

操作系统

刘振鹏 李亚平 王煜 张明 编著

TP316/395

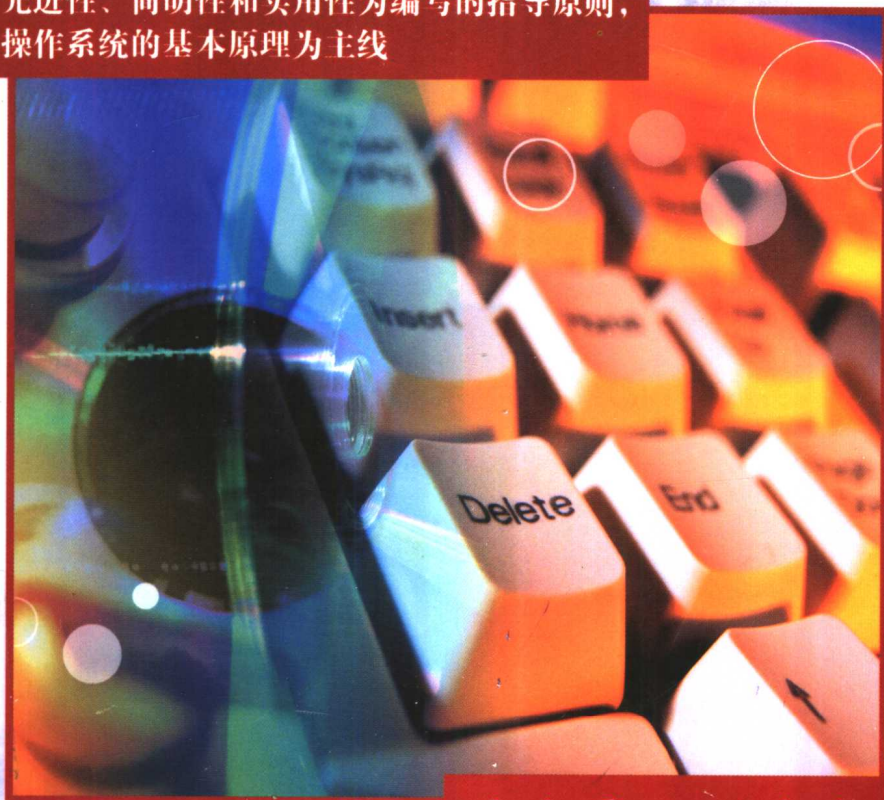


21世纪高等院校计算机教材

操作系统

刘振鹏 李亚平 王煜 张明 编著

以先进性、简明性和实用性为编写的指导原则，
以操作系统的基本原理为主线



注重原理 结合实际

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

21 世纪高等院校计算机教材

操作系统

刘振鹏 李亚平

王 煜 张 明 编著

中国铁道出版社

2003 • 北京

(京)新登字 063 号

内 容 简 介

操作系统是计算机系统必不可少的基本系统软件,是计算机专业学生的必修课。本书以先进性、简明性和实用性为编写的指导原则,以操作系统的基本原理为主线,系统地讲述了操作系统的基本概念、原理和实现技术,并以 Linux 操作系统为示例,展现了当代操作系统的本质和特点,是一本既注重基本原理,又结合实际的教学书。

本书内容丰富、结构清晰、突出基础、注重应用、强调理论与实践的结合。本教材可作为高等学校计算机科学与技术、通信工程、电子工程等电子信息类专业的教材,又适合于计算机爱好者自学,对于从事计算机应用和开发的技术人员也具有一定的参考价值。

图书在版编目(CIP)数据

操作系统/刘振鹏等编著. —北京:中国铁道出版社,2003.8

(21世纪高等院校计算机教材)

ISBN 7-113-05432-3

I. 操… II. 刘… 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 071939 号

书 名: 操作系统

作 者: 刘振鹏 李亚平 王 煜 张 明

出版发行: 中国铁道出版社(100054,北京市宣武区右安门西街8号)

策划编辑: 严晓舟 姜淑静 魏 春

责任编辑: 苏 茜 黄园园

封面设计: 孙天昭

印 刷: 化学工业出版社印刷厂

开 本: 787×1092 1/16 印张: 18 字数: 422 千

版 本: 2003 年 9 月第 1 版 2003 年 9 月第 1 次印刷

印 数: 1~5000 册

书 号: ISBN 7-113-05432-3/TP·1014

定 价: 23.00 元

版权所有 侵权必究

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社计算机图书批销部调换。

21 世纪高等院校计算机教材

编写委员会

主 任：王凤先

副主任：马胜甫 刘振鹏

委 员：（按姓氏笔划排列）

于 锋	王学军	王卫华	王春红	白乙拉
刘立民	刘清波	李亚平	李继民	张晓莉
杨怀卿	孟玉芹	赵英杰	赵新生	赵福来
徐建民	韩宪忠			

丛书策划：严晓舟 魏 春

编 辑：秦绪好 张雁芳

前 言

操作系统是计算机系统中不可缺少的基本系统软件，主要用来管理和控制计算机系统的软、硬件资源，提高其利用率，并为用户提供一个方便、灵活、安全、可靠的使用计算机的工作环境。计算机操作系统不仅是计算机有关专业的核心骨干课程，也是从事计算机应用开发人员的必不可少的知识。

本书是作者在近年来讲授《操作系统》课程的教学实践和科学研究的基础上，参考了国内外出版的各种操作系统教材，编写的一本讲述计算机操作系统原理的教科书。本书以先进性、简明性和实用性为编写的指导原则，以操作系统的基本原理为主线，系统地讲述了操作系统的基本概念、原理和实现技术，而且以 Linux 操作系统为示例，展现了当代操作系统的本质和特点，是一本既注重基本原理，又结合实际的教科书。考虑到学习和发展操作系统的需要，对于近年来国际上操作系统等领域中的新发展，也以一定的篇幅加以简单介绍。操作系统是一门实践性非常强的学科，必须对实践和应用给予必要的重视。为此，从强调应用、注重实践出发，本书以 Linux 操作系统为示例，具体分析了当代操作系统的设计思想和实现技术。

本书共分 7 章。第 1 章为操作系统引论，介绍了操作系统的概念和操作系统的形成和发展、操作系统的类型和功能，并从多个角度介绍了研究操作系统的观点。另外还介绍了操作系统的硬件环境。第 2 章为作业管理，介绍了作业管理的基本概念和功能、批处理作业管理和交互式作业管理以及系统调用的概念和处理过程。第 3 章为进程管理，介绍了进程的基本概念、进程调度，并讨论了死锁问题，在这一章中还引入了能进一步提高程序并发执行程度的多线程概念。第 4 章为存储管理，介绍了存储器管理的基本概念和功能，详尽讨论了存储器管理的机制和算法，并讨论了虚拟存储器的实现机制，对虚拟存储器的性能进行了分析。第 5 章为文件管理，介绍了文件管理的基本概念和功能，详尽讨论了文件系统的实现以及文件系统的可靠性和安全性，并对文件系统的性能问题进行了分析。第 6 章为设备管理，介绍了设备管理的基本概念和功能，讨论了设备管理的有关技术和处理过程。第 7 章为网络与分布式处理，介绍了网络服务器、进程迁移等内容，并对分布式进程管理和死锁问题进行了讨论。为了便于学习和掌握操作系统的基本知识，本书在每一章的后面还附有一定数量的习题。

本书作者提供了教材的电子讲义和书中部分习题的答案，有需要者可与中国铁道出版社计算机图书中心联系。

本书的第 1、3 章由刘振鹏编写，第 2、5 章由王煜编写，第 4、6 章由张明编写，第 7 章由李亚平编写，最后由刘振鹏、李亚平统一定稿，陈贤淑、陈晓娟、廖康良等同志参与了本书的编排工作。

本书在写作过程中，得到了许多专家的大力支持，参考了大量的文献资料，在此表示诚挚的谢意。

限于作者的水平有限，书中难免由不妥甚至错误之处，恳切希望读者予以指正。我们也会在适当时间进行修订和补充，并发布在天勤网站：<http://www.tqbooks.net>“图书修订”栏目中。

编者

2003 年 8 月

目 录

第1章 引 论	1
1-1 操作系统的概念	2
1-1-1 计算机系统	2
1-1-2 什么是操作系统	3
1-1-3 操作系统的目标	5
1-2 操作系统的形成与发展	5
1-2-1 操作系统的形成	5
1-2-2 操作系统的进一步发展	12
1-2-3 推动操作系统发展的主要动力	17
1-3 研究操作系统的几种观点	17
1-3-1 软件的观点	18
1-3-2 计算机系统资源管理的观点	18
1-3-3 进程的观点	18
1-3-4 用户与计算机硬件系统之间接口的观点	19
1-3-5 虚拟机观点	19
1-3-6 服务提供者观点	19
1-4 操作系统的功能与特征	20
1-4-1 操作系统的功能	20
1-4-2 操作系统的服务	25
1-4-3 操作系统的特征	27
1-5 操作系统的硬件环境	28
1-5-1 中央处理器 (CPU)	28
1-5-2 存储系统	31
1-5-3 缓冲技术	33
1-5-4 中断技术	33
1-5-5 I/O 控制	40
1-5-6 时钟	41
1-6 习题	42
第2章 作业管理和用户接口	43
2-1 概述	44
2-1-1 作业的基本概念	44
2-1-2 操作系统与用户接口	44
2-2 批处理作业的管理	46



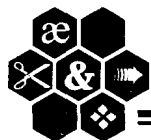
2-2-1	作业控制语言和作业控制说明书	46
2-2-2	作业控制块和作业表	47
2-2-3	作业的建立	48
2-2-4	批处理作业的调度	49
2-2-5	作业的终止与撤消	55
2-2-6	作业的状态	55
2-3	交互式作业的管理	56
2-3-1	联机命令	56
2-3-2	命令的接收和解释执行	58
2-3-3	终端作业的管理	59
2-4	系统调用	60
2-4-1	处理器状态	60
2-4-2	系统调用的概念	60
2-4-3	系统调用的处理过程	62
2-5	习题	63
第3章	进程管理	65
3-1	进程的引入	66
3-1-1	程序顺序执行	66
3-1-2	程序并发执行	66
3-1-3	多道程序设计	68
3-2	进程	68
3-2-1	进程的概念	69
3-2-2	进程的基本状态及其转换	70
3-2-3	进程控制块 PCB	72
3-2-4	进程控制	74
3-3	进程间的相互作用	78
3-3-1	进程间的联系	78
3-3-2	利用软件方法解决进程互斥问题	81
3-3-3	利用硬件方法解决进程互斥问题	83
3-3-4	信号量机制	84
3-3-5	经典进程同步问题	88
3-3-6	管程机制	93
3-4	进程通信	97
3-4-1	进程通信的类型	98
3-4-2	直接通信和间接通信	99
3-4-3	消息缓冲队列通信机制	100
3-5	进程调度 (CPU 调度)	101
3-5-1	调度的基本概念	101

3-5-2	进程调度算法	103
3-5-3	进程调度的时机和过程	105
3-6	死锁	107
3-6-1	产生死锁的原因和必要条件	107
3-6-2	预防死锁	111
3-6-3	避免死锁	112
3-6-4	检测死锁	113
3-6-5	死锁的解除	114
3-7	线程的基本概念	115
3-7-1	线程的引入	115
3-7-2	线程的定义和属性	116
3-7-3	线程与进程的比较	117
3-7-4	线程的实现机制	118
3-8	Linux 的进程管理	119
3-8-1	Linux 的进程结构与进程控制	119
3-8-2	Linux 进程间通信	120
3-8-3	Linux 的核心进程调度	129
3-9	习题	135
第 4 章	存储管理	139
4-1	概述	140
4-1-1	存储体系	140
4-1-2	存储管理的目的	140
4-1-3	存储管理的任务	140
4-1-4	各种存储管理方式	142
4-2	连续存储管理方式	142
4-2-1	单一连续分配	142
4-2-2	分区分配	143
4-3	覆盖技术与交换技术	147
4-3-1	覆盖技术	147
4-3-2	交换技术	148
4-4	分页存储管理方式	148
4-4-1	基本思想（工作原理）	148
4-4-2	动态地址变换	149
4-4-3	快表	150
4-4-4	分配与回收	151
4-5	分段存储管理方式	151
4-5-1	基本思想（工作原理）	151
4-5-2	动态地址变换	152



4-5-3	存储保护	152
4-5-4	分页和分段的主要区别	153
4-6	段页式存储管理方式	153
4-6-1	基本原理	153
4-6-2	地址变换	153
4-7	虚拟存储器	154
4-7-1	概述	154
4-7-2	分页虚拟存储管理	155
4-7-3	分段虚拟存储管理	158
4-8	Linux 的内存管理	159
4-8-1	Linux 存储管理的重要数据结构	160
4-8-2	页表的管理	161
4-8-3	页面分配和回收	162
4-8-4	页面换入	163
4-8-5	换出与丢弃页面	163
4-8-6	页面错误的处理	164
4-8-7	页面 Cache	165
4-8-8	Linux 的 swap Cache	165
4-8-9	内核 Cache 的管理	166
4-9	习题	168
第 5 章	文件管理	169
5-1	概述	170
5-1-1	文件和文件系统	170
5-1-2	文件的分类	170
5-2	文件的结构和存取方式	171
5-2-1	文件的存取方式	172
5-2-2	文件的逻辑结构	172
5-2-3	存储介质	174
5-2-4	文件的物理结构	176
5-2-5	记录式的成组分解	180
5-3	文件目录	182
5-3-1	文件控制块	182
5-3-2	文件目录结构	183
5-3-3	目录查找和目录的改进	184
5-4	文件系统的实现	186
5-4-1	打开文件表	186
5-4-2	外存空间管理	186
5-5	文件的使用	189

5-5-1	主要操作	189
5-5-2	文件共享	192
5-6	文件系统的安全性和数据一致性	193
5-6-1	防止人为因素造成的文件不安全性	194
5-6-2	防止自然因素或系统因素造成的文件不安全性	196
5-6-3	文件系统的数据一致性	196
5-7	磁盘调度	198
5-7-1	磁盘输入输出时间	199
5-7-2	磁盘的移臂调度	199
5-7-3	磁盘的优化分布	200
5-8	Linux 的文件系统	201
5-8-1	Linux 文件系统的结构	201
5-8-2	Linux 文件类型	202
5-8-3	Linux 文件系统的目录	202
5-8-4	Linux 文件的查找	203
5-8-5	Linux 文件的操作	203
5-8-6	Linux 文件的共享	206
5-8-7	Linux 文件目录操作	207
5-8-8	Linux 文件的一致性处理	207
5-8-9	Linux EXT2 文件系统	208
5-8-10	EXT2 位示图和 1 节点图	209
5-8-11	Linux 高速缓存	210
5-9	习题	211
第 6 章	设备管理	213
6-1	概述	214
6-1-1	设备的分类	214
6-1-2	设备管理的目标和任务	214
6-2	I/O 硬件特点	215
6-2-1	设备组成	215
6-2-2	设备接口	215
6-2-3	设备控制器	216
6-2-4	直接存储器访问 (DMA)	217
6-2-5	通道	218
6-2-6	缓冲技术	220
6-3	I/O 软件的组成	223
6-3-1	I/O 软件的目标	223
6-3-2	中断处理程序	224
6-3-3	设备驱动程序	225
6-3-4	独立于设备的软件	227



6-3-5 用户空间的 I/O 软件.....	228
6-4 设备分配	229
6-4-1 设备分配中的数据结构	229
6-4-2 设备独立性	230
6-4-3 设备分配	230
6-5 虚拟设备	232
6-5-1 Spooling 技术.....	232
6-5-2 共享打印机	233
6-6 Linux I/O 设备管理	234
6-6-1 Linux 中的设备文件	234
6-6-2 Linux 的设备驱动程序	234
6-6-3 Linux 的中断处理	239
6-7 习题	240
第 7 章 网络与分布式处理	243
7-1 分布式系统概述	244
7-1-1 通信结构	244
7-1-2 开放式系统互连通信结构	247
7-2 网络服务器	251
7-2-1 服务器的结构	252
7-2-2 磁盘和文件服务器	252
7-2-3 文件高速缓冲存储器的一致性问题的	253
7-2-4 打印机服务器	254
7-2-5 调制解调器服务器	254
7-3 进程迁移	255
7-3-1 进程迁移机制	255
7-3-2 迁移处理	256
7-4 分布式进程通信	257
7-4-1 信息传送机制	257
7-4-2 远程过程调用	258
7-4-3 确定分布式系统的全局状态	259
7-5 分布式进程管理	262
7-5-1 分布式互斥	262
7-5-2 分布式算法	265
7-6 分布式进程死锁问题	267
7-6-1 资源分配中的死锁	268
7-6-2 消息通信中的死锁	269
7-7 习题	272
参考文献	273

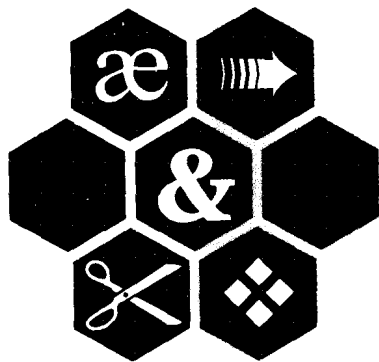
1

引 论

随着计算机技术的迅速发展,以软件为核心的信息产业对人类经济、政治、文化产生了深远的影响。信息化水平的高低,已成为衡量一个国家综合国力的重要标志。信息的收集、处理和服务是信息产业的核心内容,它们都离不开软件。在信息技术中,微电子是基础,计算机及通信设施是载体,而软件是核心。软件是计算机的灵魂,没有软件就没有计算机应用,也就没有信息化。信息社会需要众多千变万化的软件系统,因此,对于软件的研究和开发,成为当前一个极其重要的方面。在众多的软件系统中有一类非常重要的软件,为建立更加丰富的应用环境奠定了重要基础,它就是操作系统。

操作系统是配置在计算机硬件上的第一层软件,是对硬件系统的第一次扩充。操作系统在计算机系统中占据着重要的地位,其他所有的软件如汇编程序、编译程序、数据库管理系统等系统软件以及大量的应用软件,都将依赖于操作系统的支持,取得它的服务。

本章知识要点:



- 操作系统的概念
- 操作系统的形成与发展
- 研究操作系统的几种观点
- 操作系统的功能与特征
- 操作系统的硬件环境



1-1 操作系统的概念

1-1-1 计算机系统

计算机系统就是按人的要求接收和存储信息，自动进行数据处理和计算，并输出结果信息的系统。计算机是脑力的延伸和扩充，是现代科学的重大成就之一。

1. 计算机系统组成

计算机系统由硬件（子）系统和软件（子）系统组成。前者是借助电、磁、光、机械等原理构成的各种物理部件的有机组合，是系统赖以工作的实体。后者是各种程序和文件，用于指挥全系统按指定的要求进行工作。自 1946 年第一台电子计算机问世以来，计算机技术在元件器件、硬件系统结构、软件系统、应用等方面，均有惊人的进步，已广泛用于科学计算、事务处理和过程控制，日益深入社会各个领域，对社会的进步产生深刻影响。

现代计算机不再简单地被认为是一种电子设备，而是一个十分复杂的硬、软件结合而成的整体。

图 1-1 是一般的计算机系统的层次结构：最下面是硬件系统，是进行信息处理的实际物理设备；最上面是使用计算机的人，即各种各样的用户，人与硬件系统之间是软件系统，大致可分为系统软件、支撑软件和应用软件 3 层。系统软件是最靠近硬件的一层，其他软件一般都通过系统软件发挥作用；支撑软件一般用于支持实际应用系统的开发和运行；应用软件则是特定应用领域相关的软件。

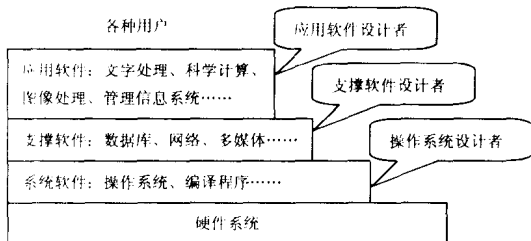


图 1-1 计算机系统的层次结构

2. 计算机的硬件

计算机硬件是指计算机系统中由电子、机械和光电元件等组成的各种计算机部件和计算机设备。这些部件和设备依据计算机系统结构的要求构成的有机整体，称为计算机硬件系统。

硬件系统是计算机系统快速、可靠、自动工作的基础。计算机硬件就其逻辑功能来说，主要是完成信息变换、信息存储、信息传送和信息处理等功能，它为软件提供具体实现的基础，计算机硬件系统主要由运算器、主存储器、控制器、输入输出控制系统、辅助存储设备等功能部件组成。

3. 计算机的软件

软件是计算机系统上的程序和有关系的文件。程序是计算任务的处理对象和处理规则的描述；文件是为了便于了解程序所需的资料说明。程序必须装入机器内部才能工作，文件一般是给人看的，不一定装入机器。程序作为一种具有逻辑结构的信息，精确而完整地描述计算任务中的处理对象和处理规则。这一描述还必须通过相应的实体才能体现。记载上述信息的实体就是硬件。

软件是用户与硬件之间的接口界面。使用计算机就必须针对要解决的问题拟定算法，用计算机所能识别的语言对有关的数据和算法进行描述，即必须编写程序和软件。用户主要是通过软件与计算机进行交互。软件是计算机系统上的指挥者，它规定计算机系统的工作，包括各项计算任务内部的工作内容和工作流程，以及各项任务之间的调度和协调。软件是计算机系统结构设计的重要依据。为了方便用户，在设计计算机系统时，必须全面考虑软件与硬

件的结构, 以及用户的要求和软件的要求。

按照应用的观点, 软件可分为系统软件、支撑软件和应用软件 3 类。

(1) 系统软件

位于计算机系统中最靠近硬件的一层。其他软件一般都通过系统软件发挥作用。它与具体的应用领域无关, 如编译程序和操作系统等。编译程序把程序设计人员用高级语言书写的程序翻译成与之等价的、可执行的低级语言程序; 操作系统则负责管理系统的各种资源、控制程序的执行。在任何计算机系统的设计中, 系统软件都要优先考虑。

(2) 支撑软件

即支撑其他软件的编制和维护的软件。随着计算机科学技术的发展, 软件的编制和维护代价在整个计算机系统中所占的比重很大, 远远超过硬件。因此, 支撑软件的研究具有重要意义, 直接促进软件的发展, 当然, 编译程序、操作系统等系统软件也可算作支撑软件。20 世纪 70 年代中期和后期发展起来的软件支撑环境, 可看成现代支撑软件的代表, 主要包括数据库管理系统、各种接口软件和工具组。三者形成整体, 协同支撑其他软件的编制。

(3) 应用软件

特定应用领域专用的软件, 例如字处理程序。

系统软件、支撑软件以及应用软件之间既有分工又有结合, 是不可以截然分开的。

4. 计算机组织

计算机的运算器、存储器、控制器、输入设备和输出设备等主要功能部件的相互连接和相互作用, 借以实现机器指令级的各种功能和特性。从最基本的功能和作用原理来说, 计算机是在控制器的全面控制下, 接收经数字化编码的输入信息(程序和数据), 把它存放在存储器中, 根据程序的要求对数据进行快速运算, 产生结果数据输出。因此, 可以把运算器、存储器、控制器、输入输出设备看成是一台计算机的逻辑组成中最基本的功能部件。现代计算机的物理组成要比这个逻辑组成复杂得多, 实际上每种功能部件可能不止一个, 有些分布于全机, 有些相互结合在一起。

计算机系统结构作为从程序设计者角度所看到的计算机属性, 在计算机系统的层次结构中处于机器语言级; 而计算机组织作为计算机系统结构的逻辑实现和物理实现, 其任务就是围绕提高性能价格比的目标, 实现计算机在机器指令级的功能和特性。研究和建立各功能部件间的相互连接和相互作用, 完成各个功能部件内部的逻辑设计等是逻辑实现的内容; 把逻辑设计深化到元件、器件级, 则是物理实现的内容, 有时把前者称为计算机组织, 把后者称为计算机实现。但是, 随着集成电路规模的日益增大, 这两步实现的内容很难分开, 因此将它们统称为计算机组织。

1-1-2 什么是操作系统

1. 操作系统的地位

计算机系统是由硬件和软件两部分构成的。操作系统属于软件中的系统软件, 操作系统是紧挨着硬件的第一层软件, 是对硬件功能的首次扩充, 其他软件则是建立在操作系统之上的。通过操作系统对硬件功能进行扩充, 并在操作系统的统一管理和支持下运行各种软件。

因此, 操作系统在计算机系统中占据着一个非常重要的地位, 它不仅是硬件与所有其他软件之间的接口, 而且任何数字电子计算机, 从微处理器到巨型计算机都必须在其硬件平台上加载相应的操作系统之后, 才能构成一个可以协调运转的计算机系统。只有在操作系统的指挥控制下,



各种计算机资源才能被分配给用户使用。也只有在操作系统的支撑下，其他系统软件如各类编译系统、程序库、运行支持环境才得以取得运行条件。没有操作系统，任何应用软件都无法运行。

可见，操作系统实际上是一个计算机系统中硬、软件资源的总指挥部。操作系统的性能高低，决定了整体计算机的潜在硬件性能能否发挥出来。操作系统本身的安全可靠程度，决定了整个计算机系统的安全性和可靠性。操作系统是软件技术含量最大、附加值最高的部分，是软件技术的核心，是软件的基础运行平台。

2. 操作系统的定义

综上所述，我们可给出操作系统的一个定义：

操作系统是计算机系统中的一个系统软件，是能有效地组织和管理计算机系统中的硬件和软件资源，合理地组织计算机工作流程，控制程序的执行，并向用户提供各种服务功能，使得用户能够灵活、方便、有效地使用计算机，并使整个计算机系统能高效地运行的一组程序模块的集合。

“有效”主要指操作系统在管理资源方面要考虑到系统运行效率和资源的利用率，要尽可能地提高处理器的利用率，让它尽可能少的空转，其他的资源例如内存、硬盘则应该在保证访问效能的前提下尽可能地减少浪费的空间等。

“合理”主要是指操作系统对于不同的用户程序要“公平”，以保证系统不发生“死锁”和“饥饿”的现象。

“方便”主要是指人机界面方面，包括用户使用界面和程序设计接口两方面的易用性、易学性和易维护性。

操作系统主要有以下两方面主要的作用：

(1) 操作系统要管理系统中的各种资源，包括硬件及软件资源。在计算机系统中，所有硬件部件（如中央处理器 CPU、存储器、输入输出设备）称作硬件资源；而程序和数据等信息称作软件资源。因此，从微观上看，使用计算机系统就是使用各种硬件资源和软件资源。特别是在多用户、多道程序的系统中，同时有多个程序在运行，这些程序在执行的过程中可能会要求使用系统中的各种资源。操作系统就是资源的管理者和仲裁者，由它负责在各个程序之间调度和分配资源，保证系统中的各种资源得以有效的利用。

(2) 操作系统要为用户提供良好的界面。一般来说，使用操作系统的用户有两类。一类是最终用户，他们只关心自己的应用需求是否被满足，而不在意其他情况，至于操作系统的效率是否高，所有的计算机设备是否正常，只要不影响他的使用，则一律不去关心。例如，用户在使用邮件服务器收发自己的电子邮件时，他只注意自己邮件是否快捷安全地收发，并不在意有多少用户同时使用这台邮件服务器。只要在这台邮件服务器上出现的堵塞、安全问题不影响到他的邮件收发，他不会去关心这台邮件服务器的整体状态。但是另一类用户就必须关心整个邮件服务器的工作状态，这就是邮件服务器管理员。这类用户一般称为系统用户。他必须时刻监视系统的整体运行状态，如空间的使用情况、有否发生通信堵塞、有否黑客攻击系统等。有时系统用户和最终用户可能是同一个人，比如许多使用 Windows 2000 的用户，他可能正在用 Office 2000 写一份报告，此时他是一位最终用户。一会儿，他想查看一下所使用硬盘上的 D 盘还有多少空间，是否需要删除一些不用的文件以获得更多的自由空间，此时他是一位系统用户。

操作系统必须为最终用户和系统用户这两类用户的各种工作提供良好的界面，以方便用

户的工作。典型的操作系统界面有命令行界面，如 UNIX 和 MS DOS，以及图形化的操作系统界面，典型的则是 Windows。

1-1-3 操作系统的目标

目前存在着多种类型的操作系统，不同类型的操作系统其目标有所侧重。在计算机硬件上配置的操作系统目标有以下几点：

(1) 方便性。配置操作系统后可使计算机系统更容易使用。

(2) 有效性。在未配置操作系统的计算机系统中，诸如 CPU、I/O 设备等各类资源，都会经常处于空闲状态而得不到充分利用；内存及外存中所存放的数据由于无序而浪费了存储空间。配置了操作系统后，可使 CPU 和 I/O 设备由于能保持忙碌状态而得到更为有效的利用，且由于使内存和外存中存放的数据有序而节省了存储空间。此外，操作系统还可以通过合理地组织计算机的工作流程，从而进一步改善系统的资源利用率及提高系统的吞吐量。

(3) 可扩充性。随着 VLSI 技术和计算机技术的迅速发展，计算机硬件和体系结构也随之得到迅速发展，它们对操作系统提出了更高的功能和性能要求。因此，操作系统必须具有很好的可扩充性才能适应发展的要求。这就是说，操作系统应采用模块化结构，以便于增加新的功能和修改老的功能模块。

(4) 开放性。20 世纪 80 年代和 90 年代陆续出现了各种类型的计算机硬件系统。为了出自不同厂家的计算机及其设备能通过网络加以集成化并能正确、有效地协同工作，实现应用程序的可移植性和互操作性，因而要求具有统一的开放的环境，其中首先是要求操作系统具有开放性。

1-2 操作系统的形成与发展

操作系统的形成迄今已有约 50 年的时间。在 20 世纪 50 年代中期出现了第一个简单的批处理操作系统，到 60 年代中期产生了多道程序批处理系统，不久又出现了基于多道程序的分时系统。80 年代是微机和计算机局域网大发展的年代，同时也是微机操作系统和网络操作系统形成和大发展的年代。此后分布式操作系统和网络操作系统得到了大发展。

1-2-1 操作系统的形成

1. 无操作系统时的计算机系统

(1) 人工操作方式

第一代计算机时期（1945 年至 50 年代中期）没有出现操作系统。这时的计算机操作是由用户（即程序员）采用人工操作方式直接使用计算机硬件系统，即由程序员将事先已穿孔（对应于程序和数据）的纸带（或卡片）装入纸带输入机（或卡片输入机），再启动它们将程序和数据输入计算机，然后启动计算机运行。当程序运行完毕并取走计算结果后，才让下一个用户上机。这种人工操作方式有以下两个缺点：

① 用户独占全机。一台计算机的全部资源只能由一个用户独占。

② CPU 等待人工操作。当用户进行装带（卡）、卸带（卡）等人工操作时，CPU 是空闲的。

人工操作方式严重降低了计算机资源的利用率。随着 CPU 速度的提高、系统规模的扩大，人机矛盾也就变得日趋严重。



此外,随着 CPU 速度的迅速提高而 I/O 设备的速度却提高缓慢,又使 CPU 与 I/O 设备之间速度不匹配的矛盾更加突出。为了缓和此矛盾,先后出现了通道技术、缓冲技术,但未能很好地解决上述矛盾,后来引入的脱机输入输出方式获得了较为令人满意的结果。

(2) 脱机输入输出方式

为了解决人机矛盾及 CPU 和 I/O 设备之间速度不匹配的矛盾,50 年代末出现了脱机输入输出技术。该技术是指事先将装有用户程序和数据的纸带(或卡片)装入纸带(或卡片)输入机,在一台外围机的控制下把纸带(卡片)上的数据(程序)输入到磁带上。当 CPU 需要这些程序和数据时再从磁带上高速地调入内存。

类似地,当 CPU 需要输出时可由 CPU 直接高速地把数据从内存送到磁带上,然后再在另一台外围机的控制下,将磁带上的结果通过相应的输出设备输出。图 1-2 所示为脱机输入输出过程。由于程序和数据的输入输出都是在外围机的控制下完成的,或者说它们是在脱离主机的情况下进行的,故称为脱机输入输出方式;反之,在主机的直接控制下进行输入输出的方式称为联机输入输出方式。

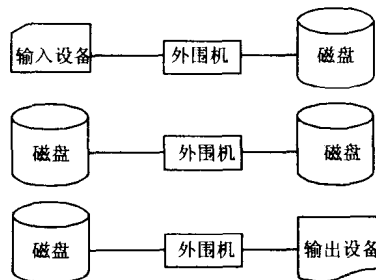


图 1-2 脱机 I/O 示意图

这种脱机 I/O 方式的主要优点是:

① 减少了 CPU 的空闲时间。装带(卡)、卸带(卡)、以及将数据从低速 I/O 设备送到高速的磁带(或盘)上,都是在脱机情况下进行的,它们不占用主机时间,从而有效地减少了 CPU 的空闲时间,缓和了人机矛盾。

② 提高 I/O 速度。当 CPU 在运行中需要数据时,是直接从高速的磁带或磁盘上将数据调入内存的,不再是从低速 I/O 设备上调入,从而大大缓和了 CPU 和 I/O 设备不匹配的矛盾,进一步减少了 CPU 的空闲时间。

2. 单道批处理操作系统

(1) 批处理系统的处理过程

早期的计算机系统非常昂贵,为了能充分地利用它,应尽量让该系统连续地运行,以减少空闲时间。为此,通常是把一批作业以脱机输入方式输入到磁带上,并在系统中配上监督程序,在监督程序的控制下使这批作业能一个接一个地连续处理。其自动批处理过程是:首先,由监督程序将磁带上的第一个作业装入内存,并把运行控制权交给该作业。当该作业的处理完成时又把控制权交还给监督程序,再由监督程序将磁带上的第二个作业调入内存。计算机系统就这样自动地一个作业一个作业地进行处理,直至磁带(盘)上的作业全部完成,这样便形成了早期的批处理系统。

由于系统对作业的处理都是成批地进行的,且在内存中始终只保持一道作业,故称为单道批处理系统。图 1-3 所示为单道批处理系统的处理流程。

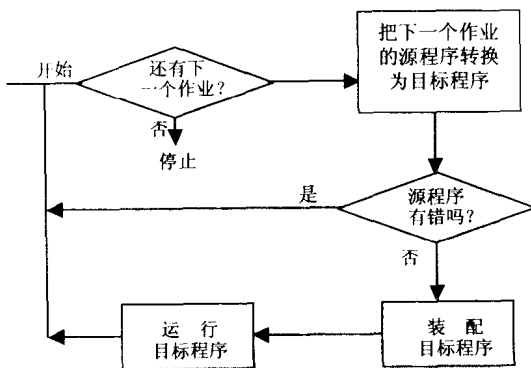


图 1-3 单道批处理系统的处理流程