

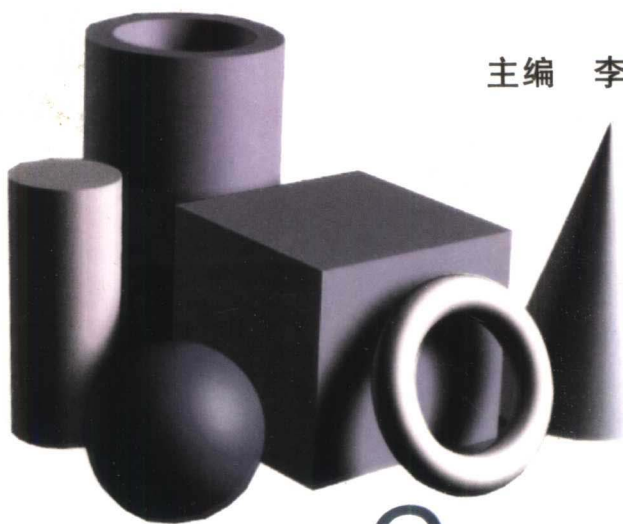
◎ 中等职业学校计算机基础教育课程规划教材

根据国家新教学大纲编写

计算机操作系统

专家审订 紧扣大纲 精解重练

主编 李自力 李建华



Computer Operating System



电子科技大学出版社

DIANZIKEJIDAXUECHUBANSHE

◎ 中等职业学校计算机基础教育课程规划教材

根据国家新教学大纲编写

计算机操作系统

Computer
Operating
System

江苏工业学院图书馆

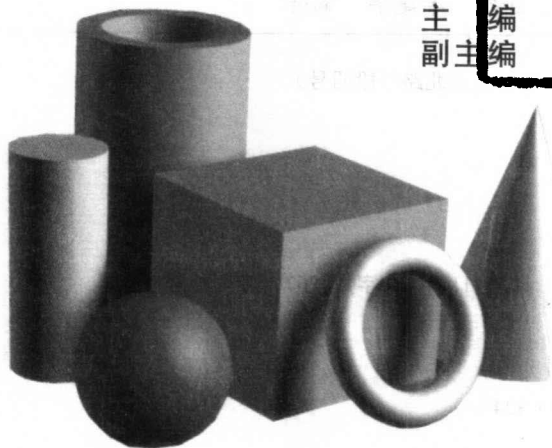
藏书章

主 编
副主 编

李 将
将 文

李 薛
博 飞

华 章



电子科技大学出版社

DIANZIKEJIDAXUECHUBANSHE

图书在版编目 (CIP) 数据

计算机操作系统/李建华编著. —成都: 电子科技大学出版社, 2004.7

ISBN 7-81094-537-8

I. 计… II. 李… III. 操作系统 IV. TP316

中国版本图书馆 CIP 数据核字 (2004) 第 061062 号

内 容 简 介

本书是根据教育部职业教育与成人教育司、教育部职业技术教育中心研究所颁布的“中等职业学校计算机及应用专业教育指导方案—计算机操作系统教学基本要求”编写的适合于中职、中专的教材。

操作系统是计算机最重要的系统软件之一。“操作系统”课程也是计算机专业非常重要的基础课程。

本教材系统地介绍了操作系统的基本结构和基本工作原理。主要内容如下: 操作系统概述、作业管理、处理机管理、进程管理、存储管理、文件管理、输入输出管理等内容。

本书严格按照大纲要求, 充分考虑到中职中专学生的认知水平, 内容安排上由浅入深, 循序渐进, 并结合大量实例, 简明扼要, 通俗易懂, 是计算机操作系统的入门教材。

计 算 机 操 作 系 统

李自力 李建华 编著

出 版: 电子科技大学出版社 (成都建设北路二段四号)

策划编辑: 付金权

责任编辑: 谢应成

发 行: 新华书店经销

印 刷: 四川南方印务有限公司

开 本: 787×1092 1/16 印张 11 字数 267 千字

版 次: 2005 年 3 月第二版

印 次: 2005 年 3 月第一次印刷

书 号: ISBN 7-81094-537-8/TP·324

定 价: 15.00 元

■ 版权所有 侵权必究 ■

◆ 邮购本书请与本社发行科联系。电话: (028) 83201635 邮编: 610054

◆ 本书如有缺页、破损、装订错误, 请寄回印刷厂调换。电话: (0833) 7641415



前 言

本书是根据教育部职业教育与成人教育司、教育部职业教育中心研究所颁布的“中等职业学校计算机及应用专业教育指导方案—计算机操作系统要求”编写的适合于中职、中专的教材。其中详细介绍了计算机操作系统的概念、基本原理和典型实现技术,着重讲述了操作系统的基本构成和基本工作原理。介绍了计算机操作系统概述、作业管理、处理机管理、进程管理、储存管理、文件管理、输入输出管理等内容。

本书结合作者多年从事教学与科研的实践经验,根据计算机操作系统课程教学的特点编写而成,概念准确,层次清晰,内容丰富,取材新颖,剪统性强,注重理论与实践的结合,并着力反映了操作系统发展过程中的新成果、新方法、新技术和新进展,全面展示了现代操作系统的基本设计原理与典型实现技术,有助于读者深入了解现代操作系统的全貌,为今后进行更深层次的大型软件系统的设计、研制与开发打下坚实的基础。

本书可作为在中等院校计算机专业学生的教材,也可供从事信息科学和计算机相关工作的工程技术人员参考。

由于作者水平有限,疏漏之处敬请各位老师和同学指正。

编 者

2005年2月

目 录

第1章 操作系统概述

1.1 操作系统的概念.....	1
1.1.1 计算机系统概述.....	1
1.1.2 什么是操作系统.....	4
1.2 操作系统的历史.....	4
1.2.1 手工操作阶段.....	5
1.2.2 早期批处理阶段.....	5
1.2.3 多道批处理系统.....	6
1.2.4 分时操作系统.....	8
1.2.5 实时系统.....	9
1.2.6 通用操作系统.....	9
1.2.7 操作系统的进一步发展.....	10
1.2.8 推动操作系统发展的动力.....	10
1.3 操作系统的基本类型.....	11
1.3.1 批处理操作系统.....	11
1.3.2 分时操作系统.....	11
1.3.3 实时操作系统.....	12
1.3.4 通用操作系统.....	12
1.3.5 个人计算机(PC机)操作系统.....	13
1.3.6 多处理器系统.....	13
1.3.7 网络操作系统.....	13
1.3.8 分布式操作系统.....	14
1.3.9 嵌入式操作系统.....	14
1.4 操作系统的功能.....	15
1.4.1 处理机管理.....	15
1.4.2 存储器管理.....	15
1.4.3 设备管理.....	16
1.4.4 文件管理.....	16
1.4.5 接口功能.....	17
1.5 现代操作系统的特征与基本概念.....	18
1.5.1 现代操作系统的特征.....	18
1.5.2 基本概念.....	19
1.6 现代主流操作系统简介.....	21

1.6.1	Windows	21
1.6.2	UNIX.....	21
1.6.3	Linux	22
1.7	研究操作系统的几种观点	22
1.7.1	计算机资源管理者的观点	23
1.7.2	用户界面的观点	23
1.7.3	进程管理的观点	23
1.7.4	虚拟机的观点 (Virtual Machine)	23
1.8	本章小结	24
	思考题	24

第 2 章 作业管理

2.1	计算机用户工作环境	25
2.1.1	用户工作环境的基本要求	25
2.1.2	建立操作系统	25
2.1.3	启动操作系统	26
2.2	用户与操作系统之间的接口	26
2.2.1	命令接口	26
2.2.2	程序接口	27
2.2.3	图形用户接口	28
2.3	作业状态与作业管理	28
2.3.1	基本概念	28
2.3.2	作业的状态	29
2.3.3	作业的管理	30
2.4	作业调度	31
2.4.1	作业调度程序及作业调度算法	31
2.4.2	算法衡量准则及常用调度算法	32
	思考题	34

第 3 章 进程管理

3.1	进程的概念	35
3.1.1	程序的顺序执行	35
3.1.2	程序的并发执行	36
3.1.3	进程的概念和特征	37
3.2	进程的状态和转换	38
3.2.1	进程的状态	38
3.2.2	进程的状态转换	39
3.3	进程的描述	40

3.3.1	进程控制块.....	40
3.3.2	进程控制块的组织.....	40
3.3.3	进程上下文.....	41
3.4	进程控制.....	42
3.4.1	进程创建.....	42
3.4.2	进程撤消.....	42
3.4.3	进程阻塞.....	43
3.4.4	进程唤醒.....	43
3.5	进程同步和互斥.....	44
3.5.1	基本概念.....	44
3.5.2	进程间互斥.....	45
3.5.3	信号量机制.....	48
3.5.4	进程间同步.....	51
3.6	进程通信.....	54
3.7	进程调度.....	56
3.7.1	进程调度概述.....	57
3.7.2	常用的进程调度算法.....	58
3.8	死锁.....	61
3.8.1	死锁及产生的原因和必要条件.....	61
3.8.2	死锁的处理.....	62
3.9	线程.....	63
	习题.....	64

第4章 存储处理

4.1	概述.....	68
4.1.1	多级存储器体系.....	68
4.1.2	存储管理的任务.....	69
4.1.3	地址重定位.....	70
4.2	连续存储空间管理.....	73
4.2.1	单一连续区分配.....	73
4.2.2	固定分区分配.....	74
4.2.3	可变分区分配.....	75
4.2.4	可重定位分区分配.....	77
4.3	分页存储管理.....	79
4.3.1	分页式存储管理.....	79
4.3.2	分页存储管理.....	80
4.3.3	两级页表和多级页表.....	83
4.3.4	存储保护.....	85
4.4	分段存储管理.....	85

4.4.1	分段式存储原理	85
4.4.2	分段式存储管理	86
4.4.3	段的共享和保护	87
4.4.4	分页和分段的区别	88
4.4.5	段页式存储管理	88
4.5	覆盖与交换技术	90
4.5.1	覆盖技术	90
4.5.2	交换技术	90
4.6	虚拟存储管理	91
4.6.1	虚拟存储器	91
4.6.2	请求分页式存储管理	93
4.6.3	请求分段式存储管理	97
	思考题	99

第5章 文件系统

5.1	文件及文件系统的概念	100
5.1.1	数据项和记录	101
5.1.2	文件	101
5.1.3	文件的类型	102
5.1.4	文件系统	102
5.1.5	文件系统的任务	103
5.2	文件结构	103
5.2.1	文件的逻辑结构	103
5.2.2	文件的存取方法	105
5.2.3	文件的物理结构	106
5.2.4	文件的存储设备	108
5.2.5	簇的概念	109
5.3	存储空间管理	110
5.3.1	空白文件目录	110
5.3.2	空白物理块链	111
5.3.3	位示图	112
5.4	文件目录管理	112
5.4.1	基本概念	113
5.4.2	目录结构	113
5.5	文件共享	115
5.5.1	文件共享概述	115
5.5.2	目录结构中的共享	116
5.6	文件存取控制	117
5.6.1	文件存取控制法	118

5.6.2 文件系统的可靠性和安全性	120
5.7 文件操作	121
5.7.1 CREATE 调用	122
5.7.2 DELETE 调用	122
5.7.3 OPEN 调用	123
5.7.4 CLOSE 调用	123
5.7.5 READ 调用	124
5.7.6 WRITE 调用	124
5.7.7 APPEND 调用	124
5.7.8 SEEK 调用	124
5.7.9 SET ATTRIBUTES 调用	124
5.7.10 RENAME 调用	125
5.8 文件系统的层次模型	125
5.8.1 层次结构的提出	125
5.8.2 Madnick 的文件系统层次模型	125
5.9 常见的文件系统	126
5.9.1 FAT16 和 FAT32 文件系统	127
5.9.2 NTFS 文件系统	127
5.9.3 其他文件系统	128
习题	129

第6章 设备管理

6.1 外部设备	133
6.1.1 外部设备的分类	133
6.1.2 设备管理的必要性	134
6.1.3 外部设备的接口	134
6.2 数据传输控制方式	135
6.2.1 程序查询方式	135
6.2.2 中断控制方式	136
6.2.3 DMA 控制方式	138
6.2.4 通道控制方式	139
6.2.5 缓冲技术	140
6.3 设备分配	140
6.3.1 基本的数据结构	140
6.3.2 设备分配方法	142
6.4 设备驱动	142
习题	142

第7章 Windows 操作系统的功能与原理

7.1 Windows 概述	144
7.1.1 Windows 的发展历程	144
7.1.2 Windows 的体系结构	145
7.1.3 API 的基本概念	147
7.1.4 DLL 的基本概念	147
7.2 Windows 处理机管理	148
7.2.1 进程的定义	148
7.2.2 进程的状态及其相互转换	148
7.2.3 进程的同步与互斥	149
7.2.4 进程的通信	150
7.2.5 线程的基本概念	150
7.2.6 处理机调度	151
7.2.7 Windows 虚拟机	151
7.2.8 虚拟机管理程序 VMM32	153
7.3 Windows 的存储管理	153
7.3.1 虚拟存储器	153
7.3.2 磁盘空间结构	154
7.3.3 Windows 的内存结构与管理	154
7.4 Windows 的文件管理	155
7.4.1 文件的共享、保护与保密	156
7.4.2 文件实现	156
7.4.3 FAT12、FAT16 及 FAT32 文件系统	157
7.4.4 CDFS 和 UDF	158
7.5 Windows 的设备管理	158
7.5.1 Windows 98 设备管理程序的结构	158
7.5.2 Windows 98 的设备驱动程序	158
7.5.3 虚拟设备驱动程序	159
7.5.4 即插即用技术的特点	159
7.5.5 Windows 其他管理功能	160
7.6 Windows 管理与维护	161
7.6.1 Winddows 98 的启动控制	161
7.6.2 Windows 98 的系统注册表的使用	162
7.6.3 操作环境的硬件设置与用户配置文件	163
7.6.4 Windows 的系统策略与故障诊断	163
习题	164

第1章 操作系统概述

【学习目标】

(1) 理解操作系统的基本概念。

(2) 理解操作系统基本功能。

(3) 理解操作系统的发展和分类。

(4) 了解现代操作系统的特征

(5) 了解操作系统的发展趋势。

电子计算机是 20 世纪人类最重大的科学技术发明之一。计算机发展到今天,从个人计算机到巨型计算机系统,毫无例外地都配置了一种或多种操作系统。很难想像用户直接使用一台没有操作系统的计算机来完成工作,那是非常困难的。那么,操作系统究竟是什么呢?它在计算机系统中的地位怎样?它具有什么样的功能?

本章作为操作系统的概述,将向读者简单介绍操作系统及其在计算机中的地位,了解操作系统的形成与发展过程,从总体上认识操作系统的功能和特征,区别多种操作系统类型,并简单介绍目前流行的三种操作系统。

1.1 操作系统的概念

1.1.1 计算机系统概述

要学习操作系统原理,首先应该了解计算机的结构,操作系统在计算机系统中的作用和地位。现代的计算机采用的是冯·诺依曼结构,计算机由硬件系统和软件系统两大部分组成。硬件构成了系统本身和用户作业的物质基础和工作环境,是软件运行的平台,而软件是对硬件功能的扩充,包括系统软件和应用软件。没有硬件,就失去了计算机系统的物质基础,软件也就无法存在;反之,若只有硬件而没有软件(这样的计算机称为裸机),那计算机是很难使用的,没有活力,也就没有多大的应用价值。只有将硬件和软件有机结合起来的计算机才具有强大的功能。因此,二者缺一不可。计算机系统结构如图 1-1 所示。



图 1-1 计算机系统结构

其中,最底层是计算机硬件,由处理机、内存存储器(简称内存)、输入/输出设备及系统

总线组成。

硬件之上是操作系统(Operating System, 简称 OS)。操作系统是计算机必须配置的最基本的系统软件, 是硬件的第一级扩充, 其主要功能之一就是隐藏底层硬件具体操作的复杂细节, 为用户使用计算机提供方便的工作环境。例如, 用户要保存文件时, 只需向计算机发出“保存(Save)”命令, 计算机就会执行这一操作, 而用户根本不用去考虑“磁头移动, 查找空闲磁盘块, 分配磁盘空间”等操作细节, 而这一切繁琐的操作细节全部由操作系统控制相应部件来完成。

操作系统之上是其他系统软件, 包括多种语言处理程序、连接装配程序、系统应用程序、多种工具软件等。

最外层是应用软件, 它们是为多种应用目的而编制的程序, 可以是购买的产品, 也可以是由用户自行开发的。例如, 文字处理、表格处理、工程计算、电子邮件收发、电子游戏等。

下面简要地考察一下计算机的硬件结构, 因为这些正是操作系统所管理的对象。

现代计算机系统由处理器(CPU)、内存和若干输入/输出设备组成, 由系统总线连接在一起, 实现彼此之间的通信。

1. 处理器

处理器是计算机的“大脑”, 它从内存中提取指令并分析执行它们。其工作周期是: 首先从内存中提取指令, 然后对指令译码——确定其类型和操作数, 最后执行指令。后续指令依此步骤进行。

每个处理器可以执行的指令集是专用的。由于访问内存、获得指令或数据字所花费的时间要比执行该指令的时间长得多, 因此, 所有的处理器都包含某些通用寄存器, 用来存放关键变量和中间结果。除此之外, 多数计算机都有几个专用寄存器, 其中之一是程序计数器, 它保存要提取的指令的内存地址。当这条指令被提取之后, 程序计数器就指向后面的一条指令。所以, 改变程序计数器的内容就改变了程序执行的流向。

另一个寄存器存放栈指针, 它指向当前内存栈的顶端。该栈中保存有关函数(过程)调用时的现场信息。程序执行过程中调用的每个函数在栈中占有一个帧面。

程序状态字 PSW 寄存器中的各个比特位代表着系统中当前的各种不同状态与信息。例如执行模式是否允许中断等。在系统调用和 I/O 中 PSW 起重要作用。

一般系统都提供核心态和用户态两种处理器执行状态, 其目的是为了保护操作系统程序(特别是其内核部分), 防止受到用户程序的损害。当执行操作系统程序时, 处理器处于核心态(也称为系统态、管理态), 它有较高特权, 可以执行所有的指令, 包括普通用户程序中不能使用的特权指令, 从而能对所有寄存器和内存进行访问, 启动 I/O 操作等。用户程序是在用户态下执行的, 它的权限较低, 只能执行指令集中非特权指令。

注意: 寄存器被广泛应用于计算机系统中, 它们与操作系统有着非常直接和密切的关系。事实上, 操作系统设计人员只有在完全掌握和了解硬件厂商所提供的各种寄存器的功能和接口之后, 才能进行操作系统设计。详细内容请参阅有关内容。

2. 存储器

计算机系统中, 存储器是最主要的组成部分之一。按照速度、容量和成本划分, 存储器构成一个层次结构, 如图 1-2 所示。

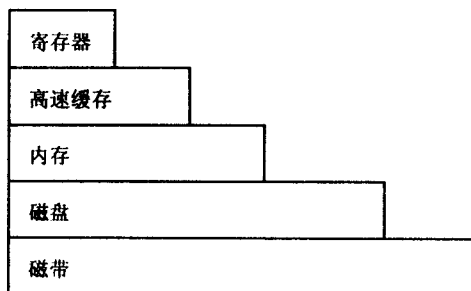


图 1-2 存储器层次结构

其中，顶层是 CPU 的内部寄存器，其速度与 CPU 一样快，所以存取时无延迟，它成本高，容量小。

然后是高速缓存(Cache)，大多由硬件控制，Cache 的速度很快，它们放在 CPU 内部或非常靠近 CPU 的地方。当程序需要读取具体信息时，Cache 硬件先查看它是否在其中，如在其中(称做“命中”)，就直接使用它；如果不在，就从内存中获取该信息，并把它放入 Cache 中，以备今后再次使用。但是 Cache 的成本很高，容量很小，一般小于 1MB。统计表明，仔细地选择 Cache 的大小和转换策略，可使命中率达到 80%~99%，从而大大提高存取性能。

高速缓存下面是内存(又称主存，也称做 RAM——随机存取存储器)，CPU 可以直接存取内存及寄存器和 Cache 中的信息，但 CPU 不能直接存取磁盘上的数据。所以，机器执行的指令及所用的数据必须预先存放在内存及寄存器和 Cache 中。然而，内存中存放的信息是易变的、暂时的，当关闭机器电源后，内存中的信息就全部丢失了。

再下一层是磁盘(一般指硬盘)，称做辅助存储器(简称辅存或外存)，它是对内存的扩展，其容量特别大，一般都达到几十 GB 以上，而且数据可以永久保存在磁盘上。磁盘只和内存交换数据。通过磁头的移动存取数据。由于是机械装置，所以，磁盘上数据的存取速度低于内存存取速度。

最下层是磁带。虽然它上面保存的数据更持久，容量更大，但它的存取速度很慢，而且不适宜进行随机存取。所以，磁带设备一般不用做辅存。它的主要用途是作为文件系统的后备，存放不常用的信息或用做系统间传送信息的介质。

除此之外，实际应用中的其他存储器还有：光盘，U 盘，软盘等。

3. I/O 设备

I/O 设备是人机交互的工具，它通常由设备控制器和设备本身两部分组成。

设备控制器协调和控制一台或多台 I/O 设备的操作，实现设备操作与整个系统操作的同步。设备控制器本身有一些缓冲区和一组专用寄存器，负责在外设和本地缓冲区之间移动数据。

设备本身的对外接口相当简单，实际上它们隐藏在控制器的后面。因而，操作系统总是和控制器打交道，而不是直接与设备作用。

I/O 设备的种类很多，因此设备控制器的类别也就很多，这就需要不同的软件——设备驱动程序来控制它们。不同操作系统上的不同控制器分别对应不同的设备驱动程序。

4. 总线

总线是连接计算机各部件的通信线路。计算机系统的总线有单总线和多总线之分。单总线是指 CPU、外设、存储器等都连接在一起的总线结构，而多总线则指把系统的 CPU 和内存分开连接，外设和外存等也用其他总线分开连接进行管理和数据传送的总线结构。显然，不同的总线结构对操作系统的设计和性能的影响也不同。

按照在总线上传送的信息所起的作用，系统总线基本上分为三种：

(1) 数据总线：这是计算机各部件之间传送数据的通道，其宽度随字长而定。如 32 位结构的数据总线应是 32 根。数据总线是双向总线，两个方向都能传送数据。

(2) 地址总线：该总线传送地址信号，它可以是存储器的地址，也可以是 I/O 设备控制器中控制寄存器或数据寄存器的地址。

(3) 控制总线：在该线上出现的信号是各个模块之间传送数据时所需的全部控制信息。

1.1.2 什么是操作系统

从上述介绍中我们不难发现，操作系统在计算机系统中的地位和作用，它是计算机硬件和其他软件以及计算机用户之间的桥梁、纽带。如果没有操作系统，用户使用计算机是相当困难的。那么，操作系统究竟是什么？对此至今也没有形成一个统一的标准化定义。因此，引入操作系统的目的可以从三方面来考察：

(1) 从系统管理人员的观点来看：引入操作系统是为了合理地组织计算机工作流程，管理和分配计算机系统硬件和软件资源，使之能为多个用户高效率地共享。

(2) 从用户的观点来看：引入操作系统是为了给用户使用计算机提供一个良好的界面，并且使用户无需了解许多有关硬件和系统软件的细节，就能方便灵活地使用计算机。

(3) 从发展的观点看：引入操作系统是为了给计算机系统的功能扩展提供支撑平台，使之在追加新的服务和功能时更加容易和不影响原有的服务与功能。

综上所述，我们可以这样来定义操作系统：

操作系统是计算机系统中的一个系统软件，它是这样一些程序的集合——它们管理和控制计算机系统硬件、软件资源；合理地组织计算机的工作流程；有效地利用这些资源为用户提供一个功能强大、使用方便并可扩展的工作环境；在计算机与用户之间起到接口的作用。

由此可见，操作系统在计算机系统中的地位是：操作系统是整个计算机系统的控制管理中心，其他所有的软件都是建立在操作系统之上的。操作系统对它们既有支配权力，同时又为它们的运行建造必备的环境。因此，在裸机之上每加了一层软件后，用户看到的就是一台功能更强的机器(通常把经过软件扩充功能后的机器称为“虚拟机”)。在裸机上安装操作系统后，就为其他软件 and 用户提供了工作环境(称为软件平台)。

1.2 操作系统的历史

为了更好地理解操作系统的基本概念、功能及特点，我们首先来回顾一下操作系统的

形成和发展的历史过程。

在计算机产生的初期,硬件技术处于起步阶段,此时操作系统并未形成,软件概念还不明确。随着硬件技术的发展,促进了软件概念的形成,也推动了操作系统的形成和发展。而软件的发展也促进了硬件的发展。

下面我们先来回顾一下计算机的发展过程。计算机的发展过程一般分为四个阶段:

1946年~二十世纪50年代末:第一代,电子管时代,无操作系统。

二十世纪50年代末~60年代中期:第二代,晶体管时代,批处理系统。

二十世纪60年代中期~70年代中期:第三代,集成电路时代,多道程序设计。

二十世纪70年代中期至今:第四代,大规模、超大规模集成电路时代,分时系统。

现代计算机正向着巨型化、微型化、并行、分布、网络化和智能化几个方向发展着。

与上述计算机的发展过程相适应,操作系统经历了:手工操作阶段(无操作系统)、早期批处理阶段、多道批处理系统、分时系统、实时系统、通用操作系统、网络操作系统、分布式操作系统等发展过程。

1.2.1 手工操作阶段

在第一代计算机时期,构成计算机的主要元器件是电子管,这种计算机非常庞大,计算机运算速度很慢,价格昂贵。没有操作系统,然而,它具有程序的概念。用户用机器语言(二进制代码)编制程序,并在上机时独占全部计算机资源。用户既是程序员,又是操作员。上机完全是手工操作:先把程序纸带(或卡片)装上输入机,然后启动输入机把程序和数据送入计算机,接着通过控制台开关或接插板启动程序运行。没有显示设备,由氖灯或数码管显示或打印机输出计算结果,用户取走打印结果并卸下纸带(或卡片)。第二个用户程序上机,如此循环办理。这种由一道程序独占机器且手工操作的情况,在计算机速度较慢时是允许的,因为此时计算机所需时间相对较长,手工操作所占比例还不很大。可见,手工操作时期,没有操作系统,计算机资源利用率很低,而且计算机的使用非常不方便。

50年代后期,计算机的运行速度有了很大提高,从每秒几千次几万次发展到几十万次上百万次。这时,手工操作的慢速度与计算机的高速度之间形成矛盾,手工操作与机器有效运行时间之比将大大地加大,这种矛盾到了不能容忍的地步。唯一的解决办法是摆脱人们的手工操作,实现作业的自动过渡。这样就出现了批处理。

1.2.2 早期批处理阶段

如上所述,在计算机发展的第一阶段用户上机时需要自己建立和运行作业,并做结束处理。由于没有任何用于管理的软件,所有的运行管理和具体操作都由用户自己承担。每个作业都由许多作业步组成,任何一步的错误操作都可能导致该作业从头开始。在当时,计算机的价格是极其昂贵的,CPU的时间是非常宝贵的,尽可能地提高CPU的利用率成为十分迫切的任务。

为了解决人工干预的问题,必须缩短建立作业(即用户的一个计算任务)和人工操作的时间。人们首先提出从一个作业转到下一个作业的自动转换方式,采用批处理的方法,操作员把用户提交的作业分类,把一批中的作业编成一个作业执行序列。每一批作业将有专

门编制的批处理程序(又称为监督程序或管理程序)自动依次运行。这个监督程序是最早的操作系统雏形。

早期的批处理分为联机批处理与脱机批处理两种类型。

(1) 联机批处理: 慢速的 I/O 设备与主机直接相连。其基本操作步骤为:

操作员有选择地把若干作业合为一批, 由监督程序先把它们存入到磁带上; 之后在监督程序的控制下使这批作业一个接一个地连续执行, 直到该批作业全部完成, 再把下一批作业输入到磁带上, 如此往复。联机批处理系统中, 作业处理是成批进行的, 并且在内存中总是只保留一道作业(所以叫单道批处理); 作业的输入、调入内存以及结果输出都在 CPU 的直接控制下进行。

虽然这种联机批处理方式解决了作业的自动转接, 减少了作业建立和人工操作时间, 但是在作业的输入和执行结果的输出过程中, CPU 仍处在停止等待状态, 这样慢速的输入输出设备和快速主机之间仍处于串行工作, 仍不能很好地利用系统资源。

(2) 脱机批处理: 这种方式的明显特征是增加一台不与主机直接相连而专门用于与外部设备打交道的卫星机(又称外围计算机), 如图 1-3 所示。

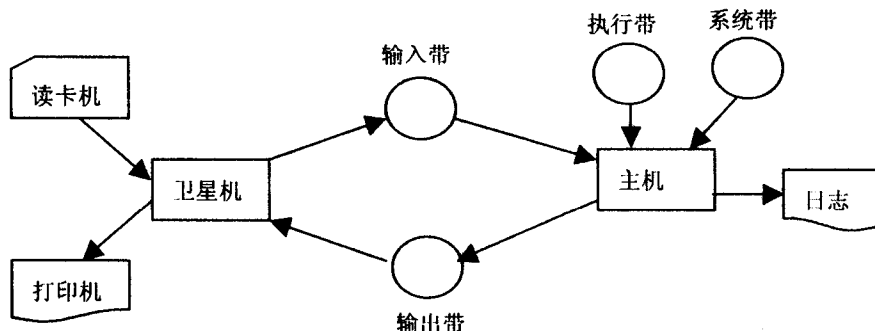


图 1-3 早期脱机批处理模型

脱机批处理的工作过程是: 卫星机把读卡机上的作业逐个送到输入磁带上; 主机只负责把作业从磁带上调入内存并运行, 作业完成后主机把计算结果记录到输出磁带上; 卫星机负责把输出磁带上的信息读出来, 交给打印机打印。

由此可见, 卫星机的工作是专门负责输入和输出, 主机则完成快速的计算任务, 二者可以并行操作。由于 I/O 设备不受主机直接控制, 所以又称为“脱机”处理。

早期的批处理系统是在解决人机矛盾和 CPU 与 I/O 设备速率不匹配的矛盾的过程中发展起来的。它的出现也促进了软件的发展, 出现了监督程序、汇编程序、编译程序、装配程序等。

1.2.3 多道批处理系统

在单道批处理系统中, 任意时刻只允许一道作业在内存中运行, 称为单道运行。图 1-4 给出了单道程序工作示例。由于作业运行期间可能需要一段时间计算, 一段时间进行输入/输出操作, 而相对于处理机的计算速度, 输入/输出的速度是非常慢的。这就意味着在一个作业的运行过程中, 处理机可能会长时间空闲。为了提高系统资源利用率和系统吞吐量,

在 20 世纪 60 年代中期出现了多道程序设计技术, 图 1-5 给出了多道程序工作示例。在单处理机系统中, 多道程序运行的特点是:

- 多道: 计算机内存中同时存放几道相互独立的程序。
- 宏观上并行: 同时进入系统的几道程序都处于运行过程中。
- 微观上串行: 实际上, 各道程序轮流使用 CPU, 交替执行。

在批处理系统中采用多道程序设计技术, 就形成了多道批处理系统。用户提交的许多作业首先存放在外存储器中, 形成作业队列, 等待运行。当需要调入作业时, 由操作系统中的作业调试程序根据其对资源的要求和一定的调度原则, 从该队列中一次选取一个或几个作业装入内存, 让它们交替运行。当某个作业运行完毕, 然后再装入一个或几个作业。这种处理方式, 在内存中总是同时存在几道程序, 系统资源得到比较充分的利用。如图 1-5 所示, 3 个程序同时驻留内存, 被处理器交替执行的过程示例。从每个程序来看, 程序获得处理器, 运行一段时间, 暂停并等待处理器, 然后又获得处理器运行, 然后再暂停并等待处理器, 如此反复, 直到结束, 退出系统。从处理器角度来看, 它依次执行程序 A 一段时间, 切换到程序 B 执行一段时间, 再切换到程序 C 执行一段时间, 再在程序 A、B、C 之间切换执行, 直到它们执行完毕, 退出系统。有时可能内存中的程序都未准备好执行, 或者它们都在进行输入/输出数据, 而使处理器等待。但与运行单个程序相比, 多道程序设计技术大大提高了处理器的利用率和系统吞吐量。这种为了提高系统吞吐量和资源利用率, 允许多个程序同时驻留内存, 使处理器在这些程序之间切换, 在一段时间内, 执行完多个程序的处理技术称为多道程序设计技术(multiprogramming)。现代操作系统大多都采用了多道程序设计技术。

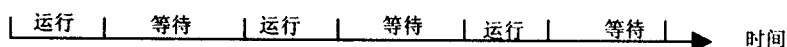


图 1-4 单道程序工作示例

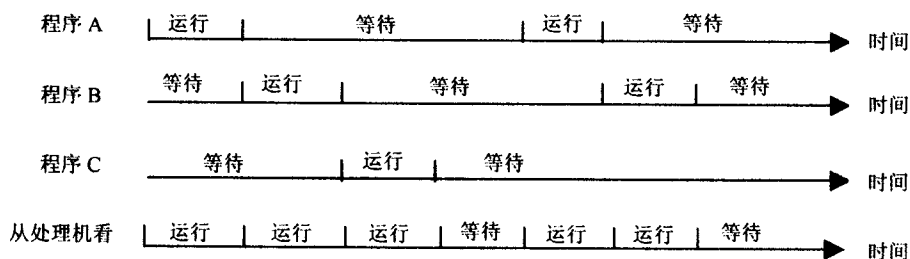


图 1-5 多道程序工作示例

多道批处理系统中, 要解决这样一些技术问题:

(1) 并行运行的程序要共享计算机系统的硬件和软件资源, 既有对资源的竞争, 又须相互同步。因此同步与互斥机制成为操作系统设计中的重要问题。

(2) 随着多道程序的增加, 出现内存不够用的问题, 提高内存的使用效率也成为关键。因此出现了诸如覆盖技术、对换技术、虚拟存储技术等内存管理技术。

(3) 由于多道程序存在于内存, 为了保证系统程序存储区和各用户程序存储区的安全可靠, 提出了内存保护的要求。