相互竞争的用户和程序之间协调冲突，保证计算机系统正常有效地运行。

1.1.3 操作系统的概念和使用操作系统的目的

没有操作系统的计算机，需要直接对计算机硬件进行操作。只有对计算机指令、操作、时序、地址和各类寄存器非常熟悉和精通的计算机专家，才能操作和使用这类计算机。

由于操作系统隐藏了对计算机硬件操作的复杂性，因此，有操作系统的计算机，用户通过操作系统使用计算机，用户不必知道更多的计算机硬件知识便能够方便地操作和使用计算机。

1. 操作系统的概念

操作系统是计算机系统中一组控制和管理计算机硬件资源和软件资源，合理地各类作

业进行调度，以方便用户使用的程序的集合。

2. 使用操作系统的目的

- 有效的管理计算机资源

操作系统采取有效管理机制，使计算机系统中的所有硬件和软件资源能够得到充分利用，在提高系统效率的同时为用户程序运行提供一个公平环境。

- 方便用户使用计算机资源

操作系统为用户提供友好的界面，使得不同国家、不同语言、不同年龄阶层的用户能够方便地使用计算机。

- 扩大计算机功能

操作系统对计算机硬件进行管理和隔离，以虚拟机的形式扩大了计算机的功能，让更多的用户可以分享计算机提供的功能。

- 构筑开放环境

操作系统的设计和构造遵循国际标准有：开放的通信标准、开放的用户接口标准、开放的线程库标准、开放的用户应用标准等。这些标准使得操作系统在体系结构上可实现伸缩性和扩展性，应用程序在不同的平台之间可以平移和相互操作。

1.2 操作系统的功能和特征

只有在明确并理解操作系统功能和特征的基础上，才能完成操作系统结构设计，才能实现并合理应用操作系统，也才能提高计算机系统性能。随着计算机技术的不断发展和计算机应用的日益广泛，操作系统的功能不断增强，操作系统的特征不断完善。

1.2.1 操作系统的功能

操作系统的功能不仅体现在对系统资源进行管理上，而且体现在其为用户提供的应用上。操作系统的功能有处理器管理、存储器管理、输入/输出设备管理、文件系统和用户接口。

1. 处理器管理

处理器（CPU）是计算机的核心部件，是对计算机性能影响最大的系统资源。处理器管理是操作系统最重要的功能。传统的操作系统以进程为资源分配和处理器调度的基本单位，因此，处理器管理主要集中在进程管理上。现代操作系统在进程的基础上引入了线程，处理器调度以线程为基本单位，处理器管理关系到进程与线程的管理。

处理器管理的主要任务如下：

- 进程和线程的描述与控制

根据进程和线程的推进状况，对进程和线程的状态进行描述，控制状态之间的转化。

- 处理器调度

处理器调度分为三级调度：作业调度、中级调度和进程调度。为了最大限度地利用处理器时间并减少作业的等待时间，需要选择合理的调度算法。调度算法包括作业调度算法和进

程调度算法。

- 进程或线程的同步与互斥

并发进程或线程相互配合与互斥地使用系统资源，为了避免进程或线程之间不能相互配合或由于竞争资源而发生“死锁”，需要对进程或线程的同步与互斥进行控制。

- 死锁的检测和预防

多进程环境下，如果进程之间发生死锁，则解决死锁的方法有预防死锁、避免死锁、检测死锁和消除死锁。

- 进程之间及线程之间的通信

进程与线程是操作系统中最基本的活动因素，实现进程之间或线程之间的通信是处理器管理的基本功能之一。

2. 存储器管理

存储器管理是指对计算机内存的管理。内存是程序运行的基本环境，在多道程序环境下，存储器管理需要合理规划并分配存储器空间，实现存储器保护和从逻辑上扩充存储器。

存储器管理的主要任务如下：

- 内存规划、分配及地址映射

为了减少内存空间浪费，提高内存利用率，内存管理需要合理规划并分配内存空间。在内存空间分配上，有连续分配与离散分配两种形式。连续分配可分为单一连续分配和分区分配。单一连续分配适合单道程序环境；分区分配适合多道程序环境。离散分配有分页分配和分段分配。离散分配能够减少连续分配带来的内存“碎片”，使内存的利用率更高。内存管理除了包括各种分配方式外，还包括内存的分配算法和置换算法，以及从逻辑地址到物理地址的映射。

- 内存保护

在多进程和多用户环境下，存储器空间被划分为操作系统空间与用户程序空间，每道程序只能在自己的内存空间中运行。操作系统核心程序与用户程序之间、用户程序与用户程序之间都需要采取保护措施，做到互不越界干扰。保护措施通过越界检查实现，如果程序越界，则需要进行越界处理。

- 内存扩充

为了给用户程序提供尽可能大的存储空间，增加系统的多道度，需要对计算机的内存空间进行扩充。虚拟存储器管理以内存的离散分配为基础，从逻辑上对内存进行扩充。通过外存作为内存的补充，外存与内存之间的页面或段面换进与换出，打破了内存大小对运行程序的约束，做到可以运行更大、更多的程序。

3. 输入/输出设备管理

输入/输出设备管理是指管理与计算机相连的各类外围设备，提高设备的使用率和设备与处理器并行工作的能力，方便用户使用。

设备管理的主要任务如下：

- 输入/输出设备控制

输入/输出设备控制的主要方式有程序控制方式、中断方式、直接存储器访问（DMA）

方式和通道方式。程序控制方式效率较低，现在已经较少采用。通道方式主要用于大型计算机系统。中断方式和直接存储器访问方式最为普遍，被大部分的计算机所采用。

- 缓冲管理

为了解决慢速输入/输出设备与快速处理器之间的矛盾，为了使得输入/输出设备与 CPU 能够并行工作，在计算机的内存空间为各种设备开设了缓冲区。缓冲区分为单缓冲区、双缓冲区、循环缓冲和缓冲池四种形式。缓冲池能够提供多个设备同时使用，与单缓冲区、双缓冲区、循环缓冲相比，利用率最高。

- 设备独立性

实现设备独立性是指通过设备的逻辑名来分配和使用设备，而与设备的物理名无关。设备独立性不仅为用户和用户程序使用设备提供了方便，而且还提高了设备的利用率。

- 设备分配

设备管理通过设备控制表、控制器控制表、通道控制表及系统设备表来管理和分配设备。在设备分配中需要预防和避免由于进程竞争设备而出现的死锁问题。

- 虚拟设备

虚拟设备管理采用 spooling 技术，实现了一台物理设备成为虚拟的多台逻辑设备，满足了多个用户进程对设备的需要，达到有效提高设备利用率的目的。

- 磁盘存储器管理

磁盘作为系统的大规模存储器，能够长期、有效地存储信息。磁盘存储器管理的主要功能包括磁盘存储空间的划分，磁盘存储空间的分配和回收。

4. 文件系统

在操作系统中，存放在磁盘等外存上的信息总是以文件的形式存在。文件系统管理指对文件进行组织和访问控制，实现文件的按名存取，为用户提供方便的文件存取访问和可靠的文件共享，实现文件保护。

文件系统的主要任务如下：

- 对文件结构进行组织和目录管理

提供文件的物理组织方法和逻辑组织方法，实现文件的目录管理。

- 提供文件的存取访问

实现文件的按名存取，为用户提供方便的文件系统接口，便于用户对文件进行操作。

- 实现文件的存储空间管理

根据文件的组织方式合理地分配文件存储空间，有效地管理文件存储空间，实现用户访问文件的快速和有效性。

- 实现文件的共享和保护

提供文件的共享和保护，做到用户对文件的访问权限控制。

5. 用户接口

为了方便用户使用计算机，操作系统为用户提供了程序接口、命令接口和图形接口。

- 程序接口

在用户编程应用上，操作系统为程序员提供系统资源调用函数，达到程序员在程序中方

便使用系统资源的目的。

- 命令接口

对系统的管理和应用，操作系统提供一套系统命令供系统管理员和用户使用。相对于图形接口，命令接口需要用户熟悉命令结构和命令形式。

- 图形接口

对系统的管理和应用，操作系统以图形窗口方式提供给系统管理员和用户使用。图形接口操作简单、直观。

总之，在操作系统的诸多功能中，核心功能是进程管理和内存管理，设备管理属于操作系统资源管理中的具体实例，文件系统是存储设备管理的映像，用户接口是操作系统提供资源的具体形式。这些功能之间的关系如图 1.3 所示。

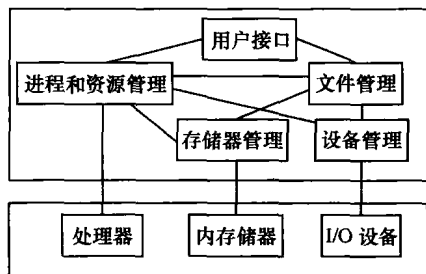


图 1.3 操作系统的功能及其之间的关系

1.2.2 操作系统的特征

操作系统是计算机资源的组织和管理者，在经历了近半个世纪的发展历程后，操作系统的特征逐渐形成，并主要体现在并发性、共享性、虚拟性和异步性上。

1. 并发性

并发性是指两个或两个以上的事件在同一时间段内发生。并发性体现了操作系统同时处理多个活动事件的能力。

对只有一个处理器的系统，在一个时间段内，可以同时运行多个进程，实现多进程并发。这些并发的进程体现为：宏观上同时执行，微观上任何时刻只有一个在执行。

并发性是操作系统最重要的特征，它能够减少计算机中各部件由于相互等待而造成的计算机资源浪费，改善资源利用率，提高系统的吞吐量。

并发性的实现比较复杂，需要解决进程之间的运行切换、进程内容保护、相互依赖进程之间的同步关系、进程资源分配的协调等问题。

多任务系统是在一段时间内能够同时执行多个任务的计算机系统。多任务系统中的多个任务并发执行。

2. 共享性

共享性是指计算机系统资源能够被并发执行的多个进程共同使用。操作系统对这些资源进行合理的调配和管理，使得并发执行的多个进程能够合理地共享这些资源，达到节约资源，提高系统效率的目的。

资源共享有两种实现方式：互斥共享和同时共享。

互斥共享资源是指一段时间内只允许一个进程访问，多个进程需要互相排斥使用的资源。当一个进程正在访问互斥共享资源时，其他需要访问资源的进程必须等待，只有资源被占有进程使用完或释放后，才允许访问。在系统中允许互斥访问的资源有硬件资源和软件资源，如打印机、磁带机、缓冲区、数据和表格等。

同时共享是指同一时间段内允许多个进程同时访问的资源，如可重入代码、磁盘等资源。

同时共享是一种宏观上的“同时”，从微观上看，如果只有一个 CPU，则一次只能处理一个进程，多个进程只能是交替访问资源。正如对磁盘的访问一样，虽然一个时刻只能有一个进程访问磁盘，但是一个时间段内，可以有多个进程交替访问磁盘。不需要等待一个进程对磁盘访问结束后，其他进程才能对磁盘进行访问。

实现资源共享需要解决的问题有资源分配优化、信息保护、存取控制、进程之间同步等。

一方面，因为进程的并发执行才会出现资源的共享；另一方面，只有系统有效管理共享资源，才能使得进程的并发顺利执行。并发和共享都是操作系统的基本特征。

3. 虚拟性

虚拟性是指操作系统通过某种技术将一个实际存在的实体变成多个逻辑上的对应体。这样的多个逻辑对应体可以为多个并发进程访问，提高了实际存在的实体的利用率。

虚拟性是操作系统管理资源的一种重要手段，其目的是为用户提供方便高效的资源利用。在计算机系统中，处理器、存储器、打印机等外围设备以及窗口、用户终端等都可以通过虚拟技术提供给并发的进程使用。

在操作系统中，与虚拟性相关的技术问题有：处理器管理、虚拟存储器管理、spooling (simultaneous peripheral operation on line) 技术等。

4. 异步性

异步性也称为随机性，是指多道程序环境中多个进程的执行、推进和完成时间都是随机的、交替的、不可预测的。多个并发的进程由于受到资源限制而不可能一贯到底，而是“走走停停”。这种异步方式的进程执行，导致的后果可能是进程执行的最终结果不可重现。

异步性会带来进程执行结果的不可重现性，使操作系统存在潜在的危机。操作系统设计和实现必须考虑异步性问题，采取一定的措施，如：信号量机制，保证进程的多次执行结果一致，不会出现不可重现性。

1.3 操作系统的发展

随着计算机硬件技术和软件技术的发展，为了充分提高计算机资源的利用率，为了给用户提供更快、更多、更好的服务，操作系统本身也在不断地向前发展。操作系统从无到有，从单道批处理系统、多道程序系统、分时系统、实时系统到今天的网络和分布式操作系统；从个人计算机系统到大型主机系统，从嵌入式系统到商务系统。在这几十年中，操作系统的发展极为迅猛。

1.3.1 早期的系统

本文所指的早期的系统是 1940~1959 年间的系统。早期计算机的所有程序都以绝对机器语言进行编写，没有程序语言，也没有汇编语言，更没有操作系统。计算机上所有运行程序的设计、构造、编程、操作和维护都由人工直接承担，计算机能做的所有工作就是完成数字计算，如：各种数学运算、函数运算等。根据管理员的时间安排，用户到计算机机房中将