



普通高等教育“十一五”国家级规划教材

普通高等学校计算机专业特色教材

# 计算机操作系统原理

(第2版)

王万森 孙卫真 温绍洁 霍其润 汪国安



高等教育出版社  
Higher Education Press

TP316/339=2

2008

普通高等教育“十一五”国家级规划教材

内容简介

普通高等学校计算机专业特色教材

# 计算机操作系统原理

(第2版)

王万森 孙卫真 温绍洁 霍其润 汪国安

高等教育出版社

ISBN 7-03-021871-0  
定价：35.00元

## 内容简介

本书是普通高等教育“十一五”国家级规划教材，全面系统地介绍操作系统的基本原理、方法与技术，并基于 Microsoft Windows Server 2003 给出相应的实验。全书分为两大部分，第一部分包括第 1~8 章，介绍操作系统概述、操作系统用户接口、进程管理、处理机调度、死锁及其对策、存储管理、设备管理和文件系统；第二部分包括第 9 章和第 10 章，介绍操作系统的进一步发展和 Windows 实验。

本书是在《计算机操作系统原理》的基础上改写而成的，语言通俗、结构合理，每章开头有学习目标，结尾有本章小结，所有练习题都附有相应的参考答案。

本书可作为高等学校本科应用型计算机及相关专业学生的教材，也可供相关专业的工程科技人员参考。

## 图书在版编目 (CIP) 数据

计算机操作系统原理 / 王万森等. — 2 版. — 北京: 高等教育出版社, 2008.5

ISBN 978 - 7 - 04 - 023312 - 4

I. 计… II. 王… 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 (2008) 第 041517 号

策划编辑 刘 艳 责任编辑 焦建虹 封面设计 张志奇 责任绘图 吴文信  
版式设计 张 岚 责任校对 张 颖 责任印制 陈伟光

出版发行 高等教育出版社

社 址 北京市西城区德外大街 4 号

邮政编码 100120

总 机 010 - 58581000

经 销 蓝色畅想图书发行有限公司

印 刷 北京奥鑫印刷厂

购书热线 010 - 58581118

免费咨询 800 - 810 - 0598

网 址 <http://www.hep.edu.cn>

<http://www.hep.com.cn>

网上订购 <http://www.landaco.com>

<http://www.landaco.com.cn>

畅想教育 <http://www.widedu.com>

开 本 787×1092 1/16

印 张 16.25

字 数 360 000

版 次 2001 年 6 月第 1 版

2008 年 5 月第 2 版

印 次 2008 年 5 月第 1 次印刷

定 价 20.60 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 23312 - 00

## 第2版前言

随着信息技术的发展,操作系统在现代计算机系统中的核心地位和在国家信息安全中的重要位置越来越受到人们的重视。这种社会认识的进步为操作系统教学提供了有利的条件和环境,同时也对操作系统教学改革提出了更高的希望和要求。现在,对操作系统认知能力和动手能力的培养已成为操作系统教学改革的两大主题。

基于这一认识,本书在改版中保留了1版书在读者定位、内容筛选和写作方式等方面的特色,更加注重对操作系统知识结构、系统方法和实际技能等方面的要求。全书可分为两大部分,第一部分为第1~8章,介绍操作系统基本原理和方法,包括操作系统概述、操作系统用户接口、进程管理、处理机调度、死锁及其对策、存储管理、设备管理和文件系统;第二部分为第9章和第10章,介绍操作系统的发展与实践,包括操作系统的进一步发展和Windows实验。

本次改版的主要变动如下:

(1) 删除1版书第9章Linux操作系统,新增第10章Windows实验。

(2) 将1版书第10章操作系统进一步发展变更为新的第9章,并删除1版书10.1节计算机系统发展对操作系统的影响以及10.3.2小节面向对象操作系统技术,增加9.1.4小节嵌入式操作系统和9.2.2小节多线程技术。

(3) 对1版书第4章和第5章的次序做了调换,新的第4章为处理机调度,第5章为死锁及其对策。在其他各章中,也有部分节和小节的变动。

(4) 在内容方面,删除或压缩一些较陈旧的方法和技术,相应地增加一些新的方法和技术。

书中除保留1版书每一章开头部分有学习目标、学习重点与难点,结尾部分有学后自测和相应的参考答案外,还在结尾部分增加本章小结,更便于读者自学。

本书由王万森主编,负责大纲的制定和统稿,参加编写工作的有王万森(第1、3、9章),孙卫真(第8、10章),孙卫真、汪国安(第6章),温绍洁(第2、5章)和霍其润(第4、7章)。此外,孙卫真也参与了统稿工作。

在这次改版中,由于地域和工作原因,编者有一定的变动,首先对1版书编者的贡献致以深深的谢意!本书编者在操作系统教学和研究中使用和阅读过不少版本的操作系统著作和教材,在此,谨向这些著作和教材的编者表示感谢!

由于编者水平有限,书中疏漏之处在所难免,恳请各位专家和读者不吝指教。邮件请发送至wansenw@126.com。

编者

2007年11月于北京

# 第 1 版前言

操作系统是现代计算机系统中必不可少的重要系统软件。随着我国信息技术和信息产业的发展,人们越来越认识到操作系统在我国国民经济和社会信息化进程中的基础性地位。确切地说,如果没有我们国家自己的操作系统,就不会有我们国家真正的信息安全。

操作系统课程是计算机专业的一门专业主干课程,主要研究操作系统的基本原理和实现方法,是计算机专业学生的知识结构中必不可少的一个重要组成部分。

本书是编者在多年操作系统教学的基础上,根据教育部师范教育司 1999 年制定的《中学教师进修高等师范本科(专科起点)教学计划》编写的。全书可分为两大部分:第一部分为第 1~8 章,全面系统地讨论了操作系统的基本原理和实现技术;第二部分为第 9 章和第 10 章,分别讨论了 Linux 操作系统和操作系统的最新进展。全书各章的安排分别为:第 1 章操作系统概述,其主要目的是让读者先对操作系统有一个概括的了解,并初步建立起操作系统的整体概念;第 2 章用户接口,介绍了操作系统向用户提供服务的各种手段;第 3 章进程管理,介绍了进程的基本概念、实现方法、控制技术以及进程互斥、同步、通信的概念与方法;第 4 章死锁,介绍了死锁的概念和各种死锁对策;第 5 章处理机管理,介绍了各种调度模型和一些重要的调度算法;第 6 章存储管理,介绍了实存和虚存管理的有关概念和主要的存储管理方法;第 7 章设备管理,介绍了 I/O 设备的管理技术和各种磁盘调度算法;第 8 章文件系统,介绍了文件系统的有关概念,文件的组织、存取、共享、保护方法以及文件目录、文件存储空间的管理方法;第 9 章 Linux 操作系统, Linux 是一个具有全部 UNIX 功能的全开放的多用户、多任务操作系统,其源代码可以直接从网上下载,再加上它的微内核结构,作为实例讨论比较合适;第 10 章操作系统的进一步发展,主要从新一代操作系统和操作系统新技术两个方面来探讨。

书中每一章的开头部分都给出了该章的学习目标和学习重点与难点,对读者学习和理解该章内容起到一定的指导作用;每章附有习题,且对所有习题均给出相应的参考答案,以便于读者自学。

本书由王万森主编,负责全书的大纲制定和统编。参加编写工作的有:王万森(第 1、10 章),沈雪明(第 3、5 章),汪国安(第 6、7 章),杨卉(第 8、9 章),涂相华(第 2、4 章)。

在编写本书之前,编者均曾主讲过不同版本的操作系统教材,这些教材不论对以往的教学,还是对本书的编写,都有很大的帮助。在此,谨向这些教材的编者深表谢意!

彭玉禄审阅了全书,并提出了许多宝贵意见,在此,谨表示衷心感谢!

由于编者水平所限,书中疏漏与错误之处在所难免,恳请各位专家和读者批评指正。

编者

2001 年 3 月

# 目 录

<b>第 1 章 操作系统概述</b> .....	1
1.1 操作系统的定义 .....	1
1.1.1 操作系统的虚拟机观点 .....	1
1.1.2 操作系统的资源管理观点 .....	2
1.1.3 操作系统的用户服务观点 .....	3
1.1.4 操作系统的进程管理观点 .....	3
1.2 操作系统的形成过程 .....	3
1.2.1 人工操作阶段 .....	3
1.2.2 单道批处理阶段 .....	4
1.2.3 多道程序系统阶段 .....	6
1.2.4 进一步发展阶段 .....	6
1.3 操作系统的基本类型 .....	7
1.3.1 多道批处理系统 .....	7
1.3.2 分时系统 .....	9
1.3.3 实时系统 .....	10
1.3.4 通用操作系统 .....	11
1.4 操作系统的特征 .....	12
1.5 操作系统的服务 .....	13
1.5.1 操作系统的公共服务 .....	13
1.5.2 操作系统的服务方式 .....	14
1.6 操作系统的功能 .....	15
1.6.1 用户接口 .....	15
1.6.2 处理机管理 .....	16
1.6.3 存储管理 .....	17
1.6.4 设备管理 .....	18
1.6.5 文件管理 .....	18
1.7 操作系统的结构 .....	19
1.7.1 整体式系统 .....	19
1.7.2 层次式系统 .....	20
1.7.3 客户/服务器系统 .....	20

本章小结 .....	21
学后自测 1 .....	21
参考答案 1 .....	21
<b>第 2 章 操作系统用户接口</b> .....	23
2.1 用户接口概述 .....	23
2.1.1 程序的执行和作业的概念 .....	23
2.1.2 作业管理概述 .....	24
2.2 命令接口 .....	25
2.2.1 脱机命令接口 .....	25
2.2.2 联机命令接口 .....	26
2.3 程序接口 .....	27
2.3.1 系统调用概述 .....	27
2.3.2 系统调用类型 .....	28
2.3.3 系统调用过程 .....	29
2.4 图形接口 .....	30
本章小结 .....	30
学后自测 2 .....	31
参考答案 2 .....	31
<b>第 3 章 进程管理</b> .....	32
3.1 进程的基本概念 .....	32
3.1.1 进程的引入 .....	32
3.1.2 进程的定义和特征 .....	35
3.2 进程的描述 .....	36
3.2.1 进程控制块 .....	36
3.2.2 进程控制块的组织方式 .....	38
3.3 进程状态及其转换 .....	39
3.3.1 进程的基本状态 .....	39
3.3.2 进程基本状态的转换 .....	40
3.4 进程控制 .....	40
3.4.1 进程控制机构 .....	41

3.4.2 进程控制原语 .....	41	4.3.5 优先级调度算法 .....	83
3.5 进程互斥与同步 .....	43	4.3.6 多级反馈队列调度算法 .....	85
3.5.1 互斥的概念 .....	43	4.4 实时系统调度 .....	87
3.5.2 进程同步的概念 .....	45	4.4.1 实现实时调度的基本条件 .....	87
3.5.3 互斥与同步机制 .....	46	4.4.2 实时调度的分类 .....	88
3.5.4 用 PV 原语实现互斥与 同步的方法 .....	47	4.4.3 常用的几种实时调度算法 .....	89
3.5.5 用 PV 原语解决经典的 互斥与同步问题 .....	50	本章小结 .....	91
3.6 进程间通信 .....	55	学后自测 4 .....	92
3.6.1 进程间通信的类型 .....	55	参考答案 4 .....	93
3.6.2 消息传递 .....	56	<b>第 5 章 死锁及其对策</b> .....	97
3.6.3 消息传递的实例——消息 缓冲通信 .....	58	5.1 死锁的概念 .....	97
3.7 线程的概念 .....	60	5.1.1 资源 .....	97
3.7.1 线程的引入和定义 .....	60	5.1.2 死锁定义 .....	98
3.7.2 线程的属性和状态 .....	61	5.2 死锁的产生及对策 .....	99
3.7.3 进程与线程的比较 .....	62	5.2.1 死锁产生的原因 .....	99
本章小结 .....	63	5.2.2 死锁产生的必要条件 .....	100
学后自测 3 .....	63	5.2.3 死锁的描述 .....	101
参考答案 3 .....	66	5.2.4 处理死锁的方法 .....	103
<b>第 4 章 处理机调度</b> .....	72	5.3 死锁的预防 .....	103
4.1 调度的层次与分类 .....	72	5.3.1 破坏互斥条件 .....	104
4.1.1 调度的层次 .....	72	5.3.2 破坏占用并等待条件 .....	104
4.1.2 调度的分类 .....	74	5.3.3 破坏不剥夺条件 .....	105
4.2 调度算法的设计目标和性能 指标 .....	76	5.3.4 破坏环路等待条件 .....	105
4.2.1 设计目标 .....	76	5.4 死锁的避免 .....	106
4.2.2 性能指标 .....	77	5.4.1 系统状态的安全性 .....	106
4.3 调度算法 .....	79	5.4.2 银行家算法 .....	108
4.3.1 先来先服务调度算法 .....	79	5.4.3 银行家算法举例 .....	110
4.3.2 短作业(进程)优先 调度算法 .....	80	5.5 死锁的检测和解除 .....	112
4.3.3 高响应比优先调度算法 .....	81	5.5.1 利用资源分配图 检测死锁 .....	113
4.3.4 时间片轮转调度算法 .....	82	5.5.2 利用死锁检测算法 检测死锁 .....	115
		5.5.3 死锁的恢复 .....	117
		本章小结 .....	119
		学后自测 5 .....	119

01 参考答案 5 .....	120	581 6.5.5 段式存储管理的优缺点 .....	151
<b>第 6 章 存储管理</b> .....	122	6.6 段页式存储管理 .....	152
01 6.1 存储管理的基本概念 .....	122	81 6.6.1 段页式存储管理的	
110 6.1.1 存储器的层次 .....	122	基本思想 .....	152
410 6.1.2 地址重定位 .....	123	081 6.6.2 段页式存储管理的地址	
010 6.1.3 存储管理的目的和功能 .....	125	变换过程 .....	153
910 6.1.4 存储分配方式 .....	126	581 6.6.3 段页式存储管理的	
81 6.2 存储管理技术 .....	127	优缺点 .....	154
190 6.2.1 单一连续区管理 .....	127	08 本章小结 .....	154
400 6.2.2 固定式和可变式分区存储		学后自测 6 .....	154
管理 .....	127	08 参考答案 6 .....	156
710 6.2.3 可重定位分区存储管理 .....	132	<b>第 7 章 设备管理</b> .....	157
050 6.2.4 分区的存储保护 .....	133	11 7.1 概述 .....	157
31 6.3 虚拟存储系统的基本概念 .....	134	401 7.1.1 I/O 设备的种类 .....	157
6.3.1 虚拟存储器 .....	134	7.1.2 设备控制器 .....	158
822 6.3.2 实现虚拟存储器的		401 7.1.3 I/O 控制方式 .....	159
基本原理 .....	134	001 7.1.4 设备管理的功能 .....	164
70 6.4 分页存储管理 .....	135	507 7.2 设备管理的实现技术 .....	165
6.4.1 分页存储管理的		701 7.2.1 中断技术 .....	165
实现原理 .....	135	801 7.2.2 缓冲技术 .....	166
6.4.2 简单分页存储管理 .....	136	000 7.2.3 设备分配技术 .....	168
001 6.4.3 请求式分页存储管理 .....	138	107 7.3 设备管理软件的层次 .....	172
6.4.4 请求式调页存储管理的		500 7.3.1 用户层 I/O 软件 .....	173
页面置换算法 .....	140	800 7.3.2 与设备无关的 I/O 软件 .....	173
6.4.5 页面抖动和工作集模型 .....	143	500 7.3.3 设备驱动程序 .....	174
000 6.4.6 页面保护 .....	146	400 7.3.4 中断处理程序 .....	175
6.4.7 页式存储管理的优缺点 .....	146	107 7.4 虚拟设备 .....	175
10 6.5 请求分段存储管理 .....	146	800 7.4.1 SPOOLing 概念 .....	175
6.5.1 段式存储管理的		800 7.4.2 SPOOLing 系统实现 .....	176
基本思想 .....	146	500 7.4.3 虚拟设备实例 .....	177
110 6.5.2 段式存储管理的		7.5 磁盘驱动调度 .....	177
实现原理 .....	147	100 7.5.1 磁盘结构 .....	177
6.5.3 段的共享与保护 .....	150	800 7.5.2 磁盘调度算法 .....	178
6.5.4 分段与分页的区别		80 本章小结 .....	181
与联系 .....	151	00 学后自测 7 .....	181



参考答案 7 .....	182	<b>第 9 章 操作系统的进一步发展</b> .....	210
<b>第 8 章 文件系统</b> .....	184	9.1 操作系统新类型 .....	210
8.1 文件系统概述 .....	184	9.1.1 微型计算机操作系统 .....	210
8.1.1 文件的概念 .....	184	9.1.2 网络操作系统 .....	212
8.1.2 文件系统的概念 .....	186	9.1.3 分布式操作系统 .....	214
8.2 文件的逻辑结构及其存取 方法 .....	187	9.1.4 嵌入式操作系统 .....	216
8.2.1 文件的逻辑结构 .....	188	9.2 操作系统新技术 .....	217
8.2.2 文件的逻辑存取方法 .....	189	9.2.1 微内核技术 .....	218
8.3 文件的物理结构及其存取 方法 .....	189	9.2.2 多线程技术 .....	221
8.3.1 文件的物理结构 .....	189	本章小结 .....	224
8.3.2 文件的物理存取方法 .....	193	学后自测 9 .....	224
8.4 文件存储空间管理 .....	194	参考答案 9 .....	225
8.4.1 文件存储空间管理的基本 方法 .....	194	<b>第 10 章 Windows 实验</b> .....	226
8.4.2 成组链 .....	196	10.1 Windows Server 2003 基础 .....	226
8.5 文件目录管理 .....	197	10.1.1 Windows Server 2003 进程和线程的实现 .....	226
8.5.1 文件目录的概念 .....	197	10.1.2 Windows Server 2003 进程状态及进程调度 .....	227
8.5.2 文件目录的基本结构 .....	198	10.1.3 Windows Server 2003 内存管理 .....	228
8.5.3 文件路径名 .....	200	10.1.4 Windows Server 2003 文件系统 .....	229
8.5.4 便于共享的目录结构 .....	201	10.1.5 Windows Server 2003 设备管理 .....	230
8.6 文件系统的安全性 .....	202	10.1.6 客户和服务服务器之 间的差异 .....	230
8.6.1 系统安全性保护 .....	203	10.2 Windows Server 2003 的 注册表管理机制和实验 .....	231
8.6.2 存取控制权限保护 .....	203	10.3 Windows Server 2003 进程和 线程实验 .....	235
8.6.3 保密 .....	204	10.4 线程调度 .....	243
8.7 文件的操作与使用 .....	205	10.5 内存实验 .....	244
8.7.1 文件创建和删除 .....	205	10.6 系统安全实验 .....	249
8.7.2 文件打开和关闭 .....	206	<b>参考文献</b> .....	251
8.7.3 文件的读和写 .....	207		
8.7.4 文件的硬链接和解 除硬链接 .....	207		
本章小结 .....	208		
学后自测 8 .....	208		
参考答案 8 .....	209		

# 第1章 操作系统概述

## 学习目标

- 了解操作系统的形成过程,建立操作系统的整体概念。
- 熟悉操作系统的基本类型和服务方式。
- 掌握操作系统的定义、特征和功能。

## 学习重点与难点

- 操作系统的各种观点。
- 操作系统的基本类型和特征。
- 操作系统的服务和功能。

操作系统(operating system)是现代计算机系统的重要组成部分。现在,无论是巨型计算机还是微型计算机,也无论是个人计算机系统还是计算机网络系统,如果离开了操作系统,都将无法运行。学习和掌握操作系统的基本原理和方法是计算机科学技术发展的需要,可以帮助人们从宏观上理解计算机系统的动态工作过程,为计算机研究和应用提供必要的理论和技术基础。

为使读者能尽快地对操作系统的整体概念有一个初步了解,本章先从不同的角度对操作系统进行概括的介绍。

## 1.1 操作系统的定义

目前,对所有使用过计算机的人来说,操作系统已经是一个习以为常的概念。然而,要给操作系统下一个精确的定义却并非易事。几十年来,人们从不同角度对操作系统有过许多不同的解释,但至今仍无一个统一的定义。为了说明什么是操作系统,下面从诸多不同解释中选出了4种最有代表性的观点。

### ▶▶ 1.1.1 操作系统的虚拟机观点

这种观点认为,操作系统是添加在硬件上的第一层软件,是对硬件功能的首次扩充和直接延伸。

众所周知,现代计算机系统由软件和硬件两大部分组成。软件部分又可大致分为系统软件和应用软件两大类。系统软件用来管理计算机本身及应用软件;应用软件用来完成用户所要求的实际任务。硬件部分是指未配置任何软件的计算机(简称裸机)。

如果计算机系统中没有配置操作系统,让用户直接使用裸机,那么就必须要对中央处理器、主存储器、时钟、终端、磁盘和其他输入/输出(I/O)设备的管理细节全部交给用户去完成,这将使程序设计和计算机应用变得非常复杂和困难。正是为了避免这种实现上的复杂性给程序设计和计算机应用所带来的极大麻烦,人们才不断在裸机上添加软件,延伸裸机功能,构造基于裸机但又比裸机更易于用户理解和编程的虚拟计算机,简称虚拟机(virtual machine)。添加在裸机上的软件就是操作系统。可见操作系统是对硬件系统功能的首次扩充。

除操作系统以外的其他系统软件,如命令解释程序、编译程序、编辑程序等,都位于操作系统之上。尽管它们也都属于系统软件,但与操作系统不同。操作系统是在核心态(kernel mode)或称管态(supervisor mode)下运行的一部分软件,硬件保护它免受用户程序的破坏。而解释程序、编译程序、编辑程序等则都在用户态(user mode)下运行,它们不受硬件的保护。例如,用户可以自主修改一个编译程序,但无权自编一个磁盘中断处理程序。因为磁盘中断处理程序是操作系统的一部分,是受硬件保护的,硬件阻止用户对它进行修改。至于应用软件,则处在所有系统软件之上。计算机系统的这种抽象层次结构如图1-1所示。

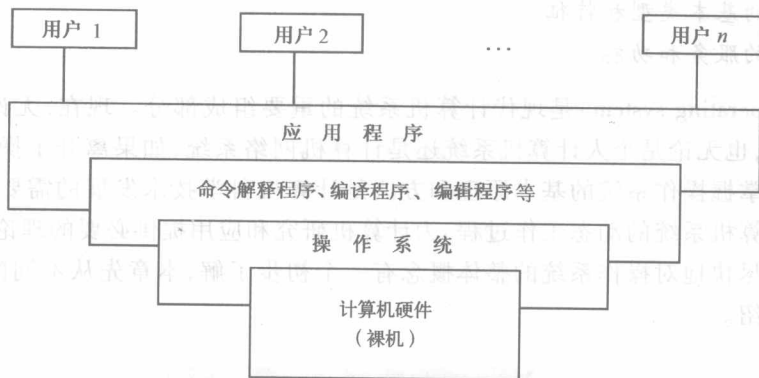


图 1-1 计算机系统的抽象层次结构

按照这种观点,其他系统软件和应用软件全部都建立在操作系统的基础之上,操作系统实际上是向它们隐藏了硬件和硬件操作细节的真相,把它们对裸机的使用转换成了对虚拟机的使用。

### ▶▶ 1.1.2 操作系统的资源管理观点

这种观点认为,操作系统是管理计算机系统资源的软件,负责控制和管理计算机系统的全部资源,确保这些资源能被高效合理地使用,确保系统能够有条不紊地运行。

从资源管理的角度看,计算机系统的资源大致可分为硬件资源和软件资源两大类。硬件资源是计算机硬件系统的总和,包括中央处理器(CPU)、内存储器、辅助存储器、I/O设备等。软件资源是系统中各种程序和数据总和,这些程序和数据均以文件的形式保存在计算机系统中。从资源管理的观点出发,根据操作系统所管理的资源类型,可把整个操作系统分为处理机管理、存储器管理、设备管理、文件管理和用户接口5大部分。其中,处理机管理负责CPU的运行和分

配;存储器管理负责内存存储器的分配、回收、保护与扩充;设备管理负责 I/O 设备的分配、回收与控制;文件管理负责文件存储空间和文件信息的管理;用户接口提供操作系统与用户之间的交互手段。这些管理共同构成了操作系统的资源管理功能。

### ▶▶ 1.1.3 操作系统的用户服务观点

这种观点认为,操作系统是一个为用户服务的大型程序。从用户服务的角度看,当计算机配置了操作系统之后,用户就不再直接使用计算机硬件,而是利用操作系统所提供的命令和服务去使用计算机。或者说,用户在操作系统的帮助下能方便、快捷、安全、可靠地操纵计算机硬件以及运行自己的程序。这就把操作系统看成了用户与计算机之间的一个接口。这一接口为用户提供了两种使用操作系统的方式:命令方式和系统调用方式。所谓命令方式,是指操作系统提供了一组联机命令,用户可通过键盘或鼠标使用这些命令,并直接操纵计算机系统。所谓系统调用方式,是指操作系统提供了一组系统调用,用户可在自己的程序中通过调用相应的系统调用来操纵计算机。

### ▶▶ 1.1.4 操作系统的进程管理观点

上述几种观点实际上都是一种静态观点,这些观点均没有揭示程序在系统中运行的本质过程和相互联系。实际上,操作系统调用当前程序运行是一个动态过程,并具有并发性。所谓并发可简单地理解为,在同一个时间段内有多个程序或多个程序段在同时执行。这种并发性能够提高系统资源的利用率,但却可能导致不同程序或程序段之间的相互制约。为了更好地描述程序的动态过程和对系统资源的有效管理,从而引入了进程的概念。

所谓进程,可简单地理解为并发程序的执行。用进程观点来研究操作系统就是要从并发程序执行的过程来讨论操作系统。这种观点能使人们更清楚地认识系统中的各种资源管理程序是如何动态实现其管理功能的。

综合上述各种观点,可以给出操作系统的解释:操作系统是控制和管理计算机硬件和软件资源、合理组织计算机工作流程以及方便用户的程序集合。

## 1.2 操作系统的形成过程

操作系统同其他任何事物一样,也有一个产生与发展的过程。为了帮助读者理解操作系统,下面先对操作系统的形成过程做一个简单的回顾。

### ▶▶ 1.2.1 人工操作阶段

在 20 世纪 50 年代以前,第一代计算机中是没有配置操作系统的,计算机只是由控制台控制的一个庞大的物理机器。当时,人们使用计算机的过程可大致描述为:首先由程序员将其编好的程序从纸带或卡片机上装入内存,然后通过控制台上的按钮或开关启动程序执行,最后当程序运行完毕时,取下纸带和运算结果,开始下一个用户程序,依次重复上述过程。这种人工操作方式

存在以下3个主要问题。

(1) 资源独占。当一个用户开始操作后,计算机中的全部资源都归该用户所有,直到程序运行结束或出现错误时才会把资源转让给下一个用户。

(2) 串行工作。人的手动操作与计算机的运行以及 CPU 与 I/O 设备之间都是按串行方式工作的。

(3) 人工干预。计算机是在人的直接联机干预下进行工作的。

可见,人工操作方式严重降低了计算机资源的利用率,形成了所谓的人工操作方式与机器利用率的矛盾,简称人机矛盾。对于早期的计算机来说,这种矛盾尚不突出,原因是那时的计算机本身拥有的资源数量较少,计算速度也较慢,一般程序所需的计算时间要比其他操作时间(如装卸带、卡、I/O 等)长。但是,随着 CPU 速度的提高、存储容量的增大和外部设备的增多,人机矛盾变得越来越尖锐,使得这种方式到了非改不可的地步。

### ►► 1.2.2 单道批处理阶段

为了解决上述问题,人们自然首先想到的是如何摆脱从一个用户程序过渡到另一个用户程序时的人工干预,使其转换能自动进行。这就产生了由计算机对一批用户程序进行自动处理的所谓批处理技术。在批处理方式中,一个用户程序及其所需要的数据和操作命令的总和被称为一个作业。批处理技术出现于20世纪50年代末期的第二代计算机中,它又可以分为早期批处理、脱机批处理和执行系统等阶段。

#### 1. 早期批处理

早期批处理方式是把若干个用户作业集中起来组成一批作业,并在内存中放置一个监督程序,由监督程序来负责实现对这批作业的处理和从一个作业过渡到另外一个作业的自动转换。

在这种处理方式下,先由操作员把一批用户作业的卡片叠放到读卡机上,然后由监督程序开始对该批作业进行自动处理。监督程序的处理过程是:先把读卡机上的这批作业全部输入到磁带上,然后按照某种策略从该批作业中选择一个作业调入内存,对其进行汇编或编译,并把汇编或编译结果装入内存执行,运行结束后输出其计算结果。当第一个作业全部完成后,监督程序会自动选择下一个作业运行。重复上述过程,直到该批作业全部完成为止。这样,在监督程序处理第一批作业的同时,操作人员可以将第二批作业的卡片叠放到读卡机上。当监督程序处理完第一批作业后,便可以自动地从读卡机上输入和处理第二批作业。这样,监督程序就可以不停地对一批批、一个个作业进行处理,从而实现作业之间的自动转换,解决了人工操作阶段所存在的人工干预问题和人工操作与计算机之间的串行工作问题。

虽然这种处理方式提高了系统的处理能力,但作业的 I/O 和 CPU 的计算仍然是串行的。也就是说,作业信息由卡片送到磁带,再由磁带调入内存,以及计算结果在打印机上输出等,都是由 CPU 来处理的。这种 CPU 和 I/O 设备之间的串行工作方式大大降低了程序的执行速度。

#### 2. 脱机批处理

为解决早期批处理所存在的问题,人们在批处理中引入了脱机 I/O 技术,形成了脱机批处理

方式。这种处理方式的处理过程如图 1-2 所示,它在早期批处理的基础上增加了一台功能较差的处理机。原来的那台处理机称为主处理机(简称主机),专门负责计算工作,不直接与 I/O 设备打交道。新增加的这台处理机称为卫星机,专门负责 I/O 工作。

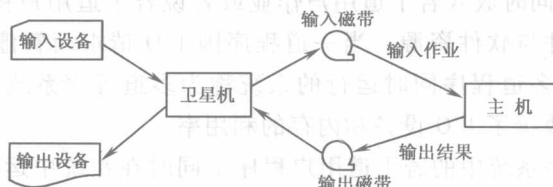


图 1-2 脱机批处理

在图 1-2 中,卫星机负责把读卡机上的作业输入到输入磁带上和把输出磁带上的计算结果从输出设备上输出;主机直接从输入磁带上调入作业执行,并把计算结果送到输出磁带上。这样,作业的输入和输出工作完全与主机脱离,并且,卫星机的工作与主机的工作是同时进行的。因此,这种批处理方式解决了 CPU 与 I/O 设备之间的串行工作问题。

虽然批处理方式提高了系统的处理能力,但却带来了必须解决的保护问题。例如,监督程序、系统程序和用户程序之间是通过相互调用的方式来进行转移的,这样,当目标程序企图执行一条非法指令时,整个系统就会停顿下来。另外,若程序陷入死循环,则整个系统也无法进行下去。更为严重的是它无法防止用户程序破坏监督程序的问题,存在系统被破坏的潜在危险。

### 3. 执行系统阶段

从 20 世纪 50 年代末到 60 年代初,硬件方面获得了两个重要进展:一是通道的引入,二是中断技术的出现。所谓通道,实质上是一个功能单一、结构简单的 I/O 处理机,它独立于 CPU,并直接控制外部设备与内存进行数据传输,可代替上述卫星机的工作。所谓中断,是作为外部设备向 CPU 的“汇报”手段提出来的,即在 I/O 结束或在硬件发生某种故障时,由相应硬件向 CPU 发出一个信号,使 CPU 停止正在执行的操作,而转去执行为处理该信号而设置的程序,中断处理完毕后 CPU 再回到原来的断点继续执行。

通道和中断技术的出现使硬件具有了较强的并行工作能力。为了获得 CPU 和外部 I/O 设备在执行时间上的重叠,必须对原来的监督程序进行扩充,提供中断处理程序和 I/O 控制程序,这样就把原来的监督程序扩大到了执行系统。可见,执行系统的程序可包括三大类:I/O 控制程序、中断处理程序和管理程序。

这时,执行系统的程序比较庞大,若将其全部放到内存,则会大大减少用户程序的可用空间。为此,可仅让那些所有程序都要用到的中断处理程序和 I/O 控制程序常驻内存,而将其他部分放在外存。常驻内存的那部分程序称为执行程序。

执行系统虽然较好地解决了 CPU 和 I/O 操作的并行问题,但内存中仍仅能存放一个程序,并且当这个程序因等待 I/O 而不能继续执行时,CPU 必须处于空闲状态。

### ▶▶ 1.2.3 多道程序系统阶段

为解决执行系统所存在的问题并提高 CPU 的利用率,又引入了多道程序技术。多道程序技术的主要思想是在内存中同时放入若干道用户作业或者说若干道用户程序,并允许它们交替执行,共享系统中的各种硬件与软件资源。当一道程序因 I/O 请求而暂停执行时,CPU 便转去执行另一道程序。这种允许多道程序同时运行的系统称为多道程序系统。多道程序系统不仅使 CPU 得到了充分利用,还改善了 I/O 设备和内存的利用率。

从宏观上看,多道程序系统中的若干道用户程序是同时在系统中运行的,而从微观上看,这些程序则是在轮流使用 CPU,只是由于 CPU 的运行速度很快,人们感觉不到而已。多道程序的出现标志着操作系统的形成,最早出现的多道程序系统是多道批处理系统,之后又出现了分时系统、实时系统等。

虽然多道程序系统有效地提高了系统资源利用率,但实现多道程序则需要妥善解决下述一系列问题。

(1) 内存的分配和保护。在内存空间中同时驻留了多道程序,应为每道程序分配自己的内存空间,使它们既不因相互重叠而丢失信息,又不因某道程序出现异常而破坏其他程序。

(2) 处理机的管理和分配。多道程序共同使用一个处理机,必将引起各道程序对处理机的争夺,系统要协调它们之间的关系,既能使那些紧急的程序优先获得处理机,又能使各道程序都有得到处理机的机会。

(3) I/O 设备的管理和分配。计算机系统中的 I/O 设备数量一般都少于多道程序所需要的设备总量,必将引起各道程序对 I/O 设备的争夺。对此,系统应该能够进行协调,并能为各道程序分配相应的 I/O 设备。

(4) 文件存储空间的组织和管理。通常,在辅助存储器上以文件形式存放着大量的有用信息,为提高文件存储空间的利用率,加速对信息的检索速度,系统应对它们进行组织和管理。同时,为方便用户使用文件,系统还应该提供存储和检索文件信息的手段。

在多道系统中,解决上述问题的一组程序的集合构成了操作系统。

### ▶▶ 1.2.4 进一步发展阶段

在操作系统形成之后,随着计算机、网络等技术的快速发展以及各种应用的普及和深入,操作系统也在进一步发展,并呈现了一种多元化的局面,出现了微型计算机操作系统、网络操作系统、分布式操作系统及嵌入式操作系统等新的操作系统类型。

#### 1. 微型计算机操作系统

微型计算机操作系统是指配置在微型计算机上的操作系统。目前,比较流行的微型计算机操作系统有单用户多任务和多用户多任务两种类型。所谓单用户多任务微型计算机操作系统,是指只允许一个用户上机但允许把程序分为若干个任务并发执行的操作系统,例如 Microsoft 的 Windows 系列个人用户版操作系统;所谓多用户多任务微型计算机操作系统,是指允许多个用户通

过各自的终端使用同一台计算机,共享系统中的各种资源,而每个用户程序又可进一步分为多个任务并发执行的操作系统,例如 SUN 公司的 Solaris 系列操作系统和源代码公开的 Linux 操作系统等。

## 2. 网络操作系统

网络操作系统是配置在计算机网络系统上的一种操作系统。在网络系统中,各个计算机的硬件特性可能不同,数据表示格式及其他方面的要求也可能不同,为能进行有效通信,它们之间必须建立相应的约定,这些约定被称为协议或规程。从这种意义上讲,网络操作系统是使网络上各计算机能方便而有效地共享网络资源,并为网络用户提供所需的各种服务的软件和有关规程的集合。网络操作系统除了应具有通用操作系统所具有的处理机管理、存储管理、设备管理和文件管理外,还应具有高效可靠的网络通信能力和各种有效的网络服务功能。

## 3. 分布式操作系统

分布式操作系统是配置在分布式计算机系统上的一种操作系统。分布式系统是在 20 世纪 80 年代发展起来的,由多台计算机组成的一种特殊的计算机网络。在硬件连接、拓扑结构和通信控制等方面与计算机网络基本相同,且都具有数据通信和资源共享的功能。分布式系统与计算机网络的主要区别表现在操作系统方面;而分布式操作系统和网络操作系统的最主要区别是系统结构对用户的透明性。在网络操作系统中,用户通过网络服务器的名称来获得资源,系统结构对用户是不透明的;在分布式操作系统中,用户只需通过服务的名称即可获得资源,系统结构对用户是透明的。

## 4. 嵌入式操作系统

嵌入式操作系统是一种支持嵌入式应用的操作系统,是嵌入式系统极为重要的组成部分,通常包括与硬件相关的底层驱动软件、系统内核、设备驱动接口、通信协议、图形界面、标准化浏览器等。嵌入式操作系统大多用于机电设备、仪器等专用控制方面,并具有十分广泛的应用和发展前景。

# 1.3 操作系统的基本类型

操作系统的类型可以有多种不同的分类方法,最常用的方法是按照操作系统的用户服务方式,将其分为多道批处理系统、分时系统和实时系统 3 种基本类型。

## ▶▶ 1.3.1 多道批处理系统

多道批处理系统实际上是把批处理技术和多道程序技术相结合的产物,出现于 20 世纪 60 年代初期,目前仍在使用的。

### 1. 多道批处理系统的运行方式

多道批处理系统的运行方式如图 1-3 所示。操作员把用户提交的作业卡片放到读卡机上,系统通过 SPOOLing(含意是外部设备同时联机操作,将在第 7 章讨论)输入程序及时地把这些作业送入外存的磁盘输入井,形成后备作业队列。作业调度程序根据系统的当前情况和后备作业



的特点,按照一定的调度原则选择一个或几个搭配合理的作业,装入内存准备运行。所谓搭配合理,主要是指作业的选择应既有利于提高资源利用率,又能满足不同作业用户在等待时间方面的要求。内存中的多个作业交替执行。当某个作业完成时,系统把该作业的计算结果交给 SPOOLing 输出程序准备输出,并回收该作业的全部资源。重复上述过程,使诸作业一个接一个地流入系统,经过处理后又一个接一个地退出系统,形成一个源源不断的作业流。

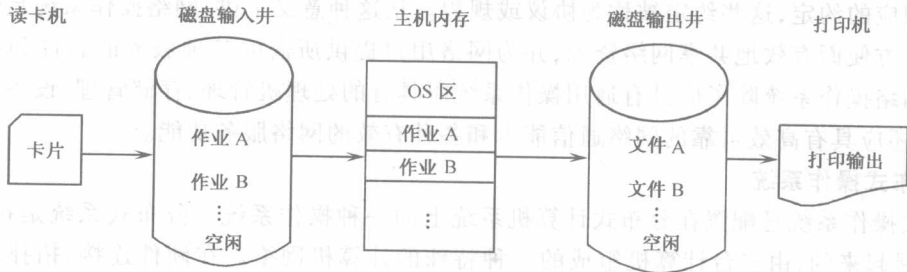


图 1-3 多道批处理系统的工作方式

## 2. 多道批处理系统的特征

多道批处理系统具有如下特征。

(1) 多道性。在内存中可以同时驻留多道程序,并允许它们同时执行,从而有效地提高系统的资源利用率和系统吞吐量。

(2) 无序性。多个作业完成的先后顺序与它们进入内存的先后顺序无关,即先进入内存的作业可能较后甚至最后完成,而后进入内存的作业则有可能先完成。

(3) 调度性。作业从提交给系统开始直至完成,需要经过以下两次调度。

① 作业调度。这是指按一定的作业调度算法从外存的后备作业队列中选择若干个作业调入内存。至于作业调度的详细讨论,将放在第 4 章。

② 进程调度。这是指按一定的进程调度算法从已在内存的作业中选择一个,将处理机分给它,使之执行。至于进程调度,也将在第 4 章讨论。

## 3. 多道批处理系统的优缺点

与单道批处理系统相比,多道批处理系统的主要优点如下。

(1) 资源利用率高。由于在内存中装入了多道程序,它们共享资源,使得系统中的各种资源得到了更充分的利用。

(2) 系统吞吐量大。系统吞吐量是指系统在单位时间内完成的总工作量。能提高系统吞吐量的主要原因有两个方面:第一,CPU 和其他资源保持忙碌状态;第二,仅当作业完成或无法运行时才进行切换,系统开销小。

多道批处理系统的主要缺点如下。

(1) 平均周转时间长。作业周转时间是指从作业进入系统开始,直至完成并退出系统为止所经历的时间。平均周转时间是指一批作业中所有作业周转时间的平均值。在多道批处理系统