



普通高等教育规划教材

操作系统教程

武伟 主编

 机械工业出版社
CHINA MACHINE PRESS



普通高等教育规划教材

操作系统教程

主 编 武 伟
副主编 邹姝稚
参 编 刘镇宇 周彩英 张成姝
主 审 胡金初



机械工业出版社

本书介绍了操作系统的基本原理,剖析了典型操作系统 Linux 和 Windows 2000。主要内容包括:处理器管理、存储器管理、I/O 设备管理、文件系统、Linux 和 UNIX 剖析及 Windows 2000 剖析。各章均配有适量习题,有助于读者领会和掌握已学过的有关知识。

本书在编写过程中力求做到:一要结合实际、突出应用,二要便于教学,注意内容的先进性。为保证实例的系统性,把 Linux、Windows 2000 作为两章单独写出,同时保证这两章的层次性,以和前面的章节配合。

本书可作为高等学校计算机科学与技术专业本科教材,也可作为非计算机专业的教学参考书。

图书在版编目(CIP)数据

操作系统教程/武伟主编. —北京:机械工业出版社,2004.1
普通高等教育规划教材
ISBN 7-111-13775-2

I. 操... II. 武... 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(2003)第 124572 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)
策划编辑:王小东 责任编辑:苏颖杰 版式设计:冉晓华
责任校对:张晓蓉 封面设计:饶薇 责任印制:施红
北京铭成印刷有限公司印刷·新华书店北京发行所发行

2004 年 2 月第 1 版第 1 次印刷

787mm × 1092mm $1/16$ · 15.5 印张 · 381 字

定价:22.00 元



凡购本书,如有缺页、倒页、脱页,由本社发行部调换
本社购书热线电话(010)68993821、88379646
封面无防伪标均为盗版

普通高等教育应用型人才培养规划教材 编审委员会委员名单

- 主任：刘国荣 湖南工程学院
- 副主任：左健民 南京工程学院
- 陈力华 上海工程技术大学
- 鲍 泓 北京联合大学
- 王文斌 机械工业出版社
- 委员：(按姓氏笔画排序)
- 任淑淳 上海应用技术学院
- 何一鸣 常州工学院
- 陈文哲 福建工程学院
- 陈志强 华北航天工业学院
- 陈 峻 扬州大学
- 苏 群 黑龙江工程学院
- 娄炳林 湖南工程学院
- 梁景凯 哈尔滨工业大学(威海)
- 童幸生 江汉大学

计算机科学与技术专业分委员会委员名单

- 主任：黄陈蓉 南京工程学院
副主任：吴伟昶 上海应用技术学院
委员：（按姓氏笔画排序）
 汤 惟 江汉大学
 沈 涛 扬州大学
 陈文强 福建工程学院
 肖建华 湖南工程学院
 邵祖华 浙江科技学院
 靳 敏 黑龙江工程学院

序

工程科学技术在推动人类文明的进步中一直起着发动机的作用。随着知识经济时代的到来,科学技术突飞猛进,国际竞争日趋激烈。特别是随着经济全球化发展和我国加入 WTO,世界制造业将逐步向我国转移。有人认为,我国将成为世界的“制造中心”。有鉴于此,工程教育的发展也因此面临着新的机遇和挑战。

迄今为止,我国高等工程教育已为经济战线培养了数百万专门人才,为经济的发展作出了巨大的贡献。但据 IMD1998 年的调查,我国“人才市场上是否有充足的合格工程师”指标排名世界第 36 位,与我国科技人员总数排名世界第一形成很大的反差。这说明符合企业需要的工程技术人员特别是工程应用型技术人才市场供给不足。在此形势下,国家教育部近年来批准组建了一批以培养工程应用型本科人才为主的高等院校,并于 2001、2002 年两次举办了“应用型本科人才培养模式研讨会”,对工程应用型本科教育的办学思想和发展定位作了初步探讨。本系列教材就是在这种形势下组织编写的,以适应经济、社会发展对工程教育的新要求,满足高素质、强能力的工程应用型本科人才培养的需要。

航天工程的先驱、美国加州理工学院的马·卡门教授有句名言:“科学家研究已有的世界,工程师创造未有的世界。”科学在于探索客观世界中存在的客观规律,所以科学强调分析,强调结论的惟一性。工程是人们综合应用科学(包括自然科学、技术科学和社会科学)理论和技术手段去改造客观世界的实践活动,所以它强调综合,强调方案优缺点的比较并做出论证和判断。这就是科学与工程的主要不同之处。这也就要求我们对工程应用型人才的培养和对科学研究型人才的培养应实施不同的培养方案,采用不同的培养模式,采用具有不同特点的教材。然而,我国目前的工程教育没有注意到这一点,而是:①过分侧重工程科学(分析)方面,轻视了工程实际训练方面,重理论,轻实践,没有足够的工程实践训练,工程教育的“学术化”倾向形成了“课题训练”的偏软现象,导致学生动手能力差。②人才培养模式、规格比较单一,课程结构不合理,知识面过窄,导致知识结构单一,所学知识中有一些内容已陈旧,交叉学科、信息学科的内容知之甚少,人文社会科学知识薄弱,学生创新能力不强。③教材单一,注重工程的科学分析,轻视工程实践能力的培养;注重理论知识的传授,轻视学生个性特别是创新精神的培养;注重教材的系统性和完整性,造成课程方面的相互重复、脱节等现象;缺乏工程应用背景,存在内容陈旧的现象。④老师缺乏工程实践经验,自身缺乏“工程训练”。⑤工程教育在实践中与经济、产业的联系不密切。要使我国工程教育适应经济、社会的发展,培养更多优秀的工程技术人员,我们必须努力改革。

组织编写本套系列教材,目的在于改革传统的高等工程教育教材,建设一套富有特色、有利于应用型人才培养的本科教材,满足工程应用型人才培养的要求。

本套系列教材的建设原则是:

1. 保证基础,确保后劲

科技的发展,要求工程技术人员必须具备终生学习的能力。为此,从内容安排上,保证

学生有较厚实的基础,满足本科教学的基本要求,使学生日后具有较强的发展后劲。

2. 突出特色, 强化应用

围绕培养目标,以工程应用为背景,通过理论与工程实际相结合,构建工程应用型本科教育系列教材特色。本套系列教材的内容、结构遵循如下9字方针:知识新、结构新、重应用。教材内容的要求概括为:“精”、“新”、“广”、“用”。“精”指在融会贯通教学内容的基础上,挑选出最基本的内容、方法及典型应用;“新”指在将本学科前沿的新进展和有关的技术进步新成果、新应用等纳入教学内容,以适应科学技术发展的需要。妥善处理传统内容的继承与现代内容的引进。用现代的思想、观点和方法重新认识基础内容和引入现代科技的新内容,并将这些按新的教学系统重新组织;“广”指在保持本学科基本体系下,处理好与相邻以及交叉学科的关系;“用”指注重理论与实际融会贯通,特别是注入工程意识,包括经济、质量、环境等诸多因素对工程的影响。

3. 抓住重点, 合理配套

工程应用型本科教育系列教材的重点是专业课(专业基础课、专业课)教材的建设,并做好与理论课教材建设同步的实践教材的建设,力争做好与之配套的电子教材的建设。

4. 精选编者, 确保质量

遴选一批既具有丰富的工程实践经验,又具有丰富的教学实践经验的教师担任编写任务,以确保教材质量。

我们相信,本套系列教材的出版,对我国工程应用型人才培养质量的提高,必将产生积极作用,会为我国经济建设和社会发展作出一定的贡献。

机械工业出版社颇具魄力和眼光,高瞻远瞩,及时提出并组织编写这套系列教材,他们为编好这套系列教材做了认真细致的工作,并为该套系列教材的出版提供了许多有利的条件,在此深表衷心感谢!

编委会主任

湖南工程学院院长

刘国荣教授

前 言

操作系统是计算机系统中最重要系统软件，它是围绕着如何提高计算机资源利用率和改善用户界面的友好性而形成、发展和不断成熟的。进入 21 世纪以来，我国的高等学校和 IT 界，对操作系统的关心和重视程度是前所未有的，因为人们普遍认识到，计算机操作系统是整个 IT 领域中的重要基础。要构建现代化的、稳固而又可靠的信息技术大厦，不掌握计算机操作系统原理是不行的。

本书针对培养技术型人才的特点，在注重操作系统原理的基础上，引入了操作系统发展过程中的最新技术，以 Linux 和 Windows 2000 操作系统为实例，剖析了现代操作系统所采用的最新技术。

本书是作者在多年教学实践和科学研究的基础上，参阅了大量国内外操作系统教材后，编写的一本适用于本科计算机科学与技术专业的教材。其编写思路及特点如下：

1) 以最新主流操作系统 Linux 和 Windows 2000 作为实例，从操作系统原理的角度对其作了详尽的介绍，并在讲授原理时注重理论联系实际。

2) 为了使学生对 Linux、Windows2000 操作系统有比较系统的了解，本书改变其他教科书的编排顺序，将其作为两个章节单独列出，同时也保证了这两章的层次性。授课教师也可把这两章拆开，分配到前面各章讲解。

3) 根据作者的教学经验，对于比较难于理解的部分，本书均以实例引出，语言深入浅出，使学生能够从简单的实例入手，比较容易地掌握操作系统的内部工作原理。

4) 本书配有适量经过精选的习题，以帮助读者检验和加深对内容的理解。

本书参考教学时数为 70~80 学时，要求先修课程为“数据结构”、“汇编语言”、“Pascal 语言”和“计算机组成原理”。

本书的内容是按照理工科院校计算机科学与技术专业的教学大纲编写的。对于非计算机专业的本科专业及非本科的计算机专业，可适当删减内容。

全书共分 12 章，第 1、11 章由刘镇宇编写，第 5、9 章由周彩英编写，第 10 章由张成姝编写，第 3、4、6 章由邹姝雅编写，第 2、7、8 章由王伟编写，第 12 章由王伟、张成姝编写，王伟任主编，负责全书的统一编排定稿。

上海师范大学胡金初副教授任主审，对本书进行了认真的审阅，并提出了修改意见。在此，谨向胡金初老师表示诚挚的谢意！

由于时间仓促，加之作者水平所限，书中错误和不足之处在所难免，敬请各位读者批评指正。同时，也恳请各位读者一旦发现错误，及时与编者联系，以便尽快更正，编者将不胜感激。编者的电子邮件地址是：weiwu_008@163.com。

编 者

目 录

序 前言

第 1 章 引论	1
1.1 操作系统的作用和定义	1
1.1.1 操作系统的作用	1
1.1.2 操作系统的定义	2
1.2 计算机系统的硬件资源	2
1.3 操作系统的发展	3
1.4 操作系统的功能和特性	6
1.4.1 操作系统的功能	6
1.4.2 操作系统的特性	7
1.5 操作系统的分类	7
1.5.1 批处理系统	8
1.5.2 分时操作系统	8
1.5.3 实时操作系统	9
1.5.4 网络操作系统	10
1.5.5 分布式操作系统	11
1.6 操作系统的结构模型	11
1.6.1 整体式模型	11
1.6.2 层次式模型	12
1.6.3 客户/服务器系统模型	12
习题 1	13
第 2 章 用户与操作系统的接口	15
2.1 作业控制级接口	15
2.1.1 作业	15
2.1.2 作业的类型	15
2.1.3 联机用户接口	16
2.1.4 脱机用户接口	17
2.2 程序级接口	19
2.2.1 用户态和核心态	19
2.2.2 特权指令和访管指令	19
2.2.3 系统调用	20
2.2.4 系统调用的使用和执行过程	21
习题 2	21

第 3 章 进程的描述与控制	22
3.1 程序执行方式	22
3.1.1 程序顺序执行	22
3.1.2 程序并发执行	23
3.2 进程的描述	25
3.2.1 进程的定义	25
3.2.2 进程的特性	25
3.2.3 进程控制块	26
3.3 进程的状态	27
3.3.1 进程的基本状态	27
3.3.2 进程的挂起	28
3.4 进程控制	29
3.4.1 内核	29
3.4.2 微内核	30
3.4.3 进程控制原语	30
3.5 线程	31
3.5.1 线程的引入	32
3.5.2 线程的定义	32
3.5.3 线程的状态	32
3.5.4 线程和进程的比较	33
3.5.5 线程的分类	33
习题 3	35
第 4 章 进程通信	36
4.1 进程的同步与互斥	36
4.1.1 进程合作	36
4.1.2 共享资源	37
4.1.3 与时间有关的错误	37
4.1.4 临界资源与临界区	39
4.1.5 同步机构设计准则	40
4.2 互斥的软件方法	40
4.3 硬件指令机制	43
4.3.1 测试与设置技术	43
4.3.2 TS 指令	44
4.3.3 利用 TS 实现进程互斥	44
4.4 信号量(Semaphore)机制	45

4.4.1 整型信号量	45	6.1.2 死锁的必要条件	75
4.4.2 记录型信号量	46	6.1.3 死锁的对策	76
4.4.3 AND型信号量集机制	49	6.2 死锁预防	77
4.5 经典进程同步问题	49	6.3 死锁避免	78
4.5.1 生产者-消费者问题	49	6.3.1 安全和不安全状态	78
4.5.2 哲学家就餐问题	51	6.3.2 利用银行家算法避免死锁	79
4.5.3 读者-写者问题	51	6.4 死锁的检测	81
4.6 进程通信	53	6.5 死锁的解除	83
4.6.1 共享存储区系统	53	6.6 死锁的综合处理	84
4.6.2 管道通信系统	54	习题6	84
4.6.3 消息传递系统	54	第7章 实存储管理技术	86
习题4	58	7.1 存储管理的基本概念	86
第5章 处理器调度	60	7.1.1 存储管理要解决的问题	86
5.1 三级调度的概念	60	7.1.2 存储管理的分类	86
5.1.1 作业的状态及其转换	60	7.1.3 地址重定位	87
5.1.2 调度的层次	61	7.2 连续分配存储管理方式	89
5.1.3 作业和进程的关系	62	7.2.1 单一连续分配方式	89
5.2 作业调度	63	7.2.2 固定分区存储管理方式	90
5.2.1 作业调度的功能	63	7.2.3 可变分区存储管理方式	90
5.2.2 作业调度的目标与性能衡量	63	7.3 离散分配存储管理方式	97
5.3 进程调度	65	7.3.1 分页存储管理方式	97
5.3.1 进程调度的功能	65	7.3.2 分段存储管理方式	103
5.3.2 进程调度的方式	65	7.3.3 段页式存储管理方式	106
5.3.3 进程调度的时机	66	习题7	109
5.4 常用的调度算法	67	第8章 虚拟存储管理技术	111
5.4.1 先来先服务调度算法	67	8.1 虚拟存储器的基本概念	111
5.4.2 短作业(进程)优先调度算法	67	8.1.1 局部性原理	111
5.4.3 时间片轮转调度算法	68	8.1.2 虚拟存储器	112
5.4.4 优先权调度算法	69	8.2 请求分页式存储管理方式	112
5.4.5 最高响应比优先调度算法	69	8.2.1 请求分页式存储管理的基本 概念	112
5.4.6 多级队列调度算法	70	8.2.2 页面分配策略	113
5.4.7 多级反馈队列调度算法	70	8.2.3 页面调入时机	115
5.5 实例分析: UNIX 进程调度	71	8.2.4 页面置换算法	115
5.5.1 调度时机	71	8.2.5 请求分页系统的性能分析	119
5.5.2 调度标记设置	71	8.3 请求分段式存储管理方式	121
5.5.3 优先数计算	71	8.3.1 请求分段式存储管理的基本 概念	121
5.5.4 调度的实现	72	8.3.2 分段共享与保护	122
习题5	72	习题8	123
第6章 死锁	74		
6.1 死锁的基本概念	74		
6.1.1 死锁产生的原因	74		

第 9 章 设备管理	125	10.3.2 单级目录结构	158
9.1 I/O 系统的组成	125	10.3.3 二级目录结构	159
9.1.1 I/O 系统的结构	125	10.3.4 多级目录结构	160
9.1.2 I/O 设备概述	126	10.3.5 目录查询技术	160
9.1.3 设备控制器	127	10.4 文件存储空间的管理	161
9.1.4 I/O 通道	127	10.4.1 空白文件目录	161
9.2 I/O 控制方式	128	10.4.2 空白块链	162
9.2.1 程序 I/O 方式	128	10.4.3 位示图	162
9.2.2 中断驱动 I/O 方式	129	10.4.4 MS-DOS 的盘空间管理	163
9.2.3 直接存储器存取方式	129	10.4.5 UNIX 文件存储空间的管理	163
9.2.4 I/O 通道方式	130	10.5 文件的共享	164
9.3 缓冲管理	130	10.5.1 目录结构中的共享	165
9.3.1 缓冲的引入	130	10.5.2 打开文件结构中的共享	166
9.3.2 缓冲区及其管理	131	10.5.3 管道(Pipe)	167
9.4 设备分配	133	10.6 文件系统的安全性	169
9.4.1 设备分配中的数据结构	133	10.6.1 文件的存取控制	169
9.4.2 设备分配策略	135	10.6.2 分级安全管理	171
9.4.3 设备独立性	136	10.6.3 文件的转储和恢复	173
9.4.4 独占设备分配方法	136	习题 10	174
9.4.5 SPOOLing 技术	138	第 11 章 Linux 操作系统	176
9.5 磁盘 I/O	139	11.1 Linux 概述	176
9.5.1 磁盘性能概述	139	11.1.1 Linux 的历史和特点	176
9.5.2 数据的组织	140	11.1.2 用户账号和系统安全	177
9.5.3 磁盘访问时间	140	11.1.3 终端用户界面简介	177
9.5.4 磁盘调度算法	141	11.1.4 Shell 程序设计	180
9.6 I/O 软件的组成	143	11.1.5 系统调用	180
9.6.1 中断处理程序	143	11.2 进程管理	180
9.6.2 设备驱动程序	146	11.2.1 Linux 进程状态	180
9.6.3 与设备无关的系统软件	147	11.2.2 task_struct 结构	181
9.6.4 用户空间的 I/O 软件	148	11.2.3 Linux 进程的创建	182
习题 9	149	11.2.4 进程使用的虚拟内存	184
第 10 章 文件系统	151	11.2.5 Linux 的进程调度	184
10.1 文件系统概述	151	11.2.6 Linux 的进程通信	186
10.1.1 文件和文件系统	151	11.3 存储管理	188
10.1.2 文件的分类	152	11.3.1 80386 体系结构的存储管理 功能	188
10.1.3 文件系统的功能和基本操作	153	11.3.2 进程的虚存管理和物理内存 分配	191
10.2 文件的结构	154	11.3.3 按需调页和内存页的使用	191
10.2.1 文件的逻辑结构	154	11.4 文件系统	192
10.2.2 文件的物理结构	154	11.4.1 ext2 文件系统	192
10.3 文件目录	157	11.4.2 虚拟文件系统	196
10.3.1 文件控制块	157		

11.4.3	VFS 索引节点缓存、目录缓存 和缓冲区缓存	199	12.3.3	Windows 2000 中线程的实现	211
11.5	设备管理	200	12.3.4	Windows 2000 的线程调度	212
11.5.1	设备和设备特殊文件	200	12.4	Windows 2000 的存储管理	216
11.5.2	设备驱动程序的动态安装	201	12.4.1	Windows 2000 的虚拟地址空间	216
11.5.3	设备驱动程序	201	12.4.2	页面状态	216
11.5.4	设备文件的存取权限和系统 安全	203	12.4.3	地址转换机制	217
第 12 章	Windows 2000	204	12.4.4	页面目录、页表和 PTE 的结构	218
12.1	Windows 2000 的历史	204	12.4.5	页面调度策略与工作集管理	219
12.1.1	MS-DOS	204	12.4.6	Windows 2000 的存储保护	220
12.1.2	Windows 95/98/Me	204	12.5	Windows 2000 的设备管理	220
12.1.3	Windows NT	205	12.5.1	Windows 2000 的 I/O 设计目标	220
12.1.4	Windows 2000	205	12.5.2	Windows 2000 的 I/O 结构	220
12.2	Windows 2000 的体系结构	206	12.5.3	Windows 2000 的设备驱动程序	221
12.2.1	系统的要求与设计目标	206	12.5.4	Windows 2000 的 I/O 处理方法	223
12.2.2	Windows 2000 的体系结构	207	12.6	Windows 2000 的文件系统	227
12.3	Windows 2000 的进程管理与处理器 调度	211	12.6.1	Windows 2000 文件系统格式	227
12.3.1	Windows 2000 的进程和线程模型 总述	211	12.6.2	NTFS 文件系统	228
12.3.2	Windows 2000 中进程的实现	211	12.6.3	NTFS 的磁盘结构	230
			12.6.4	NTFS 的可靠性	232
			12.6.5	NTFS 的安全性	233
			参考文献	234	

第 1 章 引 论

计算机系统包括硬件和软件两个部分,操作系统(Operating System, OS)是最重要的系统软件,它扩充了硬件功能,并提供软件运行环境,实现了应用软件和硬件设备的连接。本章从操作系统的作用、发展历史、类型和功能等方面来进行介绍。

1.1 操作系统的作用和定义

1.1.1 操作系统的作用

1. 操作系统是系统软件

计算机系统可以看成是一个由硬件和软件组成的有层次结构的系统。最底层是硬件,除实际设备外,机器语言也属于硬件部分,硬件部分通常称为裸机。用户直接编程来控制硬件是很麻烦的,也容易出错。为此在硬件基础上加一层软件,由它控制和管理硬件,起到隐藏硬件复杂性的作用。呈现在用户面前的是经过“包装”的虚拟机,与裸机相比要抽象,可以使用户容易理解,使用方便。操作系统就是这层软件,是最接近硬件的一层软件。

计算机系统软件通常分为系统软件和应用软件两大类。系统软件用于计算机的管理、维护、控制和运行,系统软件除操作系统外,还有命令解释程序、编辑程序、编译程序和连接装配程序等通用程序。应用软件则指的是为解决某一特定问题而编制的程序。操作系统是最重要的系统软件。

在所有的软件中,一部分软件的运行又要以另一部分软件的存在为基础,新增加的软件可以看作是原来那部分软件的扩充与完善。例如,电子邮件收发软件和网络浏览器软件就是工作在网络通信协议软件基础上的。由此可见,软件也分成若干层,每增加一个软件层后就变成一个功能更强的机器,通常把扩充后的“新机器”称为“虚拟机”。虚拟机的性质可以随应用的不同而变化。例如,民航售票系统所用的计算机特性与化工生产过程控制计算机的特性是不同的,就是因为它们配备了不同的操作系统。计算机硬件和软件的层次关系如图 1-1 所示。

2. 操作系统是系统资源管理者

如同 CPU 是硬件的核心一样,操作系统是软件的核心。从计算机系统资源管理的角度看,引入操作系统是为了合理地组织计算机的工作流程,管理和分配计算机系统硬件和软件资源,最大限度地提高计算机系统的利用率。其主要功能是对处理器、存储器、I/O 设备和信息资源进行有效的管理。可见,操作系统是计算机系统的资源管理者。

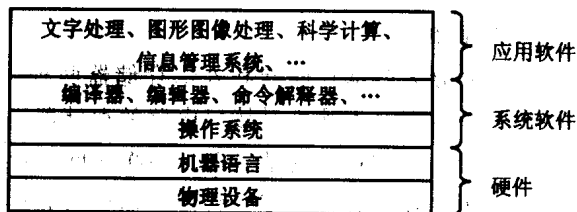


图 1-1 计算机硬件和软件的层次关系

3. 操作系统是用户与计算机硬件系统

之间的接口

操作系统的一个重要任务是使计算机便于用户使用。操作系统处于用户与计算机硬件之间，它将计算机硬件系统中各种复杂的运算和操作指令按功能进行汇总，以模块程序方式完成。它为用户提供操作系统中某个软件模块和计算机硬件资源一起完成的某个操作或运算。系统提供的接口有两类：

(1) 作业级接口 操作系统提供一组联机命令，用户可以通过键盘键入有关的命令，以组织和控制自己的作业运行。

(2) 程序级接口 操作系统提供一组系统调用，即操作系统中的某个功能模块，用户可在应用程序中通过调用相应的系统调用来操纵计算机。

由此可见，用户不直接与计算机硬件系统打交道，所以说操作系统是用户和计算机硬件系统之间的接口。

1.1.2 操作系统的定义

综上所述，操作系统可以定义为：操作系统是直接控制和管理计算机系统硬件和软件资源，合理地组织计算机工作流程，便于用户使用的程序的集合。

从操作系统的上述定义可以看出，操作系统的目标有两点，首先，操作系统要方便用户使用，一个好的操作系统应提供给用户一个易于理解、简洁方便的用户界面，这里的“用户”既包括计算机系统的最终用户，又包括计算机系统的管理员，也包括编写应用程序的程序员；其次，操作系统应尽可能地使系统中的各种资源得到最充分的利用。

后面各章所讨论的各种实现技术、算法都围绕着这两个主要目标。

1.2 计算机系统的硬件资源

由于操作系统是直接控制和管理计算机系统硬件的一层软件，在后面的章节中读者将看到操作系统如何做到隐藏硬件细节，为用户提供简单易用的接口。在这里，我们先讨论一下计算机系统的主要硬件。

1. 处理器(CPU)

CPU是计算机的“心脏”，它由运算器和控制器两部分组成。控制器是计算机的控制中心，负责指令的执行；而运算器是计算机的数据加工和处理的场所，用来完成算术和逻辑运算。

现代计算机系统大都是多道程序系统，计算机系统中CPU的数目往往少于正在运行程序的数目。只有一个处理器的计算机系统，称单处理器系统；有多个处理器的系统，称多处理器系统。

2. 存储器

存储器分为内部存储器和外部存储器两种。内部存储器(也叫内存或主存)是指CPU可直接存取其中数据，不须经过接口电路的存储器。内部存储器中每个存储单元都有一个编号，称为存储地址，也称物理地址。CPU每次访问内存时必须先指定存储单元地址。在多道程序系统中，由于内存中会存放多个程序，需要有硬件对其加以保护，使它们相互隔离。外存储器包括磁盘、光盘和优盘等，CPU要经过I/O设备控制器才可访问。与内存相比，外存存取速度要慢，但容量大，每位存储量的价格低于内存。

3. 输入输出设备(外围设备或 I/O 设备)

I/O 设备一般由机械部分和控制电路部分组成,控制电路部分即接口电路,通常又称控制器。在微机中,控制器做成印制电路板卡即适配卡,插在计算机内,经系统总线与 CPU 相连。适配卡与设备之间的接口通常采用标准接口(或事实上的工业标准),如 IDE、SCSI 和 RS-232C 等。操作系统对设备的控制实际上是将命令和数据按一定顺序写入控制器指定的寄存器中,包括从寄存器中读取设备的状态。

计算机系统应用于各种场合,配备各种 I/O 设备,这些设备的信息处理速度差异很大。输入输出设备大致分为两类:块设备和字符设备。块设备是以数据块为单位存取的,块的大小固定,每块都有编号,可按块存取;字符设备则以字符为单位存取,没有块结构。

1.3 操作系统的发展

回顾操作系统的发展历史,我们可以看到,操作系统是随着计算机技术的发展和计算机的应用范围越来越广泛而发展的。为了更清楚地把握操作系统的实质,了解操作系统的发展是很有必要的。因为操作系统的许多基本概念和实现技术,都是在操作系统的发展过程中出现并逐步得到发展,进而走向成熟的。了解操作系统的发展历史,有助于我们更深刻地认识操作系统基本概念的内含含义。

1. 第一代计算机(1945~1955):电子管和插件板

20 世纪 40 年代中期,美国科学家采用冯·诺依曼体系,使用真空管成功地建造了第一台电子数字计算机。这个机器使用了上万个电子管,占据了几个房间,然而其运算速度还比不上现在最便宜的微型计算机。

在早期,程序设计全部采用机器语言,通过在一些插件板上的连线来控制计算机。没有程序设计语言(甚至没有汇编语言),当然也没有操作系统。使用机器的方式是程序员提前预约一段机时,到机房将他的插件板插到计算机里,在接下来的几个小时里计算机全部归他一个人使用。在这段时间里,上万个真空管中没有烧坏的,就很幸运了。当时使用计算机都是一些数值计算方面的项目。

2. 第二代计算机(1955~1965):晶体管和批处理系统

20 世纪 50 年代,晶体管的发明使整个状况发生很大改变。计算机已经很可靠,计算机可以长时间运行。但这些价格昂贵的机器只有少数大公司、重要的政府部门或大学才能买得起。

程序员在使用计算机之前,要把源程序、原始数据以及执行步骤(用计算机系统规定的作业控制语言描述)都穿成卡片,再将卡片盒提交给操作员,由操作员启动计算机,按照卡片上的指示执行计算任务。这时引入了作业的概念,作业是指用户要求计算机处理的一个相对独立的任务。一个作业可以分成几个必须顺序处理的工作步骤。例如,一个用高级语言编写的源程序提交给计算机执行,通常经过以下几个步骤:编辑、编译、连接装入、运行。这些步骤是顺序执行的,一个步骤执行的结果产生出下一个步骤要用的文件。每一个作业步骤都要执行一个程序。

计算机运行完当前作业后,其计算结果从打印机上输出,然后操作员再令机器读入另一个作业。如果使用不同的语言编译器,操作员要把它读入计算机。

由于当时计算机非常昂贵，人们自然想到要提高机器的利用率，于是就有了批处理系统。早期的批处理系统采用脱机输入输出方式：输入输出操作都是通过磁带进行的。操作员使用一台相对便宜的专门用于输入输出的卫星机将卡片上的用户作业信息存到输入的磁带(输入带)上，然后把磁带从卫星机的磁带机上取下，并装到主机的磁带机上，再执行称为监督程序的软件，从磁带上读入第一个作业并运行，其输出写到另一盘磁带上(输出带)。一个作业结束后，监督程序自动读入下一个作业并运行。当一批作业完全结束后，操作员取下输入和输出磁带，将输入磁带换成下一批作业，把输出磁带拿到一台卫星机上进行脱机打印。这时，监督程序管理着用户作业的运行，还控制磁带机的输入输出操作。整个过程如图 1-2 所示。

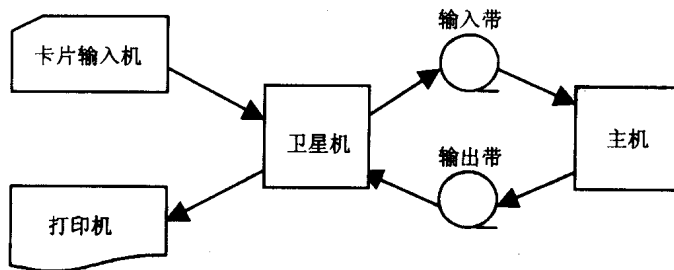


图 1-2 脱机输入、输出方式示意图

到了 20 世纪 60 年代初期，计算机硬件有了新的进展，出现了通道和中断技术。通道是一种专用的 I/O 处理器，它能控制一台或多台 I/O 设备工作，执行 I/O 设备与内存之间的数据传输。它一旦被启动就能独立于 CPU 运行，因此 CPU 与通道、CPU 与 I/O 设备可以并行工作。中断则是指，当主机接到外部信号(如 I/O 设备操作完成信号)时，主机立即停止正在执行的程序，转而处理中断事件的相应程序。处理完毕后，主机再回到原来程序被中断的位置继续执行。

借助通道、中断技术，输入输出工作由主机控制下的通道来完成。监督程序发展成为执行程序。执行系统对其他的系统程序和应用程序起到控制指挥作用，而原来的监督程序与它们是相互调用关系。执行系统常驻内存，其他程序在其指挥下工作，增强了系统的保护能力。

典型的批处理操作系统有 FMS 和 IBSYS (IBM7094 机配备的操作系统)。

3. 第三代计算机(1965 ~ 1980)：集成电路芯片和多道程序

早期的批处理系统仍是单道顺序处理作业，每次只有一个作业调入内存运行。这样可能出现两种情况：当运行以计算为主的作业时，输入输出量少，I/O 设备空闲时间多；而当运行以输入输出为主的作业时，CPU 又有较多空闲。于是，多道程序设计的思想应运而生。

多道程序设计的主要思想是，在内存中同时存放若干道用户作业，这些作业交替地运行。当一个作业由于 I/O 操作未完成而暂时无法继续运行时，系统就把 CPU 切换到另一个作业，从而使另一个作业在系统中运行。因此，从宏观上看，若干个用户作业，或者说若干道程序是同时在系统中运行的。图 1-3 所示为多道程序在系统中运行的情况。

由图 1-3 可见，当作业 A 运行一段时间需要进行 I/O 操作时，控制转向操作系统。操作系统在启动作业 A 的 I/O 操作之后不等其完成即切换至作业 B，于是作业 A 的 I/O 操作与作业 B 的运行是同时进行的。经过一段时间，当作业 B 也要求 I/O 操作时，CPU 才不得不处于