

高等学校教材

计算机 操作系统原理



李永锡 周兴社 编

西北工业大学出版社

TP316
L39

278317

高等学校教材

计算机操作系统原理

李永锡 周兴社 编



西北工业大学出版社

内 容 提 要

本书介绍计算机操作系统的原理与实践。全书共分八章。第一章概要介绍操作系统的概念、功能、发展史及分类。第二章到第六章以资源管理观点讲述操作系统的各个主要组成部分。包括：进程管理、作业管理、存贮管理、设备管理、文件管理等，这几章的内容在安排上，以论述基本原理、基本算法和设计技巧为主，并结合典型实例，介绍操作系统的实践。第七章介绍实时任务管理、系统初启和生成及操作系统维护。第八章介绍操作系统性能分析和评价的方法。

本书为高等学校计算机科学和工程类专业的教材，亦可供有关学科的技术人员参考。

JS/40/16



高等学校教材

计算机操作系统原理

李永锡 周兴社 编

责任编辑 柴文强

*

西北工业大学出版社出版

(西安市友谊西路127号)

陕西省新华书店发行

空军导弹学院印刷厂印装

*

开本787×1092毫米 1/16 印张13.5 324千字

1987年11月第一版 1987年11月第一次印刷

印数1—15000册

ISBN 7-5612-0016-1/TP·5

统一书号：15433·061

定价2.25元

前 言

计算机操作系统在现代计算机系统中的地位愈来愈重要，它已成为计算机系统的“中枢神经”。所以，《计算机操作系统原理》课程是计算机系的一门专业骨干课。我们依据以前的教学实践，并结合目前微型计算机系统迅速发展特点，编写了这本教材。在编写过程中，遵循了以下几条宗旨：

一、以讲授计算机操作系统原理为主

作为一门科学，计算机操作系统原理课程有它自己的体系结构。我们依据资源管理观点论述了计算机操作系统的基本原理、基本思想和基本算法。不拘泥分析一个具体的操作系统，而旨在结合目前流行的几个典型系统加深对基本原理的理解和应用。

二、以微机的分时和实时系统为主

计算机操作系统按其使用环境来说，有批量处理、分时、实时系统三种类型。并且，大、中型计算机系统以及小、微型计算机系统均配备了相应的操作系统。鉴于微型计算机系统更为普及，实时、分时系统日趋广泛，本教材侧重于介绍个人计算机操作系统、实时和分时操作系统。

三、由浅入深，循序渐进

本书在介绍基本原理时，尽量做到由浅入深，循序渐进，以便读者学习。举例中，总是先介绍MS-DOS，iRMX-86系统的简单思想，再叙述UNIX、VAX/VMS等系统的复杂结构和算法。

本教材共分八章，依资源管理观点逐一介绍操作系统的几大部分。

第一章介绍了操作系统的概念、发展过程及其分类。突出了分时和实时操作系统的概念和特点。

第二章首先介绍了进程的概念，进而再介绍进程管理的各个环节。以进程调度、进程通信和死锁对策为重点，读者从中可学习并发程序的设计技术。

第三章主要介绍了作业调度和作业控制两大问题，侧重于联机命令的使用及处理。

第四章内容结合各种实例，由简到繁地介绍了存贮管理的各种方案及其特点。

第五章主要介绍设备分配、I/O缓冲技术、通道管理等基本问题。鉴于目前磁盘是主要的外存设备，因而着重介绍了磁盘系统的工作原理和磁盘调度算法。

第六章结合MS-DOS、UNIX等系统，比较全面地介绍了文件结构、目录管理、文件操作使用的原理等。

第七章论述了操作系统中几个较为零碎，但非常重要的问题，如系统初启和生成，系统结构分析和维护。

第八章重点介绍随机排队模型在操作系统性能分析与评价中的应用。该章作为选学内容。

在该教材中基本原理叙述与具体系统分析混为一体，使读者难以对具体系统有一整体概念，课程小结时可予以弥补。

本书第三、五、六、八章由李永锡副教授编写，第一、二、四、七章由周兴社讲师编写。西安交通大学计算机系施鸿宝教授、沈钧毅副教授审阅了全书，并提出了有益的意见。林树根同志描画了部分图表。编者在此表示谢意。

由于编者水平有限，书中错误之处在所难免，敬请读者批评指正。

编 者

一九八六年十月

目 录

第一章 概述	1
§ 1.1 操作系统的基本概念	1
§ 1.2 操作系统发展史	3
§ 1.3 操作系统的分类	6
习 题.....	12
第二章 进程管理	14
§ 2.1 进程及其表示	14
§ 2.2 进程控制	24
§ 2.3 进程调度	31
§ 2.4 进程通信	37
§ 2.5 死锁及其对策	52
习 题.....	59
第三章 作业管理	61
§ 3.1 概述	61
§ 3.2 作业调度	62
§ 3.3 作业控制	68
§ 3.4 UNIX系统Shell 程序简介.....	72
习 题.....	75
附录A 连接的基本原理.....	75
附录B MS-DOS 下用键盘命令控制运行一个FORTRAN 程序.....	76
第四章 存储器管理	78
§ 4.1 单一连续区管理	78
§ 4.2 分区管理	81
§ 4.3 覆盖与对换技术	90
§ 4.4 页面式管理	95
§ 4.5 段式管理	106
§ 4.6 段页式管理-MVS 的虚拟存贮	114
习 题.....	117
第五章 设备管理	119
§ 5.1 引言	119

§ 5.2	外设与主机的I/O控制方式	120
§ 5.3	设备的分配和伪脱机	122
§ 5.4	块设备的物理特性	125
§ 5.5	多进程访问磁盘的调度算法	128
§ 5.6	缓冲区技术	132
§ 5.7	具有通道的 I/O 系统管理	137
§ 5.8	用系统调用命令使用外设举例	145
	习 题	148
第六章	文件系统	149
§ 6.1	概述	149
§ 6.2	磁盘文件的物理结构	152
§ 6.3	磁盘存贮空间的管理	155
§ 6.4	文件目录和目录文件	158
§ 6.5	文件目录的检索	162
§ 6.6	文件的使用	165
§ 6.7	文件卷的安装和拆卸	167
§ 6.8	文件的保密	169
§ 6.9	MS-DOS文件管理的功能调用	170
	习 题	172
第七章	操作系统的其它问题	173
§ 7.1	实时时钟和实时任务管理	173
§ 7.2	系统初启	179
§ 7.3	系统生成	181
§ 7.4	操作系统的结构分析	185
§ 7.5	操作系统的维护	188
	习 题	193
第八章	排队论在操作系统中的应用	195
§ 8.1	排队论基础	195
§ 8.2	状态函数和状态概率	196
§ 8.3	到达间隔的分布和服务时间的分布	197
§ 8.4	批处理系统和前端处理系统模型分析	199
§ 8.5	多用户分时系统模型的分析	202
§ 8.6	多I/O通道下多道程序系统模型分析	205
	参考文献	209

第一章 概 述

§ 1.1 操作系统的基本概念

一、什么叫操作系统

我们知道，现代计算机系统是一个相当复杂的系统，即使是微型计算机系统也不例外。它们都是由丰富的软、硬件资源组成。系统中用内存贮器保存正在执行或即将执行的程序，并保存所处理的数据；以中央处理机(CPU)完成各种运算和系统控制；由各种输入输出设备实现I/O信息传输；使用磁盘等外存装置保存大量的临时性或永久性信息。此外，系统中还有众多的软件资源为用户提供服务：各种编译程序将相应的高级语言源程序翻译成目标程序；文本编辑程序提供了联机输入和修改源程序的手段；连接、汇编和查错程序也是必不可少的；通常还包括应用软件包和标准过程库等等。面对如此复杂的系统，一个普通的用户会感到束手无策的，因为，他首先要了解并熟悉硬件的复杂细节，才能使用系统。如果设置一个自动化的管理机构，由该机构组织和管理各种硬件资源的使用，实现各类软件资源的查找和调用，就可以为用户提供很大方便。操作系统(Operating System)就扮演了这一重要角色。它是计算机系统中主要的系统软件之一，起着管理系统的作用。目前，不论是大、中型计算机系统，还是微、小型计算机系统，乃至个人计算机系统都配备了相应的操作系统。特别是在软、硬件结合日趋密切，应用领域日趋广泛的今天，操作系统在计算机系统所处的举足轻重的地位，就更加突出地显示出来了。在购买机器，使用系统，开发新功能的过程中，首当其冲的是要了解 and 熟悉操作系统。

综上所述，操作系统是现代计算机系统中一个必不可少的关键组成部分，它是管理系统资源的程序的集合。其任务是合理地组织计算机工作流程，有效地组织诸资源协调一致地工作以完成各种类型的复杂任务，从而达到充分发挥资源效率，尽力方便用户使用之目的。

二、操作系统的功能

关于如何研究操作系统有多种观点，如资源管理观点，分层观点，以及虚拟机观点等。这些观点彼此并不矛盾，只是代表了对同一事物的不同看法。每种观点都有助于理解和分析操作系统。下面我们依据操作系统是一资源管理器的观点来论述和研究操作系统的基本功能。

人们通常把计算机系统资源分成四类：中央处理机、内存贮器、外部设备(CRT、打印机、磁盘等)以及程序和数据。这样分类是因为它们的特征有着固有的差别，分别研究更能深入和细致。把操作系统看作资源管理器，那么，它的功能就可归纳为处理机管理，内存管理、外部设备管理和文件管理。文件管理是对存放在磁盘上的程序和数据进行管理。磁盘作为外部设备的管理是指对磁盘接口、读、取操作实施管理。除了上述四项功能之外，操作系统还应该向用户提供使用它自己的手段，这就构成了作业管理的主要内容。

这些资源管理之间并非彼此完全独立，它们之间存在着微妙的相互依赖关系。

对于每一种资源的管理，操作系统必须做好以下几件事，

- (1)记录该资源的使用情况；
- (2)实施某种策略以决定谁获得资源，何时获得，获得多少；
- (3)分配资源；
- (4)回收资源。

中央处理机是计算机系统中最宝贵的硬件资源。为了提高它的利用率，采用了多道程序技术。如果一个运行的程序因等待某一条件而不能运行下去时，就把处理机占用权转交给另一个可运行的程序；或者出现了一个比当前占用处理机的程序更重要的可运行程序时，就把占有权从占有者那里抢过来，以便合理地为用户服务。协调这些程序之间的关系，就是处理机管理的任务。

多个程序共享内存贮器时，系统既要保证它们彼此隔离，互不侵扰，又要保证能在一定条件下调用。再则，当用户程序需要的内存量超过系统目前可用的内存空间时，系统必须根据各个程序的运行状态把未占用CPU者及时地调出内存，运行前又从外存调入内存。这些均是存贮器管理的职责。

现代计算机系统一般采用中断和通道技术把各种外部设备和中央处理机连接起来。因此，外设管理的任务就是实现各种设备的具体管理，使得一般用户不必了解详细的接口技术，就可方便地使用设备。

内存贮器是计算机系统的工作存贮器。它主要由随机存贮器组成，因此，易于丢失信息，另外，由于受到种种限制，其价格较为昂贵，而容量有限，因此，大部分的用户程序和数据，操作系统以及其它系统程序都要存放在外存介质上。目前普遍使用的是磁盘设备。存放在磁盘上的一组信息必须唯一地标识，以便受到正确的访问控制，而且，系统必须有条有理地组织如此大量的信息，使得用户能以文件形式方便地访问它们。这就是文件管理的职责。操作系统中，有关这部分的程序称为文件系统。计算机系统应用于各种信息处理的今天，要求文件系统的功能尽量灵活和丰富，进而发展成为数据库系统。

如前所述，操作系统已经把系统中的软、硬件管理起来了。而用户怎样通过操作系统达到使用计算机系统的目的呢？具体地说，用户怎样能把自己的程序和数据交给系统，又怎样通知系统他所计划的执行过程呢？操作系统如何把用户的要求交给系统去处理，又如何把结果送给用户呢？为此，操作系统必须提供本身与用户之间的接口。这正是作业管理的主要任务。

三、操作系统的工作环境

我们从计算机系统的层次结构来描述操作系统的工作环境。图1·1表示了层次化的系统结构：

硬件(CPU, 内存、接口、外部设备)是操作系统赖以生存的物质基础。纯硬件只是裸体机的形象，是一种低级的工作环境，它仅给用户指令系统的原始功能。操作系统在硬件指令系统基础之上增加了其它许多“指令”供用户使用，从而扩充了计算机的功能。操作系统在裸体机上运行，而其它程序在扩充后的机器上运行。

从编译程序设计者角度看，他们只注重于词法与句法分析，代码生成与优化等问题，

关于编译程序运行时的内存管理，输入输出等则可直接调用操作系统的相应功能实现。其它服务程序的设计也不例外。各种编译程序和服务程序开发成功后，交操作系统统一管理。

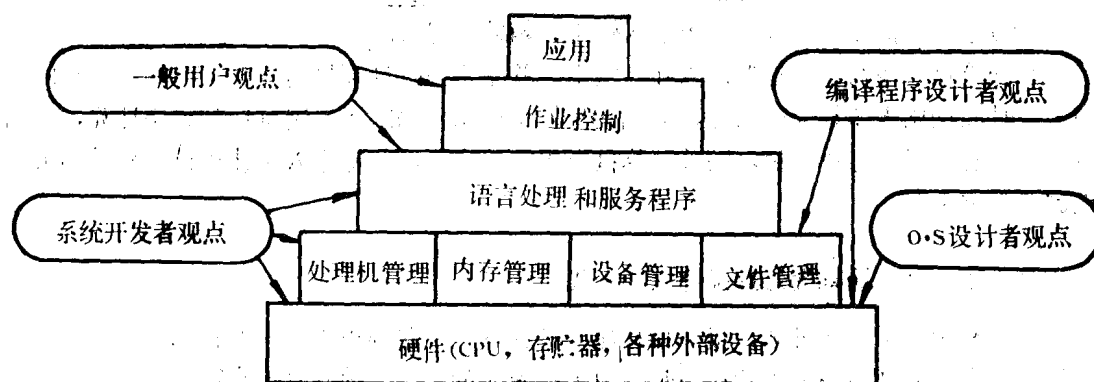


图 1.1 层次化系统结构

若从机房管理者、系统操作员的角度看待操作系统，那么它是计算机系统工作流程的自动而高效的组织者，计算机软、硬件资源的合理协调的管理者。它减轻了操作管理人员的负担，提高了计算机系统的使用价值。

若从一般用户的角度来看待操作系统，那么它是一台比裸体机功能更强，服务质量更高，使用更为方便的虚拟机器，它是用户和系统之间的一个界面，用户通过操作系统使用计算机系统。

§ 1.2 操作系统发展史

操作系统和计算机硬件一样，也有自己的形成和发展过程。我们以“代”描述它。众所周知，计算机硬件是以器件的发展为标志，并经历了四代：第一代是电子管，第二代是晶体管，第三代为中、小规模集成电路，第四代是大规模集成电路。每代的进化都是以不断减少成本、体积、功耗，增加速度和存贮容量为追求目标的。与此同时，操作系统也经历了以下几代。

- 第零代 手工操作
- 第一代 早期批处理系统
- 第二代 执行系统和多道程序系统
- 第三代 通用操作系统
- 第四代 完善和发展

操作系统的进化是以提高资源利用率，方便用户使用为追求目标而不断发展的。下面简要介绍各代的特征。

一、第零代 手工操作（40年代中—50年代初）

这一代，计算机问世不久，还没有操作系统。用户以使用机器语言和符号语言的手编

程序为主，操作员用“拨开关”方式控制机器的运行，以“看绿灯”为途径，了解系统运行情况。这种方式存在着三个突出的缺点：

- (1) 当用户使用计算机时，就独占了该机的所有资源。
- (2) 用户要熟悉机器各部分的细节，使用不便，容易出错。
- (3) 手工联机操作，人工干预输入输出，辅助时间长。

在计算机发展初期，这些缺点并不突出。首先是当时计算机速度慢，CPU等待时间与工作时间的比率不算大；另外，早期的计算机应用仅局限于个别部门的科学计算；使用者多是熟悉机器的专业人员。随着计算机应用日益广泛，非专业人员使用计算机的不方便性就非常突出了。

二、第一代 早期批处理系统 (50年代)

为了减少手工操作时间，提高价格昂贵的计算机资源利用率，人们首先想到的是从一个用户程序过渡到另一个用户程序时，力求摆脱人的干预，使其变成自动化。这就是50年代的批处理的基本思想。

在早期的批处理系统中，操作员收集若干个用户程序，按先后次序通过输入机读入内存缓冲区，再存到磁带上。然后，管理程序从磁带上逐一读入内存进行汇编和编译，进而连接成一个目标程序开始执行，处理完后又读入另一程序，直至处理完磁带上所有用户程序。照此工作，系统在管理程序控制下，实现了程序之间交接的自动化，缩短了操作时间。为了进一步提高这种能力，有些早期批处理系统还专门设置了一台小型卫星机来承担I/O设备的管理任务。该卫星机只与外部设备接通，不和CPU直接连接，因而称之为脱机批处理系统，如图1·2所示。脱机批处理的特点是卫星机与主机并行工作，分工不同，使主机摆脱了慢速的输入输出操作。

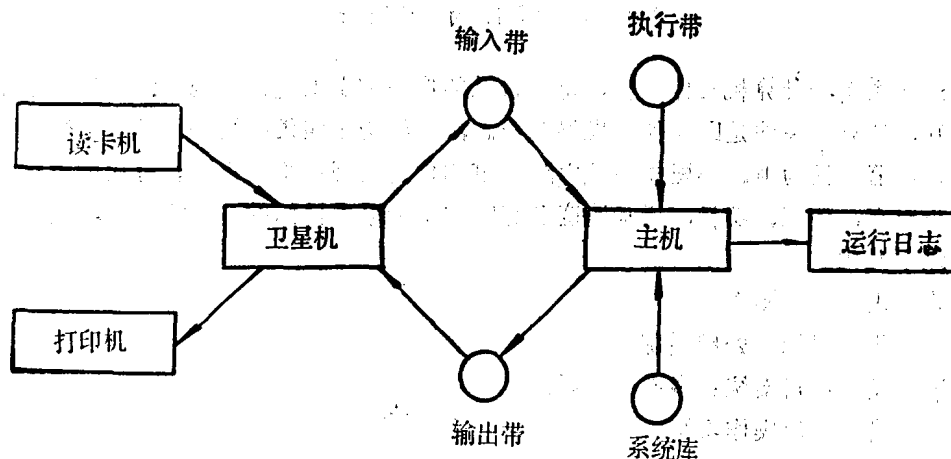


图 1·2 脱机批处理系统

早期批处理的产生也促进了一些系统软件的发展。例如，输入输出标准程序，子程序库、装配程序、运行日志及记账程序都是在那时形成的。这一代的系统虽然大大缩短了手工操作时间，但由于没有分时及多道程序能力，系统利用率仍无很大提高。

三、第二代 执行系统和多道程序系统 (60年代早期)

第二代操作系统的特征采用了多道程序设计技术,以提高系统利用率。它的前驱是执行系统。执行系统来源于60年代初引入的中断和通道技术的发展。

中断技术基本解决了主机与外设并行工作的问题,消除了因外部设备速度慢而使得主机查询等待的严重浪费现象。当时,在大型计算机系统中,也开始采用通道技术。所谓通道,实质上是一台功能单一、结构简单的I/O处理机,它独立于CPU,并直接控制外部设备,与内存进行数据传送。CPU启动通道后,通道就可独立工作,完成外部设备的启动、关闭以及信息传输等操作,只有在传送完规定的的数据后,它才向CPU发出一次中断请求。由此可见,采用通道技术大大减轻了主机处理输入输出及管理外部设备的负担,使主机集中时间完成复杂的计算任务。有关的通道控制、中断处理以及设备驱动程序的集合,在当时统称为执行系统。

不论是早期批处理,还是执行系统,它们都是单道顺序地执行程序,即一个用户程序执行完后,才允许启动另一用户程序。我们称之为单道程序系统。这种系统经常出现资源使用不充分和不均衡的现象,影响了系统效率。因此,在执行系统出现后不久,人们想到,若能随时地在内存同时驻留几道程序,每当现运行的程序因某种原因不能继续运行时,则将另一道程序投入运行。这样,几道程序共享系统资源,可使CPU内存及外部设备得到更加充分的利用。这就是通常所说的多道程序系统的基本思想。

在此阶段,主要解决了系统保护,存贮分配和简单的地址翻译等问题。多道程序带来了系统管理的复杂性,所产生的新问题还有待于解决。

四、第三代 通用操作系统 (60年代中—70年代中)

1964年,IBM公司的360系列机所配置的操作系统OS/360问世了。它标志着第三代操作系统的开始。这一代操作系统基本解决了多道程序设计所带来的种种新问题。例如,各道程序的调度,内存空间的分配、保护及扩充以及外部设备管理和分配等等。另外,为了使操作人员和一般用户能够控制程序的执行过程和了解系统运行情况,系统提供了一套作业控制语言及操作命令。以上功能按资源管理观点,可归纳分类为处理机管理、存贮器管理、设备管理、文件管理、以及作业管理。从而完善了操作系统的功能。

该时期内,计算机硬件也为操作系统设计提供了不少方便条件。例如,操作系统,尤其是它的核心部分是指挥整个计算机系统全局工作的关键,它需要严格保护,防止被用户程序破坏。为此,计算机硬件为中央处理机规定了两种工作状态,即系统态(核心态,管态)和用户态(目态)。当CPU执行操作系统程序代码时,它就被置为系统态。在该状态下,处理机可执行全部硬件指令。用户程序一般在用户态下运行,若要请求操作系统服务时,只能通过专用的访管指令,且经严格检查才可进入系统态。用户态下不能执行特权指令。所谓特权指令是指那些只允许操作系统执行而不允许用户程序使用的指令。如停机指令,开、关中断指令,I/O指令都是特权指令。另外,还有专用的界寄存器,基址寄存器等都是硬件对操作系统的支持。

由于这一代操作系统的设计强调通用性,以适应不同工作环境的需要,并适应同一系列不同档次的配置,所以,它们规模庞大,结构复杂。使得用户学习和掌握较为困难,可

维护性和可靠性也令人担忧。尽管如此，这一代仍是操作系统发展史上的重要阶段。

五、第四代 完善和发展 (70年代中—现在)

近十年来，人们设计出了丰富多彩的第四代操作系统，以适应计算机系统结构的不断发展和计算机应用日趋广泛的需要。这一代的典型代表有下述三种操作系统。

1. 个人计算机操作系统

随着微处理机的问世和飞速发展，名目繁多的个人计算机(Personal Computer)应运而生，今以APPLE机和IBM-PC机居于首位。它们广泛使用在各行各业的科学计算、办公室自动化、数据采集与处理诸方面。个人计算机强调“个人”的特点，它的操作系统规模较小，不讲究共享资源，注重使用方便。多以磁盘文件管理为主体，并提供一组键盘命令供用户使用。这一类操作系统的典型代表是CP/M和MS-DOS。CP/M是八位个人计算机的操作系统，并由CP/M派生了若干各有特点的微机操作系统。MS-DOS又称为PC-DOS，它是为IBM-PC机配置的操作系统。MS-DOS吸收了CP/M和更高级的操作系统的优点，功能较齐全。最近推出的一些个人机操作系统提供了更加灵活方便的人机接口，如“菜单”命令，“窗口”显示等等。

2. 局部网络操作系统

若干个人计算机系统配备通讯接口，以某种拓扑结构互连，或与大型计算机系统相接，组成了形式多样的信息处理网络。如办公室自动化系统，企业管理网络等。系统内的各个计算机共享系统信息资源和昂贵的设备，增强了应用能力，方便了用户。这样的局部网络操作系统依据层次结构，保证机间通讯，实现信息共享和保护诸功能。

3. 分布式操作系统

分布式计算机系统是当代计算机系统结构的重要研究方向之一。它把复杂的任务按功能分布在多个处理机上执行。而整个系统由统一的操作系统实施管理。分布式操作系统要实现并行任务分配、并行进程通信、分布控制机构、分散资源管理等功能。

局部网络操作系统和分布式操作系统在理论和实践上还不成熟，尚处于深入探讨和初步总结阶段。故本教材只涉及单机操作系统。

§ 1.3 操作系统的分类

到目前为止，我们还是笼统地讲述操作系统。实际上，一类操作系统与另一类操作系统相比，它们都有各自的特性和本质的差别。主要体现在以下几方面。第一，系统功能。如前所述，操作系统有五大部分功能，而具体系统有所侧重，它们提供诸功能的深度和精细度、自我管理以及支持程序开发的程度有很大差别。第二，基本结构。一个依据外部事件驱动程序运行为特点的系统与一个管理较长运行时间的批量程序为特点的系统相比，两者的结构全然不同。第三，支援硬件和应用环境。不同类的操作系统需要不同的支援硬件，从而运行在不同的应用环境中。因此，我们有必要将操作系统依据某一准则进行分类，以便深入研究各自的特点，掌握基本原理。

我们可以直接依据硬件系统的规模将操作系统分成大系统和小系统两类。大型计算机系统资源丰富，价格昂贵，所配的操作系统的充分发资源利用率为设计的基本出发点，

以系统调度和管理为主要任务。该类系统强调自我管理,尽量减少人工干预,并希望提供各种访问技术以及诸如数据库等高级服务,追求通用性。在具体设计技术上,一般具有设备无关性,动态分配以及丰富的人机接口等特点。小型计算机系统资源种类少,管理简单,相应的操作系统设计趋向于单一的访问技术及专门的用途。例如操作系统仅允许终端访问以实现会话式程序开发,且只可运行较小程序。操作系统仅提供文件管理及有限的数据查询功能,对用户的支持较少,需要操作人员辅助控制操作。

通常,人们依据操作系统在用户面前的使用环境及访问方式将它分成以下三类:

批量处理系统 (Batch Processing Operating System)

分时操作系统 (Time-sharing Operating System)

实时操作系统 (Real-time Operating System)

随着微型计算机系统日益普及,分时系统和实时系统愈来愈广泛。下面将对这两类系统进行较详细的描述和分析。

一、批量处理操作系统

首先介绍作业的概念。通常,一个用户程序是以某种算法语言或汇编语言描述的多个程序段的有序集合,我们称之为源程序。当源程序输入给系统时,必须以某种命令形式告诉该程序所采用的语言名称以及用户的控制意图。一般将用户程序和所需的数据及命令统称为作业。这是静态的观点。

在批量处理系统中,内存驻留多个作业,并在外存存放若干后备作业。当某些作业处理结束或因某种原因无法继续运行而暂停时,系统根据一定的调度原则,从后备作业中选中几个作业调入内存运行。如此操作,直至处理完全部作业。显然,这样的批量系统属于多道程序系统,它以力图使作业流程自动化,充分发挥资源利用率为主要目标。该系统在作业运行过程中不允许用户干预,全部受系统控制。因而吞吐量大,资源利用率高。目前,各计算中心的大、中型计算机系统一般都是批量处理系统或兼有这一功能。

对于用户来说,批量处理系统越来越不便于使用。具体表现在以下两方面:

- 1.用户一旦把作业交给系统,就失去了控制作业运行的能力。虽然可以通过一定方式把运行中可能出现的情况告诉系统,而在大多数情况下很难预料到每个细节。特别是当用户调试一个新程序时,他们是无法作出这种预料的。因此,用户还是希望能干预作业的运行。

- 2.用户从提交作业到获得结果常常需要等待数小时甚至几天时间。然而,他的作业可能只需几分钟或几秒的运行时间。而且,几经反复才能获得正确结果,拉长了解题周期。

随着软件开发量愈来愈庞大的趋势,程序设计人员很难容忍批量处理系统的上述缺点。他们对联机操作怀着留恋之情。联机操作可使程序员亲自观察程序运行情况,发现错误,及时纠正。在一定意义上讲,这能够发挥程序设计人员的主观能动性。从而促使了分时系统的问世和发展。

二、分时操作系统

1.分时系统的概念

所谓分时,顾名思义,是指对时间进行分享。在计算机硬件中,自从有了与CPU并

行工作的通道，就有了分时的概念。如 CPU 与通道分时访问内存；通道与通道分时使用内存；同一通道相连的外部设备又分时使用通道。但这些分时都是通过硬件线路实现的。而分时系统的“分时”意指若干用户分享使用同一台计算机，这种分时主要通过操作系统程序所实现。分时系统具有如下基本特征：在一台计算机周围挂上若干台终端设备，系统为用户提供一分时终端命令集，多个用户通过自己占用的终端设备直接联机使用计算机。因为终端上的用户以会话方式工作，人们又称分时系统为多用户交互式系统。

在分时系统中，分享时间的基本单位叫做时间片。一般是几十毫秒~几百毫秒。CPU 每经过一个时间片，就把控制转向操作系统，系统则轮流地把相继的一个个时间片分给每个用户程序，由于时间片相对于人的感觉来说很短，一般在1,2秒内可对用户作出反应，从而给每个用户形成一种“独占计算机”的感觉。

2. 分时系统的软、硬件接口

分时系统的硬件接口主要是多路转接器和通讯处理机。多路转接器是一个扫描机构，负责控制中心计算机与若干终端设备之间的信息传送。每个终端通过各自的通信线与多路转接器相连接。多路转接器则借助于I/O通道与中心计算机及其主存相连。多路转接器也可以经通讯处理机间接与中心计算机连接。设置通讯处理机可以减轻中心计算机的负担，使得它能集中完成系统控制和任务处理。上述分时系统的结构如图1·3(a、b)所示。目前，人们已用微处理机和标准接口芯片完成多路转接器与通讯处理机两者的功能。

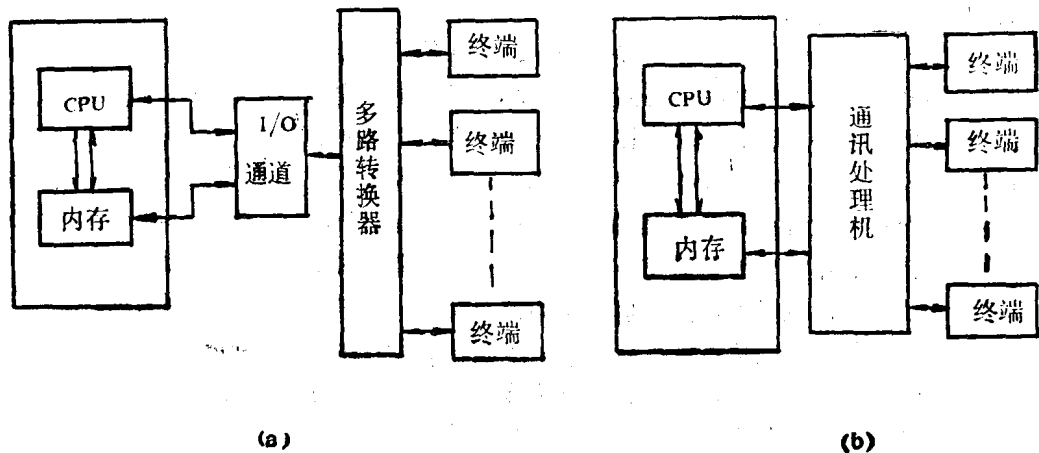


图 1·3 分时系统的结构

多路转接器依次扫描每个终端设备。当扫描到字符时，它就把该字符及标识该终端线的编号一起传送给中心计算机。然后，依次扫描下一终端设备，查看该终端是否处于激励状态，是否要传送字符等。类似地，对于输出而言，当通道把中心计算机的输出信息送给多路转接器时，也送来相应终端的编号，以便多路转接器正确地将信息送往指定终端。综上所述，多路转接器的输入工作流程如图1·4所示。

分时系统的接口软件主要是指直接与多路转接器有关的软件部分。它负责管理所有通讯线，接收每一个字符并把它们汇集成一组完整的信息（如一条语句，一条命令）进行初步装配，保存在内存或文件中。

3. 分时终端的操作过程

各分时系统的终端命令符号及显示形式互有差异，而用户在终端上的操作步骤大致相同。基本上依下述步骤进行：

(1) 接通 用户首先按下指定的键向中心计算机发出信号，标志该终端请求工作。系统收到信号后，在终端上输出接通时间及提示符等引导信息，表示允许用户开始人-机对话。

(2) 通名 用户看到允许对话的符号后键入用户名或作业名。系统判别用户名后，若允许，则显示“请输入保密口令”的提示。用户随后输入口令。

(3) 核实口令 系统将已保存的口令与输入的口令作比较。口令相符时，就输出终端命令提示符，表示允许工作。

(4) 正式工作 用户看到命令提示符后，输入键盘命令、程序或数据。系统根据命令的功能进行相应处理，并向用户报告成功或失败的回答信息。用户根据回答信息，作出下一步操作的决定。如此联机操作，人-机对话，进行工作。

(5) 退出系统 用户完成自己的任务时，要输入退出系统命令。系统停止扫描该终端，并保存有关信息以保证信息的完整性。

4. 分时系统的特征及优点

从上述介绍，我们可以概括出分时系统的主要特点如下：

(1) 同时性 若干用户同时使用一台计算机。从微观上看，各用户轮流使用计算机，从宏观上讲，各用户并行工作。

(2) 独立性 各用户之间彼此独立地工作，各自独占一台终端，互不干涉。

(3) 及时性 用户在可允许的时间间隔内得到响应。这里顺便指出，分时系统的响应时间是用户通过终端发出命令到系统进行应答所需的时间。它是衡量分时系统性能的主要指标。

(4) 交互性 用户可根据系统对自己请求的响应结果，继续向系统提出新的要求。

从以上特点中不难分析出分时系统的优点。第一，促进了计算机应用的普及。有了分时系统，无力或不必购买计算机系统的用户仅需配备一台终端设备就可方便使用计算机系统。例如，学校里的分时系统可供若干名学生同时上机实习。第二，显著提高了开发、调试程序的效率。用户联机使用计算机，边调试，边思考，边修改。可以大大缩短调试程序的周期。因而，分时系统是程序设计人员比较理想的开发环境。第三，充分利用系统资源。计算机系统中的各种软、硬件资源由若干分时用户所共享，从而提高了利用率。

最早的分时操作系统是美国麻省理工学院(M·I·T)在1963年研制的CTSS(The Compatible Time-Sharing System)，它是为IBM7904机所配备的操作系统。CTSS的成功运

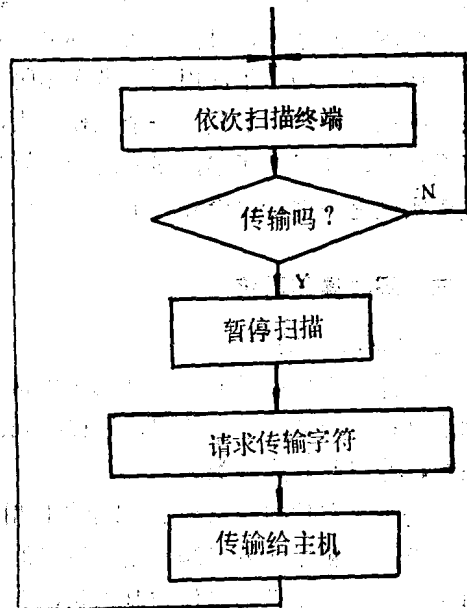


图 1·4 多路转换器输入工作流程

行证明了分时系统的潜在优势；1969年，M·I·T又设计了第二代分时操作系统 MULTICS (Multiplexed Information and Computing Service)。MULTICS用高级语言PL/I的前趋EPL编程，用CTSS作为开发环境研制成功的；此外，IBM的TSS/360当时也是一个主要的分时系统；目前，优秀的分时操作系统应首推UNIX，UNIX系统是Bell实验室于1969年—1970年用两个多人年研制成功的。UNIX的设计吸取了MULTICS等分时系统的成功经验，形成了一个规模不大，功能较强的系统。目前，它已被移植到许多计算机系统上，并派生了丰富多样的变种UNIX系统以满足各种需要。本教材将以UNIX作为主要实例之一进行讨论和分析。

三、实时操作系统

1. 实时系统的概念

实时计算机系统是随着计算机应用领域的日益广泛而出现的。我们知道，计算机的主要用途是科学计算，信息处理，过程控制。而后两方面均具有一定的实时要求，则统称实时系统，并以此分成两类。

第一类是实时控制系统。把计算机用于导弹发射、飞机飞行等军事方面的自动控制时，要求计算机尽快处理测量系统测得的数据，及时地对飞机或导弹进行控制，或将有关信息通过显示终端提供给操作控制人员，以便迅速作出相应决策；把计算机用于石油生产、钢铁冶炼、机械加工等工业控制中，要求计算机及时收集和处理采样的现场信息，进而控制有关的执行机构，使得某些参数，如温度、压力、流量按一定规律变化或保持不变，从而达到提高质量，增加产量之目的。用计算机组成自动化交通管理系统，实验数据收集和监测系统，也可以获得确切数据，提高效率。

第二类是实时信息处理系统。把计算机用于飞机订票查询班机航线及票价事务，也可以组成军事指挥系统，资料查询系统，情报处理系统等。在这些系统中，计算机接收来自终端设备的服务请求，在很短的时间内对用户作出正确回答。

上述实时系统都是由实时操作系统加以管理和协调的。所谓“实时”表示立即、现在的意思。具体含义是指计算机要能及时响应外部事件的请求，并以足够快的速度完成对事件的处理，还控制所有实时设备和实时任务协调一致地运行。这里要特别指出，实时系统是以管理时间的精确度决定着它对实时环境的适应能力。诚然，各种系统均有某种程度的实时响应能力，而它们对实时的要求有着很大差异。例如，对某个事务处理系统来说，0.2秒的响应时间就可构成实时响应；而0.2秒却不能满足某些飞行器控制系统的实时要求，它甚至期望比0.2秒还快几百倍的实时响应时间。

在实时信息处理系统中，也常配备多个终端设备。多个终端设备有可能同时发出询问请求服务。为了适应这种情况，这类实时操作系统也采用分时使用计算机来处理这些实时性询问。因此，我们要注意这类实时系统与前述的分时系统之间的差别。

第一，实时系统的应用程序是预先已经编制好了的，在工作过程中不许改变这些程序。所以它只允许终端操作员键入如何控制的命令，不需要重新编排程序。实时系统中终端装置仅仅是一台询问的装置，用户可通过它请求系统处理实时信息。而分时系统是通过终端装置，向用户提供一台由他随便使用的计算机，不仅可以请求服务，而且可以通过终端设备自由地输入和修改程序。