



计算机应用与软件专业领域应用型人才培养系列教材

操作系统

(第2版)

徐宗元 主编



高等教育出版社
HIGHER EDUCATION PRESS

计算机应用与软件专业领域应用型人才培养系列教材

操作系统

(第2版)

徐宗元 主编

高等教育出版社

内容提要

本书是“计算机应用与软件专业领域应用型人才培养系列”教材之一,本书遵照 Operating Systems CC2001 Report 的要求,以 Windows2000 和 UNLX/Linux 为例介绍操作系统的原理。真正与考证接轨,教材内容涵盖了计算机软件专业技术资格和水平考试中有关操作系统的内容,并且在习题中选择了历年的相关考试试题。本书为了强调技术、突出实践,每章在相关理论内容述说后,均安排有相关实验。本书配套完备的教学解决方案,免费为任课老师提供教学指导光盘、习题答案、实验源程序。

本书适用于高等院校应用型人才培养,也可供高等职业教育计算机以及广大计算机爱好者学习使用。

图书在版编目(CIP)数据

操作系统/徐宗元主编. —2 版. —北京:高等教育出版社, 2005.3

(计算机应用与软件技术专业领域应用型人才培养系列教材)

ISBN 7-04-016314-4

I. 操… II. 徐… III. 操作系统—教材
IV. TP316

中国版本图书馆 CIP 数据核字(2005)第 019118 号

责任编辑 司马镭 封面设计 吴 昊 责任印制 蔡敏燕

书 名 操作系统(第 2 版)
主 编 徐宗元

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 100011
总 机 010-82028899
传 真 021-56965341

购书热线 010-64054588
021-56964871
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
<http://www.hepsh.com>

排版校对 南京理工出版信息技术有限公司
印 刷 江苏如皋市印刷有限公司

开 本 787×1092 1/16
印 张 17
字 数 414 000

版 次 2000 年 7 月第 1 版
2005 年 5 月第 2 版
印 次 2005 年 5 月第 1 次
定 价 24.00 元

凡购买高等教育出版社图书,如有缺页、倒页、脱页等质量问题,请在所购图书销售部门联系调换。

版权所有 侵权必究

前 言

操作系统是计算机系统最重要的系统软件,作为一门课程,它是计算机及有关专业的必修课。为了适应高等职业教育的发展,跟上计算机技术不断更新的潮流,我们在多年教学和应用实践的基础上,参阅了许多有关操作系统的著作,并针对高等职业教育培养技术型人才的特点编写了本教材。

本教材的特点如下:

(1) 为了强调技术、突出实践,在介绍操作系统原理的基础上,各章中穿插了许多实验的内容,以便读者能够通过上机操作加深对概念的理解。

(2) 为了反映当代操作系统发展的最新技术,在介绍操作系统原理时,以当前流行的、有发展潜力的微机操作系统 Windows 98/NT 为实例,使得教学内容能跟上技术发展和应用的潮流。

(3) 为了便于学习,在每章中都提供了一些范例和大量的习题,从而帮助读者加深对内容的理解。

(4) 为了满足读者参加计算机软件专业技术资格和水平考试的需要,教材的内容涵盖了初级程序员级、程序员级和高级程序员级中有关操作系统的内容,并且习题中的许多选择题也选自历届考试试卷。

本教材从资源管理的角度来介绍操作系统。在教学中特别要注意安排足够的实验课时数,以突出高职教学的特点。对于教材中的原理部分,可以根据课时和学生的实际情况作适当的裁剪。

本教材共分 6 章,第三、四章(1~6 节)由李巍编写,其余部分由徐宗元编写,张航参加了部分实验的编写,全书由徐宗元统稿。

由于时间仓促,加之作者水平有限,必然存在错误和不妥之处,恳请读者批评指正!

编 者

2000 年 5 月

第 2 版前言

操作系统是计算机系统中最关键的系统软件;同时作为一门课程,它是计算机及相关专业的必修课。为了满足培养应用型人才的需要,我们在 2000 年编写了《操作系统》这本教材。虽然自 2000 年本教材第 1 次印刷出版后,我们根据教材在教学中的使用情况和读者的反馈,在后续的印次中不断作了一些修改。并且在 2003 年的修订中附上了至 2003 年的程序员、高级程序员考试试题,并增加了一些关于 Linux 操作系统的知识。但近年来,飞速发展的计算机技术,特别是操作系统的更新换代,使得该书中的有些内容已陈旧。为了满足当今教学的需求,我们对原书作了改版。

第 2 版最大的特点是结合当前主流的微机操作系统 Windows 2000 和 Linux/UNIX 分析讲解操作系统原理,并在讲解相关理论内容后,配上相应实验。实验强调实际的应用,包括与操作系统内核有关的 API 和系统调用的应用,计算机系统的性能测试等。

本书概念叙述简单、清楚、明了,结合原理详尽讲解例题,并配有大量帮助理解概念的习题,以便于担任该课程教学的教师的教学和学生的自学,这些都是本书在各版次中坚持的特点。

本书从资源管理的角度来介绍操作系统,全书共分 5 章,主要由徐宗元编写,李巍参加第三、四章部分编写,徐辰昊等参加实验的编写,全书由徐宗元统稿。

有关教学的辅导材料(习题答案、实验参考程序、PowerPoint 课件等)在高等教育出版社的相关网站(www.hepsh.com)上提供。由于编著者的水平有限,本教材中难免有错误和不妥之处,恳请读者批评指正,email 联系地址为 xu_zongyuan@163.com。

编 者

2005 年 1 月

目 录

第一章 引论	1	3.5 Windows 2000 的内存管理 ...	131
1.1 操作系统概述	1	3.6 实验和习题	142
1.2 计算机硬件系统概述	3	第四章 设备管理	156
1.3 操作系统类型	5	4.1 设备管理概述	156
1.4 操作系统的特征和功能	14	4.2 I/O 控制方式	158
1.5 用户与操作系统的接口	15	4.3 缓冲技术	161
1.6 实验和习题	19	4.4 I/O 系统软件	165
第二章 进程管理	27	4.5 设备的分配	170
2.1 进程的引入和描述	27	4.6 磁盘 I/O	172
2.2 进程控制	31	4.7 Windows 2000 I/O 系统	177
2.3 进程同步	38	4.8 实验和习题	187
2.4 进程通信	54	第五章 文件系统	200
2.5 调度	60	5.1 文件系统概述	200
2.6 进程死锁	69	5.2 文件的组织	201
2.7 操作系统结构	76	5.3 文件的目录和管理	208
2.8 实验和习题	81	5.4 文件存储空间的管理	220
第三章 存储管理	105	5.5 NTFS 文件系统	227
3.1 存储管理概述	105	5.6 安全和保护	232
3.2 存储器的连续分配方式	108	5.7 实验和习题	243
3.3 存储器的离散分配方式	112	参考文献	262
3.4 虚拟存储器管理技术	120		

第一章 引 论

操作系统(operating system, 简记为 OS)是一组控制和管理计算机硬件和软件资源、合理地组织计算机工作流程以及方便用户使用的程序集合,它是最重要的系统软件,本章多方面地介绍了操作系统的作用、发展史、类型、特征和功能。

1.1 操作系统概述

1.1.1 什么是操作系统

1. 从计算机系统组成观点——操作系统是系统软件

计算机系统由硬件系统和软件系统两大部分组成:硬件系统包括处理器、存储器和输入/输出(I/O)设备,软件系统包括系统软件和应用软件。系统软件用于计算机软硬件资源的管理、维护、控制和运行以及对运行的应用软件进行翻译、装入等服务工作。系统软件分成三类:操作系统、语言处理系统和常用的例行服务程序,操作系统是第一层系统软件。应用软件是指那些为某类应用需要而设计的程序或为解决某个特殊问题而编制的程序。

2. 从资源管理观点——操作系统是系统资源管理者

硬件和软件都称为计算机系统资源。操作系统是资源管理程序,它用于控制和管理计算机硬件和软件资源。它主要的功能也是对计算机资源——处理器、存储器、I/O 设备等硬件和软件信息进行有效管理,相应的操作系统的模块是 CPU 管理(主要是进程管理)、存储器管理、I/O 设备管理和文件系统。

3. 从软件分层、扩充机器的观点——操作系统是扩充裸机的第一层系统软件

计算机系统中硬件和软件、软件各部分之间的关系是层次结构关系。不包括软件的计算机硬件系统通常称为裸机,一台裸机的功能即便再强,它往往也是不便于用户使用的,功能上相对来说也是有局限性的,而软件则在硬件基础之上对硬件的性能加以扩充和完善。例如,要进行数据的输入/输出,在裸机上用户必须使用机器语言编写上千条指令来完成该操作,而如果在裸机上覆盖一层 I/O 设备管理软件,用户便可利用它所提供的几条 I/O 命令来完成数据的输入/输出。在裸机上添加一层软件可以看成是对原裸机的扩充和完善,二者结合变成了功能更强大、使用更方便的机器,这个新的功能更强的机器被称为虚拟机。进而,可以在虚拟机上增添新的软件,得到功能更加强大的机器,即一个裸机可以被多次扩充成功能逐步增加的虚拟机。而操作系统是在裸机上扩充的第一层软件。

操作系统是紧靠裸机的软件,它的功能需要硬件支持。为了改善系统的性能,操作系统中与硬件密切有关的部分软件通过微程序设计技术转化为硬件,这一过程也就是通常所说的固化,而这些具有软件功能的硬件称为固件。例如,IBM 公司生产的 AS/400 在操作系统与硬件之间配了层微码层。微码层对 CISC 和 RISC 不同指令集设置不同的特许内码,同时又设置了技术独立于机器的接口,使操作系统可以独立于硬件,便于操作系统在不同机器上的移植。

4. 从服务用户的观点——操作系统是用户与裸机之间的接口

操作系统的一个重要任务是使计算机方便于用户的使用。操作系统处于用户和计算机硬件系统之间,操作系统中将计算机硬件系统中各种复杂的运算和操作指令(例如三角函数计算和数据 I/O 等)按功能进行汇总,以模块程序方式完成,而提供的是操作系统中某个软件模块和计算机硬件资源一起完成某个操作或运算的接口。系统提供的接口有两类:一类是作业级接口,它提供一组键盘命令,供用户组织和控制自己的作业运行;另一类是程序级接口,它提供一组系统调用,即操作系统中某个模块功能,供用户程序和其它程序调用。用户不直接与计算机硬件系统打交道,所以操作系统是用户和计算机系统的接口。

1.1.2 操作系统的历史

为了提高计算机资源的利用率和方便用户的使用,随着计算机硬件的不断更新换代和体系结构的不断发展,操作系统经历了一系列革命性的变迁,一般认为操作系统发展已经历了五代。

1. 第零代(20 世纪 40 年代末期—20 世纪 50 年代初期)无操作系统

此阶段的计算机操作方式是由程序员采用人工操作方式直接使用计算机硬件系统,即由程序员将事先已穿孔的纸带(或卡片)装入纸带输入机(或卡片读入机),再通过控制台开关板键启动程序运行,程序运行完毕打印结果输出,卸下纸带(或卡片)后才能让另一用户使用计算机。

2. 第一代(20 世纪 50 年代中期—20 世纪 50 年代末期)初级批处理系统

第一代操作系统是为了减少人工操作时间和作业转换时间,从而提高 CPU 利用率而设计的。由于有了较高速和大容量的磁带机和外围 CPU,通过脱机 I/O 方式,将许多作业依次由低速输入设备(纸带机)通过与主机脱机的外围机输入到较高速的磁带机上,这样在主机系统上作业已成批组合,可以依次调入系统运行,实现作业自动转接和自动成批处理作业,这个阶段的操作系统称为初级单道批处理系统。

3. 第二代(20 世纪 60 年代初期—20 世纪 60 年代中期)多道程序设计共享系统

由于硬件技术的重大进步、通道技术的引进和中断技术的发展,计算机系统能支持并行操作,操作系统进入多道程序设计系统阶段,在多道程序系统中若干用户同时驻在内存来共享 CPU 和 I/O 设备。这一代操作系统包括多道批处理系统、采用交互方式的分时系统和以提高瞬时响应时间为特征的实时系统。

4. 第三代(20 世纪 60 年代中期—20 世纪 70 年代中期)多模式系统

这一代计算机系统是通用系统,像 IBM-360 号称 360°面向各种类型的用户。这一代操作系统是多模式系统,即一个系统同时支持批处理、分时处理、实时处理和多重处理。在一个多重处理系统中一个计算机系统使用了多个 CPU 以增强计算机的处理能力。

5. 第四代(20 世纪 70 年代中期—现在)网络操作系统和分布式操作系统

计算机网络通过通信设施将地理上分散的、具有自治功能的多个系统互连起来,网络操作系统实现在计算机网络上进行信息交换、资源共享和互操作等功能。分布式操作系统将地域上分散的各系统互连成一个具有整体功能的系统,并可将一个任务分布在各系统上运行,实现分布式处理。

1.2 计算机硬件系统概述

1.2.1 计算机硬件系统结构

1. 大、中、小型计算机的硬件系统结构

从小型计算机到大型计算机的硬件系统多由中央处理机(CPU)、I/O 处理机(又称通道)、存储器和 I/O 设备组成。图 1-1 所示的是一个典型中型计算机(IBM-370)的硬件系统结构,这类计算机以存储器为中心,CPU 和各种通道都与存储器相连,CPU 执行的程序和数

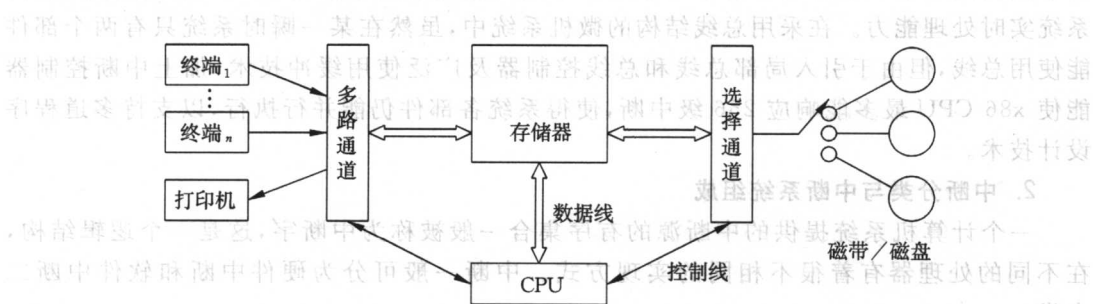


图 1-1 IBM-370 的硬件系统结构

2. 微型计算机的典型硬件系统结构

微型计算机(以下简称微机)是以总线为纽带来构成计算机系统的,CPU、存储器、I/O 设备(包括外存磁盘、磁带)都挂接在总线,如图 1-2 所示。CPU 与存储器、CPU 与 I/O 设备之间经过总线来交换信息,而存储器与 I/O 设备之间的信息交换则要通过直接存储器访问控制器(DMAC)到总线再进行信息交换。

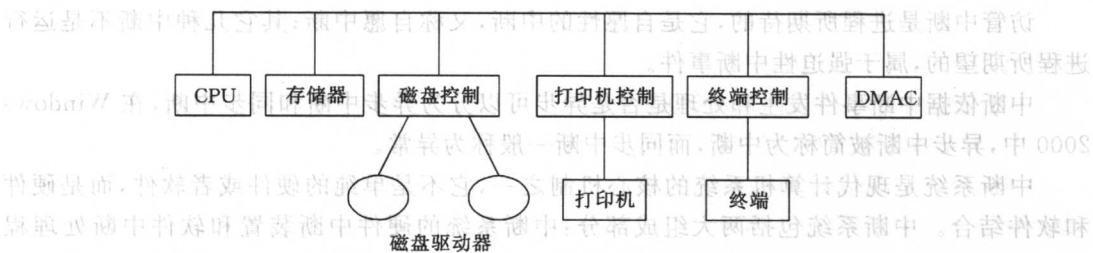


图 1-2 微机的总线型系统结构

1.2.2 现代操作系统的硬件基础

1. 支持并行操作的硬件

现代计算机系统为了提高计算机的资源利用率,采用多道程序设计技术,它在主存中存

放多道用户的作业,计算机系统硬件支持作业并行操作,使多道作业同时处于运行状态共享系统资源。在现代批处理系统中支持并行操作的硬件是通道和中断。通道是可以与 CPU 并行操作的专门用来控制 I/O 设备数据传送的处理机。当 CPU 需要输入数据时,只要发命令给通道去完成,通道进行 I/O 操作和 CPU 操作可并行异步执行,在通道完成 I/O 操作后,它用中断方式向 CPU 报告其完成 I/O 的情况。中断是指 CPU 对系统中发生的异步事件的响应,异步事件是指无一定时序关系的随机发生的事件。中断的含义是当这些异步事件发生后,打断了 CPU 对当前程序的执行,而转去处理该异步事件,即执行该事件的中断处理程序,直到处理完了之后,再转向原程序的中断点(发生中断时正在执行的程序的暂停点)继续执行。处理器暂停当前程序转而处理中断的过程称为中断响应,中断处理结束之后恢复原来程序的执行被称为中断返回。中断技术的应用范围不断扩大,它使得通道具有中断主机工作的能力,以支持多道程序设计技术,它作为所有要打断 CPU 正常工作并要求其去处理某一事件的一种常用手段。中断的应用能充分利用 CPU 的使用效率,提高系统实时处理能力。在采用总线结构的微机系统中,虽然在某一瞬时系统只有两个部件能使用总线,但由于引入局部总线和总线控制器及广泛使用缓冲技术,加上中断控制器能使 x86 CPU 最多能响应 256 级中断,使得系统各部件仍能并行执行,以支持多道程序设计技术。

2. 中断分类与中断系统组成

一个计算机系统提供的中断源的有序集合一般被称为中断字,这是一个逻辑结构,在不同的处理器有着很不相同的实现方式。中断一般可分为硬件中断和软件中断两大类。

硬件中断又可分为硬件异常/硬件故障中断(例如电源故障、内存奇偶校验错、总线错误)、I/O 中断(I/O 设备和通道完成数据传输操作引起的中断)和外部中断(例如时钟中断、控制台中断)。

软件中断又可分为软件异常/程序中断(例如被零除、定点操作数溢出、地址越界、虚存管理中缺页)和访管中断(系统调用)是用户程序在运行中请求操作系统为其提供服务而执行一条“访管指令”所引起的中断,又称软件中断。在 PDP-11 机上访管指令是 tarp 指令,在 x86 机上访管指令是 INT 指令。

访管中断是进程所期待的,它是自愿性的中断,又称自愿中断;其它几种中断不是运行进程所期望的,属于强迫性中断事件。

中断依据中断事件发生和处理是否是异步可以分为异步中断和同步中断,在 Windows 2000 中,异步中断被简称为中断,而同步中断一般称为异常。

中断系统是现代计算机系统的核心机制之一,它不是单纯的硬件或者软件,而是硬件和软件结合。中断系统包括两大组成部分:中断系统的硬件中断装置和软件中断处理程序。中断装置负责捕获中断源发出的中断请求,并以一定的方式响应中断源,然后将处理器的控制权移交给特定的中断处理程序。中断处理程序则负责辨别中断类型并根据请求做出相应的操作。

3. Windows 2000 的中断类型和优先级

Windows 2000 的中断调度程序将硬件中断级映射到由操作系统识别的中断请求级(IRQL)的标准集上(表 1-1)。

表 1-1 中断请求级

级	用途	来源
31	高优先级	硬件中断
30	电源故障	
29	处理器内部中断	
28	时钟中断	
27	配置文件	
26	设备 n	
$\vdots$	$\vdots$	
3	设备 1	
2	延迟过程调度(DPC/Dispatch)	
1	异步过程调度(APC)	软件中断
0	低优先级	正常线程执行

IRQL 按优先级排列中断,并进行中断服务,较高优先级的中断服务可以抢占较低优先级的中断服务。IRQL 从高往低到设备 1 都是为硬件中断保留的。延迟过程调度(DPC/Dispatch)和异步过程调度(APC)级中断是内核和设备驱动器产生的软件中断。低优先级实际上并不是真正的中断级,在该级上执行普通的线程,并允许发生所有的中断。

IRQL 设置决定了每个处理器可以接收的中断。当程序运行时,它可以提高或降低处理器的 IRQL 来屏蔽一些事件。如果中断源的 IRQL 高于当前中断设置,则它的中断可以中断该处理器;如果中断源的 IRQL 等于或低于当前中断设置,则它的中断将被封锁或“屏蔽”,直到一个正在执行的程序降低了 IRQL。当产生中断时,陷阱处理程序提高处理器的 IRQL 直到与中断源所指定的 IRQL 相同,这可以保证服务于该中断的处理器不会被同级或较低级的中断抢先。被屏蔽的中断将被阻挡,直到 IRQL 降低。可见改变处理器的 IRQL 对操作系统具有很重要的影响。

1.3 操作系统类型

1.3.1 批处理系统

1. 初级单道批处理系统

为了解决人工操作(无操作系统)严重降低了计算机资源的利用率的问题,即解决 CPU 等待人工操作和高速 CPU 与低速 I/O 间矛盾等问题。在 20 世纪 50 年代末引入脱机 I/O 技术,该技术利用一台外围机,脱离主机先将低速输入设备(如纸带机)的数据输入到较高速、大容量的输入设备(如磁带)上。同时引入批处理技术,它利用磁带容量较大的特点,先将一批作业按序以脱机方式输入到磁带上。主机在监督程序控制下,根据卡片机读入控制作业操作信息,逐个从磁带上读入作业到内存,调入汇编/编译程序将源程序翻译成目标代

码,连接装配后执行,结果输出打印。执行完毕后再调入下一个作业,系统就这样自动地一个一个作业地进行处理,直至完成。这就是早期单道批处理系统(图 1-3),它实现了作业自动转换,从而减少了作业建立和人工操作时间。

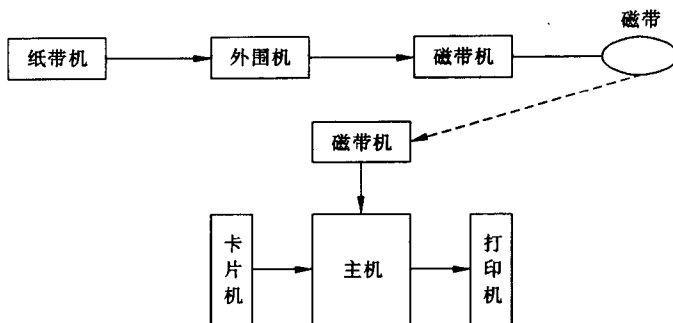


图 1-3 单道批处理系统

2. 现代多道批处理系统

在早期的单道批处理系统中,内存仅有一道作业,这使系统中仍有较多的空闲资源,致使系统的性能较差。为了进一步提高资源利用率,在硬件采用通道和中断技术支持并行操作的情况下,引入了多道程序设计技术,由此引入多道批处理系统。

多道程序设计技术是在计算机内存中同时存放几道相互独立的程序,这几道程序都处于运行过程中,它们先后开始了各自的运行,但都未运行完毕。多道程序在宏观上并行执行,而在微观上多道程序在某个部件(如 CPU、I/O)上是串行,即多道程序轮流地使用部件,交替执行。由于计算机和部件支持并行操作,所以多道程序可能在各部件上平行执行。

例如,一个有一个 CPU 和两个 I/O 设备(磁盘和磁带)的二道批处理系统运行情况如图 1-4 所示。

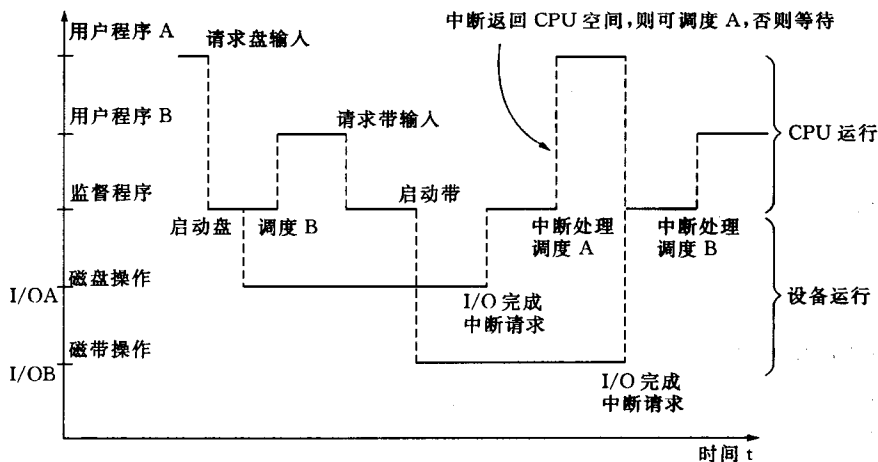


图 1-4 二道批处理系统运行图

从图 1-4 中可看出 CPU 与 I/OA(磁盘)、I/OB(磁带)设备可并行工作,而对系统的某个部件如 CPU,两个程序 A、B 和监督程序(操作系统)都是串行,即交替运行。程序在

CPU 与 I/O 设备间切换, CPU 运行哪个程序都由监督程序控制, 并根据一定的算法调度, 对 I/O 设备调度采用非抢占方式(即一旦将该 I/O 设备分配给某进程, 那就要等该程序 I/O 完成后再进行分配); 而 CPU 调度可采用两种方式: 非抢占方式和抢占方式。如采用抢占方式调度 CPU, 当一个优先级更高的程序要使用 CPU 时, 就要暂停正在 CPU 上运行的优先级较低的程序, 让优先级高的程序运行完毕后再运行原程序。

图中开始时程序 A、B 都在内存中, 先运行程序 A。当程序 A 要请求盘输入时, 转由 CPU 运行监督程序来启动盘, 启动成功后 I/OA(磁盘)进行程序 A 盘输入。这时 CPU 空, CPU 运行监督程序来调度程序 B, 接着 CPU 运行程序 B(此时 CPU 在运行程序 B, 而 I/OA(磁盘)进行程序 A 盘输入)。程序 B 运行一段时间后请求带输入时, 转由 CPU 运行监督程序来启动磁带, 启动成功后 I/OB(磁带)进行程序 B 磁带输入(此时 I/OA(磁盘)进行程序 A 盘输入, I/OB(磁带)进行程序 B 磁带输入, 而 CPU 空操作)。CPU 空操作到 I/OA(磁盘)结束程序 A 盘输入, 产生中断。转由 CPU 运行监督程序的中断处理程序处理程序 A 中断, 处理后调度程序 A, 接着运行程序 A。在程序 A 运行时, I/OB(磁带)结束程序 B 磁带输入, 产生中断。CPU 暂停正在 CPU 上运行的程序 A 转运行监督程序的中断处理程序处理程序 B 中断, 处理后根据一定的算法调度程序 A 或 B, 这里调度程序 B, 接着运行程序 B。在 1.6.2 节中将介绍一个多道程序调度运行实例来增强对多道程序设计的认识。

在现代批处理系统中引入缓冲技术, 在外存设置输入缓冲区(输入井)和输出缓冲区(输出井)。同时又引入 SPOOLing 技术(又称假脱机 I/O 或联机 I/O), 该技术采用通道代替外围机, 它与 CPU 并行操作, 控制外设完成 I/O 任务。在该系统中用户提交的作业由输入设备经 SPOOLing 系统联机输入到外存磁盘的输入井内, 并排成一个后备作业队列, 然后由作业调度程序按一定的算法从后备队列中选择若干个作业调入内存, 它们在内存中按一定调度规则共享 CPU 和系统中各种 I/O 资源, 以达到提高资源利用率和系统吞吐量的目的。SPOOLing 系统实现了对 I/O 设备的虚拟, 只要输入设备空闲, 系统可预先将输入数据从设备传输到输入井中供用户程序随时读取; 系统可将用户程序随时输出的数据先送到输出井中, 待输出设备空闲时再执行数据输出操作。

由于现代批 CPU 系统中多道程序共享 CPU、内存和 I/O 设备, 所以现代多道批处理系统要保证多道程序互不干扰、高效并发执行。

典型的多道批处理系统是 IBM OS/360 中的任务数固定多道程序系统(MFT)和任务数可变多道程序系统(MVT)。

1.3.2 分时系统

1. 分时系统原理

分时系统的形成和发展的动力是用户的需要。用户希望自己能直接控制程序的运行和及时进行人机交互, 以便能方便地修改程序的错误。同时又希望一台计算机能为多位用户共享, 解决当时(20 世纪 60 年代)计算机昂贵而缺乏的矛盾, 方便用户使用。

分时系统是指在一台主机上连接了多个带有显示器和键盘的终端, 它允许多个用户同时通过终端以交互方式使用并共享主机中的资源。所谓分时是指多个用户分时共享使用同一台计算机, 也就是说把计算机的系统资源(尤其是 CPU)进行时间上分割, 即将整个工作时间分成一个个的时间片, 每个时间片分给一个用户使用, 这样将 CPU 工作时间分别提供

给多个用户使用,每个用户依次轮流使用每一个时间片。

2. 分时系统的实现

(1) 采用调进/调出的单道分时系统

20 世纪 60 年代初期,由美国麻省理工学院建立的第一个分时系统 CTSS 便属于单道分时系统,它采用调进/调出技术来实现。在该系统内存中只驻留一道程序,其余程序都在外存上。每当内存中一个程序运行一个时间片后,便调至外存(称为调出)。再从外存中选一个程序装入内存(称为调进)并运行一个时间片,依此方法使所有的作业都能在一规定的时间内轮流运行一个时间片,这样便能使用户都能与自己的作业交互。由于单道分时系统只有一道作业驻留在内存,在多个作业轮流运行的过程中,每个作业往往可能被频繁地调进调出多次,开销大,故使系统的性能较差。

(2) 基于多道程序设计的多道分时系统

由于现代计算机系统采用大容量内存,并采用内存与磁盘对换技术来扩充内存,现代分时系统引入了多道程序设计技术,该技术在内存中存放多道作业并排成一个队列,依次轮流地获得一个时间片来运行。由于切换的作业就在内存,不要花费大量时间用于调进调出,故多道分时系统具有较好的系统性能。作业调度算法也不限于时间片不变,次序不变的简单轮转。像美国贝尔实验室最早研制的 UNIX 操作系统,它对系统中各进程采用可抢占的动态优先级调度算法,也实现了各作业分时共享计算机资源,但时间片大小和轮转次序都在变化。

批处理系统和分时系统都是多道程序系统,是属于处理用户作业的系统。系统本身没有要完成的作业,它只是起着管理调度系统资源,向用户提供服务的作用,这类系统可以说是“通用系统”。

3. 分时系统的特征

(1) 多路性

允许在一台主机上同时联接多台联机终端,系统按分时原则为每个用户服务。宏观上是多个用户同时工作,共享系统资源;而微观上则是每个用户作业轮流运行一个时间片。多路性即同时性,它提高了资源利用率,从而促进了计算机更广泛的应用。

(2) 独立性

每个用户各占一个终端,彼此独立操作,互不干扰。因此用户会感觉到就像他一人独占主机。

(3) 交互性

用户可通过终端与系统进行广泛的人机对话。用户可以请求系统提供多方面服务,如文件编辑、数据处理和资源共享等。

(4) 及时性

用户的请求能在很短时间内获得响应,此时间间隔是以人们所能接受的等待时间来确定的,通常为 1 秒~2 秒。

4. 分时系统的重要指标——响应时间

响应时间是分时系统的重要指标,它是用户发出终端命令到系统开始作出响应之间的时间间隔。假设分时系统中用户数为 n ,每个用户的运行时间片为 q ,则系统的响应时间为

$$T = n \cdot q$$

每个用户分到的时间片 q 由两部分组成,用于切换的时间 T (切换)和真正用于处理的时间 T (处理)。

分时系统中时间片 q 的选择是一个复杂和关键的任务,如时间片选得过大,造成响应时间不变时用户数减少,或造成用户数不变时响应时间过长;而时间片选得过小又会造成在一个时间片内切换开销相对增加,一个进程相对要花费更多的时间片才能运行结束,一个进程在系统中的周转时间大大增长。最佳的时间片值应既能使分时用户得到好的响应时间,同时又要使在一个时间片内切换开销相对较小,可忽略片内切换的开销。

5. 典型的分时系统

1969年由美国贝尔实验室 Ritchie 和 Thompson 开发的 UNIX 系统是典型的分时系统,UNIX 系统是一个多用户、多任务操作系统。

1.3.3 实时系统

实时系统的目标是为了缩短系统的响应时间,对随机发生的外部事件能作出及时响应并对其进行处理。

1. 实时系统分类

根据使用任务的不同,实时系统分成两类:

(1) 实时控制系统

此类实时控制系统主要用于生产过程的自动等控制、实验数据自动采集、武器的控制(包括火炮自动控制、飞机自动驾驶、导弹的制导系统等)。这类系统随机发生的外部事件是并非由于人工启动和直接干预而引起的。系统的响应时间由发生事件的外部设备所决定,可以快到毫秒数量级。

(2) 实时信息处理系统

这类系统主要用于实时信息处理,像飞机订票系统、情报检索系统。这类系统随机发生的外部事件是由人工通过终端启动,并连续对话而引起的,这类系统的响应时间主要是用户能接受的秒数量级。

2. 实时系统特征

与批处理系统和分时系统相比实时系统具有以下特征:

(1) 快速的响应时间

实时系统是为了提高系统响应时间而设计的操作系统,特别是实时控制系统,对外部事件的响应要十分及时迅速。外部事件往往以中断方式通知系统,系统有较强的中断处理能力,实时系统的设计也以“事件驱动”方式来设计。

(2) 有限的交互能力

批处理系统和分时系统是向用户提供作业服务的“通用系统”,实时系统(如实时信息处理系统)一般是专用系统,它能提供人机交互方式,但用户只能访问系统中某些特定的专用服务程序,不能像分时系统一样向终端用户提供多方面服务。

(3) 高可靠性

批处理系统和分时系统虽也要求系统可靠,但相比之下实时系统对系统的可靠性要求更高。因为实时系统处理的对象往往是重要的经济或军事目标,同时又是现场直接控制处理。任何差错都可能带来巨大经济损失,甚至是无法预料的灾难性后果。因此实时系统中

往往采用双机系统、多级容错措施来保证系统和数据的安全。

1.3.4 微机操作系统

随着大规模集成电路的发展,微机产生了,配置在微机上的操作系统称为微机操作系统。早期的微机是个人独立使用的,而现代的微机可作为上网工具。常用的微机操作系统有 MS Windows、OS/2、SCO UNIX、Linux、Mac OS、Solaris 等。

1. MS Windows

为了克服 MS-DOS 字符界面使用不便的缺点,Microsoft 在 MS-DOS 基础上推出了易学易用的图形用户界面的 MS Windows。Windows 是一个多任务的操作系统。Windows 有多个版本,早期有 Windows 3.0/3.1/3.2,接着有 Windows 95、Windows 98、Windows NT、Windows 2000、Windows XP、Windows Server 2003。

Windows 95/98 全面实现了 32 位内存访问,将 Microsoft 网络并入到 Windows 系统中,可以通过 Microsoft Network 访问 Internet,以获得广泛的信息资源。同时它改变了早期 Windows 的外观,引入“即插即用”功能等许多先进技术。Windows 95/98 是一个多任务、多线程、全 32 位的操作系统,它支持 32 位和 16 位两种应用程序和驱动程序,具有安全向下兼容的特点。Windows 98 又进一步将 Internet 的应用软件纳入系统,使 Windows 向着公用网络操作系统的方向向前发展了一大步。

Windows NT 是 Microsoft 公司推出的 32 位网络操作系统,它主要面向高性能微机、工作站和多处理器服务器,用户管理整个网络,但它为了新功能而牺牲了兼容性。Windows 2000 系列是 Windows NT 4.0 系列的换代产品,它分成四个产品:Windows 2000 Professional、Windows 2000 Server、Windows 2000 Advanced Server 和 Windows 2000 Datacenter Server。

Windows XP 把消费型和商业型操作系统融合为统一系统代码,它既适合家庭用户使用,也适合商业用户使用。Windows XP 有三个版本,以满足家庭和工作中的计算需要。Windows XP Home Edition 有良好的数字媒体平台,适合家庭用户和游戏爱好者。Windows XP Professional 是为商业用户设计的,有高级别的可扩展性和可靠性。Windows XP 64-Bit Edition 可满足最高级别的专业技术工作站用户的需要。

自 2003 年起,Microsoft 将其服务器产品统一在 Windows Server System 品牌下,这是一个全面、集成而且具有交互操作能力的服务器基础结构,提供了一个公用体系结构,Windows Server System 的基础是 Windows Server 2003 操作系统。

2. OS/2

OS/2 是 IBM 公司推出的微机操作系统,OS/2 是一个真正的 32 位单用户多任务操作系统。它性能非常完善,包括了某些大型机才具有的代码共享和动态链接等。

3. SCO UNIX

SCO 公司将运行于大、中、小型机上的 UNIX 操作系统移植到微机上形成 SCO UNIX,SCO UNIX 是多用户多任务操作系统。

4. Linux

Linux 操作系统是一个遵循标准操作系统界面(Portable Operating System Interface,简称为 POSIX)的标准免费操作系统,具有 UNIX BSD 和 UNIX SYS V 的扩展特性。它的

版权所有者是芬兰籍的 Linus B. Torvalds 和其他开发人员,并且遵循 GNU 通用公共许可证(GNU General Public License,简称为 GPL)声明。Linux 可以在基于 Intel 处理器的微机上运行,它可以将一台普通的微机变成功能强劲的 UNIX 工作站。

Linux 有一个基本内核,一些组织和厂商将内核与应用程序、文档包装起来,再加上安装、设置和管理工具,就构成了直接供一般用户使用的套件。Linux 版本分为两部分:内核版本和发行套件版本。内核版本指的是在 Linus 领导下开发的系统内核的版本号,2004.10 Linux 的最新稳定内核版本号为 2.6.9。发行套件常见的有 Slackware、Red Hat Linux、Debian 等。Red Hat Linux 具有支持多种硬件平台(如 Intel、Sparc 和 Alpha 平台)、软件安装、配置、升级、维护简单,系统管理工具使用方便等优点,是初学者的最佳选择。

中国政府于 2001 年发布的指导性文件《当前优先发展的高技术产业化重点领域指南》中确定基于 Linux 的操作系统平台及其集成应用环境软件是我国高技术发展的重点领域之一。

5. Mac 操作系统

Mac 操作系统是 Apple 公司推出的微机操作系统,广泛用于图形、图像处理。

1.3.5 网络操作系统

1. 网络操作系统的定义

计算机网络可以定义为一些互连的自主计算机系统的集合,所谓自主计算机是指计算机具有独立处理能力,而互连则表示计算机之间能够实现通信和相互合作,按网络所覆盖的地理范围大小可把计算机网络分成广域网(Wide Area Network,简称为 WAN)和局域网(Local Area Network,简称为 LAN)两类。

早期的网络操作系统实际上是在原自主计算机操作系统之上附加上具有实现网络访问功能的模块,因此通常将网络操作系统(Network Operating System,简称为 NOS)定义为: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,为网络用户提供所需各种服务的软件和有关规程的集合。

2. 网络操作系统的功能

网络操作系统应具有下述的功能:

- (1) 高效、可靠的网络通信;
- (2) 对网络中共享资源(如硬盘、打印机等)有效的管理;
- (3) 提供电子邮件、文件传输、共享硬盘、打印机等服务;
- (4) 网络安全管理;
- (5) 提供互操作能力。

3. 网络操作系统的模式

网络操作系统具有以下两种工作模式:

- (1) 客户/服务器(Client/Server,简称为 C/S)模式

客户/服务器的模式是流行的网络工作模式,它具有分布处理和集中控制的特称。这种模式网络可分为以下两大类。

- 服务器 它是网络的控制中心,其任务是向客户提供一种或多种服务,服务器可有多种类型,如提供文件/打印服务的文件服务器,提供数据库服务的数据库服务器