



操作系统 原理及应用

CAOZUO XITONG YUANLI JI YINGYONG

张元 杨晓文 熊风光 张静 编著



国防工业出版社
National Defense Industry Press

TP316.81/324

2013

操作系统原理及应用

张元 杨晓文 熊风光 张静 编著

ISBN 978-7-111-31111-1

定价：39.00元

7-111-31111-1

00000-000-000-000

00000-000-000-000

00000-000-000-000

00000-000-000-000

RFID

北方工业大学图书馆



C00340112

国防工业出版社

· 北京 ·

内 容 简 介

本书主要讲述操作系统的基本原理,并以 Linux 系统为例对其原理部分进行实例分析。全书共分 8 章。第 1 章介绍了操作系统的引入、概念、发展过程、特征及当前常用的操作系统;第 2 章阐述了进程的概念、控制管理、进程同步及经典进程同步问题、进程通信、线程的引入及分类和 Linux 系统进程管理;第 3 章阐述了计算机系统中对资源的管理、死锁的发生及解决方式;第 4 章介绍了计算机系统中用户任务的执行管理的调度概念及评价准则、批处理系统的调度算法、分时系统的调度算法、实时系统调度算法和 Linux 系统调度算法;第 5 章介绍了内存管理的存储器分类、内存的连续分配、内存的离散分配、虚拟存储器的实现和 Linux 系统的内存管理;第 6 章介绍了设备管理的 I/O 控制管理方式、缓冲管理、I/O 软件、设备分配、设备处理、虚拟设备、磁盘存储器 I/O 管理和 Linux 系统的设备管理;第 7 章阐述了文件的分类、逻辑结构和物理结构分类、外存空间管理、目录管理、文件共享与安全及 Linux 系统的文件管理;第 8 章介绍了操作系统的安全性、安全策略、安全保护机制和 Linux 系统的安全机制。

图书在版编目(CIP)数据

操作系统原理及应用 / 张元等编著. —北京:国防工业出版社, 2013.7

ISBN 978-7-118-08938-7

I. ①操... II. ①张... III. ①Linux 操作系统-高等学校-教材 IV. ①TP316.89

中国版本图书馆 CIP 数据核字(2013)第 168490 号

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100048)

北京奥鑫印刷厂印刷

新华书店经售

*

开本 787×1092 1/16 印张 19 字数 470 千字

2013 年 7 月第 1 版第 1 次印刷 印数 1—2000 册 定价 32.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777

发行邮购:(010)88540776

发行传真:(010)88540755

发行业务:(010)88540717

前 言

计算机系统包括计算机软件系统和计算机硬件系统,操作系统是计算机系统的核心系统软件,它是配置在计算机硬件之上的第一层软件,是对硬件系统的首次扩充。作为计算机硬件系统与软件系统的接口,用户依靠操作系统提供的服务完成特定应用。操作系统处在计算机硬件层之上、其他系统软件及应用软件层之下的位置,在整个计算机系统中起到承上启下的作用。操作系统对整个计算机系统的资源进行有效管理,并协调用户共享使用计算机系统中的各类资源,使得计算机系统中各类硬件设备并行工作,提高其运行性能。

操作系统课程是计算机科学与技术专业、软件工程专业和网络工程专业的核心课程。随着计算机技术的发展,各类嵌入式系统得到广泛应用,其他相关专业也相继把操作系统课程作为一门重要的必修或者选修课程。而由于操作系统在计算机系统中地位的特殊性以及抽象性特点,使得这门课程学懂、学通的难度较大。很多初学者认为该门课程理论性强,原理概念较多,仅靠死记硬背很难掌握其中的重点和难点并对其融会贯通。在这种情况下,作者结合多年的操作系统课程教学经验,将一些生活示例引入难以理解的操作系统原理部分,并以 Linux 系统为例讲述其各部分的具体管理功能实现。

全书内容由杨晓文统稿,第1章和第6章由杨晓文编写,第2章由熊风光编写,第3章由韩燮编写,第4章由况立群编写,第5章和第8章由张静编写,第7章由张元编写。

本书的编写参考了大量的相关书籍和资料(详见参考文献),在此向这些参考文献的作者表示感谢。

本书在编写过程中得到了中北大学电子与计算机科学技术学院领导、同事的大力帮助和热情支持,在此向他们表示感谢。

由于作者水平有限,疏漏与错误在所难免,恳请各位专家和读者批评指正。

编 者

2013年5月

目 录

第 1 章 操作系统概述	1
1.1 操作系统的地位、作用与目标	1
1.1.1 操作系统的地位	1
1.1.2 操作系统的作用	2
1.1.3 操作系统的目标	4
1.2 操作系统的形成与发展	5
1.2.1 操作系统的产生与形成	5
1.2.2 操作系统的定义	10
1.2.3 作业的概念	11
1.2.4 操作系统的发展	13
1.3 进程的引入	18
1.3.1 前趋图	18
1.3.2 程序的顺序执行	19
1.3.3 程序的并发执行	19
1.3.4 引入进程	20
1.4 操作系统的特征	21
1.5 操作系统的功能	23
1.5.1 处理机管理功能	23
1.5.2 存储器管理功能	25
1.5.3 设备管理功能	26
1.5.4 文件管理功能	27
1.5.5 用户接口	28
1.5.6 操作系统的安全性	28
1.6 流行操作系统简介	29
1.6.1 Windows 系列操作系统	29
1.6.2 UNIX 操作系统	31
1.6.3 Mac 操作系统	35
1.6.4 Linux 操作系统	36
习题	38
第 2 章 进程管理	42
2.1 进程的描述	42

2.1.1	进程的概念	42
2.1.2	进程控制块	44
2.1.3	进程的状态及其转换	47
2.1.4	Linux 中进程描述	51
2.2	进程控制	57
2.2.1	进程的创建	59
2.2.2	进程的撤销	60
2.2.3	进程的阻塞与唤醒	61
2.2.4	进程的挂起与激活	63
2.2.5	程序的执行过程	64
2.3	线程	64
2.3.1	线程的引入	64
2.3.2	线程的概述	65
2.3.3	线程的分类	67
2.4	进程同步	68
2.4.1	资源与进程	68
2.4.2	进程同步机制	71
2.4.3	信号量机制	73
2.4.4	信号量机制的应用	77
2.5	经典的进程同步问题	82
2.5.1	生产者—消费者问题	82
2.5.2	哲学家进餐问题	86
2.5.3	读者—写者问题	89
2.5.4	管程机制	92
2.6	进程通信	96
2.6.1	高级通信概述	96
2.6.2	进程的直接通信与间接通信	98
2.6.3	基于消息缓冲通信方式	99
2.7	Linux 进程管理	101
2.7.1	Linux 进程控制	101
2.7.2	Linux 系统中进程通信	105
习题		106
第3章	资源分配与死锁	113
3.1	资源管理与分配	113
3.1.1	资源管理的目的和任务	113
3.1.2	资源分配方式及策略	113
3.1.3	资源分配图	114
3.2	死锁概述	115
3.2.1	死锁的定义	115

3.2.2	死锁产生的原因	115
3.2.3	死锁产生的必要条件	118
3.2.4	处理死锁的方法	118
3.3	死锁处理方式——预防死锁	119
3.4	死锁处理方式——避免死锁	121
3.4.1	引入安全状态	121
3.4.2	银行家算法的数据结构	123
3.4.3	银行家算法的实现	123
3.4.4	安全性算法检查	124
3.4.5	银行家算法举例	125
3.5	死锁处理方式——检测与解除死锁	128
3.5.1	死锁的检测	128
3.5.2	死锁的解除	130
习题		131
第4章	任务调度管理	134
4.1	调度概述	134
4.1.1	作业处理概述	134
4.1.2	调度概述	135
4.1.3	调度算法的评价准则	138
4.2	批处理系统调度算法	141
4.2.1	先来先服务调度算法	142
4.2.2	短作业(进程)优先调度算法	144
4.2.3	高优先权优先调度算法	145
4.3	分时系统调度算法	147
4.4	实时系统调度算法和多处理机调度	150
4.4.1	实时系统调度	150
4.4.2	常用的实时调度算法	151
4.4.3	多处理机调度	153
4.5	Linux 系统调度	153
习题		155
第5章	存储器管理	159
5.1	存储器概述	159
5.1.1	计算机的存储体系	159
5.1.2	内存管理的功能	161
5.2	连续分配存储管理方式	165
5.2.1	单一连续分配方式	165
5.2.2	固定分区分配方式	166
5.2.3	可变分区分配方式	167

5.2.4	可重定位分区分配方式	171
5.2.5	伙伴系统	172
5.3	离散分配存储管理方式——基本分页存储管理方式	173
5.3.1	基本思想	174
5.3.2	地址变换机构	175
5.3.3	多级页表和反置页表	178
5.3.4	页的共享和保护	180
5.3.5	基本分页存储管理方式特点	181
5.4	离散分配存储管理方式——基本分段式存储管理方式	182
5.4.1	分段系统的基本原理	182
5.4.2	段的共享和保护	185
5.4.3	分段存储管理方式和分页存储管理方式的区别	186
5.4.4	段页式存储管理方式	186
5.5	虚拟存储器	188
5.5.1	常规存储器特征	189
5.5.2	局部性原理	189
5.5.3	虚拟存储器的基本原理	189
5.5.4	虚拟存储器的容量	190
5.5.5	虚拟存储器的特征	190
5.6	请求分页存储管理方式	190
5.6.1	请求分页系统的基本原理	191
5.6.2	页表机制	191
5.6.3	缺页中断机构	192
5.6.4	地址变换机构	193
5.6.5	内存分配策略	193
5.6.6	调页策略	195
5.6.7	页面置换算法	196
5.6.8	请求分页系统性能分析	201
5.7	请求分段存储管理方式	203
5.7.1	段表机制	203
5.7.2	缺段中断机构	203
5.7.3	段的共享与保护	204
5.8	Linux 虚拟存储器管理	205
5.8.1	Linux 内存管理概述	205
5.8.2	分段和分页机制	206
5.8.3	Buddy 算法	209
5.8.4	虚存空间映射	210
5.8.5	交换机制	211
习题		211

第 6 章 设备管理	216
6.1 I/O 设备管理概述	216
6.1.1 设备管理的功能	216
6.1.2 I/O 系统结构	217
6.1.3 I/O 设备分类	217
6.1.4 I/O 设备应用原理	219
6.2 I/O 控制方式	221
6.2.1 程序 I/O 控制方式	221
6.2.2 中断驱动控制方式	222
6.2.3 直接存储器访问 I/O 控制方式(DMA I/O 控制方式)	223
6.2.4 I/O 通道控制方式	224
6.3 缓冲管理	227
6.3.1 缓冲的引入	227
6.3.2 缓冲区分类	228
6.4 I/O 软件	234
6.4.1 I/O 软件的层次	234
6.4.2 中断处理程序	234
6.4.3 设备驱动程序	235
6.4.4 设备独立性	236
6.5 设备分配	237
6.5.1 设备分配的安全性	237
6.5.2 设备分配策略	238
6.5.3 设备分配	238
6.6 虚拟设备	242
6.6.1 虚拟设备的引入	242
6.6.2 SPOOLing 系统的基本原理	242
6.6.3 共享打印机的实现	243
6.7 磁盘存储器管理	243
6.7.1 磁盘存储器结构	243
6.7.2 磁盘调度算法	244
6.7.3 磁盘高速缓存	247
6.8 Linux 系统设备管理	247
习题	248
第 7 章 文件管理	252
7.1 文件概述	252
7.1.1 文件分类	252
7.1.2 文件系统结构	253
7.2 文件的结构	255

7.2.1 文件的逻辑结构	255
7.2.2 文件的物理结构	256
7.2.3 外存空间管理	261
7.3 文件目录管理	264
7.3.1 目录管理功能	264
7.3.2 文件控制块	264
7.3.3 目录管理	265
7.4 文件的共享与安全性	268
7.4.1 文件共享	268
7.4.2 文件的安全性	269
7.5 Linux 文件系统	270
7.5.1 Linux 系统文件的物理结构	270
7.5.2 Linux 系统的文件系统	271
习题	278
第 8 章 操作系统的安全与保护	282
8.1 安全性概述	282
8.1.1 安全问题	282
8.1.2 安全威胁	283
8.1.3 安全操作系统的功能	283
8.2 操作系统的安全策略	284
8.2.1 隔离保护策略	284
8.2.2 分层保护策略	284
8.2.3 操作系统安全的评估和标准	285
8.3 安全保护机制	285
8.3.1 内存保护机制	286
8.3.2 访问控制技术	289
8.3.3 数据加密技术	290
8.3.4 认证机制	291
8.4 Linux 系统的安全机制	292
习题	293
参考文献	294

第1章 操作系统概述

本章知识点串联分析:

计算机系统由硬件系统和软件系统两大部分构成,操作系统是对计算机硬件系统的首次扩充,是计算机系统各类资源的管理者。操作系统的目标是为用户提供比计算机硬件系统功能更强、使用更方便的计算机系统;依据不同阶段操作系统的设计结构,设计并实现对处理机、存储器、I/O 设备及文件进行管理的软件集合。

本章根据计算机系统的组成,分析操作系统在计算机系统中的地位以及发展的目标,总结其在计算机系统中的作用,并给出操作系统的定义;简述自计算机诞生起配置其上的操作系统软件的发展过程,并给出其特征;最后分析操作系统软件的功能,并简要介绍目前常用的几类操作系统软件。

1.1 操作系统的地位、作用与目标

1.1.1 操作系统的地位

目前,任何一台计算机都配置了一种或多种操作系统软件。计算机系统由硬件系统和软件系统两大部分构成。硬件系统包括计算机系统中看得见、摸得着的所有物质设备,即由运算器、存储器与控制器组成的计算机主机和完成输入与输出功能的外部设备,这五部分硬件系统相互配合、协同工作,为计算机进行数据处理提供了物质基础。未配置任何软件的计算机称为裸机,由于计算机硬件只能识别 0、1 这样的机器代码,即只有少部分采用机器代码编写程序的人员能使用仅包括硬件系统的计算机裸机,为了让更多的群体都能使用计算机硬件设备,必须对其添加软件。软件是在计算机硬件系统的基础上,对硬件系统功能的扩充和完善。软件一般是为计算机自动运行和工作服务的全部技术和各种程序,即包括程序以及开发、使用和维护程序所需的文档的信息集合。只有为硬件配置了各种软件之后,计算机才能更好地工作,才能称为一个真正的计算机系统。计算机系统的组成如图 1-1 所示。

计算机系统的简单工作原理为:首先由输入设备接收外界信息(程序和数据),控制器发出指令将数据送入(内)存储器,然后向内存储器发出取指令命令;在取指令下,程序指令逐条送入控制器。控制器对指令进行译码,并根据指令的操作要求,向存储器和运算器发出存数、取数命令和运算命令,经过运算器计算并把计算结果存在存储器内;最后在控制器的取数和输出命令的作用下,通过输出设备输出计算结果。

计算机系统的层次结构如图 1-2 所示。硬件层提供基本的可计算性资源,按照用户的程序需求接收和存储信息、处理数据并输出运算结果。操作系统层是硬件层之上的第一层软件,负责管理和控制计算机硬件并对其进行扩充和完善,主要完成资源的调度与分配、信息的存取与保护、程序执行的协调与控制等工作,将计算机硬件层与上层软件进行分离,为其运行提供

计算机系统是为用户提供服务的,计算机系统所完成的任何工作,都是为了满足用户的计算或处理需求。用户在操作系统的帮助下,能够方便、快捷、安全和可靠地操作计算机硬件及其他软件运行自己的程序(包括自己编写的程序和应用软件程序)。因此,引入操作系统能更好地为用户提供服务,提供服务的软件接口如图 1-3 所示,主要通过以下三种方式使用计算机。

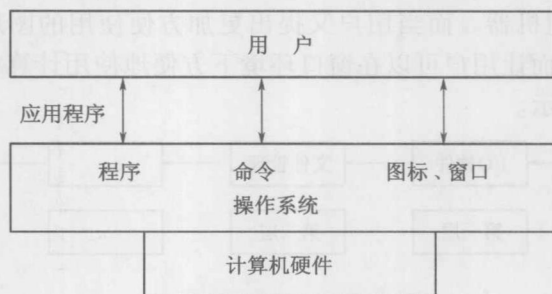


图 1-3 操作系统提供服务的软件接口

① 命令接口。这是应用用户较常用的操作系统接口,应用用户可以直接通过键盘输入命令来获得操作系统的服务。MS-DOS 及 Linux 的终端就是典型的命令接口的操作系统。

② 程序接口。这是应用程序开发人员开发应用程序过程中采用的接口方式,应用程序开发人员通过系统调用,在自身开发的程序中调用操作系统的服务,依据用户需求完成特定功能。

③ 图标、窗口接口。该接口方式实际上是对命令接口方式的一个扩展,通过对命令的封装,让用户可以通过更为直观的图标、窗口方式,利用鼠标操作,应用系统的各项功能及各种应用软件,这样使得计算机的操作更为简单、方便和生动有趣。Windows 就是典型的图形用户接口的操作系统。

(2) 操作系统作为计算机系统资源的管理者。

在计算机系统中通常有各种各样的计算机软硬件系统,可归纳分类为处理器、存储器、I/O设备和文件(保存的程序及数据)。为了合理地组织计算机工作流程,管理和分配计算机系统的以上资源,使之能够为更多的用户共享使用而引入操作系统。因此,操作系统是计算机系统中软件系统与硬件系统的管理者,负责各类资源的分配、回收等工作。操作系统开发人员应从总体角度考虑操作系统对各类资源的管理功能,并对其进行开发。

在操作系统对各类资源进行管理的同时,还需考虑在一个计算机系统中多个用户共享使用某类资源时可能发生冲突,为了更好地管理好这些共享资源(打印机、共享变量或链表等)的使用,操作系统必须监督各类资源的使用情况,并对请求使用进行管理,协调各个用户对共享资源的使用,避免发生冲突,即操作系统还是竞争资源的仲裁者。

(3) 操作系统是计算机系统功能的扩充。

仅有计算机硬件系统的计算机称为裸机,由于其仅能识别 0、1 代码,故在对其运用过程中需要用户编写其能识别的 0、1 代码的机器指令才能顺利执行,因此,即使硬件裸机功能再强,也必定难以使用。为了方便用户使用计算机硬件裸机资源,通过在计算机硬件上覆盖软件的方式隐蔽硬件细节,将计算机硬件的复杂性与用户进行隔离,从而让用户看到一台比硬件裸机功能更强、使用更加方便的计算机系统。

例如,为了方便使用 I/O 设备,为硬件增加了具有 Read 和 Write 命令的接口,帮助用户

完成数据的输入和输出。通常覆盖上软件的计算机系统称为扩充机器或虚机器,它向要执行的用户作业提供了一组对硬件管理和控制的抽象接口,用户作业通过该抽象接口使用计算机,而无需了解硬件接口的应用细节,从而使用户更容易地使用计算机硬件系统。为了方便用户使用文件系统,可以在现有扩充机器的基础上再覆盖上一层用于对文件进行管理的软件集合,并为用户提供一组对文件进行存取及保护的命令接口,此时用户已得到功能更强、使用更加方便的虚机器。而当用户又提出更加方便使用的图形界面时,还可以覆盖一层图标、窗口软件,从而让用户可以在窗口环境下方便地使用计算机,形成一台功能更强的虚机器,如图1-4所示。

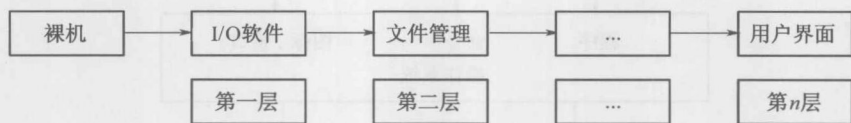


图1-4 虚机器构成

由此得出,操作系统是覆盖在计算机硬件裸机上的多层软件的集合,实现了对计算机硬件裸机的多层次抽象,它们不仅隐藏了硬件的复杂应用细节,而且丰富了系统的功能。另外,对计算机硬件的第一层抽象接口还可以被更高层次的抽象功能所调用,从而减少复杂的重复工作,使更高层次的抽象功能更加强大,用户使用起来也更加方便。

1.1.3 操作系统的目标

操作系统是对计算机硬件裸机的扩充,在计算机发展的不同阶段,对其扩充相应的软件,完善相应的功能,在不同阶段的各种类型的操作系统发展过程中各有侧重,总结其发展的目标如下:

1. 有效性

在计算机发展初期,计算机硬件裸机非常昂贵,操作系统最重要的目标无疑是有效性,即充分利用计算机硬件裸机上的各类资源,提高各类资源的利用率,从而提高计算机系统的运行效率。

在未配置操作系统的计算机硬件系统中,例如CPU和I/O设备不能同时工作,没有对内存及外存进行有效管理,不能合理安排程序进入内存存放的位置等,不能充分利用资源现象很严重;配置了操作系统后可以让CPU和I/O设备同时工作,而且有效管理内存及外存中程序的存放位置,即充分利用内外存空间。另外,操作系统中通过调度程序对要执行的程序进行合理的调度,充分提高各类资源的利用率,缩短程序在计算机上的处理时间,从而提高系统单位时间内完成的任务量。

2. 方便性

一台未配置操作系统的计算机硬件裸机仅能识别0和1的机器代码,用户如果想让自己的数据到计算机上运行并取得结果,那就要求用户必须将自己要处理的数据通过0和1这样的机器语言书写成对应的程序,再到计算机上去执行。而对于普通用户来说,通过机器语言书写一段完成指定计算功能的程序难度非常大,或者是不会写,即需要一批懂得机器语言的程序员来帮助完成机器语言程序的书写,才能去使用计算机硬件系统。可见,对于普通用户来说,使用计算机硬件裸机非常不方便,因此需要对计算机硬件裸机覆盖上可以将其他一些高级语言所编写的程序翻译成机器代码程序的编译软件,通过编译命令极大地方便了用户,从而使计

算机更加容易使用。

3. 可扩充性

随着大规模集成电路技术和超大规模集成电路技术的发展,计算机硬件体系结构、多处理机及计算机网络得到了迅速发展,进而对相应的操作系统软件体系结构也提出了更高的功能及性能需求。因此,操作系统必须具备良好的可扩充性以适应计算机硬件、体系结构及用户应用需求的发展。

4. 开放性及标准化

由于计算机网络的迅速普及,基于网络的应用也日益增多,使计算机系统的应用从封闭的单机环境转变为开放的网络环境,为使来自不同生产商的计算机硬件设备能在开放的网络环境中进行联系、通信,并且网络应用可在不同厂商的计算机之间进行移植和相互操作,因此需要操作系统软件提供统一的开放网络环境,即操作系统必须具有开放性及标准化。

开放性及标准化是指操作系统软件遵循国际标准及规范,特别是遵循开放系统互联国际标准。遵循国际标准设计的硬件及开发的软件均能彼此兼容,实现相互操作。20 世纪 90 年代后新开发的系统及软件能否被广泛推广的关键因素正是其是否具有开放性。

1.2 操作系统的形成与发展

操作系统自诞生起,从最初的人工操作方式由于处理机等待人工操作而使处理机资源得不到充分利用,到单道批处理系统减少人工操作,提高了处理机资源的利用率,但由于计算机系统中的资源如 CPU 与 I/O 设备不能同时工作,即未能提高计算机系统中各类资源的利用率。为提高计算机系统中各类资源的利用率,引入了多道程序设计技术,通过允许多道程序在内存中轮流使用 CPU 及其他资源,从而充分利用计算机系统各类资源。至此真正的操作系统已经诞生。随着用户应用需求的提出、计算机硬件体系结构及网络的发展,操作系统经历了分时系统、实时系统、网络操作系统、分布式操作系统等发展过程。

1.2.1 操作系统的产生与形成

1. 人工阶段

从计算机诞生到 20 世纪 50 年代中期的计算机属于第一代计算机,机器速度慢、规模小、外设少,操作系统尚未出现。计算机的操作由程序员采用人工操作,直接控制和使用计算机硬件。程序员采用机器语言完成用户程序的书写,并将事先编好的程序和数据通过在纸带或卡片上打孔来记录,由纸带或卡片输入机将纸带或卡片携带的程序和数据输入计算机。然后启动计算机运行用户程序,程序员通过控制台上的按钮或开关来操纵和控制程序,运行完毕后,用户通过将结果穿孔到对应的纸带或者卡片上取走计算结果,才轮到下一个用户上机。

这种工作方式存在的严重缺点如下:

(1) 用户独占全机,造成计算机资源利用率不高,计算机系统效率低下。

(2) 人工干预,如装纸带或卡片、按开关等。人工操作过多,中央处理机等待人工操作时间较长,不仅浪费计算机时间,还极易发生操作错误。

这种工作方式在刚诞生的计算机上还能接受,随着计算机运算速度的提高,其缺点就暴露得更加明显了。例如,一个作业在计算机上需 1h 运行完成,作业的建立和人工干预花费

4min,那么,手工操作时间占总运行时间的6.25%;如果计算机运行速度提高10倍,那么作业的运行时间仅需6min,而手工操作仍然需要4min,这时手工操作时间占了总运行时间的67%。因此,随着计算机运算速度的提高,缩短手工操作时间即解决人机矛盾非常必要。

2. 单道批处理阶段

随着第二代计算机的产生,计算机运算速度明显提高,人机矛盾已到了不可容忍的地步。为了解决这一矛盾,只有设法减少或去掉人工干预,这样就出现了脱机批处理阶段。

当时的计算机非常昂贵,人们自然希望减少时间浪费,采用的方案就是每次提交的作业不止一个,而是一批,即批处理系统。其实现是在一台相对便宜的计算机上输入所有的作业到磁带上,如IBM 1401,它适合将卡片数据读入计算机、由计算机复制到磁带以及将计算机上的数据打印输出,但不适合进行大量的数学计算,故仅作为输入/输出机使用,而完成大量数学计算且速度较快的计算机称为主机,如IBM 7094。

各用户把自己的作业交给机房,由操作员把这批作业从输入机输入并复制到磁带上,与作业同时提交的还有对应的作业说明书,表明了作业的控制意图,然后将磁带送到主机所在机房,主机通过读取磁带上的作业完成作业的输入,并开始数学计算,计算结果存储到一个磁带上,等这批数据的执行结果都存储到磁带上后,将磁带拿到输出机完成打印输出。在此过程中一个作业一个作业的顺序执行过程需要由一段特殊的程序来控制管理,该程序由作业控制语言进行编写,能控制程序的运行,即监督程序。由于当时作业已经采用汇编语言或FORTRAN语言编写,故为了在主机上执行该作业,还必须提供对应语言的编译程序,将程序编译、链接成可执行程序。

为了执行一个作业,监督程序将解释该作业的作业说明书,若系统资源能满足其要求,则将该作业调入内存,并从外部存储器(如磁带)上输入所需要的编译程序。编译程序将用户源程序翻译成目标代码,然后由链接装配程序把编译后的目标代码及其所需的子程序装配成一个可执行的程序,接着开始执行。计算完成后输出该作业的计算结果。只有一个作业处理完毕后,监督程序才可以自动地调度下一个作业进行处理,依次重复上述过程,直到该批作业全部处理完毕。第一批程序执行结果去输出机上输出的同时,监督程序又可以开始第二批输入数据的计算工作。

由于作业的输入是在输入机而不是主机的控制下,也就是说输入/输出工作是在脱离主机的情况下进行的,故称为脱机输入/输出(Off Line I/O),脱机的批处理系统如图1-5所示。反之,如果主机直接控制输入/输出工作的称为联机输入/输出。由于联机的输入/输出存在浪费中央处理机时间等待人工输入操作的问题,故在批处理系统中采用脱机输入/输出方式更加合理。该批处理系统虽然提交了一批作业,但执行的时候还是一个作业一个作业地运行,故该批处理系统也称为单道批处理系统(Simple Batch Processing System, SBPS)。单道批处理系统的执行过程如图1-6所示。该系统的主要优点如下:

(1) 减少了CPU的等待时间。输入和输出不是在CPU的控制下完成的,即不占用CPU时间,以减少CPU的空闲等待时间。

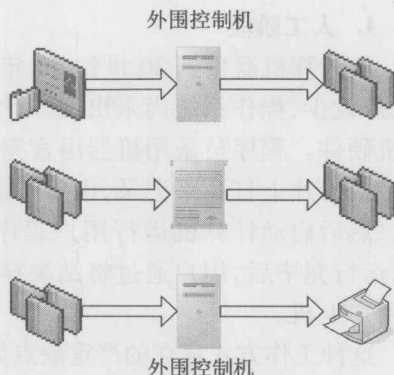


图1-5 脱机的批处理系统

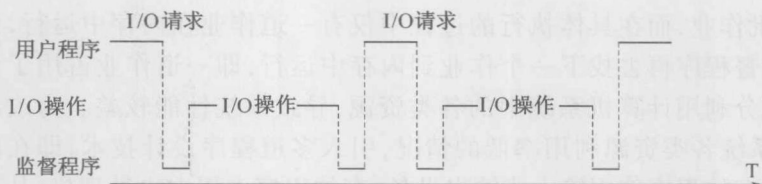


图 1-6 单道批处理系统的执行过程

(2) 提高 I/O 速度。对于 CPU 的输入与输出是指 CPU 直接从磁带读取数据及将运算结果写入磁带,很显然,比人工从读卡机输入程序与数据的速度要快很多,从而缓解了 CPU 与 I/O 设备之间速度不匹配的矛盾。

单道批处理系统是最早出现的能控制作业自动运行的软件,但还需人工干预系统中资源及设备的管理,因此,它只能算是办公系统(OS)的前身而并非是现在人们所理解的操作系统软件。尽管如此,该系统比起人工操作方式的系统已有很大进步。该系统的主要特征如下:

(1) 自动性:在不发生程序异常的情况下,保存在磁带上的整批作业自动地逐个按顺序执行,无需人工干预。

(2) 顺序性:磁带上的每道作业按照提交顺序进入内存,每道作业的完成顺序与它们进入内存的顺序一致,即先到达的作业先完成。

(3) 单道性:在单道批处理系统中,由监督程序控制每道作业的运行,每次仅允许一道作业进入内存,该程序运行完成或异常结束时,再由监督程序寻找下一程序进入内存运行。

单道批处理系统的出现改善了 CPU 和外设的使用情况,实现了作业的自动定序、自动过渡,从而使整个计算机系统的处理能力得以提高。但仍存在着许多缺陷,如卫星机与主机之间的磁带装卸仍需人工完成,操作员需要监督机器的状态等。如果一个程序进入死循环,系统就会踏步不前,只有当操作员提出请求,要求终止该作业,删除它并重新启动,系统才能恢复正常运行。当目标程序执行一条引起停机的非法指令时,机器就会错误地停止运行。此时,只有操作员进行干预,即在控制台上按启动按钮后,程序才会重新启动运行。并且,由于系统没有任何保护自己的措施,无法防止用户程序破坏监督程序和系统程序,即系统保护的问题亟待解决。

3. 多道批处理系统

在 1964 年之后,随着第三代计算机的诞生即计算机由集成电路进行设计生产,其特征是用集成电路 IC(Integrated Circuit)代替了分离元件,集成电路是把多个电子元器件集中在几平方毫米的基片上形成的逻辑电路。第三代计算机的基本电子元件是每个基片上集成几个到十几个电子元件(逻辑门)的小规模集成电路和每片上几十个元件的中规模集成电路。第三代计算机已开始采用性能优良的半导体存储器取代磁芯存储器,运算速度提高到每秒几十万到几百万次基本运算,在存储器容量和可靠性等方面都有了较大的提高。同时,计算机软件技术的进一步发展,尤其是操作系统的逐步成熟是第三代计算机的显著特点。最有影响的是 IBM 公司研制的 IBM 360 计算机系列,该系列计算机体积更小,功耗更低,速度更快。

1) 多道程序设计技术

单道批处理系统所提交的一批作业在监督程序的监督下按照顺序依次执行,在执行的过程中监督程序让其中的一个程序经过编译、链接,装入内存运行,运行结束后再将使用权交给监督程序,由监督程序查找是否存在下一个作业,继续循环执行。很显然,在此过程中