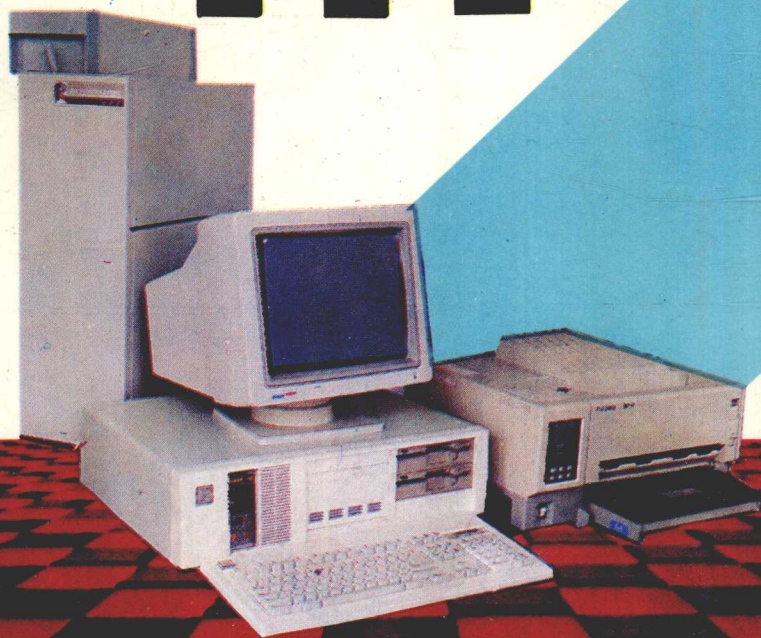


计算机操作系统

曹 聪 主编



兰州大学出版社

计算机操作系统

曹 聪 主编

兰州大学出版社

(甘)新登字第 08 号

内 容 提 要

本教材介绍计算机最重要的系统软件——操作系统。以操作系统的基本原理、基本概念和基本算法为主线,按其组成分别讨论操作系统的各个管理系统及其关联衔接。全书共分为七章:第一章介绍操作系统的基础知识和操作系统的类型。第二章起到第六章,分别讲述处理机管理、作业管理、内存管理、设备管理和文件管理。最后第七章概略介绍了操作系统的结构设计方法。为便于读者自学,书末附有全部习题答案,便于启发和加深巩固学习成果。

本书可供计算机专业本科生学习,适当删去部分内容(目录或正文中带“*”号处)即可供非计算机专业本科生及计算机专业专科生学习。由于书中采用了深入浅出、分散难点的方法,所以特别适用于自学者学习。

计算机操作系统

曹 聪 主编

兰州大学出版社出版

(兰州大学校内)

兰州大学印刷厂印刷 甘肃省新华书店发行
开本:787×1092 毫米 1/16 印张:20

1994 年 8 月第 1 版 1994 年 11 月第 2 次印刷
字数:493 千字 印数:5001—10000 册

ISBN7-311-00766-6/T·11 定价:19.80 元

序 言

当代计算机科技领域在日新月异地发展和更新换代,我国国民经济各个领域和社会活动各个方面应用日趋普遍。然而,与我国经济建设的要求相比,在计算机技术的研究和系统的利用与开发方面,尚需做更多的努力,继续进一步提高其资源利用率,增强系统处理能力,以适应国际上计算机科学迅速发展的形势。

伴随计算机科学技术的发展,逐步形成了计算机系统中最主要的一个系统软件——操作系统,这是计算机开发、教学等各个方面不可缺少的主要组成部分。在教学方面,操作系统已是计算机专业的基础主干课程之一,为培养既有扎实理论基础,又具备娴熟的技术能力的计算机专业人材,使之能够较好地掌握应用和开发计算机资源,讲授教师对于教材编排和处理是一个关键。现在国内外操作系统教材、讲义和出版物不少,内容在理论或实践方面各有侧重,在这样的基础上,如果能够引各家之长,编写一套计算机操作系统教材,使学生简明易懂,掌握课程内涵,将会达到良好的教学效果。

以曹聪为主编,陈丽华参与编写的教材《计算机操作系统》正是基于这样的设想,经过几年努力,在多次试讲的基础上写成的一本有特色的教材。两位作者结合多年来的教学实践,征求了学习过操作系统课程的各个层次、不同专业的大学本科、专科毕业生的意见,并与校内同行进行大量切磋编写而成的。因此给我的印象是:明显深化了操作系统的算法设计概念,成为整个教材的骨干主线,并围绕这条主线,讲述操作系统的基本概念,基本原理,基本算法,从而使读者能够系统地理解《计算机操作系统》。各章节条理分明,层次有序,对于各章的主要问题、难点,讲述清楚,举例恰当,易于各层次大学生学习;在讲述方法上由浅入深,分条分项依次归纳,再附以习题练习,能够巩固学习成果,为学生将来开发、使用计算机系统软件奠定牢靠的基础;同时,作者还注意搜集整理了近年来新概念,新设计思路和新算法,引荐进一步深造的参考书目。总而言之,这本教材的编写,对大学本科及专科学习计算机的学生,大有可取之处,对于社会上研究、使用计算机的技术人员,也不无参考价值。

由于《计算机操作系统》的续篇《操作系统分析和使用》,尚在编写安排,建议作者以UNIX系统分析为主要线索,介绍当今世界较为新颖实用的若干操作系统实例分析,充实教材内容,深入浅出地介绍使用方法,以适应不同层次教学需要,形成较为完整的一套教材。

李 康

1994年3月28日

前 言

现代计算机系统,通常拥有数量可观的软硬件资源,提高这些资源的利用率和增加系统处理能力的需求,就形成和发展了操作系统——现代计算机系统中必不可少、至关重要的系统软件。由于近年来,计算机在科学技术、国民经济及社会活动中的广泛应用和不断发展更新,作为基础软件的操作系统,其理论和应用,相应同步地迅速发展。为适应计算机科学发展的需要,掌握使用计算机和学习操作系统,已成为不可分割的一体。此外在计算机应用方面,已渗透于各行各业、乃至千家万户的今天,选择购买计算机,开发应用计算机,也必须有一定的操作系统基础。为此而编写了这套操作系统教材,内容包括两个方面:《计算机操作系统》、《操作系统分析与使用》。

本教材《计算机操作系统》是主编多年对本科生、专科生在讲授《计算机操作系统》教学实践经验的基础上,对《计算机操作系统》庞杂、散乱、深奥、难懂的内容,进行了梳理,以深入浅出,分散难点,强化算法设计思想为主线贯穿全书。力图通过分析操作系统的基本概念、基本原理、基本算法,使读者对操作系统有一个全面系统的理解,其次对操作系统中出现的新概念、新算法和新的设计思想给予概略介绍,使读者对其新课题、新方向有所了解。为便于加深巩固学习成果和启发自学理解,书末附有全部习题答案供参考。

本教材内容不仅适合于计算机专业本科生学习,也可将教材内容中带“*”部份删去(主要是基本算法及其实现方法等),即适应非计算机专业本科生和计算机专业专科生学习,同时也便于计算机自学者学习和使用计算机。

为便于使用、分析和开发操作系统,今后准备将大、中、小、微机操作系统中的典型实例(如 UNIX 系统)分析和使用方法,并加上实习用的开发系统,另编成册,作为《计算机操作系统》的续篇,定名《操作系统分析与使用》编成出版。

这本《计算机操作系统》中,第一章阐明了有关操作系统中的基本定义和操作系统的类型;第二章主要讲述操作系统中最基本的问题——进程及有关内容:进程控制、同步与互斥、进程通讯及死锁;第三章介绍作业管理,包括作业组织和运行,作业调度及用户与操作系统之间的接口;第四章阐述存储管理,从软件、硬件存储保护等角度分别讲述了各种不同的存储管理方法;第五章介绍了设备管理,叙述了与设备管理有关的设备管理技术和一个模拟控制系统;第六章讲述文件管理,从文件的组织、存储、控制、使用等方面详加阐明;第七章概略介绍操作系统的结构设计方法。每章最后附有习题(答案见书末)。

本书由曹聪主编,除陈丽华参与编写了习题答案部分和第一章的§3,第二章的§3中部分内容外,其他各章及剩余全由主编本人编写完成,并统一修编定稿。

在编写过程中,兰州大学副校长,计算机科学系李廉教授给予指导,审阅了全稿并写序,在此致以衷心感谢。此外兰州大学计算机科学系刘莉等同学协助绘图校对,王汉峰等

同学协助誊写,在此一并致谢。

限于编者业务水平,加之时间短,编写仓促,错误和缺点在所难免,恳祈专家、学者及读者不吝赐教,以资进一步完善这套教材,使之成为教学服务。

编者

1993年12月于兰州大学

目 录

第一章 操作系统引论	(1)
§1 引言	(1)
一、计算机和计算机系统	(1)
二、什么是操作系统	(1)
§2 操作系统的分类	(2)
一、几个基本术语	(2)
二、成批处理系统	(3)
三、多道程序系统	(5)
四、分时系统	(7)
五、实时系统	(8)
§3 微机操作系统	(9)
一、单用户微机操作系统	(9)
二、多用户操作系统	(9)
§4* 计算机网络和网络环境下的操作系统	(10)
一、什么是计算机网络	(10)
二、计算机网络的发展和类型	(10)
三、网络环境下的操作系统	(11)
§5* 分布式计算机和分布式操作系统	(11)
一、多处理机系统	(11)
二、分布式系统	(12)
三、分布式操作系统	(13)
§6 小结	(13)
习题	(14)
第二章 处理机管理	(15)
§1 进程	(15)
一、资源与资源共享	(15)
二、程序的顺序执行	(15)
三、程序的并行执行及特征	(16)
四、进程	(18)
五、进程状态和状态转换	(19)
六、进程控制块(PCB)	(22)
§2 进程的互斥与同步	(23)
一、进程制约	(23)

二、临界区.....	(24)
三、同步与互斥.....	(25)
四*、软件解决方法	(25)
五、同步机构.....	(28)
六、进程互斥.....	(32)
七*、AND 同步机制	(33)
八、进程同步.....	(34)
§3 进程通信	(40)
一、利用共享存储器进行通信.....	(40)
二、消息缓冲通信.....	(41)
三、利用共享文件通信——管道通信方式.....	(45)
§4 进程控制.....	(45)
一、进程树.....	(45)
二、进程状态和队列.....	(46)
三*、控制原语	(49)
四*、几个实例	(56)
§5 进程调度.....	(58)
一、进程调度的功能.....	(58)
二、进程调度方式.....	(59)
三、进程调度的引起原因.....	(59)
四、进程调度算法.....	(60)
五*、进程调度过程	(63)
§6 死锁.....	(64)
一、什么是死锁.....	(64)
二、死锁产生的原因和条件.....	(65)
三、死锁的预防和避免.....	(68)
四、死锁的检测.....	(71)
五、死锁解除.....	(74)
习题	(75)
第三章 作业管理	(79)
§1 用户与操作系统之间的接口.....	(79)
一、几个概念.....	(79)
二、用户和操作系统之间的接口.....	(79)
§2 命令接口.....	(80)
§3 程序接口.....	(84)
一*、系统调用及执行过程	(84)
二、系统调用命令类别.....	(85)
§4 作业的组织运行.....	(86)

一、作业的组织.....	(86)
二、作业的运行.....	(87)
§ 5 作业管理和调度.....	(88)
一、作业状态和状态转换.....	(88)
二、作业的输入与输出方式.....	(90)
三、作业调度功能和调度性能的衡量.....	(91)
四、单道批处理系统的几种调度算法.....	(92)
五、多道程序环境中的作业调度.....	(94)
习题.....	(99)
第四章 存储管理.....	(100)
§ 1 概述.....	(100)
一、存储管理的目的、功能.....	(100)
二、几个有关的概念.....	(100)
三、存储分配的方式和策略.....	(102)
四、重定位.....	(102)
五、虚拟存储器概念的引入.....	(104)
§ 2 单一连续区分配.....	(105)
§ 3 分区管理.....	(106)
一、固定分区.....	(106)
二、可变分区.....	(108)
三、多重分割.....	(120)
四、可重定位分区分配.....	(122)
五、存储保护.....	(126)
六、分区分配的优缺点.....	(128)
§ 4 覆盖和交换技术.....	(128)
一、覆盖管理.....	(128)
二、交换技术.....	(131)
§ 5 简单分页存储管理.....	(133)
一、实现原理.....	(133)
二、硬件支持和实现.....	(133)
三、高速缓冲存储器.....	(135)
四、软件支持和存储保护.....	(136)
五、优缺点分析.....	(136)
§ 6 请求分页管理系统.....	(137)
一、实现原理.....	(137)
二、地址变换和页面中断处理.....	(137)
三、软件支持和管理.....	(139)
四、优缺点分析.....	(144)

§7 分段存储管理	(144)
一、分段的地址空间	(144)
二、实现原理	(145)
三、保护措施	(147)
四、段式虚拟存储系统	(147)
五、分段管理的主要优缺点	(149)
§8 段页式存储管理	(152)
一、实现原理	(152)
二、管理算法	(153)
三、优缺点分析	(155)
习题	(155)
第五章 设备管理	(157)
§1 概论	(157)
一、设备管理的任务和功能	(157)
二、设备的类型	(158)
§2 通道和设备的控制与操作	(158)
一、I/O 的控制和操作	(159)
二、通道结构和类型	(164)
三、通道指令和通道程序	(167)
四、通道利弊	(168)
§3 中断及其处理	(170)
一、中断概念	(170)
二、中断向量和优先级	(170)
三、程序状态字	(171)
四、中断类型	(172)
五、中断过程	(173)
§4 缓冲技术	(173)
一、为什么要引入缓冲	(173)
二、缓冲技术的实现	(174)
§5 设备分配程序	(182)
一、设备管理中的数据基	(182)
二、设备分配原则	(184)
三、设备分配程序	(185)
§6* 设备驱动程序	(188)
一、设备的启动	(188)
二、I/O 请求的完成	(189)
三、驱动程序之例	(189)
§7 设备处理程序	(190)

§ 8* 一个输入输出控制系统	(193)
§ 9 一、设备管理数据基	(193)
§ 10 二、Request 过程	(194)
§ 11 三、Attach 和 Detach 过程	(196)
§ 12 四、Release 过程	(197)
§ 13 五、Read/Write 过程	(199)
§ 14 六、I/O 过程	(199)
§ 15 习题	(204)
第六章 文件管理	(205)
§ 1 文件管理综述	(205)
§ 1.1 一、文件概念	(205)
§ 1.2 二、文件类型	(205)
§ 1.3 三、文件系统	(206)
§ 2 文件结构和存取方法	(208)
§ 2.1 一、文件的逻辑结构	(208)
§ 2.2 二、文件的存取方法	(209)
§ 3 文件的物理组织	(212)
§ 3.1 一、存储器的物理分块	(212)
§ 3.2 二、文件的物理结构	(213)
§ 3.3 三、文件类型、存取方法与文件存储器的关系	(218)
§ 4 文件控制块和文件目录	(219)
§ 4.1 一、文件控制块	(219)
§ 4.2 二、简单文件目录	(220)
§ 4.3 三、二级文件目录	(221)
§ 4.4 四、多级文件目录	(222)
§ 5 文件的使用	(223)
§ 5.1 一、建立文件	(224)
§ 5.2 二、打开文件	(225)
§ 5.3 三、读文件和写文件	(225)
§ 5.4 四、关闭文件	(226)
§ 5.5 五、添加命令	(226)
§ 5.6 六、修改记录	(226)
§ 5.7 七、拷贝文件	(226)
§ 5.8 八、“搬家”命令	(226)
§ 5.9 九、撤消文件	(227)
§ 6 文件的存取控制	(227)

一、存取控制矩阵	(227)
二、存取控制表和用户权限表	(228)
三、口令和密码	(228)
§ 7 文件存储器的管理	(229)
一、文件存储器的主要参数	(229)
二、空闲文件目录	(230)
三、空闲块链	(230)
四、位示图	(231)
五、成组链接法	(231)
§ 8 文件系统的一般模型	(235)
一、文件系统的组织	(235)
二、文件系统的层次模型	(236)
§ 9 UNIX 文件系统概述	(240)
习题	(242)
第七章 操作系统的结构设计方法简介	(244)
一、大型软件开发设计问题	(244)
二、模块接口法	(244)
三、层次结构法	(245)
四、以管程为工具的结构设计方法(P、C、M 法)	(246)
附录 习题参考答案	(248)
参考书目	(307)

第一章 操作系统引论

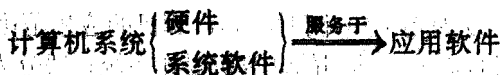
§1 引言

一、计算机和计算机系统

随着人类社会的不断进步和发展,科技水平的不断提高,象征着人类文明、进步的电子计算机也以惊人的速度不断渗透到人类生产和生活的各个领域。

从1946年第一台电子计算机诞生之日距今仅40多年的时间,计算机已经历了四代演变。由当初每秒3次的运算速度发展到如今每秒数十亿次。存储容量由最初的几K个字节发展到至今的以兆为单位计量。因而现代计算机已不仅仅是从本身物理设备来完成各种复杂问题的解决,而是需要其它内容配合使用的集体。

现代计算机系统是包括软件、硬件两部分内容的一个集合名词。硬件,是指包括中央处理机(Central Processing Unit, 简称CPU)、主存储器(Main Storage)、外围设备(包括辅助存储器和各类输入/输出设备)等的物理设备集合(也称裸机)。软件,是指程序及研制、使用、维护程序的所有资料等。软件又可分为系统软件和应用软件两部分。系统软件,是所有供用户使用的为解决用户使用计算机而编制的程序和资料。例如:操作系统、编译程序、汇编程序、及各种服务性程序。应用软件,是为解决某个特定问题而编制的程序和资料。于是计算机系统就是由硬件和软件组成,而它们又服务于应用软件。



通常我们把计算机系统简称为计算机。

二、什么是操作系统

计算机系统的硬件配置可以因为机器用途不同而各有千秋。但它至少必备的设备包括CPU、主存和输入/出设备。要使其高效、高速运行,就要对CPU的利用问题加以关注,内存容量的充分利用,数据信息的安全、保密是计算机使用的必要前提。另外,为提高计算机的使用效率,允许多个用户同时使用计算机,从而使许多用户竞争使用系统资源,诸如CPU、主存、输入/出设备等,为此,需要协调、分配这些资源。如:谁占有、占有多长时间等。所有这些问题的解决构成了对计算机系统的管理、控制工程。这个管理控制

工程是通过一个大型系统软件来实现的。这就是操作系统 Operating System(简称 OS)。

操作系统:是控制和管理计算机硬件和软件资源,合理组织计算机工作流程以及方便用户使用计算机的系统软件。

对操作系统的含义,我们还可以从以下几个不同观点来阐述:

1. 科普观点:操作系统是一个计算机系统的管理指挥机构或控制中心。它按设计者制定的各种调度管理策略,来组织管理整个计算机系统,使之能高效高速运行,以实现设计者的意愿。

2. 功能观点:操作系统是计算机的资源管理系统,由它负责对计算机的全部软硬件资源进行分配、控制、调度和回收。

3. 用户观点:操作系统是一个比裸机功能强、服务质量高、使用方便灵活的虚拟机,它实际上是用户与裸机间的一个界面,用户是通过它来使用计算机的。

4. 管理者观点:操作系统是计算机工作流程自动高效的组织者,系统软硬件资源的合理而协调的管理者,从而成为解放操作员劳动力和提高计算机使用价值的强大机构。

5. 软件观点:操作系统是一大型系统软件,是由程序和数据结构组成的集合。

综上所述,犹如人的大脑一样,操作系统在计算机中所处地位举足轻重。若不了解它,使用计算机时,颇感不便,修改、维护计算机更是无从下手。至于选择机器也必感力不从心。

对操作系统的研究,我们先从资源管理的观点入手。在系统中,一般将资源管理分为四大类:CPU管理(这又包括进程和作业管理两部分)、存储管理、外设管理和文件管理。本书将从这四类五种管理入手,介绍操作系统实现原理,并结合一些典型例子予以说明。

§ 2 操作系统的分类

计算机的发展经历了电子管时代(1946年—1959年)、晶体管时代(1959年—1965年)、集成电路时代(1965—1970年),以及大规模集成电路时代(1971年—至今)的四个阶段。硬件结构的发展,促使软件也随之发展、产生。操作系统作为系统软件的一部分,也是在这段时间中应运而生、发展起来的,同时随着计算机的应用范围日益广泛,操作系统也分门别类,各有所长。到目前为止,通用的计算机操作系统一般分为批处理系统,分时处理系统和实时处理系统三大类。近几年来又出现了网络操作系统和分布式操作系统。

在介绍各类操作系统之前,先引入几个术语。

一、几个基本术语

作业(Job):要求计算机系统按指定步骤对初始数据进行处理并得到结果的这样一系列工作集合。

作业步(Jobstep):是由完成作业中某一相对独立事件的程序和数据构成,并由命令定义之。从逻辑上,作业是由作业步组成。作业分系统作业和用户作业两大类。一般常将用户作业简称为作业。

裸机:仅含硬件的计算机,也称物理机器。

虚拟机(Virtual computer):配置了系统软件的计算机系统,也称为逻辑机器。

资源(Resource):一个作业可以使用的、以及它们运行所依赖的软硬件的集合,称之为这个作业要求的资源的集合。

联机处理(On-line processing):外设与 CPU 相联、并在 CPU 的直接控制下进行数据处理的工作方式。

脱机处理(Off-line processing):外设不与 CPU 直接相联、并不受 CPU 的直接控制进行数据处理的工作方式。

单机系统(Monoprocessor system):具有单个中央处理机的计算机系统。

多机系统(Multicomputer system):具有多个中央处理机或运算器的计算机系统。

二、成批处理系统

早期的电子计算机,由于速度慢、容量小,用户只能自己手编所有指令,并对机器拥有全部使用权,因而并无操作系统。50 年代初期,为了平顺作业之间的转接,减少在计算机系统上放置(建立时间)和取下(拆卸时间)作业所浪费的时间,人们开始关心并研究出最早的操作系统。

为协调由于硬件发展(运行速度提高、存储容量增大)而带来的严重的人机矛盾和机器利用率下降问题,人们改进操作方式。将作业由慢速输入设备输入到辅助存储器(简称辅存)上,继而再由辅存调入内存运行,并将其运行结果输出在辅存上,再分别由慢速输出设备输出。如图 1-1 所示。这种系统一次只能运行一个作业,所提高的机器效率完全来源于节省手工操作时间。为了对系统中的作业和计算机资源进行监督,又相应配置了监督程序(称之为管理程序 Monitor),这就形成了操作系统的前身和核心。

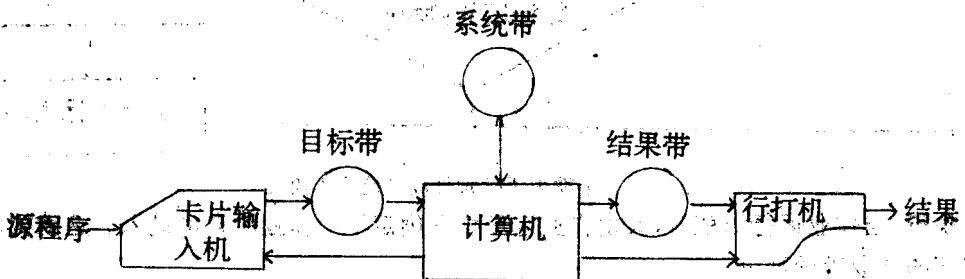


图 1-1 早期单道批量处理系统工作流程

早期的单道批量处理系统配置如图 1-2 所示。这是一种联机处理方式。其目的在于使整个作业流自动地、顺序地运行,以节省人工操作时间、提高机器利用率。整个系统的系统程序包括编译程序、汇编程序、装入程序等,全部装入系统磁带,作业流则由卡片机输入到目标磁带上存储起来,然后一个作业一个作业地顺序运行,当作业都是短小作业时,这种批量处理结果能获得较好的效果。它的基本工作流程如图 1-3 所示。

对这种批处理系统的控制是由监督程序实现的。它替代操作员对作业进行控制和管理,用户将作业的有关信息用某种命令方式提供给系统,系统则通过命令解释程序将命令

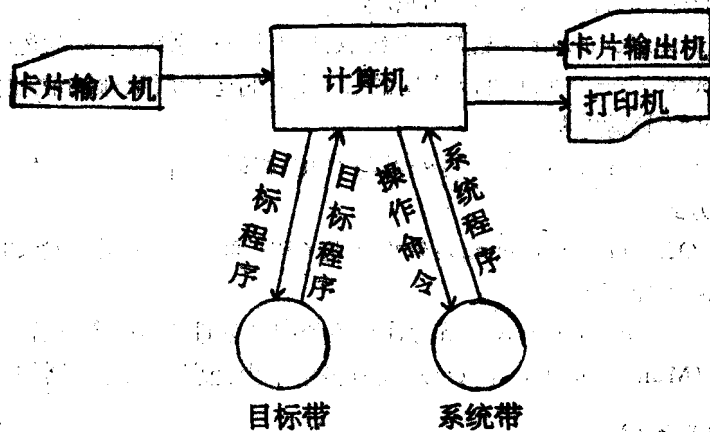


图 1-2 早期单道批量处理系统配置

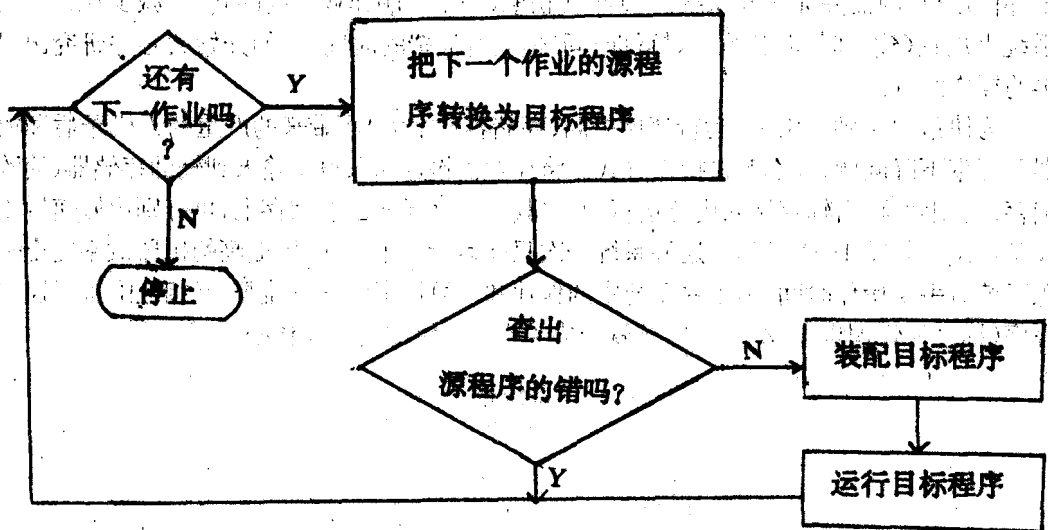


图 1-3 早期单道批量处理系统工作流程

解释执行，当作业运行结束时，又由作业定序程序确定下一作业的自动进入。这就构成监督程序的作业管理部分。

此外，由于系统程序的调入/出、以及用户程序从辅存到主存的调入，使监督程序的内存分配程序必须对内存空间予以统筹安排，为进入的作业准备一切必要的内存容量。

对于用户的外部信息和计算机内的内部信息交换，一般情况下，是由作业的读/写控制程序自动控制进行的，除非在异常情况下才需人工干预。对于用户需要保留的信息则由文件管理程序将其存入辅存。此外，文件管理程序还为用户提供信息和检索信息的手段。至于具体的输入输出操作则是由设备处理程序来实现的。

由此可见，监督程序初步具备了作业控制，监督管理和设备管理等重要功能，是操作系统的重要雏型。

此外,由于 CPU 速度的提高,人们又将输入/出工作由原来完全依赖于 CPU 而改为由一小型处理机来完成。亦即,引入一卫星机专门用于对输入输出设备的控制。而输入设备则将信息传送到辅存中去(这种专门用于存放输入信息的辅存称之为输入井),CPU 则由辅存将源程序和数据取入内存加工处理,其结果存入专门用于存放结果信息的辅存(称为输出井)中去。这些结果信息又由卫星机控制输出设备陆续输出。见图 1-4 示。这种单道批量处理系统装置与前述的不同之处在于前述装置(参见图 1-2)是联机处理方式,而此装置(图 1-4)是脱机处理方式。这种方式一方面硬件成本提高了(因为增添了小型处理机,尽管小型处理机价格低于 CPU),但另一方面却由于 CPU 完全摆脱了对输入/出的控制而可以高效运行。这比联机处理方式的运行效率提高得很多。

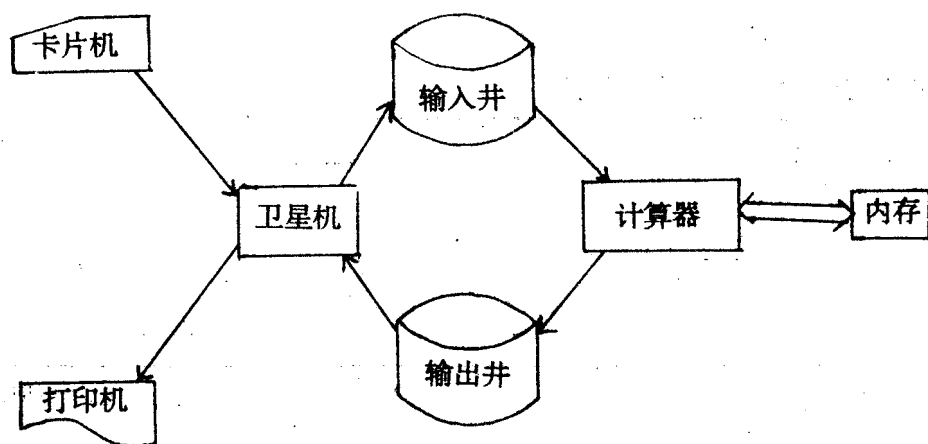


图 1-4 单道批量处理系统又一装置

总之,批量处理系统使大量作业收容存放在辅助存储器上,CPU 便可源源不断地对作业加工处理,从而减少了作业进入和退出计算机的操作时间。

三、多道程序系统

单道批量处理系统虽然大大减少了人工操作的时间,但由于每次仅有一个作业运行,其系统利用率仍不高。参见图 1-5,可以看出,CPU 在很大一段时间里处于空闲状态。至于输入/输出操作则更是如此。此外,不同作业对内存的需求量也各有不同,所以单道情况下的内存利用率也不高。还有,单道运行作业在运行中一旦运行出错,便须立即暂停,等待人工干预。无疑,这又是对 CPU 的浪费。这样一来,提高 CPU、内存、外设利用率就构成了对多道程序的引入之因。

所谓“多道程序系统”是指内存中同时放多个作业,它们分时共用一台计算机资源的系统。这样在多个作业运行时,依照一定原则,使得 CPU、各种外设得以充分利用,同时内存也得以充分利用。

图 1-6 就是三道程序情况下的运行之例。

情况分析:显然无论 CPU 还是外设,其利用都是比较充分的。这是因为多道程序的