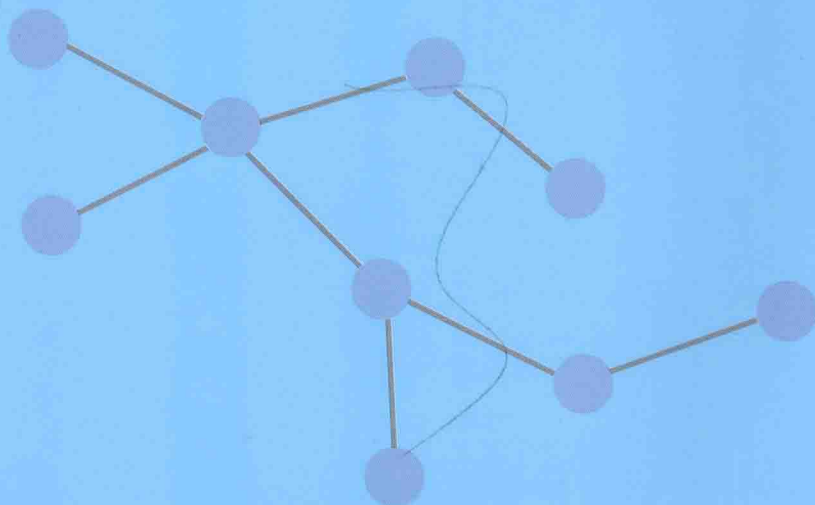


高等院校计算机任务驱动教改教材

操作系统实用教程

姜庆玲 杨 云 主 编

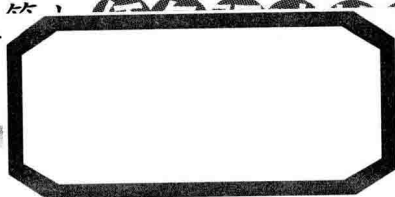


清华大学出版社



高等院校计

教材



操作系统实用教程

姜庆玲 杨 云 主 编



清华大学出版社
北 京

内 容 简 介

操作系统是计算机系统必备的核心系统软件,是计算机从业人员必须掌握的基本知识,是计算机相关专业的学生必修的专业基础课程。

本书在编写上突出内容的实用性和系统性。第1章重点介绍操作系统的定义、功能、分类、特性以及操作系统接口等知识;第2章介绍两种典型操作系统(Windows 和 Linux)的发展历史、基本概念和体系结构;第3~7章介绍操作系统的基本原理和技术,包括进程管理、作业管理、进程间的制约关系、存储管理、设备管理、文件管理等。在每章的最后两节,结合两种典型操作系统,介绍本章涉及的操作系统原理和技术在实际操作系统软件中的应用,进一步加深学生对知识点的理解,提高应用水平。

本书可作为应用型本科院校和高职高专院校计算机相关专业操作系统课程的教材和参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

操作系统实用教程/姜庆玲,杨云主编.--北京:清华大学出版社,2015

高等院校计算机任务驱动教改教材

ISBN 978-7-302-41509-1

I. ①操… II. ①姜… ②杨… III. ①操作系统—高等学校—教材 IV. ①TP316

中国版本图书馆 CIP 数据核字(2015)第 209100 号

责任编辑:张龙卿

封面设计:徐日强

责任校对:刘 静

责任印制:何 芊

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795764

印 刷 者:三河市君旺印务有限公司

装 订 者:三河市新茂装订有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:15

字 数:344 千字

版 次:2015 年 10 月第 1 版

印 次:2015 年 10 月第 1 次印刷

印 数:1~2500

定 价:32.00 元

前言

操作系统是计算机系统必备的核心系统软件,是计算机从业人员必须掌握的基本知识,是计算机相关专业的学生必修的专业基础课程。因此,掌握操作系统的基本原理和技术对于学习后面的专业课以及工作中的实际应用都大有帮助。

本书主要面向应用型本科院校和高职高专院校计算机相关专业的学生,秉承“以应用为主体、培养实践能力”的指导思想。理论知识以够用为准,做到重点突出,详略得当,并以两种典型操作系统——Windows 和 Linux 为例,说明操作系统基本原理和技术在实际操作系统中的实现和使用方法,以培养学生的应用能力。

本书主要介绍操作系统的基本功能、基本原理和设计技术。全书共分 7 章。

第 1 章重点介绍操作系统的定义、功能、分类、特性以及操作系统接口等知识。

第 2 章介绍两种典型操作系统——Windows 和 Linux 的发展历史、基本概念和体系结构。

第 3~7 章详细介绍操作系统的原理和技术,包括进程管理、作业管理、进程间的制约关系、存储管理、设备管理、文件管理等。

本书具有以下特点。

(1) 理论适度,重在应用。理论知识不求面面俱到,做到重点突出,分析透彻,配以实例分析,突出理论在实际操作系统中的应用。

(2) 循序渐进,启发思考。按照具体技术的发展轨迹依次引入不断完善的技术,比如,在讲解存储管理技术时,先分析前一种技术的不足,引导学生分析产生的原因,再针对原因寻求解决方法,从而很自然地引出后一种技术。既提高了学生的学习兴趣,又提高了学生分析问题、解决问题的能力。

(3) 每章前面都含有“学习目标”,指出本章知识点的学习要求,可以作为学生自主学习的评价指标;每章最后都含有“习题”,可供学生学习和课下练习使用。

本书的教学时数以 60~80 学时为宜,教师可以根据实际学时数、学

生情况等自行调整进度和教学内容。

本书由姜庆玲、杨云担任主编,姜庆玲对全书进行了统稿。具体编写分工如下:杨云编写第1章,李国明、张晖、李宪伟、马立新、牛文琦、郭娟、刘芳梅、王春身、张亦辉、徐莉编写第2章,姜庆玲编写文前部分及第3~7章。

本书配套资源丰富,提供课堂教学课件(PPT 文件)和习题答案,可登录出版社网站下载。

尽管我们尽了最大努力,书中难免存在疏漏与不足之处,敬请读者批评指正并提出宝贵意见。

作者 E-mail: yangyun90@163.com, Windows & Linux 教师交流群: 189934741。

编者

2015 年 6 月

目 录

第 1 章 操作系统概述	1
1.1 操作系统的定义与功能	1
1.1.1 操作系统的形成	1
1.1.2 操作系统的定义	5
1.1.3 操作系统的功能	7
1.2 操作系统的分类	11
1.2.1 单用户操作系统	11
1.2.2 多道批处理系统	12
1.2.3 分时操作系统	15
1.2.4 实时操作系统	17
1.2.5 网络操作系统	18
1.2.6 分布式操作系统	19
1.3 操作系统的特征和性能指标	20
1.3.1 操作系统的特征	20
1.3.2 操作系统的性能指标	22
1.4 操作系统的接口	23
1.4.1 操作系统用户接口概述	23
1.4.2 操作系统的三种接口	23
练习题	27
第 2 章 常用操作系统概述	29
2.1 Windows 操作系统	29
2.1.1 Windows 的基本概念	29
2.1.2 Windows 的版本	30
2.1.3 Windows 的体系结构	32
2.2 Linux 操作系统	33
2.2.1 Linux 的基本概念	33
2.2.2 Linux 的版本	34

2.2.3 Linux 的体系结构	35
练习题	36
第3章 处理机管理	37
3.1 进程的引入	37
3.1.1 程序的顺序执行	37
3.1.2 多道程序设计的引入	38
3.1.3 程序的并发执行	39
3.2 进程	39
3.2.1 进程的定义	39
3.2.2 进程的特征	40
3.2.3 进程的基本状态	41
3.2.4 进程的三个组成部分	43
3.2.5 进程控制块	44
3.3 进程的调度与管理	45
3.3.1 进程控制块队列	45
3.3.2 进程的调度	46
3.3.3 进程的管理	53
3.4 作业调度	55
3.4.1 作业与作业管理	55
3.4.2 作业的调度算法	57
3.5 Windows 的进程及其调度	61
3.5.1 Windows 的进程和线程	61
3.5.2 Windows Server 2008 的线程调度	65
3.6 Linux 的进程及其调度	71
3.6.1 Linux 的进程	71
3.6.2 Linux 的进程调度	74
练习题	76
第4章 进程间的制约关系	79
4.1 进程间的制约关系	79
4.1.1 进程同步	80
4.1.2 进程互斥	82
4.2 信号量机制	82
4.2.1 信号量机制简介	82
4.2.2 进程制约关系的实现	83
4.2.3 资源分配与信号量	86
4.2.4 样例分析	87

4.3	死锁	93
4.3.1	死锁的基本概念	93
4.3.2	死锁的预防	94
4.3.3	死锁的避免	95
4.3.4	死锁的检测与解除	98
4.4	高级进程通信	99
4.4.1	共享存储器机制	100
4.4.2	消息传递机制	100
4.4.3	管道通信机制	103
4.5	Windows 中的进程通信	103
4.5.1	微软标准进程间通信技术的发展	103
4.5.2	Windows Server 2008 的通信技术	104
4.6	Linux 中的进程通信	108
	练习题	110
第 5 章	存储管理	114
5.1	存储管理概述	114
5.1.1	存储管理的功能	114
5.1.2	存储管理的方式	116
5.1.3	地址重定位	116
5.2	固定分区存储管理	118
5.2.1	单一连续分区存储管理	119
5.2.2	固定分区存储管理	121
5.3	可变分区存储管理	123
5.3.1	基本思想	123
5.3.2	内存空间的分配与回收	124
5.3.3	地址重定位与存储保护	129
5.3.4	管理特点	129
5.4	页式存储管理	129
5.4.1	基本思想	130
5.4.2	内存空间的分配与回收	130
5.4.3	地址重定位与存储保护	132
5.4.4	管理特点	137
5.5	段式和段页式存储管理	138
5.5.1	段式存储管理	138
5.5.2	段页式存储管理	140
5.6	请求分页式存储管理	142
5.6.1	虚拟存储器	142

5.6.2	请求分页式存储管理的基本思想	142
5.6.3	缺页中断	143
5.6.4	页面淘汰算法	145
5.6.5	管理特点	148
5.7	Windows 中的存储管理	148
5.7.1	Windows Server 2008 的地址重定位	148
5.7.2	Windows Server 2008 的内存管理	150
5.8	Linux 中的存储管理	152
5.8.1	Linux 操作系统中的虚拟存储空间	152
5.8.2	Linux 的页面调度策略	152
5.8.3	Linux 的交换文件和交换分区	153
	练习题	155
第 6 章	设备管理	158
6.1	设备管理概述	158
6.1.1	设备分类	158
6.1.2	设备管理的目标与功能	162
6.1.3	I/O 设备所需的资源	163
6.1.4	设备处理程序	164
6.2	输入/输出的处理步骤	165
6.2.1	I/O 请求的提出	165
6.2.2	对 I/O 请求的管理	166
6.2.3	I/O 请求的具体实现	166
6.3	设备分配与调度	167
6.3.1	管理设备时的数据结构	167
6.3.2	设备分配策略	170
6.3.3	设备分配步骤	172
6.3.4	独享设备的分配	172
6.3.5	共享磁盘调度	174
6.4	数据传输方式	178
6.4.1	设备控制器	178
6.4.2	程序直接控制方式	179
6.4.3	中断方式	180
6.4.4	直接存储器存取(DMA)方式	181
6.4.5	通道方式	182
6.5	设备管理技术	184
6.5.1	I/O 缓冲技术	184
6.5.2	虚拟设备与 SPOOLing 技术	185

6.6	Windows 中的设备管理	187
6.6.1	Windows Server 2008 设备管理综述	187
6.6.2	Windows Server 2008 单层驱动程序的 I/O 处理	190
6.6.3	Windows Server 2008 两级中断处理过程	191
6.7	Linux 中的设备管理	193
6.7.1	Linux 设备管理概述	194
6.7.2	Linux 对字符设备的管理	195
6.7.3	Linux 对块设备的管理	196
	练习题	198
第 7 章	文件管理	201
7.1	文件管理的基本概念	201
7.1.1	文件	201
7.1.2	文件系统	203
7.2	文件的结构与文件目录	204
7.2.1	文件的逻辑结构	204
7.2.2	文件的物理结构	205
7.2.3	文件的存取	207
7.2.4	文件目录	207
7.3	文件存储空间的管理	210
7.3.1	位示图	211
7.3.2	空闲区表	211
7.3.3	空闲块链	211
7.3.4	成组链接法	212
7.4	文件的使用	213
7.4.1	文件的操作	213
7.4.2	文件的共享	214
7.4.3	文件的保护	215
7.4.4	按名存取的实现	216
7.5	Windows 中的文件管理	217
7.5.1	Windows 文件管理概述	217
7.5.2	Windows 文件管理的实现技术	219
7.6	Linux 中的文件管理	223
7.6.1	Linux 文件管理概述	223
7.6.2	Linux 文件管理的实现技术	223
	练习题	227
	参考文献	230

第 1 章 操作系统概述

操作系统(Operating System, OS)是现代计算机系统中不可缺少的基本系统软件。操作系统管理和控制计算机系统中的所有软、硬件资源,是计算机系统的灵魂和核心。除此之外,它还为用户提供计算机提供一个方便灵活、安全可靠的工作环境。

本章主要介绍操作系统的发展历史,给出操作系统的定义,介绍操作系统的五大功能以及常见的几种操作系统。

本章学习要点

- ◇ 了解操作系统的历史、操作系统基本类型和研究操作系统的几种观点。
- ◇ 掌握操作系统的定义。
- ◇ 掌握操作系统的五大管理功能。
- ◇ 掌握几种有代表性操作系统的特点。

1.1 操作系统的定义与功能

1.1.1 操作系统的形成

1. 人工操作阶段

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,计算机速度慢、规模大、外设少,操作系统尚未出现。由程序员采用手工方式直接控制和使用计算机硬件,程序员使用计算机语言编程,并将事先准备好的程序和数据穿孔在纸带或卡片上,从纸带或卡片输入机将程序和数据输入计算机。然后,启动计算机运行,程序员可以通过控制台上的按钮、开关和氛灯来操纵及控制程序,运行完毕,取走计算输出的结果,才轮到下一个用户上机操作。

随着时间的推移,汇编语言产生,在汇编系统中,数字操作码被记忆码代替,程序按固定格式的汇编语言书写。系统程序员预先编制一个汇编程序,它把用汇编语言书写的“源程序”解释成计算机能直接执行的计算机语言格式的目标程序。稍后,一些高级程序设计语言出现,FORTRAN、ALGOL 和 COBOL 语言分别于 1956 年、1958 年和 1959 年设计完成并投入使用,进一步方便了编程。

执行时需要把汇编程序或编译系统以及源程序和数据都穿在卡片或纸带上,然后装入和执行。其大致过程如下。

- (1) 人工把源程序用穿孔机穿在卡片或纸带上。
- (2) 将准备好的汇编程序或编译系统装入计算机。
- (3) 汇编程序或编译系统读入人工装在输入机上的穿孔卡片或穿孔带。
- (4) 执行汇编过程或编译过程,产生目标程序,并输出目标卡片或纸带。
- (5) 通过引导程序把装在输入机上的目标程序读入计算机。
- (6) 启动目标程序执行,从输入机上读入人工装好的数据卡片或数据带。
- (7) 产生计算结果,把执行结果从打印机上或卡片机上输出。

用上述方式算题比直接用计算机语言前进了一步,程序易于编制和易读性好,汇编程序或编译系统可执行存储分配等辅助工作,从而在一定程度上减轻了用户的负担。但是计算机的操作方式并没有多大改变,仍然是在人工控制下进行程序的装入和执行。

人工操作方式存在以下严重缺点。

- 用户独占资源。用户一个个、一道道地串行算题,上机时独占了全机资源,造成计算机资源利用率不高,计算机系统效率低下。
- 人工干预较多。要求程序员装纸带或卡片、按开关、看氛灯等。手工操作多,不但浪费处理机时间,而且也极易发生差错。
- 计算时间拉长。由于数据的输入、程序的执行、结果的输出均是联机进行的,因而,每个用户从上机到下机的时间拉得非常长。

这种人工操作方式在慢速的计算机上还能容忍,随着计算机速度的提高,其缺点就更加暴露出来。譬如,一个作业在每秒 1 万次的计算机上,需运行 1 个小时,作业的建立和人工干预花了 3 分钟,那么,手工操作时间占总运行时间的 5%;当计算机速度提高到每秒 10 万次,此时,作业运行时间仅需 6 分钟,而手工操作不会有多大变化,仍为 3 分钟,这时手工操作时间占了总运行时间的 50%。由此看出缩短手工操作和人工干预时间十分必要。此外,随着 CPU 速度迅速提高,I/O 设备速度却提高不多,导致 CPU 与 I/O 设备之间的速度不匹配,矛盾越来越突出,需要妥善解决这些问题。

2. 管理程序阶段

早期批处理系统借助于作业控制语言变革了计算机的手工操作方式。用户不再通过开关和按钮来控制计算机执行,而是通过脱机方式使用计算机,通过作业控制卡来描述对作业的加工控制步骤,并把作业控制卡连同程序、数据一起提交给计算机的操作员,操作员收集到一批作业后,一起把它们放到卡片机上输入计算机。计算机上则运行一个驻留在内存中的执行程序,以对作业进行自动控制和成批处理,自动进行作业转换以减少系统空闲和手工操作时间,其工作流程如下:执行程序将一批作业从纸带或卡片机输入磁带上,每当一批作业输入完成后,执行程序自动把磁带上的第一个作业装入内存,并把控制权交给作业。当该作业执行完成后,执行程序收回控制权并调入磁带上的第二个作业到内存执行。计算机在执行程序的控制下就这样连续地一个作业一个作业地执行,直至磁带上的作业全部做完。这种系统能实现作业到作业的自动转换,缩短作业的准备和建立时间,减少人工操作和干预,让计算机尽可能地连续运转。

早期批处理系统中,一开始作业的输入和输出均是联机的,联机 I/O 的缺点是速度

慢,I/O 设备和 CPU 仍然串行工作,CPU 时间浪费相当大,为此,在批处理中引进脱机 I/O 技术。除主机外,另设一台辅机,该机仅与 I/O 设备打交道,不与主机连接。输入设备上的作业通过辅机输到磁带上,这叫脱机输入;主机负责从磁带上把作业读入内存执行,作业完成后,主机负责把结果输出到磁带上,这叫脱机输出;然后,由辅机把磁带上的结果信息在打印机上打印输出。这样,I/O 工作脱离了主机,辅机和主机可以并行工作,大大加快了程序的处理和数据的输入及输出,这称作脱机 I/O 技术,这比早期联机处理系统提高了处理能力。

3. 多道程序设计

20 世纪 60 年代初,有两项技术取得了突破:中断和通道。这两种技术结合起来为实现 CPU 和 I/O 设备并行工作提供了基础,此时,多道程序的概念才变成了现实。

多道程序设计(Multiprogramming)是指允许多个程序(作业)同时进入一个计算机系统的内存储器并启动进行交替计算的方法。也就是说,计算机内存中同时存放了多道(两个以上相互独立的)程序,它们均处于开始和结束点之间。从宏观上看是并行的,多道程序都处于运行过程中,但都未运行结束;从微观上看是串行的,各道程序轮流占用 CPU,交替地执行。引入多道程序设计技术的根本目的是提高 CPU 的利用率,充分发挥计算机系统部件的并行性,现代计算机系统都采用了多道程序设计技术。引入多道程序设计的好处:一是提高了 CPU 的利用率;二是提高了内存和 I/O 设备的利用率;三是改进了系统的吞吐率;四是充分发挥了系统的并行性。其主要缺点是延长了作业周转时间。

注意:多道程序设计系统与多重处理系统(Multiprocessing System)有差别,后者是指配置了多个物理 CPU,从而能真正同时执行多道程序的计算机系统。当然要有效地使用多重处理系统,必须采用多道程序设计技术;反过来,多道程序设计不一定要有多重处理系统支持。多重处理系统的硬件结构可以多种多样,如共享内存的多 CPU 结构、网络连接的独立计算机结构。虽然多重处理系统增加了硬件,但却换来了提高系统吞吐量、可靠性、计算能力和并行处理能力的好处。

实现多道程序设计必须妥善地解决三个问题:存储保护与程序浮动;处理器的管理和分配;系统资源的管理和调度。

在多道程序设计的环境中,内存储器为几道程序所共享,因此,硬件必须提供必要的手段,使得在内存储器中的各道程序只能访问它自己的区域,以避免相互干扰。特别是当一道程序发生错误时,不致影响其他的程序,更不能影响系统程序,这就是存储保护。同时,由于每道程序不是独占全机,这样不能事先规定它运行时将放在哪个区域,所以程序员在编制程序时无法知道程序在内存储器中的确切地址。甚至,在运行过程中,一个程序也可能改变其运行区域,所有这些都要求一个程序或程序某一部分能随机地从某个内存储器区域移动到另一个区域,而不影响其执行,这就是程序浮动,或称地址重定位。此外,多道程序共存于内存,会引起内存容量不足,因此,内存扩充也成为操作系统必须解决好的问题。

在多道程序设计系统里,如果系统仅配置一个物理处理器,那么,多个程序必须轮流占有处理器,这涉及处理器调度问题。为了说明一个程序是否占有或可以占有处理器,可以把程序在执行中的状态分成三种。当一个程序正占有处理器运行时,就说它是处于运

行状态(运行态);当一个程序在等待某个事件发生时,就说它处于等待状态(等待态);当一个程序等待的条件已满足可以运行而未占用处理器时,则说它处于就绪状态(就绪态),所以,一道程序在执行中总是处于运行、就绪、等待三种状态之一。一道程序在执行过程中,它的程序状态是变化的,从运行态到等待态的转换是在发生某种事件时产生的。这些事件可能是由于启动外围设备输入、输出而使程序要等待输入、输出结束后才能继续下去;也可能是在运行中发生某种故障使程序不能继续运行下去等。从等待态转换成就绪态是在等待的某个事件完成时产生的。例如,程序甲处于等待外围设备传输完毕的等待状态,当传输结束时,程序甲就从等待态转为就绪态。从运行态也能转变为就绪态。例如,当程序乙运行时发生了设备传输结束事件,而当设备传输结束后,使得程序甲从等待态转变为就绪态;假定程序甲的优先级高于程序乙,因此,让程序甲占有处理器运行,这样程序乙就从运行态转为就绪态。

在多道程序设计系统里,系统的资源为几道程序所共享,上面谈到的处理器就是一例。此外,如内存储器、外围设备以及一些信息资源等也需要按一定策略去分配和调度,要解决好多道程序共享系统硬、软件资源的竞争与协调。

4. 操作系统的形成

第三代计算机的性能有了更大提高,机器速度更快,内外存容量增大,I/O 设备数量和种类增多,为软件的发展提供了有力支持。如何更好地发挥硬件功效,如何更好地满足各种应用的需要,这些都迫切要求扩充管理程序的功能。

中断技术和通道技术的出现使得硬部件具有较强的并行工作能力,从理论上来说,实现多道程序系统已无问题。但是,从半自动的管理程序方式过渡到能够自动控制程序执行的操作系统方式,对辅助存储器性能的要求增高。这个阶段虽然有个别的磁带操作系统出现,但操作系统的真正形成还期待着大容量高速辅助存储器的出现。大约 20 世纪 60 年代中期以后,随着磁盘的问世,相继出现了多道批处理操作系统、分时操作系统和实时操作系统,这时标志着操作系统正式形成。

计算机配置操作系统后,其资源管理水平和操作自动化程度有了进一步提高,具体表现如下。

- 操作系统实现了计算机操作过程的自动化。批处理方式更为完善和方便,作业控制语言有了进一步发展,为优化调度和管理控制提供了新手段。
- 资源管理水平有了提高,实现了外围设备的联机同时操作(即 SPOOLing),进一步提高了计算机资源的利用率。
- 提供虚存管理功能,由于多个用户作业同时在内存中运行,在硬件设施的支持下,操作系统为多个用户作业提供了存储分配、共享、保护和扩充的功能,导致操作系统步入实用化。
- 支持分时操作,多个用户通过终端可以同时联机地与一个计算机系统交互。
- 文件管理功能有改进,数据库系统开始出现。
- 多道程序设计趋于完善,采用复杂的调度算法,充分利用各类资源,最大限度地提高计算机系统效率。

促使操作系统不断发展的主要动力有以下五个方面。

(1) 器件快速更新换代。微电子技术是推动计算机技术飞速发展的“引擎”，每隔 18 个月其性能要翻一番。推动微机快速更新换代，它由 8 位机、16 位机发展到 32 位，当前已经研制出 64 位机，相应的微机操作系统也就由 8 位微机操作系统发展到 16 位、32 位微机系统，而 64 位微机操作系统也已研制出来。

(2) 计算机体系结构不断发展。硬件的改进导致操作系统发展的例子很多，内存管理支撑硬件由分页或分段设施代替寄存器后，操作系统中便增加了分页或分段存储管理功能。图形终端代替逐行显示终端后，操作系统中增加了窗口管理功能，允许用户通过多个窗口在同一时间提出多个操作请求。引进中断和通道等设施后，操作系统中引入了多道程序设计功能。计算机体系结构的不断发展，有力地推动着操作系统的发展，例如，计算机由单处理机改进为多处理机系统，操作系统也由单处理机操作系统发展到多处理机操作系统和并行操作系统；随着计算机网络的出现和发展，出现了分布式操作系统和网络操作系统。随着信息家电的发展，又出现了嵌入式操作系统。

(3) 提高计算机系统资源利用率的需要。多用户共享一套计算机系统的资源，必须千方百计地提高计算机系统中各种资源的利用率，各种调度算法和分配策略相继被研究与采用，这也成为操作系统发展的一个动力。

(4) 让用户使用计算机越来越方便的需要。从批处理到交互型分时操作系统的出现，大大改变了用户上机、调试程序的环境；从命令行交互进化到 GUI 用户界面，操作系统的界面变得更加友善。

(5) 满足用户的新要求，提供给用户新服务。当用户要求解决实时性应用时，便出现实时操作系统；当发现现有的工具和功能不能满足用户需要时，操作系统往往要进行升级换代，开发新工具，加入新功能。

1.1.2 操作系统的定义

操作系统的出现、使用和发展是近四十年来计算机软件的一个重大进展。尽管操作系统尚未有一个严格的定义，但一般认为：操作系统是管理系统资源、控制程序执行、改善人机界面、提供各种服务、合理组织计算机工作流程和为用户使用计算机提供良好运行环境的一种系统软件。

计算机发展到今天，从个人机到巨型机，无一例外都配置一种或多种操作系统，操作系统已经成为现代计算机系统不可分割的重要组成部分，它为人们建立各种各样的应用环境奠定了重要基础。配置操作系统的主要目标可归结为以下方面。

- 方便用户使用：OS 通过提供用户与计算机之间的友善接口来方便用户使用。
- 扩大机器功能：OS 通过扩充改造硬件设施和提供新的服务来扩大机器功能。
- 管理系统资源：OS 有效管理好系统中所有硬件软件资源，使之得到充分利用。
- 提高系统效率：OS 合理组织好计算机的工作流程，以改进系统性能和提高系统效率。
- 构筑开放环境：OS 遵循有关国际标准来设计和构造，以构筑出一个开放环境。

其含义主要是指：遵循有关国际标准(如开放的通信标准、开放的用户接口标准、开放的线程库标准等)；支持体系结构的可伸缩性和可扩展性；支持应用程序在不同平台上的可移植性和可互操作性。

计算机系统包括硬件和软件两个组成部分。硬件是所有软件运行的物质基础，软件能充分发挥硬件潜能和扩充硬件功能，完成各种系统及应用任务，两者互相促进、相辅相成、缺一不可，计算机系统的层次结构如图 1-1 所示。

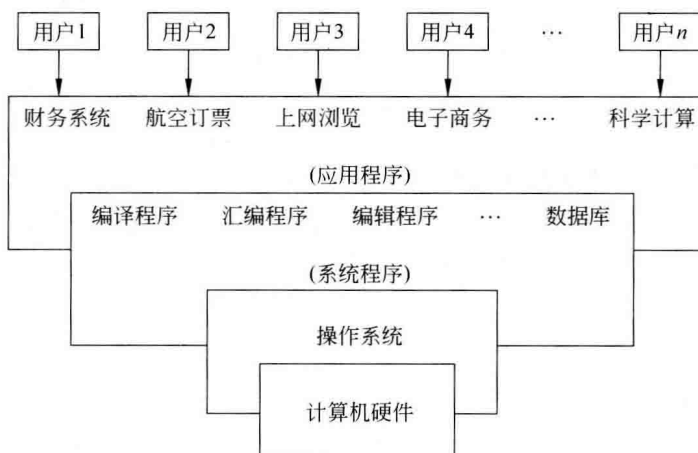


图 1-1 计算机系统的层次结构

硬件层提供了基本的可计算性资源，包括处理器、寄存器、存储器，以及各种 I/O 设施和设备，是操作系统和上层软件赖以工作的基础。操作系统层通常是最靠近硬件的软件层，对计算机硬件作首次扩充和改造，主要完成资源的调度和分配、信息的存取和保护、并发活动的协调和控制等许多工作。操作系统是上层其他软件运行的基础，为编译程序和数据库管理系统等系统程序的设计者提供了有力支撑。系统程序层的工作基础建立在操作系统改造和扩充过的机器上，利用操作系统提供的扩展指令集，可以较为容易地实现各种各样的语言处理程序、数据库管理系统和其他系统程序。此外，还提供种类繁多的应用程序，如连接装配程序、库管理程序、诊断排错程序、分类/合并程序等供用户使用。应用程序层解决用户特定的或不同应用需要的问题，应用程序开发者借助于程序设计语言来表达应用问题，开发各种应用程序，既快捷又方便。而最终用户则通过应用程序与计算机系统交互来解决他的应用问题。

软件包括应用软件和系统软件。应用软件用于解决各种具体的应用问题，如财务软件用于财务管理，办公软件用于处理公务，网络软件用于网络应用。系统软件为各种应用提供使用计算机环境和访问支持。

系统软件主要包括操作系统和数据库系统等。在系统软件中，操作系统是基础，操作系统为其他系统软件提供支持与服务，其他的系统软件可以看成运行在操作系统之上的应用软件或应用程序。

计算机的价值通过计算机应用体现。程序员、应用程序和终端用户及本地用户应用通过操作系统应用计算机，操作系统为他们提供运行平台。操作系统的作用如下。

- 直接位于计算机硬件之上,为计算机的应用提供接口。
- 提供通用的计算机服务,与专用的应用领域无关。
- 实现资源管理服务,为不同的应用提供共享资源。

操作系统作为计算机资源的管理者,能够协调和指挥计算机的各个组件按照一定的计划协同工作,有序地控制计算机中的处理器、存储器和输入/输出设备的分配,在相互竞争的用户和程序之间协调冲突,保证计算机系统正常有效地运行。

1.1.3 操作系统的功能

操作系统是用户与计算机硬件之间的接口。可以认为操作系统是对计算机硬件系统的第一次扩充,用户通过操作系统来使用计算机系统。换句话说,操作系统紧靠着计算机硬件并在其基础上提供许多新的设施和能力,从而使得用户能够方便、可靠、安全、高效地操纵计算机硬件和运行自己的程序。例如,改造各种硬件设施,使之更容易使用;提供原语和系统调用,扩展计算机的指令系统;而这些功能到目前为止还难以由硬件直接实现。操作系统还合理组织计算机的工作流程,协调各个部件有效工作,为用户提供一个好的运行环境。经过操作系统改造和扩充过的计算机不但功能更强,使用也更为方便,用户可以直接调用操作系统提供的各种功能,而无须了解许多软硬件本身的细节,对于用户来讲操作系统便成为它与计算机硬件之间的一个接口。

操作系统为用户提供了虚拟计算机(Virtual Machine)。许多年以前,人们就认识到必须找到某种方法把硬件的复杂性与用户隔离开来,经过不断地探索和研究,目前采用的方法是在计算机裸机上加上一层又一层的软件来组成整个计算机系统,同时,为用户提供了一个容易理解和便于程序设计的接口。在操作系统中,类似把硬件细节隐藏并把它与用户隔离开来的情况处处可见,例如: I/O 管理软件、文件管理软件和窗口软件向用户提供了越来越方便地使用 I/O 设备的方法。由此可见,每当在计算机上覆盖一层软件,提供一种抽象,系统的功能便增加一点,使用就更加方便一点,用户可用的运行环境就更加好一点。所以,当计算机上覆盖操作系统后,可以扩展基本功能,为用户提供一台功能显著增强,使用更加方便,安全可靠性好,效率明显提高的机器,对用户来说好像可以使用的是与裸机不同的虚拟计算机。

操作系统是计算机系统的资源管理者。在计算机系统中,能分配给用户使用的各种硬件和软件设施总称为资源。资源包括两大类: 硬件资源和信息资源。其中,硬件资源分为处理器、存储器、I/O 设备等; I/O 设备又分为输入型设备、输出型设备和存储型设备; 信息资源则分为程序和数据等。操作系统的重要任务之一是对资源进行抽象研究,找出各种资源的共性和个性,有序地管理计算机中的硬件、软件资源,跟踪资源使用情况,监视资源的状态,满足用户对资源的需求,协调各程序对资源的使用冲突; 研究使用资源的统一方法,为用户提供简单、有效的资源使用手段,最大限度地实现各类资源的共享,提高资源利用率,从而使得计算机系统的效率有很大提高。

资源管理是操作系统的一项主要任务,而控制程序执行、扩充机器功能、提供各种服务、方便用户使用、组织工作流程、改善人机界面等都可以从资源管理的角度去理解。下