



全国高职高专计算机新创规划教材

操作系统与上机指导

CAOZUO XITONG
YU SHANGJI ZHIDAO

石海霞 王震江◎主编

QUANGUO GAOZHI GAOZHUAN JISUANJI XINCHUANG GUIHUA JIAOCAI



中国科学技术出版社
CHINA SCIENCE AND TECHNOLOGY PRESS

全国高职高专计算机新创规划教材

操作系统与上机指导

石海霞 王震江 主编

中国科学技术出版社
CHINA SCIENCE AND TECHNOLOGY PRESS
· 北 京 ·
BEIJING

图书在版编目 (CIP) 数据

操作系统与上机指导/石海霞, 王震江主编. —北京: 中国科学技术出版社, 2006. 8
(全国高职高专计算机新创规划教材)

ISBN 7 - 5046 - 4466 - 8

I. 操... II. ①石... ②王... III. 操作系统 - 高等学校: 技术学校 - 教材
IV. TP316

中国版本图书馆 CIP 数据核字 (2006) 第 088933 号

自 2006 年 4 月起本社图书封面均贴有防伪标志, 未贴防伪标志的为盗版图书。

内容介绍

《操作系统》是计算机专业的一门专业必修课。

本书包含两大部分: 第一部分为理论教学部分, 第二部分为实验部分。其中理论部分分别介绍了操作系统的进程管理、存储管理、设备管理、文件管理。每章围绕教学目的安排了相应的实训内容和适量的习题。实验部分列出了 (DOS、Windows、Linux) 操作系统的常见操作。

我们在编写时从高职高专学生的需求和理解能力出发, 以普遍性为原则, 目的是使本书成为一本能够让读者掌握操作系统相关知识的简单而又实用的教程。

中国科学技术出版社出版

北京市海淀区中关村南大街 16 号 邮政编码: 100081

电话: 010 - 62103210 传真: 010 - 62183872

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

科学普及出版社发行部发行

北京迪鑫印刷有限公司印刷

*

开本: 787 毫米 × 1092 毫米 1/16 印张: 12.875 字数: 330 千字

2006 年 8 月第 1 版 2006 年 8 月第 1 次印刷 定价: 18.20 元

(凡购买本社的图书, 如有缺页、倒页、
脱页者, 本社发行部负责调换)

全国高职高专计算机新创规划教材

编委会

主 任 颜 实

副主任 刘加海

委 员 (以姓氏拼音为序)

蔡向东	陈 胤	陈永东	陈月波	崔恒义	丁芝芳
范伟继	方锦明	方玉燕	黄云龙	金栋林	康震群
孔美云	李智庆	李天真	李永平	李连弟	黎浩宏
厉 毅	梁钜汎	楼程伟	马尚风	欧阳江林	
齐幼菊	沈丽梅	沈素娟	沈凤池	石海霞	孙 霞
王 勇	吴 坚	熊卫民	徐晓明	严小红	余再东
余根墀	俞伟新	詹重咏	张锦祥	朱 炜	

秘书长 熊盛新

教材编辑办公室

主 任 林 培

编 辑 孙卫华 程安琦 符晓静 彭慧元 甘丹红

《操作系统与上机指导》编写人员

主 编 石海霞 王震江
副 主 编 张 雁 黄 磊
编写人员 于化强 孙文方

策划编辑 林 培 孙卫华
责任编辑 孙卫华
封面设计 鲁 筱 杨 军
责任校对 林 华
责任印制 安利平

前 言

操作系统是计算机及应用专业的一门主干专业课，是计算机系统中重要的系统软件。通过本课程的学习，要使学生全面、系统地掌握计算机操作系统的基本概念、基本技术和基本方法；掌握计算机操作系统的组成和各部分的功能；了解计算机操作系统的发展特点、设计的基本原理和方法；对常用操作系统（DOS、Windows、Linux）能进行基本的操作和使用。

本书主要内容为：操作系统引论、进程管理、存储器管理、设备管理、文件管理及常用操作系统的实验部分。本书的主要特点是通俗易懂、由浅入深。在理论部分强调易懂性，在实验部分强调普遍性。

本书可作为高职高专计算机专业和网络专业的教材，也可以作为自学用书或参考资料。

本书由石海霞老师和王震江老师担任主编，于化强老师担任副主编，同时张雁老师和孙文方老师也参编了大量的内容。

由于编者水平的有限，加之时间的仓促和工作的繁忙，书中不妥之处在所难免，恳请广大读者不吝指教，我们将不胜感激。

编 者

2006 年 6 月

目 录

第一章 操作系统引论	(1)
第一节 操作系统的定义	(1)
第二节 操作系统的功能	(1)
第三节 操作系统的特征	(3)
第四节 操作系统的类型	(4)
第二章 进程管理	(12)
第一节 进程的概念与特征	(12)
第二节 进程的描述	(14)
第三节 进程和程序的关系	(18)
第四节 进程的互斥与同步	(19)
第五节 进程的调度与管理	(31)
第六节 进程的通信	(38)
第七节 死 锁	(42)
第八节 作业管理	(53)
第三章 存储管理	(68)
第一节 存储管理的基本知识	(68)
第二节 内存管理技术	(70)
第三节 存储管理方法	(72)
第四节 页式存储管理	(76)
第五节 请求页式存储管理	(81)
第六节 段式存储管理	(89)
第七节 段页式存储管理	(94)
第八节 高速缓冲存储器	(100)
第四章 设备管理	(108)
第一节 设备管理的概述	(108)
第二节 通道	(111)
第三节 输入输出的处理步骤	(114)
第四节 中断与中断处理	(116)
第五节 数据传输方式	(119)
第六节 直接存储器存取 (DMA) 方式	(121)
第五章 文件管理	(125)

第一节 文件概述	(125)
第二节 文件的逻辑结构和物理结构	(127)
第三节 文件存储空间的管理	(132)
第四节 文件的目录	(133)
第五节 文件的共享与安全	(137)
第六节 文件的操作	(140)
第六章 Linux 概要	(143)
第一节 什么是 Linux	(143)
第二节 Linux 的发展	(143)
第三节 Linux 的特征	(144)
第四节 Linux 与其他操作系统的区别	(145)
第五节 Linux 的组成	(146)
第六节 Linux 的安装、进入和退出	(153)
DOS 实验部分	(161)
Linux 操作实验部分	(178)
Windows 2000 实验部分	(190)
参考文献	(198)

第一章 操作系统引论

本章重点

- 操作系统的定义
- 操作系统的功能
- 操作系统的特征
- 操作系统的分类

本章难点

- 操作系统是如何定义的
- 操作系统的功能
- 分时系统、实时系统

计算机是人类社会 20 世纪最伟大的创造之一，自 1946 年在美国诞生第一台计算机至今的短短 60 年中，其技术得到了突飞猛进的发展。操作系统的形成和发展是和计算机发展密不可分的，操作系统始终是计算机科学和计算机工程的重要研究领域。一个新的操作系统往往汇集了计算机发展过程中的重要研究成果和技术。

第一节 操作系统的定义

计算机由硬件系统和软件系统组成，它们构成了一个完整的计算机系统。操作系统是在裸机上加载的第一层软件，是对计算机硬件系统功能的首次扩充。从用户的角度看，计算机系统配置了操作系统后，由于操作系统隐蔽了硬件的复杂细节，用户会感到机器使用起来更简单、更容易了。通常就说操作系统为用户提供了一台功能经过扩展的机器，或叫“虚拟机”，因为现实生活中并不存在具有这种功能的真实机器，它只是用户的一种感觉而已。从计算机系统的角度看，由于操作系统的组织与管理，系统中的各种硬件、软件资源得到了更有效的利用，机器的工作流程更为合理。因此，操作系统是现今计算机系统中不可缺少的一个系统软件。

操作系统的定义为：“操作系统是一个程序、系统软件，是用来控制和管理计算机硬件和软件资源、合理有效地组织计算机系统的工作，为用户提供一个方便并且可扩展的工作环境，从而起到连接计算机和用户的接口作用”。

第二节 操作系统的功能

从资源管理的角度看，操作系统应该具有四个方面的功能：处理机管理、存储管理、设备管理、文件管理。此外，为了方便用户使用计算机，还必须向用户提供一个方便的用户接口。这四部分相互配合，协同工作，实现对计算机系统的资源管理和控制程序的执行。下面就以上四个方面对操作系统的功能做一个简略的说明。

一、处理机管理

中央处理机（CPU）是计算机系统中的一个举足轻重的资源。用户程序进入内存后，只

有通过 CPU，才能真正得以运行。为了提高 CPU 的利用率，系统必须采用多道程序设计技术，使内存中同时有几个用户作业程序存在。这样一来，当一个程序因等待某事件（如输入/输出）的完成而暂时放弃使用 CPU 时，操作系统就可以把它重新分配给其他要运行的作业程序使用，从而提高它的利用率。

处理机管理的主要工作如下：

(1) 记录系统中当前每个作业程序的状态。这样，在需要对 CPU 重新进行分配时，就在候选的程序中选取。

(2) 指定处理机调度策略，这是挑选候选程序时应遵循的原则。

(3) 实施 CPU 分配（也就是处理机调度），以便让获得 CPU 的作业程序真正投入运行。

处理机管理总是把 CPU 分配给参与 CPU 竞争的那些作业程序使用。那么，究竟哪些作业有资格参与对 CPU 的竞争，就涉及作业管理的问题。

实行多道程序设计可以提高 CPU 的利用率，但这需要“适度”。内存中可运行的作业程序多了，参与系统资源竞争的对手也就多了，就肯定会影响系统效率的发挥。所以在有的操作系统中，就采用处理机的两级调度：第一级是作业调度，涉及作业管理；第二级是处理机调度，属于处理机管理。

作业管理的主要工作如下：

(1) 记录提交给系统的作业（一般存放在磁盘中）的状态，以及对系统资源的需求信息。

(2) 制定作业调度策略，在需要从“作业井”中的候选作业里选择作业进入内存，参与对 CPU 的竞争。

(3) 为用户提供一个好的系统使用环境，以便有效组织自己的工作流。

二、存储管理

存储器是计算机的记忆装置。在计算机系统中，存储器可以分为内存储器（也称主存储器）和外存储器（也称辅助存储器）两种。内存储器（简称内存）速度快，CPU 可以直接访问，用于存放计算机当前正在运行的程序和数据；外存储器（简称外存）速度相对较慢，由于 CPU 不能直接对它进行访问，因此一般作为内存的延伸或后援，存放暂时不用的程序和数据。

由于多道程序运行时竞争的存储资源是内存，所以操作系统中的存储管理是针对内存而言的。也就是说，存储管理的主要对象是内存储器，其主要工作如下：

(1) 记住内存各部分的使用情况，哪些为已经分配，哪些为空闲待分配。

(2) 制定内存的分配策略，实施内存的具体分配和回收。

(3) 保证内存中各独立作业程序的安全，互不侵扰。

(4) 解决“作业程序比内存大时，也能得到正确的运行”的存储扩充问题。

三、设备管理

计算机系统中，除了处理机和内存储器外，全都是设备管理的对象，主要是一些输入输出设备和外存储器。由于外部设备品种繁多，性能千差万别，因此设备管理是操作系统中最为复杂庞大的部分。

设备管理的主要工作如下：

(1) 记录各类设备的使用状态，按各自不同的性能特点进行分配和回收。

(2) 为各类设备提供相应的设备驱动程序、启动程序、初始化程序以及控制程序等, 保证输入输出操作的顺利完成。

(3) 利用中断、通道等技术, 尽可能地使 CPU 与外部设备、外部设备与外部设备之间并行工作, 以提高整个系统的工作效率。

(4) 根据不同的设备特点, 采用优化策略, 使对具体设备的使用更趋合理和有效。

四、文件管理

程序与数据都是以文件的形式存放在外存储器(如硬盘、软盘)上的, 是计算机系统的软件资源。用户总是通过文件的名称对所需要的文件进行访问, 这就是所谓的“按名存取”方式。为了满足用户的这种需求, 操作系统文件管理的主要工作如下:

(1) 维持一个目录表, 里面登记有每一个文件的名称和有关信息。这样, 当用户通过文件名来访问某文件时, 就可以通过查目录表找到它的目录项, 从而完成所需的读/写操作。

(2) 由于文件都存放在外存, 因此要随时记录外存里文件存储空间的使用情况, 了解哪些已经分配, 哪些为空闲待分配。

(3) 制定文件存储空间的分配策略, 实施具体的分配和回收。

(4) 确保存放在外存中文件的安全、保密和共享。

(5) 提供一系列文件使用命令, 以使用户能对文件进行存取、检索和更新等操作。

第三节 操作系统的特征

操作系统作为一种系统软件, 区别于其他软件, 有着其独特的特征。

一、并发性

并发和并行很相似, 但它们是有区别的。并发就是指两个以上的事物在同一时间间隔内发生。并行是指两个或两个以上的事物在同一时刻发生。操作系统的并发性具体体现在以下两个方面: 用户程序与用户程序之间的并发执行; 用户程序和操作系统程序之间的并发执行。

在单处理机环境里, 某一时刻, CPU 上仅能运行一个程序, 程序是交替执行的。但多道程序设计技术的引入, 使得在一个小的时间间隔内多个程序并发执行成为可能, 所以准确地说, 宏观上讲, 多个程序是同时向前推进的; 微观上讲, 程序之间是交替执行的。

应该注意的是, 无论是在什么样的环境中, 这里指的并发都是在一个操作系统统一控制下的并发。在两个独立的操作系统控制下的计算机中, 运行在其上的程序是并行的, 但这不是我们研究的并发性。

二、共享性

共享是指计算机中的各种资源可以供在其上运行的程序共同享用。这种共享是在操作系统的统一控制下实现的。共享有两种方式: 互斥共享方式和共同访问方式。

并发和共享是一对孪生兄弟, 是操作系统中两个最基本的特征, 它们互为存在条件。一方面资源的共享是以程序的并发执行为条件的。若系统不允许程序并发执行, 也就不存在共享的问题。另一方面, 只有操作系统提供对资源共享的可能, 才能使程序真正做到并发执行。

三、虚拟性

在操作系统中,所谓的虚拟是指通过某种技术手段把一个物理实体变成多个逻辑上的对应物,物理实体是实际存在的,而逻辑实体是虚拟的,是用户的一种感觉。

比如,在多道程序设计下的分时系统中,虽然只有一个 CPU,但每个用户都认为有一个 CPU 专门在为其服务。也就是说,利用多道程序设计技术把一个物理上的 CPU 变成了多个逻辑上的 CPU,这些逻辑上的 CPU 也称为虚处理机。

同样,SPOOLing 技术把一台物理上的输入/输出设备变成了多个逻辑上的输入/输出设备。

综上所述,操作系统的虚拟性主要是通过分时使用的方法实现的。所有逻辑设备的工作量之和必然等于实际物理设备的工作量。

四、不确定性

操作系统是在一个不确定的环境中运行,也就是说,人们不能对目前所运行程序的行为做出判断。因为在多道程序环境下,进程的执行是“走走停停”的,在内存中的多个程序,何时运行?何时暂停?以怎样的速度向前推进?每个程序需要多少时间才能完成?都是不可预知的。我们无法知道运行着的程序会在什么时候做什么事情,因而一般来说无法确切地知道操作系统正处于什么样的状态。但是,这并不能说不能很好地控制操作系统,这种不确定性是允许的。不管怎样,只要在相同的环境下,一个程序无论执行多少次,其运行结果是相同的。但是它执行的时间可能不同,因为它每次执行时操作系统要处理的状况可能是不同的。

第四节 操作系统的类型

一、批处理操作系统

批处理操作系统(下称批处理系统)是一种基本的操作系统类型。在该系统中,用户的作业(包括程序、数据及程序的处理步骤)被成批地输入到计算机中,然后在操作系统的控制下,用户的作业自动地执行。

(一) 批处理系统的形式

批处理系统在计算机的产生和发展过程中曾经起过很重要的作用,有两种批处理系统的形式:单道批处理系统和多道批处理系统。

1. 单道批处理系统

单道批处理系统(Simple Batch System)是在早期的计算机系统上实现的,因为当时的计算机系统非常昂贵,为了充分地利用它,尽量让其连续地运行,减少计算机的空闲时间,通常将一批作业以脱机的形式输入到磁带上,然后在系统的控制下使这批作业连续自动地执行。

自动处理过程是这样的,首先在系统的控制下将磁带上的第一个作业装入内存,并将控制权交给作业,使之运行。当作业处理完成后再把控制权交还给系统,系统再将磁带上的第二个作业调入内存,使第二个作业运行,如此下去,一直到磁带上的作业全部运行完毕。

由于系统对作业的处理是成批进行的,内存中始终只保持一个作业,故称为单道批处理

系统。

单道批处理系统是最早出现的一种操作系统,严格地说,它的功能比较简单,并非是我们今天所理解的操作系统,但它比人工操作方式已经有了很大的进步。单道批处理系统有以下主要特征:

- (1) 自动性。磁带上的一批作业能自动地逐一执行,无需人工干预;
- (2) 顺序性。磁带上的作业进入内存的顺序与各道作业的完成顺序完全相同,也就是说先调入内存的作业先完成;
- (3) 单道性。内存中一直只有一道作业,一个作业完成并退出内存之后另一个作业才可进入内存。

2. 多道批处理系统

多道批处理系统是在多道程序设计技术引入后产生的。在该系统中,用户提交的作业先存放在外存上并排成一个队列,该队列被称为“后备队列”。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,内存中的作业由系统为之建立进程,在进程调度程序的统一调度下,交替地执行若干个进程。

多道批处理系统比单道批处理系统有以下优点:

(1) CPU 的利用率得到提高。在单道批处理系统中,由于内存中只有一道作业,当作业发出输入/输出请求后,CPU 空闲,必须等到输入/输出完成后才能继续运行。因为输入/输出设备的速度很低,致使 CPU 的空闲时间很长。而在引入了多道程序设计之后的多道批处理系统中,内存中存放了多道作业,它们是在 CPU 上交替执行的。这样,当一个程序因为请求输入/输出暂停时,系统会调度另一道程序使其运行,从而使 CPU 处于忙碌状态,CPU 的利用率也就高了;

(2) 提高了内存和输入/输出设备的利用率。作业调度程序在调度时,可以考虑到作业的大小和内存空间的容量。如果运行的作业都比较小,在单道批处理系统中因为内存中只能有一道作业,同样也造成内存空间的浪费。而在多道批处理系统中,作业大时,内存中的作业数可以少些,当作业小时,内存中的作业可以多些,从而充分利用了内存空间,也提高了内存的利用率。同样的道理,输入/输出设备也从单道环境中的串行使用转化为多道环境下的并发执行,从而提高了输入/输出设备的利用率;

(3) 增加了系统的吞吐量。由于 CPU、内存、输入/输出在多道环境下都忙碌了起来,必然大幅度提高系统的吞吐量,从而降低了作业运行的成本。

多道批处理系统有以下主要特征:

- (1) 多道性。内存中同时驻留多道作业,使它们并发地执行。提高了资源的利用率;
- (2) 无序性。作业进入内存的顺序与各道作业的完成顺序无严格的对应关系。也就是说先调入内存的作业可能较后或最后完成;
- (3) 调度性。从作业提交到完成,需要经过两次调度:作业调度和进程调度。

(二) 批处理系统的优缺点

批处理系统早在 20 世纪 60 年代就产生了,至今仍然配置在大多数的大、中型机上,其旺盛的生命力表明它有一定的优点。主要表现在以下两个方面:

(1) 资源利用率高。在多道批处理系统中,由于内存中装有多道程序,它们共享资源,保持系统中的资源处于忙碌状态,从而使各种资源得到充分利用;

(2) 系统吞吐量大。系统吞吐量是指单位时间内所完成的总工作量。由于系统中的资源一直处于忙碌的状态,并且,仅当作业完成或出错不能运行时操作系统才更换另一个作业运行,系统的开销小。

批处理系统也有一定的缺点,主要表现在以下两个方面:

(1) 平均周转时间长。作业的周转时间是指从作业进入系统开始直至其完成并退出系统为止所经历的时间。在批处理系统中,由于作业要排队,依次进行处理,因而作业的周转时间较长,通常要几个小时,甚至几天;

(2) 无交互能力。用户一旦把作业提交直到作业完成,都不能与自己的作业进行交互,这对修改和调试程序极为不便。可能只是需要修改程序中的一个小错误,就会导致用户等上几小时,才能知道程序的运行结果。

(三) 批处理系统解决的问题

批处理系统是一种有效但十分复杂的系统。为使系统中的多道程序间协调地运行,需要解决下面的问题。

(1) 处理机管理问题。解决在多道程序之间如何分配处理机,使 CPU 既能满足各个程序运行的要求,又尽量使其处于忙碌状态,以及当 CPU 分配给某个程序后,何时收回处理机等一系列问题。

(2) 内存管理问题。要使多个程序交替地占用 CPU,这些程序应先调入内存,操作系统要为每道程序分配必要的内存空间,并且要防止某个程序出现异常时不会破坏其他程序。

(3) 输入/输出设备管理问题。系统中的各种输入/输出设备也被多个程序所共享,系统应如何分配这些输入/输出设备,如何做到既方便用户使用设备又能提高设备的利用率。

(4) 文件管理问题。用户的程序和数据以及系统的程序和数据是以文件的形式存放在系统中的,应考虑如何组织和管理这些文件,才能便于用户使用。同时还要考虑数据的安全性和一致性。

(5) 作业管理问题。对于进入系统要运行的作业,有的属于计算型的(以计算为主),有的属于 I/O 型的(以 I/O 为主)。有的作业非常重要,有的作业又要求能及时响应,要解决如何组织这些作业,才能使作业的要求都能得到满足。

二、分时操作系统

1. 分时技术的引入

让用户通过控制台直接操作、控制自己程序的运行,是用户欢迎的一种工作方式,因为在这种方式下,用户可以向计算机发出各种控制命令,使系统按自己的意图控制程序的运行。另外,在程序运行的过程中系统可以输出一些必要的信息,如报告程序的运行情况、操作结果等,以便让用户根据输出信息决定下一步的工作。

20 世纪 60 年代,产生了一种新的操作系统,这就是分时操作系统(下称分时系统)。在分时系统中,一台计算机和许多终端设备连接,每个用户可以通过终端向计算机发出命令,请求完成某项工作,而操作系统要分析从终端发来的命令,完成用户提出的要求,重复上述交互会话过程,直到用户完成预计的全部工作为止。

在分时系统中,计算机能够同时为多个终端用户服务,而且能在很短的时间内响应用户的要求。因为系统采用了分时技术,把处理机时间划分成很短的时间片轮流地分配给各个联机用户使用。如果某个作业在分配给它的时间片用完之后程序还未执行完毕,该作业就暂时

中断,等待下一轮继续执行,此时将处理机让给另一个作业使用。这样,每个用户的要求都能快速地得到响应,给每个用户的印象是,他独占一台计算机。

2. 分时系统中要解决的关键问题

为了实现分时系统,必须解决一系列的问题。其中,最关键的问题是如何使用户与自己的作业交互,即当用户在自己的终端上键入命令时系统应能及时接收、处理该命令,并将处理结果返回给用户。接着用户可键入下一条命令,此即人-机交互。应当强调指出,即使有多个用户同时通过自己的键盘键入命令,系统也应能全部地及时接收并处理。

(1) 及时接收。要及时接收用户输入的命令或数据并不困难,只需要在系统中配置一块多路卡。例如,当要在主机上连接8个终端时,需要配置8用户的多路卡。多路卡的作用是使主机能同时接收用户从各个终端上输入的数据。此外,还需为每个终端配置一个缓冲区,用来暂存用户键入的命令。

(2) 及时处理。人-机交互的关键是使用户键入命令后能及时地控制自己的作业运行或修改自己的作业。为此,各个用户的作业都必须在内存中,且应能频繁地获得处理机而运行,否则,用户键入的命令将无法作用到自己的作业上。为了实现人-机交互应该做到:

(1) 使所有的用户作业都直接进入内存;

(2) 在不长的时间内,例如3秒钟内,就能使每个作业都运行一次,这样方便使用户键入的命令获得及时的处理。

3. 分时系统的实现形式

分时系统在其历史发展过程中经历了从简单到复杂的过程,分时系统的实现形式有以下三种类型。

(1) 单道分时系统。第一个分时系统 CCTS 是 20 世纪 60 年代初,由美国麻省理工学院建立的。它属于单道分时系统。其特点是内存中只驻留一个作业(程序),其他的作业都放在外存上。每当内存中的作业运行一个时间片之后,便被调至外存(称为调出),再从外存上选一个作业装入内存(称为调入)并运行一个时间片,依此方法使所有的作业都能在一个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,系统开销大,系统的性能也较差。

(2) 具有“前台”和“后台”的分时系统。在单道分时系统中,作业调进调出时 CPU 空闲,内存中的作业在执行输入/输出时 CPU 也空闲。为了充分利用 CPU,引入了“前台”和“后台”的概念。

在具有“前台”和“后台”的系统中,内存被划分为“前台区”和“后台区”两部分,“前台区”存放按时间片调进调出内存的作业流,“后台区”存放批处理作业。当“前台”调进、调出,或前台无作业可运行时,才运行“后台区”中的作业。

(3) 多道分时系统。将多道程序设计技术引入分时系统中后,可在内存中存放多道作业,由系统将具备运行条件的所有作业排成一个队列,使它们依次轮流地获得一个时间片来运行。系统具有较好的系统性能。现代的分时系统都属于多道分时系统。

4. 分时系统的特征

分时系统是操作系统的一种类型。它一般采用时间片轮转法,使一台计算机同时为多个终端用户服务。保证每个用户的要求都能得到足够快的响应,并提供交互会话功能。分时系

系统与批处理系统的主要区别在于,所有用户都是通过联机终端直接与计算机交互,对自己的程序有一定的控制能力。当今流行的操作系统中,如 OS₂、Unix、Windows 系列使用的都是分时系统。

分时系统具有以下特征:

(1) 多路性。指众多联机用户可以同时使用一台计算机,也称同时性。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,共享系统中的资源;而微观上,则是每个用户轮流地使用计算机。

(2) 独占性。由分时操作系统采用时间片轮转法使一台计算机同时为多个终端用户服务,因此,客观效果是这些用户彼此之间感觉不到其他用户也在使用这台计算机,好像自己独占一台计算机一样。一般分时系统的响应时间控制在 3 秒内,用户就会感到满意,因为这段时间用户感觉不到等待。

(3) 交互性。用户和计算机之间进行“会话”。用户从终端键入命令,提出要求,系统收到命令后分析用户的要求并完成之,然后把运算结果通过屏幕或打印机告诉用户,用户可以根据运算结果提出下一步的要求,直到完成全部工作。

(4) 及时性。用户请求能在很短的时间内获得响应,此时,时间间隔是以人们所能接受的等待时间确定的,通常为 2~3 秒。

需要说明的是,在一个具体的计算机系统中,往往配置的操作系统是结合了批处理能力和分时能力的。它以前/后台的方式提供服务,前台以分时方式为多个联机终端服务,当终端用户较少或没有终端用户时,系统采用批处理方式处理后台的作业。

三、实时操作系统

1. 实时系统的引入

早期的计算机基本上用于科学和工程问题的数值计算。在 20 世纪 50 年代的后期,计算机开始用于生产过程的控制,形成实时操作系统(下称实时系统)。到了 60 年代中期,计算机进入第三代,机器性能得到了极大的提高,整个计算机系统的功能大大增强了,导致了计算机的应用领域越来越宽广。例如,钢铁、纺织、制药生产的过程控制,航空航天系统中的实时控制。更为重要的是计算机广泛用于信息管理,如仓库管理、医疗诊断、教学、气象、地质勘探,直到图书检索、飞机订票、银行储蓄、出版编辑等。

实时系统是操作系统的又一种类型。对外部输入的信息,实时系统能够在规定的时间内处理完毕并做出反应。“实时”二字的含义是指计算机对于外来信息能够及时处理,并在被控对象允许的范围内做出快速反应。实时系统对响应时间的要求比分时系统更高,一般要求秒级、毫秒级甚至微秒级的响应时间。

2. 实时系统的类型

电子计算机应用到实时控制中,配置上实时系统,组成各种各样的实时系统。实时系统按使用方式的不同可以分成下述两类。

(1) 实时控制系统。当把计算机用于生产过程的控制,形成以计算机为中心的控制系统时,系统要求能够实时采集现场数据,并对所采集的数据及时进行处理,进而自动地控制相应的执行机构,使某些参数能按预定的规律变化,以保证产品的质量并提高产量。类似地,也可将计算机用于武器的控制,如导弹的自动控制、飞机的自动驾驶等。

微机的最早应用之一就是提供过程控制和提供环境监控。图 1-1 是火灾报警控制系统

的简要工作原理。由一台微型计算机控制系统不断给探测器发出监测脉冲,探测器实时采集房间中的烟浓度和温度值,并进行模拟/数字信号(A/D)的转换后送到计算机控制系统中,计算机将所收到的数值按照一定的算法进行处理,判断房间是否发生火灾。若有火灾发生时,计算机控制系统就使探测器上的指示点亮,并触发相应楼层的声、光报警装置,必要时启动相应的灭火设备,这样就起到了自动报警及灭火的作用。

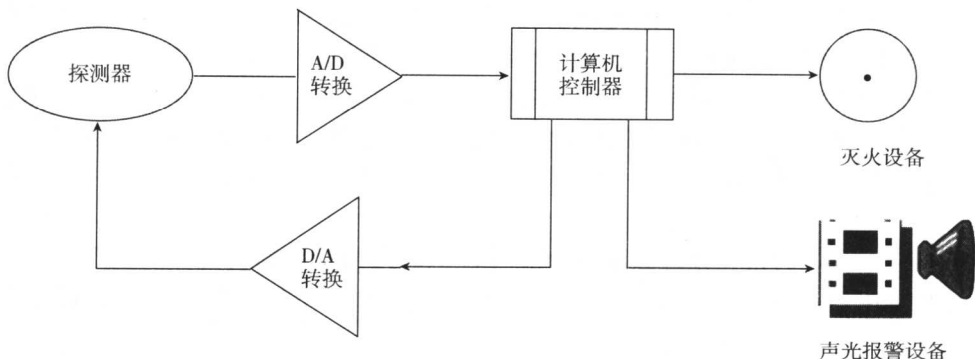


图 1-1 火灾报警控制系统

(2) 实时信息处理系统。通常,把要求对信息进行实时处理的系统,称为实时信息处理系统。该系统由一台或多台主机通过通信线路连接成百上千台远程终端,计算机接收从远程终端发来的服务请求,根据用户提出的问题,对信息进行检索和处理,并在很短的时间内为用户做出正确的回答。典型的实时信息处理系统有:飞机自动订票系统、情报检索系统等。用户可通过终端向计算机提出某种要求,而计算机处理后会通过终端将处理结果通知给用户。

3. 实时系统的特殊要求

实时系统除了应具备一般操作系统的功能外,还有下述一些特殊的要求。

(1) 高可靠性。实时系统一个重要的设计目标就是高可靠性。尤其在实时控制系统中,任何故障都可能造成难以弥补的后果。因此在实时系统中必须采用相应的硬件和软件容错技术来提高系统的可靠性。常用的技术之一是双工体制,即用两台相同的计算机并行运行,其中一台作为主机,实现实时控制或实时信息处理,另一台作为备用机,一旦主机出现故障,备用机可以接替其工作继续运行。

(2) 过载防护。系统必须设置某种防护机构,以保证系统出现过载时,仍能正常工作。当系统出现短暂的过载时,可通过配置适量的缓冲区予以平滑。当系统出现持续的过载时,要采取某种措施来防止超载。比如,一旦出现过载时可以采用拒绝新任务的方法,也可以抛弃一些不重要的任务,以保证重要的任务顺利地完成任务。

(3) 对截止时间的要求。实时系统因控制着某个外部事件,往往带有某种程度的紧迫性。对要处理的实时任务有一些是呈周期性的实时任务,即按指定的周期循环执行,以便周期性地控制某个外部事件。有一些实时任务没有周期性,但必须有一个截止时间,它又分为:开始截止时间和完成截止时间。开始截止时间要求必须在某时间以前开始执行指定的操作;完成截止时间要求某任务必须在某时间以前完成。

4. 实时系统和分时系统的比较

分时系统具有的多路性、独立性、及时性和交互性这四大特征,实时系统也同样具备,