

21世纪高等学校计算机规划教材

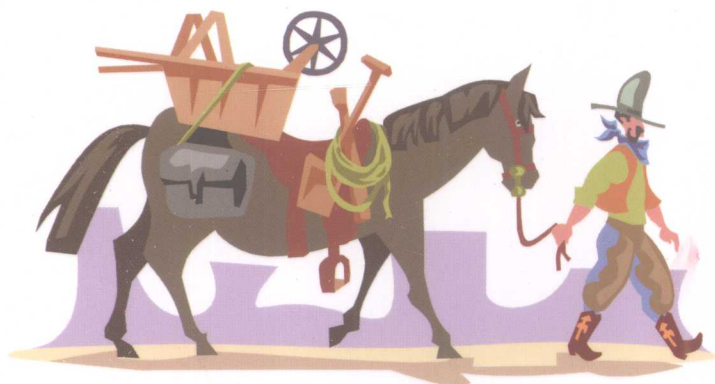
21st Century University Planned Textbooks of Computer Science

操作系统 教程

Operating System

黄刚 徐小龙 段卫华 编著

- 结构清晰 体系完整 简明扼要
- 理论教学 案例驱动 实例丰富
- 综合应用 全面培养 重在能力



高校系列

人民邮电出版社
POSTS & TELECOM PRESS

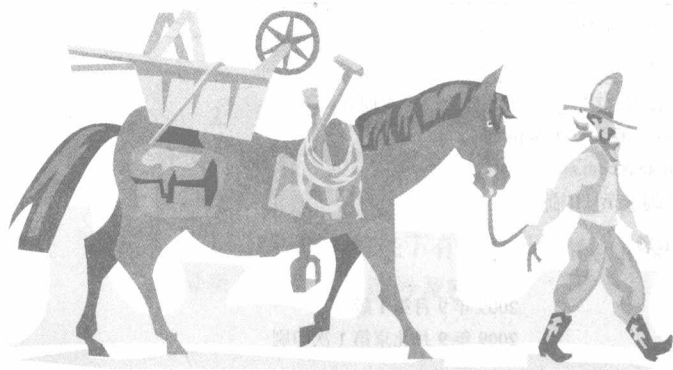
21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

操作系统 教程

Operating System

黄刚 徐小龙 段卫华 编著



高校系列

人民邮电出版社

北京

人民邮电出版社

样书

专用章

图书在版编目(CIP)数据

操作系统教程 / 黄刚, 徐小龙, 段卫华编著. —北京:
人民邮电出版社, 2009. 9
21世纪高等学校计算机规划教材. 高校系列
ISBN 978-7-115-20048-8

I. 操… II. ①黄…②徐…③段… 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(2009)第130555号

内 容 提 要

操作系统是现代计算机系统中必不可少的基本系统软件。本书讲述了计算机操作系统的基本概念、基本原理和基本算法, 并以 Linux 操作系统为实例进行介绍。

全书共分 9 章, 第 1 章介绍操作系统的概念、功能、类型及其发展; 第 2 章至第 7 章介绍操作系统对处理机管理、存储管理、设备管理和文件管理; 第 8 章介绍操作系统中网络与通信管理; 第 9 章介绍操作系统的安全。最后给出了操作系统实验。

本书可作为计算机专业和计算机相关专业的操作系统课程教材, 也可作为从事计算机工作的科技人员学习操作系统的参考书, 对报考研究生的学生也有一定的参考价值。

21 世纪高等学校计算机规划教材——高校系列

操作系统教程

-
- ◆ 编 著 黄 刚 徐小龙 段卫华
责任编辑 滑 玉
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京楠萍印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 15.75
字数: 410 千字 2009 年 9 月第 1 版
印数: 1—3 000 册 2009 年 9 月北京第 1 次印刷

ISBN 978-7-115-20048-8

定价: 27.00 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223
反盗版热线: (010)67171154

前言

计算机系统是一个软件和硬件紧密结合的统一整体,而操作系统是对于硬件功能的首次扩充,是其他系统软件和应用软件建立的基础和平台。因此,操作系统在整个计算机系统中处于承上启下的关键地位,在系统软件中,则占据着核心地位。操作系统课程是计算机教育的必修课程,作为计算机专业的核心课,不但高等院校计算机相关专业学生必须学习它,而且从事计算机行业的人员也需要深入了解它。

本书是在多年科研实践和教学实践的基础上,吸取国内外新理论和技术,参考国内外比较成熟的教材,针对计算机科学相关专业学生应掌握的知识结构需求,以培养研究型、应用型人才为出发点,参照计算机类教学指导委员会关于该课程的教学大纲进行编写的,并满足全国研究生计算机统考大纲的要求。教材内容力求由浅入深、通俗易懂、便于讲授和自学,并力图反映操作系统技术的新发展和新成果。

本教材以介绍操作系统基本概念为主,依据资源管理观点论述了操作系统的基本原理、基本结构、实现技术和运行机制。将操作系统中基本的、共性的知识介绍给读者。有了这些知识,就为读者学习和理解新型的操作系统、应用操作系统,以及在操作系统基础上开发自己的应用系统打下一个坚实的基础。为了加深理解和应用,选择了具有代表性的 Linux 操作系统作为实例贯穿全书。

全书共分 9 章。第 1 章操作系统概论,概括地介绍操作系统的基本概念、类型和功能、操作系统的形成、操作系统的结构设计;第 2 章用户界面,介绍操作系统用户界面、用户接口以及操作系统的使用;第 3 章处理机管理,首先引入进程概念,进而阐述进程管理中的进程控制,介绍线程及其实现,并讨论处理机管理中作业调度和进程调度;第 4 章并发进程的同步与互斥,介绍并发程序设计有关技术,讲解进程互斥、同步通信机制,以及进程通信和进程死锁问题;第 5 章存储管理,介绍存储管理的基本概念,引入虚拟存储器概念,从简单到复杂地讲述存储管理的各种方案及其特点;第 6 章设备管理,讲述设备控制、设备分配和设备处理等问题;第 7 章文件系统,阐述了文件结构、文件目录、文件存储空间管理以及文件使用等问题;第 8 章网络与通信管理,介绍网络体系结构、网络操作系统的工作模式,讲述网络管理、远程过程调用、网络服务等问题;第 9 章操作系统安全,介绍操作系统安全的策略、安全机制、安全评测等。最后给出操作系统的基本实验。

本书编写分工情况如下:第 1 章、第 5 章、第 9 章以及操作系统实验由黄刚编写;第 3 章、第 4 章、第 8 章由徐小龙编写;第 2 章、第 6 章、第 7 章由段卫华编写。

本书在编写过程中得到了领导的关心和支持。衷心感谢王汝传在本书出版前所做的许多建设性工作;特别感谢孙力娟百忙中仔细审阅了全书,提出了许多极为宝贵的意见。此外,感谢陈丹伟、刘志鹏、邹志强、刘茜萍、骆健、谢秋丽对教材

提出的意见和建议。在本书的编写过程中,参考和引用了国内外专家学者的相关著作以及来源于互联网等各种渠道的资料,在此一并表示衷心感谢。

由于作者水平有限,书中缺点和错误在所难免,希望广大读者批评指正。

作 者

2009 年 5 月

目 录

第 1 章 操作系统概论..... 1

1.1 操作系统概念.....1

1.1.1 操作系统的地位和作用..... 1

1.1.2 操作系统的定义..... 2

1.1.3 操作系统的功能..... 2

1.1.4 操作系统的特性..... 5

1.1.5 操作系统的性能..... 6

1.2 操作系统的形成和发展.....7

1.2.1 硬件软件的发展轨迹..... 7

1.2.2 手工操作阶段..... 9

1.2.3 早期批处理阶段..... 9

1.2.4 执行系统阶段..... 10

1.2.5 多道程序系统阶段..... 10

1.2.6 操作系统的形成..... 11

1.3 操作系统的分类.....11

1.3.1 批处理操作系统..... 12

1.3.2 分时操作系统..... 13

1.3.3 实时操作系统..... 14

1.3.4 微型计算机操作系统..... 15

1.3.5 网络操作系统..... 15

1.3.6 分布式操作系统..... 16

1.3.7 嵌入式操作系统..... 17

1.4 操作系统的结构设计.....18

1.4.1 整体式结构的操作系统..... 18

1.4.2 层次式结构的操作系统..... 18

1.4.3 虚拟机结构的操作系统..... 19

1.4.4 客户/服务器结构的操作系统..... 19

1.4.5 微内核结构的操作系统..... 19

1.5 Linux 操作系统简介.....20

1.5.1 Linux 系统的特点..... 20

1.5.2 Linux 和其他系统的区别..... 21

小结.....21

习题..... 22

第 2 章 用户界面.....23

2.1 用户工作环境.....23

2.1.1 用户环境.....23

2.1.2 系统生成和系统启动.....24

2.1.3 运行一个用户程序的过程.....25

2.2 操作系统用户界面.....26

2.2.1 操作系统的用户界面.....26

2.2.2 操作接口.....27

2.2.3 图形用户接口.....29

2.3 系统调用.....29

2.3.1 系统调用的基本概念.....29

2.3.2 系统调用的基本类型.....30

2.3.3 系统调用的实现.....30

2.4 Linux 的用户界面.....31

2.4.1 Linux 的命令行接口.....31

2.4.2 Linux 的图形方式.....32

2.4.3 Linux 的系统调用.....32

2.4.4 与系统调用有关的数据结构.....33

小结.....34

习题..... 34

第 3 章 进程管理与调度.....35

3.1 中央处理器.....35

3.1.1 CPU 的构成与基本工作方式.....35

3.1.2 特权指令和非特权指令.....36

3.1.3 处理器状态.....36

3.1.4 程序状态字.....37

3.2 中断.....37

3.2.1 中断的概念.....37

3.2.2 中断的分类	38
3.3 进程及其实现	38
3.3.1 进程定义	38
3.3.2 进程的类型和特性	39
3.3.3 进程的状态和转换	40
3.3.4 进程控制块	41
3.3.5 进程要素	42
3.4 进程的控制	43
3.4.1 进程的创建	44
3.4.2 进程的阻塞和唤醒	45
3.4.3 进程的撤销	46
3.4.4 进程的挂起和激活	47
3.5 进程切换与模式切换	47
3.5.1 执行模式	47
3.5.2 模式切换	48
3.5.3 进程切换	48
3.6 处理器调度	49
3.6.1 调度策略	49
3.6.2 调度模式	49
3.7 作业的管理与调度	50
3.7.1 作业及其调度	50
3.7.2 作业与进程	51
3.7.3 调度原则	51
3.7.4 作业调度算法	52
3.7.5 单道程序环境下的作业调度算法	54
3.7.6 多道程序环境下的作业调度算法	55
3.8 低级调度	56
3.8.1 低级调度功能	56
3.8.2 低级调度的方式	56
3.8.3 低级调度算法	57
3.9 线程及其基本概念	58
3.9.1 线程及其基本概念	58
3.9.2 线程的结构	59
3.9.3 进程与线程	60
3.9.4 线程的实现	61
3.10 Linux 进程管理	62

3.10.1 Linux 进程结构	62
3.10.2 进程调度	64
3.10.3 创建进程	65
3.10.4 Linux 的线程	66
小结	66
习题	66

第4章 并发进程 69

4.1 并发进程	69
4.1.1 顺序程序与并发进程	69
4.1.2 与时间有关的错误	70
4.1.3 进程间的联系	71
4.2 临界区管理	72
4.2.1 临界区及其使用原则	72
4.2.2 实现临界区管理的软件方法	73
4.2.3 实现临界区管理的硬件方法	75
4.3 信号量与 P、V 操作	76
4.3.1 信号量定义	76
4.3.2 P、V 操作定义	77
4.3.3 信号量的使用	78
4.3.4 信号量及 P、V 操作讨论	81
4.3.5 信号量与 P、V 操作经典问题	81
4.3.6 POSIX 信号量	84
4.3.7 Linux 中的信号量机制	87
4.4 进程间通信	92
4.4.1 进程间通信概念	92
4.4.2 进程间通信方式	92
4.4.3 Linux 中的进程间通信机制	93
4.5 死锁	100
4.5.1 死锁的基本概念	100
4.5.2 死锁的预防——解决死锁的静态方法	102
4.5.3 死锁的避免——解决死锁的动态方法	102
4.5.4 死锁的检测及解除	105
4.6 管程	106

4.6.1 管程的基本概念	106	5.6 虚拟存储管理	134
4.6.2 管程的特性	107	5.6.1 相关基本概念	135
小结	108	5.6.2 请求页式存储管理	136
习题	108	5.6.3 页面置换算法	140
第 5 章 存储管理	112	5.6.4 请求页式管理性能分析	142
5.1 存储管理中的概念	112	5.6.5 请求段式存储管理	143
5.1.1 存储管理的目的和功能	112	5.6.6 请求段页式存储管理	145
5.1.2 存储分配的方式	113	5.7 Linux 系统存储管理	145
5.1.3 重定位	113	5.7.1 Linux 中的页表	146
5.2 分区存储管理	115	5.7.2 页面的分配与回收	146
5.2.1 固定分区	115	5.7.3 页交换进程与页面的换入换出	147
5.2.2 可变分区	116	5.7.4 Linux 虚拟存储器的保护	148
5.2.3 分配和释放算法	118	小结	148
5.2.4 分区移动技术	120	习题	149
5.2.5 覆盖与交换技术	120	第 6 章 设备管理	151
5.2.6 分区的存储保护	122	6.1 设备管理的概念	151
5.2.7 分区存储管理的主要优缺点	122	6.1.1 设备的分类	151
5.3 页式存储管理	122	6.1.2 设备管理的功能和任务	152
5.3.1 页式基本原理	123	6.2 I/O 控制方式	152
5.3.2 页式管理表	123	6.2.1 程序直接查询控制方式	153
5.3.3 页式地址转换	124	6.2.2 中断方式	153
5.3.4 快表	125	6.2.3 DMA 方式	153
5.3.5 页面分配策略	126	6.2.4 通道方式	155
5.3.6 页面的共享与保护	126	6.3 缓冲技术	156
5.3.7 多级页表	127	6.3.1 单缓冲	156
5.3.8 反置页表	128	6.3.2 双缓冲	157
5.4 段式存储管理	129	6.3.3 循环缓冲	157
5.4.1 段式基本原理	130	6.3.4 缓冲池	158
5.4.2 段式地址转换	130	6.4 驱动调度技术	159
5.4.3 内存分配与释放	131	6.4.1 磁盘的物理结构	159
5.4.4 段的共享与保护	131	6.4.2 磁盘调度	160
5.4.5 段式和页式的比较	132	6.4.3 提高磁盘 I/O 速度的一些方法	162
5.5 段页式存储管理	133	6.5 设备分配	163
5.5.1 段页式基本原理	133	6.5.1 设备分配中的数据结构	163
5.5.2 段页式地址转换	134	6.5.2 设备分配过程	165

6.5.3 设备分配应考虑的因素	166	7.6.2 存取控制表	186
6.5.4 设备独立性	167	7.6.3 口令	187
6.6 虚拟设备	168	7.6.4 密码	187
6.6.1 SPOOLing 系统的组成	168	7.7 Linux 文件系统	188
6.6.2 SPOOLing 系统的工作原理	169	7.7.1 虚拟文件系统	188
6.6.3 SPOOLing 系统的应用	169	7.7.2 Linux 系统主要文件操作	189
6.6.4 SPOOLing 技术的特点	170	小结	190
6.7 Linux 设备管理	170	习题	190
6.7.1 Linux 设备管理概述	170	第 8 章 网络与通信管理	192
6.7.2 设备驱动程序概述	171	8.1 网络的体系结构	192
6.7.3 设备驱动程序框架	171	8.2 网络操作系统的工作模式及网络通信功能	195
小结	172	8.2.1 网络操作系统的工作模式	195
习题	172	8.2.2 操作系统的网络组件	196
第 7 章 文件系统	174	8.3 网络管理	198
7.1 文件系统的基本概念	174	8.3.1 网络管理概述	198
7.1.1 文件、记录和数据项	174	8.3.2 网络管理功能	199
7.1.2 文件类型	175	8.3.3 集中式网络管理技术	200
7.1.3 文件系统模型	175	8.3.4 分布式网络管理技术	201
7.2 文件的逻辑结构与存取方法	176	8.4 远程过程调用	202
7.2.1 文件的逻辑结构	176	8.4.1 远程过程调用概述	202
7.2.2 文件的存取方法	177	8.4.2 远程过程调用的操作	203
7.3 文件的物理结构与存储设备	177	8.4.3 远程过程调用的实现	203
7.3.1 常见的物理结构	178	8.5 网络服务	204
7.3.2 文件的存储设备	180	8.6 网络通信的安全性问题	207
7.4 文件存储空间的管理	180	8.6.1 网络安全概述	207
7.4.1 空闲区表法	181	8.6.2 网络安全措施	208
7.4.2 空闲链表法	181	8.6.3 网络安全服务	208
7.4.3 位示图法	182	小结	211
7.5 文件目录	182	习题	211
7.5.1 文件目录管理的基本要求	182	第 9 章 操作系统安全	213
7.5.2 文件控制块和索引结点	183	9.1 操作系统安全概述	213
7.5.3 文件目录结构	183	9.1.1 操作系统安全的重要性	213
7.5.4 文件的共享	185	9.1.2 操作系统面临的安全威胁	214
7.6 文件安全	186		
7.6.1 存取控制矩阵	186		

9.1.3 操作系统安全的目标	214	9.6.1 标识与口令安全	225
9.2 安全策略	215	9.6.2 存取控制	226
9.2.1 访问支持策略	215	9.6.3 鉴别	227
9.2.2 访问支持策略	216	9.6.4 审计	227
9.3 硬件安全机制	217	9.6.5 网络安全性	227
9.3.1 存储保护	217	9.6.6 加密	228
9.3.2 运行保护	218	9.6.7 网络监视和入侵检测	228
9.3.3 I/O 保护	218	9.6.8 备份/恢复	229
9.4 软件安全机制	219	小结	229
9.4.1 注册与登录	219	习题	229
9.4.2 存取控制	220	操作系统实验	230
9.4.3 最小特权管理	221	实验 1 Linux 系统用户接口和使用环境	230
9.4.4 可信通道	221	实验 2 Linux 进程的创建	233
9.4.5 隐蔽通道	222	实验 3 Linux 进程信号通信	235
9.4.6 安全审计	222	实验 4 Linux 进程消息通信	238
9.4.7 病毒防护	222	实验 5 进程调度模拟	239
9.5 操作系统安全评测	223	实验 6 页面置换算法模拟	240
9.5.1 操作系统安全评测方法	223	实验 7 文件系统模拟设计	241
9.5.2 美国国防部可信计算机系统 评测准则	224	参考文献	242
9.6 Linux 操作系统安全机制	225		

第 1 章

操作系统概论

随着计算机技术的发展,计算机系统的硬件和软件资源也愈来愈丰富。为了提高这些资源的利用率和增强系统的处理能力,计算机系统必须配置操作系统。操作系统是计算机系统的重要组成部分,是不可缺少的系统软件,从个人计算机到巨型计算机系统以及计算机网络都要配置一种或多种操作系统。本章介绍操作系统的基本概念,包括操作系统的地位与作用、操作系统的功能和特性、操作系统的类型等,由此引出操作系统追求的目标,以作为后续章节中操作系统解决各种问题的指导思想。

1.1 操作系统概念

1.1.1 操作系统的地位和作用

1. 地位

现代计算机系统是一个相当复杂的系统,即使是微型计算机系统也不例外,它们都是由系统硬件和系统软件这两部分组成。系统硬件是构成计算机系统所必须配置的设备,它为形成和组织一个系统提供了控制机构,是提供给操作系统的物质基础。系统软件是指计算机系统必须配备的那部分软件,它通常是对各种领域都适用的一些软件,诸如各种程序设计语言的处理程序、各种操作系统、标准程序库以及维护软件等。面对这样复杂的系统,需设置一个自动化的管理机构,由该机构组织各种硬件资源的利用,实现各类软件资源的查找和调用,以及方便用户使用计算机。操作系统(Operating System, OS)正是扮演了这一重要角色。它是计算机系统中主要的系统软件之一,起着管理系统的作用。目前,不论是大、中型计算机系统,还是微、小型计算机系统都配备了相应的操作系统。特别是在软件、硬件结合日趋密切,应用领域日趋广泛的今天,操作系统在计算机系统中所处的举足轻重的地位,就更加突出地显示出来了。

没有任何软件支持的计算机称为裸机,它仅仅构成了计算机系统的物质基础,而实际呈现在用户面前的计算机系统是经过若干层软件改造的计算机。

如图 1-1 所示展示了这种情形。由图可看出,计算机的硬件和软件、部分软件之间是一种层次结构的关系。裸机在最里层,它的外面是操作系统,经过操作系统提供的资源管理功能和方便用户的各种服务功能把裸机改造成为功能更强、使用更方便的机器,通常称之为虚拟机;而各种实



图 1-1 操作系统和软硬件的层次关系图

用程序和应用程序运行在操作系统之上，它们以操作系统为支撑环境，同时又向用户提供完成其作业所需的各种服务。

2. 作用

操作系统的作用可以从不同的观点来观察。从一般用户的观点，可把操作系统看作是用户与计算机硬件系统之间的接口；从资源管理观点，可把操作系统视为计算机系统资源的管理者。因此，操作系统有以下两个重要的作用。

① 管理系统中的各种资源

在计算机系统中，通常都包含了各种各样的硬件和软件资源，操作系统能管理和分配这些资源，合理地组织计算机工作流程，保证系统中的各种资源得以有效的利用。

② 为用户提供良好的界面

操作系统为用户使用计算机提供了良好的界面，使用户无需了解许多有关硬件和系统软件的细节，就能方便灵活地使用计算机。

3. 操作系统的设计目标

目前存在着多种类型的操作系统，不同类型的操作系统，其目标各有所不同。通常期望操作系统达到以下几个目标。

① 方便性。操作系统为用户提供良好的、一致的用户接口，用户按需要输入命令，操作系统按命令去控制程序的执行；用户也可以在程序中调用操作系统的功能模块完成相应服务，而不必了解硬件的物理特性。

② 有效性。操作系统可有效地管理和分配硬件、软件资源，合理地组织计算机的工作流程，提高系统工作效率。操作系统可扩充硬件的功能，使硬件的功能发挥得更好。操作系统使用户合理共享资源，防止各用户间的相互干扰。操作系统以文件形式管理软件资源，保证信息的安全和快速存取。

③ 可扩充性。为满足计算机硬件与体系结构的发展以及不断扩大的应用要求，操作系统应能方便地扩展新的功能。

④ 开放性。开放性指的是产品和技术之间相互连接和协作的能力。无论是硬件还是软件范畴，开放性接口都已作为一种明确的或实际的行业标准广泛应用在公开发行的文档中。例如，互联网依赖于 TCP/IP 网络协议的传播。开放性设计使操作系统必须提供统一开放的接口，使其应用在不同的系统中具有可移植性，并使不同的系统能够通过网络进行集成，从而能正确、有效的协同工作。

方便性和有效性是操作系统设计所追求的两个主要目标。这两个目标有时会发生矛盾：为求方便使用，可能要牺牲效率；同样，为了保证高效性可能影响使用的方便性，在操作系统设计时应根据计算机系统的功能和服务对象进行权衡。

1.1.2 操作系统的定义

根据前面对操作系统地位和操作系统作用的描述，可以给出操作系统的一个描述性定义：操作系统是一组用于控制和管理计算机系统中的所有资源的程序集合，其任务是合理地组织计算机的工作流程，有效地组织诸资源协调一致地工作以完成各种任务，从而达到充分发挥资源效率、方便用户使用计算机之目的。

1.1.3 操作系统的功能

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；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。

操作系统面向系统中的所有软件和硬件资源，要实现对处理机、存储器、I/O 设备、文件和网络的管理；操作系统面向用户，要提供一个良好的用户界面。从资源管理和面向用户的角度看，操作系统的功能主要包括以下几个方面。

1. 处理机管理

在操作系统所管理的系统资源中，处理机是最重要的资源。操作系统要支持多用户、多任务对处理机的共享，因此对处理机的管理成为操作系统最重要的一个组成部分。处理器管理的主要工作是处理器调度。在单用户单任务的情况下，处理器仅为一个用户的一个任务所独占，处理器管理的工作十分简单。为了提高处理器的利用率，操作系统采用了多道程序设计技术。在多道程序或多用户的情况下，组织多个作业或任务执行时，就要解决处理器的调度、分配和回收等问题。为了实现处理器管理的功能，描述多道程序的并发执行，操作系统引入了进程的概念，处理器的分配和执行都是以进程为基本单位；随着并行处理技术的发展，为了进一步提高系统并行性，使并发执行单位的粒度变细，并发执行的代价降低，操作系统又引入了线程的概念。对处理器的管理和调度最终归结为对进程和线程的管理和调度，包括：①进程控制和管理；②进程同步和互斥；③进程通信；④进程死锁；⑤线程控制和管理；⑥处理器调度，又分高级调度，中级调度和低级调度。

正是由于操作系统对处理器的管理策略不同，其提供的作业处理方式也就不同。例如，批处理方式、分时处理方式、实时处理方式等。从而，呈现在用户面前，成为具有不同处理方式和不同特点的操作系统。

操作系统所采用的处理机管理策略决定了操作系统的主要性能。

2. 存储管理

存储管理的主要任务是管理存储器资源，为多道程序运行提供有力的支撑，便于用户使用存储资源，提高存储空间利用率。存储管理的主要功能包括：①存储分配。存储管理将根据用户程序的需要分配给它存储器资源，这是多道程序能并发执行的首要条件，当然程序运行结束撤离时，还需回收存储资源。②存储共享。存储管理可设存储器中的多个用户程序实现存储资源的共享，提高存储器的利用率。③地址转换与存储保护。存储管理负责把用户的逻辑地址转换成物理地址，同时要保证各个用户程序相互隔离互不干扰，更不允许用户程序访问操作系统的程序和数据，从而，保护系统和用户程序存放在存储器中的信息不被破坏。④存储扩充。由于受到处理器寻址能力的限制，一台计算机的物理内存容量总是有限的，难以满足用户大型程序的需求，而外存储器容量大且价格便宜。存储管理还应该能从逻辑上来扩充内存存储器，把内存和外存混合起来使用，为用户提供一个比内存实际容量大得多的逻辑编程空间，方便用户的编程和使用。存储管理须提供虚拟存储来提高内存的利用率和提高进程的并发度。

操作系统对内存管理的实现在很大程度上依赖于硬件机制，操作系统必须针对不同 CPU 所提供的不同硬件机制提供统一的接口，做到在不同的硬件支持环境下得到相同的操作系统界面和系统功能，这就要求操作系统有很好的可移植性。

3. 设备管理

计算机系统所配置的外部设备是多种多样的,其工作原理、I/O 传输速度、传输方式都有很大的差异。操作系统采取统一的文件系统界面来管理外部设备,而将设备本身的物理特性交由设备驱动程序去解决,从而提高系统对多种设备的适应性。

设备管理的主要任务是管理各类外围设备,完成用户提出的 I/O 请求,加快 I/O 信息的传送速度,发挥 I/O 设备的并行性,提高 I/O 设备的利用率,以及提供每种设备的驱动程序和中断处理程序,为用户隐蔽硬件细节,提供方便简单的设备使用方法。为实现这些任务,设备管理应该具有以下功能。①提供 I/O 设备的控制与处理。设备管理利用处理机提供的中断机制来实现 I/O 设备与 CPU 的并行。②提供缓冲区的管理。I/O 设备与 CPU 是并发执行的,但 CPU 与 I/O 设备的运行速度并不匹配,因此通常需要采用缓冲技术来平滑速度上的差异。③提供设备独立性。为了方便用户,也为了提高设备利用率,操作系统采用“设备独立性”的概念,即操作系统仅向用户提供逻辑设备名,而将物理设备的分配交由操作系统控制和管理。④外围设备的分配和去配。在一个系统中配置的设备的种类和台数总是有限的,一般少于使用者的个数,设备管理需制定设备分配的策略和分配的实施。⑤实现共享型 I/O 设备的驱动调度。为了提高直接存取存储设备的效率,能够提供一种调度策略,使能按最佳次序执行要求访问的请求。⑥实现虚拟设备。为了提高独占设备的利用率,操作系统利用共享设备比如磁盘来模拟独占设备,将独占设备在逻辑上改造为共享设备,这种在逻辑上模拟独占设备的共享设备被称为虚拟设备。操作系统采用虚拟分配方式分配虚拟设备。

4. 文件管理

文件管理是对系统中信息资源的管理,信息资源包括:程序和数据。在现代计算机中,通常把程序和数据以文件形式存储在外存储器(又称文件存储器)上。文件管理是最接近用户的部分,它给用户提供了一个方便、快捷、可以共享、同时又提供保护的文件使用环境。文件管理的主要任务是对用户文件和系统文件进行有效管理,实现按名存取;实现文件的共享、保护和保密,保证文件的安全性;并提供给用户一整套能方便使用文件的操作和命令。具体来说,文件管理要完成以下任务。①提供文件逻辑组织方法。②提供文件物理组织方法。③提供文件存取方法。④提供文件使用方法。⑤实现文件的目录管理。⑥实现文件的共享和存取控制。⑦实现文件的存储空间管理。

5. 网络管理

随着计算机网络功能的不断加强,网络的应用不断深入,操作系统也应提供计算机与网络进行数据传输和网络安全防护功能。操作系统至少应具有以下网络管理功能。①网上资源管理功能。计算机网络的主要目的之一是共享资源,网络操作系统应实现网上资源的共享,管理用户对资源的访问,保证信息资源的安全性和完整性。②数据通信管理功能。计算机联网后,节点之间可以互相传送数据,按照通信协议的规定,完成网络上计算机之间的信息传送。③网络管理功能。包括:故障管理、安全管理、性能管理、配置管理等。

6. 提供良好的用户界面

操作系统是计算机与用户之间的接口,最终是用户在使用计算机,所以它必须为用户提供一个良好的用户界面。用户界面的好坏直接关系到操作系统能否得到用户认可的一个关键问题。为了使用户能灵活、方便地使用计算机和系统功能,操作系统应提供一组使用其功能的手段称用户接口,它包括两大类:程序接口和操作接口。用户通过这些接口能方便地调用操作系统功能,有效地组织作业及其工作和处理流程,并使整个系统能高效地运行。

1.1.4 操作系统的特性

采用了多道程序设计技术的操作系统具有如下特性。

1. 并发性

并发性是指两个或两个以上的事件或活动在同一时间间隔内发生。操作系统是一个并发系统，并发性是它的重要特征，操作系统的并发性指计算机系统中同时存在若干个运行着的程序，因此，它应该具有处理和调度多个程序同时执行的能力。内存中同时有多个用户程序，或内存中同时有操作系统程序和用户程序被启动交替、穿插地执行，这些都是并发性的例子。并发性能够消除计算机系统中部件和部件之间的相互等待，有效地改善系统资源的利用率，改进系统的吞吐率，提高系统效率。

并发执行的好处是能够使 CPU 与外设同时并行执行，提高了资源的利用率。并发存在的问题是：并发必然导致共享资源的矛盾；并发使多道程序执行过程的中断；同一程序在不同硬件上的调度转换；同一 CPU 上不同程序的现场切换、存储保护以及合作进程如何同步等。这些都是导致操作系统变得复杂的原因。为了更好地解决上述问题，操作系统必须提供机制和策略来进行协调，以使各个并发进程能顺利推进，并获得正确的运行结果。另外，操作系统还要合理组织计算机工作流程，协调各类硬软件设施工作，充分提高资源的利用率，充分发挥系统的并行性，这些也都是在操作系统的统一指挥和管理下进行的。

计算机系统中，并发实际上是一个物理 CPU 在若干道程序之间多路复用，这样就可以实现运行程序之间的并发，以及 CPU 与 I/O 设备、I/O 设备与 I/O 设备之间的并行，并发性的实质是对有限物理资源强制行使多用户共享以提高效率。在多处理器系统中，程序的并发性不仅体现在宏观上，而且体现在微观上，这称为并行性。并行性是指两个或两个以上事件或活动在同一时刻发生。在多道程序环境下，并行性使多个程序同一时刻可在不同 CPU 上同时执行。而在分布式系统中，多台计算机的并存使程序的并发性得到了更充分的发挥，因为，同一时刻每台计算机上都可以有程序在执行。可见并行的事件或活动一定是并发的；但反之并发的活动未必是并行的，并行性是并发性的特例，而并发性是并行性的扩展。

2. 共享性

共享指计算机系统资源可被多个并发执行的用户程序和系统程序共同使用，而不是被其中某一个程序所独占。共享有两种形式：其一是顺序共享，所有系统资源都是可顺序共享的，即一个资源可被多个用户或程序顺序使用，当一个用户或程序使用完了之后，其他用户或程序便可使用；其二是并发共享。这是一种竞争性共享，即在一段时间内有多个用户或程序在同时使用某个资源，但在某一具体时刻，只能有一个用户或程序在真正使用，其他用户或程序必须等待，而且每个用户或程序都可能分多次使用该资源。

共享性和并发性是操作系统两个最基本的特性，它们互为依存。一方面，资源的共享是因为程序的并发执行而引起的，若系统不允许程序并发执行，自然也就不存在资源共享问题。另一方面，若系统不能对资源共享实施有效管理，必然会影响到程序的并发执行，甚至程序无法并发执行，操作系统也就失去了并发性，导致整个系统效率低下。

要实现资源的并发共享，操作系统需要解决一系列问题，如资源分配、资源公用、信息保护、存取控制等。

3. 不确定性

不确定性也称异步性。在多道程序并发执行的环境中，各程序之间存在着直接或间接的联系，

程序的推进速度会受到运行环境的影响,若不能正确控制,则执行结果会因为运行环境的不同而不同。由于在多道并发环境下,程序执行的过程与其他并发程序相互之间有干扰,所以其执行过程中在何时何地由于时间片耗尽而被中断,其请求的 I/O 何时完成以及继续执行所需的消息或信号何时到来等,都具有随机性和不确定性。另外,分时系统中用户随时从终端上发出的终止执行命令,也导致操作系统具有不确定性。随机性事件是造成操作系统不确定性的基本原因,这给操作系统的设计与运行带来了很大的困难。

操作系统的不确定性是并发与共享的必然结果。为了实现并发与共享,操作系统必须有能力随时响应和正确处理各种随机事件。它不是事先规定各种事件何时发生,而是事先安排好对各种可能事件的处理,不管这些事件何时、以何种次序以及何种组合方式发生,这就是操作系统的不确定性。

4. 虚拟性

虚拟性是指操作系统中的一种管理技术,它是把物理上的一个实体变成逻辑上的多个对应物,或把物理上的多个实体变成逻辑上的一个对应物的技术。所谓虚拟是指物理上没有提供,但是逻辑上却具备的功能。在用户看来好像是物理上原来就具有的功能一样。采用虚拟技术的目的是为了

提高资源利用率和为用户提供易于使用、方便高效的操作环境。操作系统的虚拟性体现在如下两个方面。①操作系统虚拟机。操作系统运行在硬件提供的功能支持之下,为用户提供方便易用的功能界面。在用户看来,似乎是计算机硬件本身具有的功能,这就是操作系统虚拟机。操作系统采用不断扩充功能、逐层虚拟的分层结构,其每一层都依赖于下层功能的支持,同时为上层提供一个功能更加强大的支撑平台,每一层都是一个虚拟机。②操作系统对于系统的硬件和软件资源都采用了一定的虚拟机制。如虚拟处理机、虚拟内存、虚拟设备、虚拟文件系统。

虚拟与异步性是操作系统的两个重要特征。虚拟技术为共享提供了更好的条件,而并发与共享是导致不确定性的根本原因。

操作系统的并发性、共享性、不确定性、虚拟性四个特征不是相互独立的,具有密切的关系:并发和共享是操作系统的两个最基本的特征,它们又是互为存在的条件。

1.1.5 操作系统的性能

操作系统的性能如何,直接影响计算机系统的总体性能。一般从下面几个方面来评价操作系统的性能。

(1) 系统效率

操作系统的效率主要体现在:资源的利用率要高,内存利用率高,外部设备利用均衡;对用户的周转时间要短,吞吐量要大,响应时间要尽量短。

(2) 系统可靠性与安全性

一个运行可靠、安全的操作系统是大家所期待的。由于操作系统是整个计算机系统的所有硬件与软件资源的管理者,它的可靠性与安全性,直接影响着整个计算机系统的运行可靠性与安全性。此外,由于系统的使用环境复杂多变,用户的误操作也可能造成系统工作不正常。在系统运行时,操作系统能及时检测出错误,并能尽量准确地指出错误发生的原因,采取相应的措施排除错误;对错误产生的损害,能进行修复,使系统恢复正常工作,尽量减少损失,具有较好的健壮性。

(3) 系统可维护性

在软件的生命期中,其维护成本已占软件成本的 70%。操作系统和普通软件一样应具有较好

的可维护性。操作系统往往是一个庞大的软件集合，目前还无法证明软件的正确性，操作系统本身存在错误与不足在所难免，因而对操作系统进行纠错性维护是必须的。另外，随着应用领域的扩大，硬件的扩充，计算机体系结构的变化，新功能的增加，都将要求对操作系统进行修改维护，这就要求操作系统有较好的结构，完备的文档，以便于维护。操作系统是否具有较好的维护性是决定操作系统生命周期长短的重要因素。

(4) 易用性

系统提供的各种服务能方便用户使用，人们对易用性要求越来越重视，它涉及计算机系统使用的简单性、可操作性、可携带性等。操作系统的人性化设计越来越受到人们的重视。

(5) 可扩充性

操作系统的功能应该不断被加强、改进和完善。在引进新的系统组件时不应干扰现有的服务能力，从而能够适应不断发展的应用需求。

(6) 开放性

为了能够集成不同厂家生产的计算机及设备，且能正确有效地协同工作，支持应用程序的可移植性和互操作性，要求操作系统具有开放性。

1.2 操作系统的形成和发展

操作系统的形成和发展是和计算机发展密不可分的，随着计算机性能的不断提高，运行在其上的操作系统也从无到有，从简单到复杂，成为一个非常重要的系统软件。

1.2.1 硬件软件的发展轨迹

操作系统理论是在计算机的应用中诞生并成长的，它的发展与计算机硬件的发展是密不可分的。从硬件角度看操作系统的发展轨迹如表 1-1 所示。

表 1-1 从硬件角度看操作系统发展轨迹

年 代	硬 件 特 点	操作系统特点	背 景
机械计算机时代（17 世纪 ~ 20 世纪初）	① 纯机械结构，低速 ② 只能进行简单的数学运算	纯手工操作	从计算尺到差分机再到分析机发展了数百年
第一代计算机（1946 年 ~ 20 世纪 50 年代末） 电子管计算机	① 体积大，能耗高，故障多，价格贵 ② 难以普及应用	无操作系统，手工操作阶段（程序以机器码编写，载体从插件板到卡片与纸带）	1906 年发明电子管； 1946 年第一台电子管计算机 ENIAC 研制成功
第二代计算机（20 世纪 50 年代末 ~ 60 年代中期） 晶体管计算机	① 采用印刷电路 ② 稳定性与可靠性大大提高 ③ 批量生产成为可能 ④ 进入实际应用领域但数量有限	① 单道批处理系统 ② 操作系统以监督软件形式出现 ③ 任务按顺序方式处理	1947 年发明晶体管
第三代计算机（20 世纪 60 年代中期 ~ 70 年代初） 集成电路计算机	① 体积减小，性价比迅速提高 ② 小型计算机发展迅速 ③ 进入商业应用领域 ④ 尚不适合家庭应用的需求	① 涌现大批操作系统，包括多道批处理系统、分时系统和实时系统 ② 奠定了现代操作系统的基本框架	1958 年发明集成电路； 1971 年 Intel 公司发明微处理器