

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

操作系统教程

Operating System

宗大华 宗涛 编著

- 不追求深奥的理论，突出操作系统的实用性
- 不追求玄妙的抽象，注重操作系统的简明性
- 不追求内容的全面，关注操作系统的基本性



精品系列

 人民邮电出版社
POSTS & TELECOM PRESS

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

操作系统教程

Operating System

宗大华 宗涛 编著



精品系列

人民邮电出版社

北京

图书在版编目 (CIP) 数据

操作系统教程 / 宗大华, 宗涛编著. —北京: 人民邮电出版社, 2008.11

21 世纪高等学校计算机规划教材
ISBN 978-7-115-18790-1

I. 操… II. ①宗…②宗… 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 136504 号

内 容 提 要

操作系统是现代计算机系统中最重要和最基本的系统软件。本书共 9 章。第 1 章概述操作系统在计算机系统中的地位、基本功能以及分类; 第 2 章到第 5 章从资源管理的角度出发, 讲述操作系统对处理机、存储器等软硬件资源的各种管理策略; 第 6 章集中介绍资源管理中的互斥、同步问题, 以及解决办法; 第 7 章到第 9 章对 3 个典型操作系统 Windows、Linux、MS-DOS 做简略的剖析, 以求加深对操作系统原理的理解。

本书不追求深奥的理论, 而是突出实用性; 不追求玄妙的抽象, 而是注重简明性; 不追求内容的全面, 而是关注基本性。本书可作为高等院校理工科操作系统课程的教材。

21 世纪高等学校计算机规划教材

操作系统教程

-
- ◆ 编 著 宗大华 宗 涛
责任编辑 滑 玉
执行编辑 武恩玉
 - ◆ 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>
北京顺义振华印刷厂印刷
 - ◆ 开本: 787×1092 1/16
印张: 16.5
字数: 433 千字 2008 年 11 月第 1 版
印数: 1—3 000 册 2008 年 11 月北京第 1 次印刷

ISBN 978-7-115-18790-1/TP

定价: 29.80 元

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

目 录

第 1 章 操作系统概述	1	习题	38
1.1 计算机系统	1	第 3 章 存储管理	41
1.1.1 硬件与软件	1	3.1 固定分区存储管理	41
1.1.2 操作系统的形成	2	3.1.1 地址重定位	41
1.2 操作系统的定义与功能	3	3.1.2 地址的静态重定位	42
1.2.1 操作系统的定义	3	3.1.3 单一连续分区存储管理	43
1.2.2 操作系统的功能	4	3.1.4 固定分区存储管理	45
1.3 操作系统的种类	6	3.2 可变分区存储管理	47
1.3.1 批处理操作系统	6	3.2.1 可变分区存储管理的基本思想	47
1.3.2 分时操作系统	7	3.2.2 地址的动态重定位	50
1.3.3 实时操作系统	8	3.2.3 空闲区的合并	51
1.3.4 网络操作系统	9	3.2.4 分区的管理与组织方式	52
习题	10	3.2.5 空闲分区的分配算法	56
第 2 章 处理机管理	11	3.3 分页式存储管理	57
2.1 进程	11	3.3.1 分页式存储管理的基本思想	57
2.1.1 多道程序设计	11	3.3.2 分页式存储管理的地址转换	59
2.1.2 进程的定义	13	3.3.3 内存块的分配与回收	64
2.1.3 进程的特征	14	3.4 分段式及段页式存储管理	65
2.1.4 进程的基本状态	15	3.4.1 分段式存储管理的基本思想	65
2.2 进程控制块	16	3.4.2 分段式存储管理的地址转换 过程	67
2.2.1 进程的三个组成部分	16	3.4.3 存储保护与共享	68
2.2.2 进程控制块的内容	16	3.4.4 段页式存储管理的基本思想	70
2.2.3 进程控制块队列	18	3.4.5 段页式存储管理的地址转换过程	70
2.3 进程的调度与管理	19	3.5 虚拟存储与请求页式存储管理	71
2.3.1 进程调度算法	19	3.5.1 虚拟存储器的概念	71
2.3.2 进程管理的基本原语	24	3.5.2 请求分页式存储管理的基本思想	72
2.4 线程	25	3.5.3 缺页中断的处理	73
2.4.1 线程的概念	25	3.5.4 页面淘汰算法	77
2.4.2 进程与线程的关系	27	习题	83
2.5 作业调度	29	第 4 章 设备管理	86
2.5.1 用户与操作系统的两种接口	29	4.1 概述	86
2.5.2 作业与作业管理	32	4.1.1 计算机设备的分类	86
2.5.3 作业的调度算法	33		

4.1.2 设备管理的目标与功能	89	第 6 章 进程间的制约关系	135
4.2 输入/输出的处理步骤	90	6.1 进程间的制约关系	135
4.2.1 I/O 请求的提出	90	6.1.1 与时间有关的错误	135
4.2.2 对 I/O 请求的管理	91	6.1.2 竞争资源—互斥	138
4.2.3 I/O 请求的具体实现	91	6.1.3 协同工作—同步	140
4.3 设备的分配与调度算法	92	6.2 信号量与 P、V 操作	141
4.3.1 管理设备时的数据结构	92	6.2.1 信号量与 P、V 操作的定义	141
4.3.2 独享设备的分配	94	6.2.2 用 P、V 操作实现互斥	142
4.3.3 共享磁盘的调度	96	6.2.3 用 P、V 操作实现同步	143
4.4 数据传输的方式	100	6.2.4 用 P、V 操作实现资源分配	145
4.4.1 设备控制器	100	6.2.5 互斥/同步的样例分析	147
4.4.2 程序循环测试方式	101	6.3 死锁、高级进程通信	155
4.4.3 中断方式	102	6.3.1 死锁与产生死锁的必要条件	155
4.4.4 直接存储器存取方式	103	6.3.2 死锁的预防	158
4.4.5 通道方式	104	6.3.3 死锁的避免	159
4.5 设备管理中的若干技术	105	6.3.4 死锁的检测并恢复	162
4.5.1 I/O 缓冲技术	105	6.3.5 高级进程通信	164
4.5.2 虚拟设备与 SPOOLing 技术	106	习题	166
4.5.3 磁盘数据的组织技术	108	第 7 章 实例分析：Windows 2000	
习题	111	操作系统	169
第 5 章 文件管理	113	7.1 Windows 2000 的处理机管理	169
5.1 文件的结构	113	7.1.1 Windows 2000 的结构	169
5.1.1 文件与文件系统	113	7.1.2 Windows 2000 的进程和线程	171
5.1.2 文件的逻辑结构	115	7.1.3 Windows 2000 的线程调度	173
5.1.3 文件的物理结构	116	7.2 Windows 2000 的存储管理	178
5.1.4 文件的存取	120	7.2.1 Windows 2000 进程的空间	
5.2 磁盘存储空间的管理	120	布局	178
5.2.1 位示图	121	7.2.2 Windows 2000 的地址变换	
5.2.2 空闲区表	122	机构	179
5.2.3 空闲块链	123	7.2.3 Windows 2000 对内存的管理	181
5.3 文件管理与目录结构	124	7.2.4 Windows 2000 的页面调度	183
5.3.1 文件控制块与目录	124	7.3 Windows 2000 的文件管理	185
5.3.2 目录的层次结构	125	7.3.1 Windows 2000 文件系统综述	185
5.3.3 “按名存取”的实现	127	7.3.2 NTFS 的主控文件表 (MFT)	189
5.4 文件的使用	129	7.3.3 NTFS 文件和目录的结构	191
5.4.1 文件的共享	129	7.3.4 NTFS 对可恢复性的支持	192
5.4.2 文件的保护	130	7.4 Windows 2000 的设备管理	195
5.4.3 文件的操作	132	7.4.1 Windows 2000 设备管理综述	195
习题	132		

7.4.2 Windows 2000 单层驱动程序的 I/O 处理	197	习题	227
7.4.3 Windows 2000 两级中断处理过程	199	第 9 章 实例分析：MS-DOS 操作系统	230
习题	201	9.1 MS-DOS 的处理机管理	230
第 8 章 实例分析：Linux 操作系统	203	9.1.1 MS-DOS 的基本组成	230
8.1 Linux 的处理机管理	203	9.1.2 MS-DOS 的进程	232
8.1.1 Linux 的进程	203	9.1.3 MS-DOS 的作业管理	234
8.1.2 Linux 的进程调度	206	9.2 MS-DOS 的存储管理	236
8.1.3 Linux 进程间的通信——消息队列	208	9.2.1 MS-DOS 对常规内存的管理	236
8.2 Linux 的存储管理	210	9.2.2 PC 机地址的构成——“分段”表示法	239
8.2.1 Linux 的虚拟存储空间	210	9.2.3 MS-DOS 的各种内存区域	241
8.2.2 管理虚拟存储空间的数据结构	212	9.3 MS-DOS 的文件管理	242
8.2.3 管理内存空间的数据结构	213	9.3.1 MS-DOS 文件管理综述	242
8.2.4 内存区的分配和页面淘汰策略	214	9.3.2 MS-DOS 的文件分配表 (FAT)	245
8.3 Linux 的文件管理	216	9.3.3 MS-DOS 的树型目录结构	247
8.3.1 Linux 文件系统的构成	216	9.3.4 MS-DOS 文件访问的实现	248
8.3.2 Ext2 对磁盘的组织	217	9.4 MS-DOS 的设备管理	250
8.3.3 Ext2 文件的物理结构	220	9.4.1 MS-DOS 设备管理综述	250
8.3.4 虚拟文件系统 VFS 的数据结构	221	9.4.2 MS-DOS 设备驱动程序的构成	251
8.4 Linux 的设备管理	224	9.4.3 MS-DOS 对块设备的管理	252
8.4.1 Linux 设备管理概述	224	9.4.4 MS-DOS 对字符设备的管理	254
8.4.2 Linux 对字符设备的管理	225	习题	254
8.4.3 Linux 对块设备的管理	226	参考文献	257

第 1 章

操作系统概述

计算机是人类社会 20 世纪最伟大的创造。自 1946 年诞生第一台计算机至今的 60 多年里，其技术得到了突飞猛进的发展。目前，它不仅被广泛应用于科学计算、过程控制、数据处理以及军事技术等领域，而且也渗透到办公、教育、家庭等方方面面。计算机已成为社会信息化的重要支柱，成为人类文明高度发展的象征。

本章将讲述以下 3 方面的内容：

- (1) 计算机必备的系统软件——操作系统的形成过程；
- (2) 操作系统的 4 大功能；
- (3) 简述 4 类基本操作系统。

1.1 计算机系统

1.1.1 硬件与软件

一个完整的计算机系统由硬件系统和软件系统两大部分组成。

计算机硬件系统是各种物理设备的总称，是完成工作任务的物质基础。按功能划分，可以把硬件划分成 5 大块：运算器、控制器、存储器、输入设备以及输出设备，其中运算器和控制器常被称为中央处理机（CPU），如图 1-1 所示。图中带箭头的实线代表控制信号，带箭头的细虚线代表数据传输。

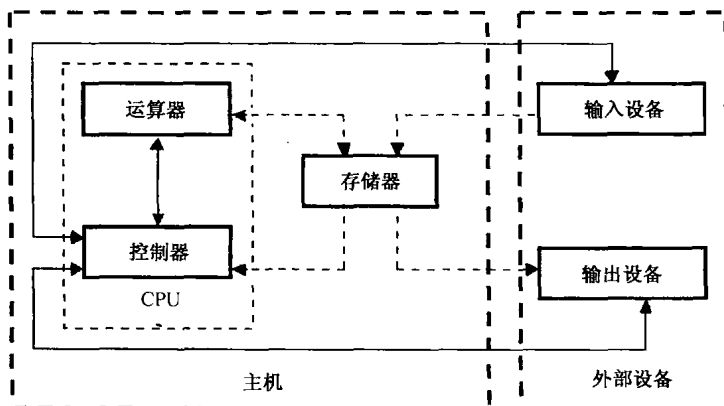


图 1-1 计算机硬件系统的组成

计算机软件系统是指程序和与程序相关的文档的集合,按功能划分,软件可分为系统软件和应用软件。系统软件是指由计算机生产厂家提供、具有通用功能的软件,如操作系统、语言处理程序(如 C 语言编译程序)、数据库管理系统以及各种完成服务功能的程序。应用软件是指为解决实际问题而研制的那些软件,涉及计算机应用的各个领域,如各种管理软件、用于工程计算的软件包、辅助设计软件以及过程控制软件等。

1.1.2 操作系统的形成

通常,把未配置任何软件的计算机称为“裸机”。如果让用户直接面对裸机,事事都深入到计算机的硬件中去,那么他们的精力就绝对不可能集中在如何用计算机解决自己的实际问题上,计算机本身的效率也不可能充分发挥出来。

举例说,要在一台 PC 上进行硬盘读操作,使用者至少应该把磁盘地址、内存地址、字节数、操作类型(读/写)等具体值装入到特定的硬件寄存器中,否则根本谈不上完成预定的输入/输出任务。实际上,对许多 I/O 设备而言,往往会要求比这更多的操作参数。在输入/输出结束后,还需要对设备返回的诸多状态加以判别。

又如,某计算机内存存储器可供用户使用的容量为 576KB,若现在装入的用户程序占用其中的 360KB,那么余下的 216KB 被闲置了。想象一下,如果能够在内存中装入多个程序,如在 216KB 中再装一个需要存储量 116KB 的程序,当第一个程序等待输入/输出完成暂时不占用 CPU 时,让第二个程序投入运行,那么,整个计算机系统的利用率就会比原来的大为提高。理由如下:

(1) 内存浪费得少了,原来浪费 216KB,现在只浪费 100KB;

(2) CPU 比原来更加忙碌了,在第一个程序等待输入/输出完成时,原来 CPU 只能够采取空转的方式来等待,现在可以让它去执行第二个程序;

(3) 在 CPU 执行第二个程序时,它与第一个程序启动的输入/输出设备呈现并行工作的态势。

可见,为了从复杂的硬件控制中脱出身来,合理有效地使用计算机系统,给用户使用计算机提供必要的方便,最好的解决办法就是开发一种软件,通过它来管理整个系统,发挥系统的潜在能力,达到扩展系统功能、方便用户使用的目的。实际应用的需要,就是“操作系统”这一软件呼之欲出的根本原因。

第一台电子管计算机出现后的若干年(1946~1958年),计算机上并没有名为“操作系统”的这种软件。那时计算机的运行速度慢,外部设备少,因此程序的装入、调试以及控制程序的运行等工作,全部由上机的人员通过按动控制台上的一排排开关和按钮来实现。这一时代的特点是人工完成上、下机操作,一台计算机被一个用户所独占。

1958年,计算机进入了晶体管时代(1958~1964年)。这时计算机的速度、存储容量、外部设备的功能和种类等都有了很大的发展,慢速的人工操作与快速的计算机处理能力之间显得很不协调,出现了所谓的“人-机矛盾”。例如,有一个程序通过 3min 的安装等手工操作后,在运算速度为 1 万次/s 的计算机上用 1h 得到了结果。这时手工操作与程序运行时间之比为 1:20。把这个程序拿到第二代速度为 60 万次/s 的机器上运行,它只需花费 CPU 时间 1min 即可得到结果。如果在这种高速的机器上仍然坚持手工操作,那么这时手工操作与程序运行时间之比为 3:1,这种比例是难以让人接受的。

正是这种“人-机矛盾”,向软件设计人员提出了“让计算机自动控制用户作业的运行,废除上、下机手工交接”的要求。为了达到这个目的,需要用户一方在编写自己程序时,还要编写“作业说明书”,详细规定程序运行的步骤,与程序、数据一起提交给系统;而系统一方需要设计一个

“管理程序”（也称监督程序），它的功能是从磁盘上读入第一个作业的作业说明书，按照它的规定控制该作业执行。这个作业运行结束后，它又从磁盘上读入第二个作业的作业说明书，继而执行之。这一过程一直进行到提交给系统的一批作业全部执行完毕时为止，如图 1-2 所示。由于这种系统一次集中处理一批用户作业，故被称为“批处理系统”，其管理程序就是现今操作系统的雏形。这个时代的特点是对一批作业自动进行处理，没有人工交接。在一个用户作业运行时，仍独占计算机。

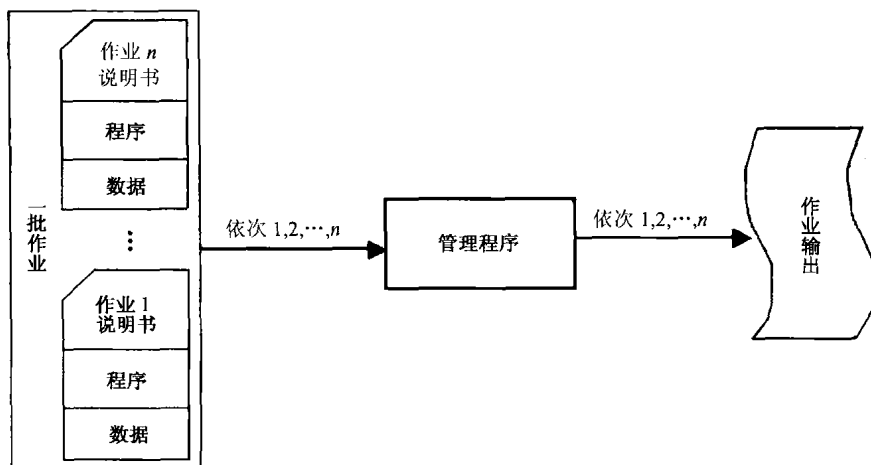


图 1-2 批处理系统示意图

1964 年后，计算机进入了集成电路和大规模集成电路时代。这时，硬件又有了长足的发展，中断和通道技术的出现，为输入/输出设备和中央处理机并行操作奠定了物质基础。另外，随着计算机应用的日益广泛，也要求进一步发展和扩大管理程序的功能，希望它能够最大限度地挖掘计算机系统本身的潜在能力。这时，人们开始把 CPU、存储器、外部设备以及各种软件都视为计算机系统的“资源”，提出不仅要合理地，而且要高效地使用这些资源。为此，在软件设计上提出了“多道程序设计”的技术，即在计算机内存中同时存放几个相互独立的程序，让它们去“共享”、去“竞争”系统中的这些资源，使系统中的各种资源尽可能地满负荷工作，从而提高整个计算机系统的使用效率。具有这种功能的软件就是“操作系统”。

操作系统可以被看作是计算机系统的核心，统管整个系统的所有资源，制定各种资源的分配策略，调度系统中运行的用户程序，协调它们对资源的需求，从而使整个系统在高效、有序的环境里工作。

1.2 操作系统的定义与功能

1.2.1 操作系统的定义

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，如图 1-3 所示。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。通常说操作系统为用户提供了一台功能经过扩展了的机器或“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，系统中的各种硬软件资源得到了更有效的利用，机器的工作流程更为合理与协调。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。

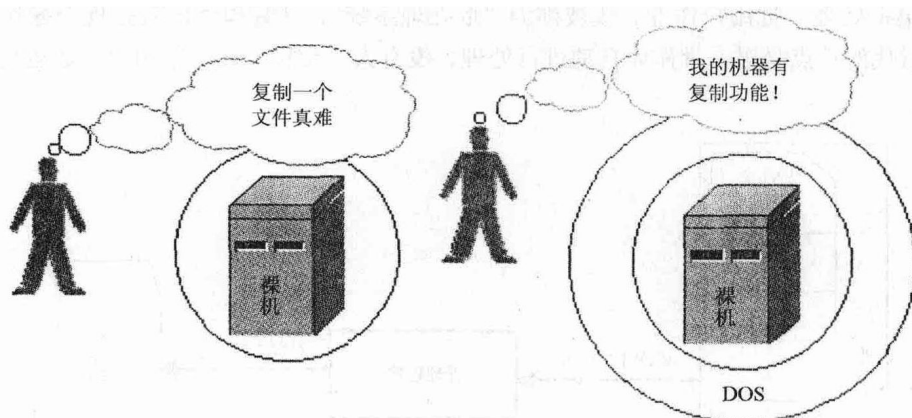


图 1-3 操作系统提供了一台虚拟机

至此，我们可以把操作系统定义为：操作系统是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计算机的一个大型程序。

1.2.2 操作系统的功能

从资源管理的角度看，操作系统应该具有 5 个方面的功能：处理机管理、存储管理、设备管理、文件管理以及作业管理。这 5 大部分相互配合，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。本书将处理机管理与作业管理合并在一起介绍。下面分 4 个方面对操作系统的功能做一个简略的说明。

1. 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，只有获得 CPU，才能真正得以运行。如前已经提及的，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 CPU 重新分配给其他可运行的作业程序使用，从而提高它的利用率。

为此，处理机管理的主要工作如下。

(1) 记住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，它是在候选程序中进行挑选时应遵循的原则。

(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来参与对 CPU 的竞争，这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。面对“僧多粥少”的局面，不控制“僧”的数量，就肯定会影响系统效率的发挥。所以在有的操作系统中，实行处理机的两级调度：第一级是作业调度，涉及作业管理；第二级才是处理机调度，属于处理机管理。

作业管理的主要工作如下。

(1) 记住提交给系统诸作业（一般存放在磁盘）的状态，以及对系统资源的需求信息。

(2) 制定作业调度策略，在需要从磁盘的候选作业中选择作业进入内存，参与对 CPU 的竞争。

(3) 为用户提供一个使用系统的良好环境，以便有效地组织自己的工作流。

2. 存储管理

存储器是计算机的记忆装置。根据速度和价格，从层次上它至少可分为内存储器和外存储器。内存储器（简称内存）是 CPU 能直接访问的存储介质，它速度快，价格昂贵。它是一种“易失寄存器”，即当断开电源时，就会丢失其中的内容。因此，内存只能用来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，价格低廉。虽然 CPU 不能直接对它进行访问，但它是“非易失寄存器”，因此一般作为内存的延伸和后援，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，存储管理的对象是内存，其主要工作如下。

(1) 记住内存各部分的使用情况，哪些已经分配，哪些为待分配。

(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，也能正确运行”的存储扩充问题。

3. 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。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) 记住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、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。

4. 文件管理

程序与数据都是以文件的形式存放在外存（如硬盘、软盘）上，是计算机系统的软件资源。用户是通过文件的名称来访问所需要的文件的，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（这就是该文件的目录项）。当用户通过文件名来访问某文件时，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，要随时记住外存上文件存储空间的使用情况，哪些已经分配，哪些为待分配。

(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) 确保存放在外存上文件的安全、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、更新等操作。

1.3 操作系统的种类

1.3.1 批处理操作系统

在讲述操作系统的形成过程时，曾提及批处理系统。在那里，要求系统配置管理程序，以保证一个个作业程序能自动地进入内存处理，此时人工不得干预。这其实就是被称为“单道批处理操作系统”的早期操作系统。“单道”的意思是指一次只让一个作业程序进入系统内存加以运行，也可以说它是一个单用户操作系统。

在单道批处理操作系统的控制和管理下，计算机系统的工作过程如下。

用户为自己的作业编写程序和准备数据，同时编写控制作业运行的作业说明书。然后将它们一并交给操作员。

(1) 操作员将收到的一批作业信息存入辅助存储器中等待处理。

(2) 单道批处理操作系统从辅助存储器中依次选择作业，按其作业说明书的规定自动控制它的运行，并将运行结果存入辅助存储器。

(3) 操作员将该批作业的运行结果打印输出，并分发给用户。

单道批处理操作系统有如下特点。

(1) **单路性**：每次只允许一个用户程序进入内存。

(2) **独占性**：整个系统资源被进入内存的一个程序独占使用，因此资源利用率不高。

(3) **自动性**：作业一个一个地自动接受处理，期间任何用户不得对系统的工作进行干预。由于没有了作业上、下机时用户手工操作耗费的时间，提高了系统的吞吐量。

(4) **封闭性**：在一批作业处理过程中，用户不得干预系统的工作。即便是某个程序执行中出现一个很小的错误，也只能等到这一批作业全部处理完毕后，才能进行修改，这给用户带来不便。

在单道批处理的基础上，引入多道程序设计技术，就产生了多道批处理操作系统。配置多道批处理操作系统的本质仍然是批处理。不同的是由于采用了多道程序设计技术，允许若干个作业程序同时装入内存，造成对系统资源共享与竞争的态势。用户为自己的作业编写程序和准备数据，同时编写控制作业运行的作业说明书，然后将它们一并交给操作员。其工作过程如下。

(1) 操作员将收到的一批作业信息存入辅助存储器中等待处理。

(2) 多道批处理操作系统中的作业调度程序从辅助存储器里的该批作业中选出若干合适的作业装入内存，使它们不断地轮流占用 CPU 来执行，使用各自所需的外部设备。若内存中有作业运行结束，再从辅助存储器的后备作业中选择对象装入内存执行。

(3) 操作员将该批作业的运行结果打印输出，分发给用户。

多道批处理操作系统有如下特点。

(1) **多路性**：每次允许多个用户程序进入内存，它们轮流交替地使用 CPU，提高了内存和 CPU 的利用率。

(2) **共享性**：整个系统资源被进入内存的多个程序共享使用，因此整个系统资源的利用率较高。

(3) **自动性**：作业处理期间任何用户不得对系统的工作进行干预。由于没有了作业上、下机时用户手工操作耗费的时间，提高了系统的吞吐量。

(4) 封闭性：在一批作业处理过程中，用户不得干预系统的工作。即便是某个程序执行中出现一个很小的错误，也只能等到这一批作业全部处理完毕后，才能进行修改。这给用户带来不便。

1.3.2 分时操作系统

将多道程序设计技术与分时技术结合在一起，就产生了分时操作系统。配有分时操作系统的计算机系统称为分时系统。

所谓分时系统，即一台计算机与多个终端设备连接，最简单的终端可以由一个显示器和一个键盘组成。每个用户通过终端向系统发出命令，请求系统为其完成某项工作。系统根据用户的请求完成指定的任务，并把执行结果返回。这样，用户可以根据运行结果，再次通过终端向系统提出下一步请求。重复这种交互会话过程，直至每个用户实现自己的预定目标。图 1-4 所示为分时系统工作过程的示意图。

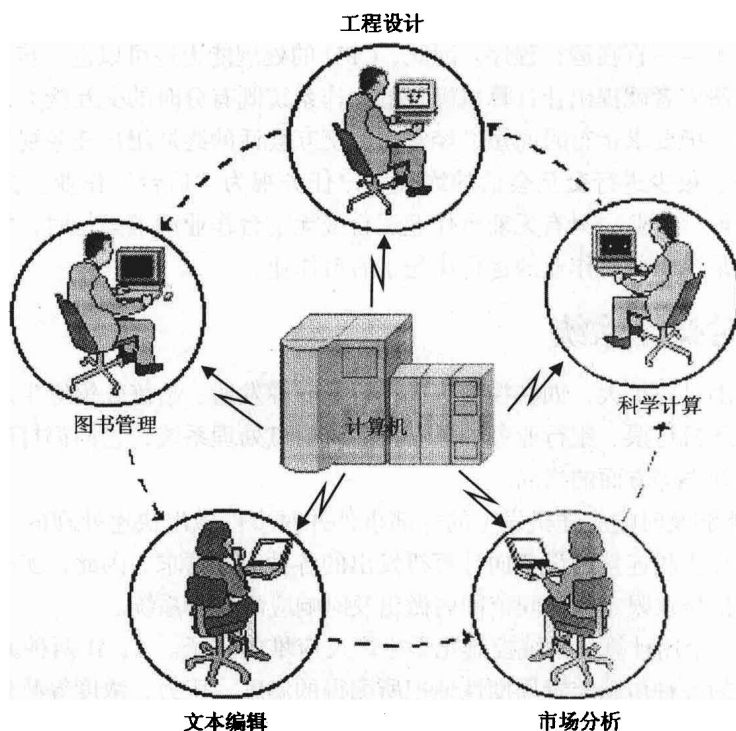


图 1-4 分时系统工作过程示意图

分时系统之所以能在较短的时间内响应用户的请求，同时为多个终端用户提供服务，主要是因为分时操作系统中采用了“时间片轮转”的处理机调度策略。这种调度策略是把处理机时间划分成一个个很短的“时间片”，对提出请求的每个联机用户终端，系统轮流分配一个时间片给其使用。若在一个时间片内，用户所请求的工作未能全部做完，就会被暂时中断执行，等待下一轮循环再继续做，让出的 CPU 被分配给另一个终端使用。由于计算机的处理速度很快，只要时间片的间隔取得适当，用户就不会感觉到从一个时间片跨越到另一个时间片之间的“停顿”，就好像整个系统全由他“独占”使用似的。例如，若时间片为 100ms，系统中有 10 个用户终端分享 CPU，并假定忽略操作系统为实现用户终端之间的切换所需耗费的时间。那么每个用户平均响应时间（即从用完一个时间片到获得下一个时间片所需的时间间隔）为 1s。这 1s 的“停顿”，用户是完全感觉不出来的。

不难看出，分时系统有如下特点。

(1) **多路性**：在一台主机上连接多个用户终端。从宏观上看，多个用户同时工作，共享系统的资源；从微观上看，各终端程序是轮流使用一个时间片。多路性提高了系统资源的整体利用率。

(2) **交互性**：用户在终端上能随时通过键盘与计算机进行“会话”，从而获得系统的各种服务，并控制作业程序的运行。交互性使用户能随时掌握自己作业程序的执行情况，为用户调试、修改以及控制程序的执行提供了极大的便利。

(3) **独立性**：每个用户在自己的终端上独立操作，互不干扰，感觉不到其他用户的存在，就如同自己“独占”该系统。

(4) **及时性**：用户程序是轮流执行 CPU 的一个时间片，但由于计算机的高速处理能力，能保证在较短和可容忍的时间内响应和完成处理用户请求。

在分时系统中，一台主机常会连接多达数十台终端，但一方面由于 CPU 处理速度极快，另一方面大部分终端并不会一直在运行程序。因此，CPU 的处理能力还可以进一步挖掘。为了改善系统性能，操作系统研究者就提出让计算机配置的操作系统既有分时的交互能力，又有批处理的能力。在这种系统中，把要求运行时间短、经常需要交互会话的终端用户任务视为“前台”作业；把要求运行时间长、很少进行交互会话的终端用户任务视为“后台”作业。操作系统总是按分时系统的方式运行前台作业，只有无前台作业运行或无前台作业请求运行时，才按批处理方式运行后台作业。也就是说，前台作业的运行优先于后台作业。

1.3.3 实时操作系统

计算机应用范围日益扩大，如在控制飞机飞行、导弹发射、冶炼轧钢等生产过程中均采用了实时控制系统，在飞机订票、银行业务中采用了实时信息处理系统，它们都打破了只把计算机用于科学计算和数据处理等方面的格局。

“实时”是指能够及时响应随机发生的外部事件并对事件做出快速处理的一种能力。而“外部事件”是指与计算机相连接的设备向计算机发出的各种服务请求。因此，实时操作系统是能对来自外部的请求和信号在限定的时间范围内做出及时响应的操作系统。

图 1-5 所示为一个用计算机系统控制化学生产反应堆的例子。A、B 两种原料通过阀门进入反应堆。反应堆中的各种传感装置周期性地 把所测得的温度、压力、浓度等测量信号传送给计算机系统。计算机中的实时操作系统及时接收这些信号，并调用指定的处理程序对这些数据进行分

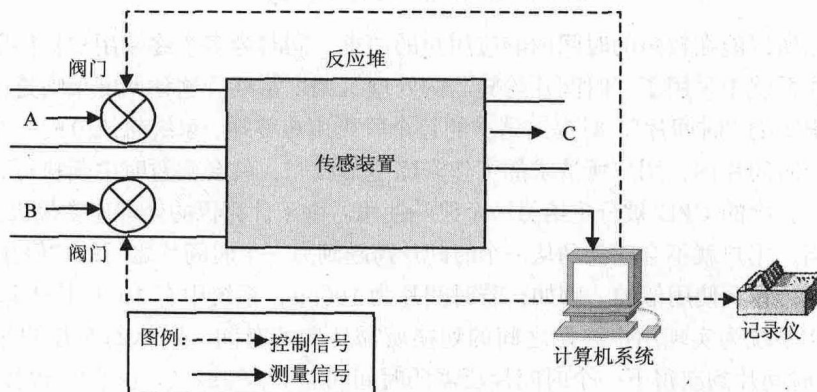


图 1-5 化学生产反应堆的控制

析, 然后给出反馈信号, 控制两种原料 A、B 的流量, 确保反应堆中的诸原料参数维持在正常范围之内。若参数超过极限允许值, 就立即发出报警, 甚至关闭反应堆, 以免事故的发生。

配置实时操作系统的计算机系统称为实时系统。这种系统有如下特点。

(1) **高及时性**: 分时系统的“及时性”, 是一种在用户可以容忍的时间内响应用户请求的时间性能。而对于实时系统而言, 这种及时性应该更高, 对外部事件信号的接收、分析处理以及给出反馈信号进行控制, 都必须在严格的时间限度内完成。

(2) **高可靠性**: 无论是实时控制系统还是实时信息处理系统, 都必须有高可靠性。例如, 对于后者, 计算机在接收到远程终端发来的服务请求后, 系统应该根据用户提出的问题, 对信息进行检索和处理, 并在有限的时间内做出正确的回答。

1.3.4 网络操作系统

计算机网络是指把地理上分散的、具有独立功能的多个计算机和终端设备, 通过通信线路加以连接, 以达到数据通信和资源共享目的的一种计算机系统。计算机网络是计算机和通信技术相结合的产物。

分时系统提供的资源共享有两个限制: 一是限于计算机系统内部; 二是限于同一地点(或地理位置很近)。计算机网络在分时系统的基础上, 又大大前进了一步。

在网络范围内, 用于管理网络通信和共享资源, 协调各计算机上任务的运行, 并向用户提供统一的、有效方便的网络接口的程序集合, 就称为网络操作系统。要说明的是, 在网络中各独立计算机仍有自己的操作系统, 由它管理着自身的资源。只有在计算机间要进行信息传递或要使用网络中的可共享资源时, 才会涉及网络操作系统。

网络操作系统有如下 4 个基本功能。

(1) **网络通信**: 为通信双方建立和拆除通信通路, 实施数据传输, 对传输过程中的数据进行检查和校正。

(2) **资源管理**: 采用统一、有效的策略, 协调诸用户对共享资源的使用, 用户使用远地资源如同使用本地资源一样。

(3) **提供网络服务**: 向用户提供多项网络服务, 如电子函件服务, 它为各用户之间发送与接收信息, 提供了一种快捷、简便、廉价的现代化通信手段。再如远程登录服务, 它使一台计算机能登录到另一台计算机上, 使自己的计算机就像一台与远程计算机直接相连的终端一样进行工作, 获取与共享所需要的各种信息。还有文件传输服务, 它允许用户把自己的计算机连接到远程计算机上, 查看那里有哪些文件, 然后将所需文件从远程计算机复制到本地计算机, 也可以将本地计算机中的文件复制到远程计算机中。

(4) **提供网络接口**: 向网络用户提供统一的网络使用接口, 以使用户能方便地上网, 使用共享资源, 获得网络提供的各种服务。

计算机网络系统有如下特点。

(1) **自治性**: 在网络中的每台计算机都有自己的内存和输入/输出设备, 安装有自己的操作系统, 因此具有很强的自治性, 能独立承担分配给它的任务。

(2) **分散性**: 系统中的计算机分布在不同的地域。

(3) **互连性**: 网络中分散的计算机及各种资源, 通过通信线路实现物理上的连接, 进行信息传输和资源共享。

(4) **统一性**: 网络中的计算机, 使用统一的网络命令。

习 题

一、填空题

1. 计算机系统由_____系统和_____系统两个部分组成。
2. 按功能划分, 软件可分为_____软件和_____软件。
3. 操作系统是在_____上加载的第一层软件, 是对计算机硬件系统功能的_____扩充。
4. 操作系统的基本功能是_____管理、_____管理、_____管理和_____管理。
5. 在分时和批处理系统结合的操作系统中引入“前台”和“后台”作业的概念, 其目的是_____。
6. 分时系统的主要特征为_____、_____、_____和_____。
7. 实时系统与分时以及批处理系统的主要区别是_____和_____。
8. 若一个操作系统具有很强的交互性, 可同时供多个用户使用, 则是_____操作系统。
9. 如果一个操作系统在用户提交作业后, 不提供交互能力, 只追求计算机资源的利用率, 大吞吐量和作业流程的自动化, 则属于_____操作系统。
10. 采用多道程序设计技术, 能充分发挥_____和_____并行工作的能力。

二、选择题

1. 操作系统是一种_____。
 - A. 通用软件
 - B. 系统软件
 - C. 应用软件
 - D. 软件包
2. 操作系统是对_____进行管理的软件。
 - A. 系统软件
 - B. 系统硬件
 - C. 计算机资源
 - D. 应用程序
3. 操作系统中采用多道程序设计技术, 以提高 CPU 和外部设备的_____。
 - A. 利用率
 - B. 可靠性
 - C. 稳定性
 - D. 兼容性
4. 计算机系统中配置操作系统的目的是提高计算机的_____和方便用户使用。
 - A. 速度
 - B. 利用率
 - C. 灵活性
 - D. 兼容性
5. _____操作系统允许多个用户在其终端上同时交互地使用计算机。
 - A. 批处理
 - B. 实时
 - C. 分时
 - D. 多道批处理
6. 如果分时系统的时间片一定, 那么_____, 响应时间越长。
 - A. 用户数越少
 - B. 内存越少
 - C. 内存越多
 - D. 用户数越多

三、问答题

1. 什么是“多道程序设计”技术? 它对操作系统的形成起到什么作用?
2. 怎样理解“虚拟机”的概念?
3. 对于分时系统, 怎样理解“从宏观上看, 多个用户同时工作, 共享系统的资源; 从微观上看, 各终端程序是轮流运行一个时间片”?