

# 操作系统原理 与实践教程

- ◆ 操作系统的概念
- ◆ 操作系统的界面
- ◆ 处理机管理
- ◆ 进程同步与死锁
- ◆ 存储管理
- ◆ 文件管理
- ◆ 设备管理
- ◆ 安全和保护



周湘贞 曾宪权 编著



清华大学出版社

高等院校计算机应用技术系列教材

# 操作系统原理与实践教程

周湘贞 曾宛权 编著

清华大学出版社

北 京

## 内 容 简 介

本书采用了理论与实践相结合的方式,全面系统地介绍了现代操作系统的基本理论和最新技术。全书共分为8章,第1章介绍了操作系统的概念、功能和特点、发展历史以及操作系统结构和设计的相关问题;第2章介绍了操作系统的生成及操作系统向用户提供的接口;第3、4章详细介绍了进程和线程的概念、同步和通信机制、调度与死锁;第5、6、7章分别介绍了操作系统的存储管理、文件管理和设备管理功能;第8章介绍了操作系统的安全和保护的问题。

本书可作为计算机科学与技术、软件工程、信息管理等专业本、专科生教材和考研、考证参考书,也可以作为从事计算机工作的科技人员进行学习和开发的参考书。

本书每章中的教学课件可以到 <http://www.tupwk.com.cn/downpage/index.asp> 网站下载。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

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。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

### 图书在版编目(CIP)数据

操作系统原理与实践教程/周湘贞,曾宪权 编著. —北京:清华大学出版社,2006.10

(高等院校计算机应用技术系列教材)

ISBN 7-302-13410-3

I. 操… II. ①周…②曾… 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(2006)第 078054 号

出 版 者:清华大学出版社 地 址:北京清华大学学研大厦

<http://www.tup.com.cn> 邮 编:100084

社 总 机:010-62770175 客 户 服 务:010-62776969

组稿编辑:胡辰浩

文稿编辑:袁建华

封面设计:王 永

版式设计:康 博

印 刷 者:北京市清华园胶印厂

装 订 者:三河市化甲屯小学装订二厂

发 行 者:新华书店总店北京发行所

开 本:185×260 印张:21.75 字数:502千字

版 次:2006年10月第1版 2006年10月第1次印刷

书 号:ISBN 7-302-13410-3/TP·8423

印 数:1~5000

定 价:30.00元

# 前 言

操作系统是计算机科学与技术专业的一门专业基础课，是计算机相关专业学生的必修课程。在知识经济蓬勃发展的今天，软件产业占据着核心地位，而操作系统是所有软件的基础，是软件的根本，是计算机系统中的核心系统软件，用于控制和管理计算机系统中的各种软硬件资源，提供了用户与计算机之间的接口，其性能直接影响到计算机系统的工作效率，所以，操作系统是计算机领域中最活跃的学科之一，其发展极为迅速。然而，正是由于操作系统在计算机中所处的特殊地位，使得这门课程的学习具有相当的难度，主要原因如下：

(1) 操作系统是一个庞大的系统软件，从而导致本课程的内容非常复杂，涉及计算机软、硬件的诸多方面。

(2) 课程内容较抽象。尽管很多用户使用计算机时都要与操作系统打交道，但是对于什么是操作系统、操作系统应有哪些功能、这些功能如何实现等问题并不是很清楚，因而对操作系统的内容倍感抽象、费解。

(3) 发展变化快。操作系统课程随着计算机的发展而不断发展，是计算机软件中变异、更新最快的软件，因而更加重了学生的学习难度。

为了解决这些问题，提高操作系统课程的教学质量，在广泛汲取国内外优秀教材和研究成果的基础上，借鉴、参照 ACM/IEEE-CS2002 和 CCC2002 中操作系统课程教学的相关内容，结合自己多年的操作系统课程的教学经验，笔者编写了本教材。整个教材根据课程的特点，结合实例来讲解理论，培养学生的实践能力和创新能力。

具体来说，本教程具有以下特点：

(1) 理论联系实际。本教材把操作系统原理和实际的操作系统实现结合起来讲解，在讲解原理之后给出实例，从而把理论与实践有机地结合起来，使抽象的理论更利于读者进行理解和消化，提高学生的学习兴趣。

(2) 实用性强。本教程结合教学内容，提供了相关的编程实例和技术，并精选了部分与课程内容相关的实验项目，每个实验项目还给出了相关的背景知识。读者通过阅读这些背景知识，基本可以完成这些项目。这些实验可以加深读者对操作系统原理的理解，另一方面也可以培养学生进行系统程序设计和分析的能力。

(3) 内容新颖。本教材在选材的过程中，根据现代操作系统的发展要求，选择最近几年出现的新思想、新概念、新技术补充到教材中，在一定程度上反映了操作系统的发展方向，从而提高学生适应迅速变化的操作系统发展的能力，了解操作系统的发展趋势，培养学生的创新能力。

(4) 适用面广。本教材做到了理论与实践的有机结合，既讲述了操作系统的基本原理，又介绍了相关的编程知识，因此，本书既可以作为教材使用，也可以作为其他工程人员的

参考书。

(5) 每章的章首给出本章的教学内容和要求, 并给出了相关的学习建议, 利于教学也利于学生的学习。

全书共分 8 章, 基本上涵盖了现代操作系统的主要内容。第 1 章介绍了操作系统的概念、功能、特征、发展历史和结构; 第 2 章分析了操作系统的生成和启动以及操作系统提供的服务和接口; 第 3~7 章分别介绍了处理器管理与调度、进程同步和死锁、存储管理、文件管理和 I/O 管理; 第 8 章叙述了操作系统与计算机系统安全和保护问题。

本书可作为普通高校计算机科学与技术及其相关专业本、专科学生的教材, 也可以作为自学参考书和考研参考书。由于各高校的不同专业教学安排要求和教学时间有一定的差别, 因此, 在教学内容上可酌情进行取舍。如果课时较充分, 可以讲授全部内容, 并安排上机实践来完成实验项目。如果课时较少, 可讲授每章的基本内容, 实验项目可作为学生的作业。另外, 为检查学习效果, 每章后面留有作业和相应的实验项目, 读者可根据实际情况有选择的使用。

本书主要由周湘贞和曾宪权编写, 其中, 第 1 章由周湘贞编写, 其余章节(第 2~8 章)由曾宪权编写, 李晓利等 8 人对文稿进行了认真阅读并提出了自己宝贵的意见。在教材的编写和出版过程中, 赵树升给予了大力的支持和帮助, 并提出了指导意见, 在此表示衷心感谢。刘春丽参与了部分内容的校对工作, 另外, 邬治国、乔丹丹、朱金玲、郑立艳、胡琼怡、许成琳、孙芳、张志云和韩垒等参加了材料收集和制作, 在此表示感谢。

当然, 由于编者水平有限, 时间紧张, 疏漏和错误在所难免。真诚希望各位读者批评和指正, 我们将不胜感激。我们的信箱: [huchenhao@263.net](mailto:huchenhao@263.net)。

作者  
2006 年 5 月

# 目 录

<b>第 1 章 操作系统概论</b> .....	1
1.1 操作系统的概念 .....	1
1.1.1 计算机系统的层次结构 .....	1
1.1.2 操作系统的作用 .....	2
1.1.3 操作系统的主要特征 .....	4
1.1.4 操作系统的功能 .....	5
1.2 操作系统的形成和发展 .....	7
1.2.1 人工操作阶段 .....	8
1.2.2 管理程序阶段——操作系统的雏形 .....	8
1.2.3 多道批处理——现代意义上的操作系统出现 .....	9
1.2.4 分时系统与实时系统的出现——操作系统步入实用化 .....	10
1.2.5 操作系统的进一步发展 .....	11
1.3 操作系统的结构 .....	15
1.3.1 操作系统的设计 .....	15
1.3.2 操作系统的结构 .....	16
1.4 Windows Server 2003 和 Linux 系统模型 .....	21
1.4.1 Windows Server 2003 操作系统模型 .....	21
1.4.2 Linux 系统的结构 .....	22
1.5 小结 .....	23
1.6 思考练习 .....	23
<b>第 2 章 操作系统的界面</b> .....	25
2.1 系统的生成与启动 .....	25
2.1.1 系统的生成 .....	25
2.1.2 系统的启动 .....	26
2.1.3 实例分析： Linux 系统启动 .....	27
2.2 操作系统提供的服务和用户接口 .....	29
2.2.1 操作系统提供的基本服务 .....	29
2.2.2 命令接口 .....	30
2.2.3 图形用户接口 .....	32
2.2.4 程序接口——系统调用 .....	34
2.3 小结 .....	39
2.4 思考练习 .....	40
2.5 实验室项目——操作系统的生成和接口的使用 .....	40
<b>第 3 章 处理器管理</b> .....	42
3.1 进程的概念 .....	42
3.1.1 程序的执行方式 .....	42
3.1.2 进程的概念 .....	43
3.1.3 进程的状态 .....	45
3.1.4 进程描述——进程控制块 .....	48
3.2 进程控制 .....	52
3.2.1 进程控制机构 .....	52
3.2.2 进程控制 .....	53
3.3 线程 .....	59
3.3.1 线程的概念 .....	59
3.3.2 线程的实现 .....	60
3.3.3 多线程模型 .....	61
3.3.4 线程池 .....	62
3.4 处理器调度 .....	66
3.4.1 处理器调度的层次 .....	67
3.4.2 进程调度 .....	68
3.4.3 选择调度算法的准则 .....	69
3.5 调度算法 .....	70
3.5.1 先来先服务 .....	70
3.5.2 短作业(进程)优先 .....	71

3.5.3	优先级调度	71	4.5.3	消息缓冲队列通信机制	120
3.5.4	轮转法调度	72	4.5.4	客户—服务器系统通信	122
3.5.5	多级队列调度	73	4.6	死锁概述	134
3.5.6	多级反馈队列调度	74	4.6.1	死锁的概念	134
3.5.7	高响应比优先调度	75	4.6.2	死锁产生的原因和 必要条件	135
3.6	多处理器调度和实时调度	76	4.6.3	死锁的描述— 资源分配图	137
3.6.1	多处理器调度	76	4.6.4	处理死锁的方法	139
3.6.2	实时调度	76	4.7	死锁的预防和避免	139
3.7	小结	82	4.7.1	死锁的预防	139
3.8	思考练习	83	4.7.2	死锁的避免	141
3.9	实验室项目二 Shell 编程	84	4.8	死锁的检测和解除	146
3.10	实验室项目三 Windows 多线程控制台程序	87	4.8.1	死锁的检测	146
<b>第 4 章</b>	<b>进程同步机制与死锁</b>	<b>92</b>	4.8.2	死锁的解除	149
4.1	进程的同步和互斥	92	4.9	小结	150
4.1.1	进程的同步	93	4.10	思考练习	151
4.1.2	进程互斥	93	4.11	实验室项目四 有限缓冲区问题	153
4.1.3	信号量机制	95	<b>第 5 章</b>	<b>存储管理</b>	<b>160</b>
4.2	经典同步问题	103	5.1	存储管理的概念	160
4.2.1	生产者—消费者问题 (有限缓冲区问题)	103	5.1.1	用户程序的处理过程	161
4.2.2	读者—写者问题	105	5.1.2	存储管理的功能	163
4.2.3	哲学家进餐问题	106	5.2	连续内存分配	167
4.2.4	理发师问题	107	5.2.1	分区管理基本原理	168
4.3	管程	109	5.2.2	分区的分配与回收	170
4.3.1	管程的基本概念	109	5.2.3	碎片问题	172
4.3.2	条件变量	111	5.3	内存不足时的管理	173
4.3.3	利用管程解决生产者 —消费者问题	111	5.3.1	覆盖	173
4.4	操作系统同步模型	113	5.3.2	交换	174
4.4.1	Windows Server 2003 中的进程同步	113	5.4	分页存储管理	176
4.4.2	Linux 中的进程同步	114	5.4.1	分页存储管理的 基本原理	176
4.5	进程通信	116	5.4.2	地址映射	178
4.5.1	进程通信的方式	116	5.4.3	页表的结构	180
4.5.2	消息传递系统	117	5.4.4	页面的共享	182

5.5 分段存储管理	182	6.4 文件系统的实现	238
5.5.1 分段存储管理的基本原理	183	6.4.1 文件系统的结构	238
5.5.2 地址转换	184	6.4.2 文件系统的实现	240
5.5.3 段的共享和保护	185	6.4.3 文件存储空间的分配	241
5.5.4 段页式存储管理	186	6.4.4 空闲空间的管理	245
5.6 虚拟存储器	190	6.5 文件的共享和保护	248
5.6.1 虚拟内存的概念	190	6.5.1 文件的共享	248
5.6.2 虚拟内存的特征	192	6.5.2 文件的保护	250
5.7 请求分页存储管理技术	192	6.5.3 文件系统的可靠性	252
5.7.1 请求分页存储管理的基本思想	192	6.6 实例研究: Windows 和 Linux 的文件系统	252
5.7.2 页面置换算法	194	6.6.1 Windows Server 2003 文件管理	252
5.7.3 请求分页系统内存块的分配	201	6.6.2 Linux 文件管理的实现	260
5.8 存储管理实例研究	203	6.7 小结	265
5.8.1 Windows Server 2003 内存管理	203	6.8 思考练习	266
5.8.2 Linux 系统的存储管理	207	6.9 实验室项目五 简单文件系统的实现	267
5.9 小结	214	<b>第 7 章 设备管理</b>	<b>272</b>
5.10 思考练习	215	7.1 设备管理的概念	272
<b>第 6 章 文件管理</b>	<b>218</b>	7.1.1 设备的分类	272
6.1 文件的概念	218	7.1.2 设备管理的功能和任务	273
6.1.1 文件及其分类	218	7.1.3 I/O 系统结构	274
6.1.2 文件的属性	220	7.1.4 设备控制器	275
6.1.3 文件的组织	221	7.2 I/O 控制方式	276
6.1.4 文件的访问方法	222	7.2.1 轮询(polling)方式	276
6.2 目录结构	224	7.2.2 中断方式	277
6.2.1 文件控制块和文件目录	224	7.2.3 DMA 方式	279
6.2.2 单级目录	226	7.2.4 通道控制方式	281
6.2.3 二级目录	226	7.3 中断技术	283
6.2.4 树形目录	227	7.3.1 中断的基本概念	283
6.2.5 目录的实现	230	7.3.2 中断的分类与优先级	284
6.3 文件和目录操作	231	7.3.3 软中断	284
6.3.1 文件操作	231	7.3.4 中断处理过程	285
6.3.2 目录操作	237	7.4 缓冲技术	287
		7.4.1 缓冲的引入	287



7.4.2 缓冲的种类 .....	287	8.1.1 计算机系统安全概述 .....	323
7.4.3 缓冲池的管理 .....	288	8.1.2 操作系统安全是 系统安全的基础 .....	324
7.5 设备分配 .....	290	8.1.3 安全威胁及其分类 .....	325
7.5.1 设备分配用数据结构 .....	290	8.2 操作系统安全策略 .....	327
7.5.2 设备分配的原则 .....	292	8.2.1 安全策略和机制 .....	327
7.5.3 设备分配技术 .....	294	8.2.2 身份认证机制 .....	328
7.5.4 SPOOLing 系统 .....	294	8.2.3 授权机制 .....	330
7.6 I/O 软件原理 .....	296	8.2.4 加密机制 .....	330
7.6.1 I/O 软件的设计 目标和原则 .....	296	8.2.5 审计 .....	331
7.6.2 I/O 中断处理程序 .....	296	8.3 操作系统的内部保护机制 .....	331
7.6.3 设备驱动程序 .....	297	8.3.1 操作系统保护层次 .....	331
7.6.4 与硬件无关的 I/O 软件 .....	298	8.3.2 内存储器的保护 .....	332
7.6.5 用户空间的 I/O 软件 .....	300	8.3.3 面向用户的访问控制 .....	332
7.7 磁盘调度和管理 .....	301	8.3.4 面向数据的访问控制 .....	333
7.7.1 磁盘的结构 .....	301	8.4 访问控制机制 .....	333
7.7.2 磁盘调度 .....	302	8.4.1 保护域 .....	334
7.7.3 磁盘管理 .....	306	8.4.2 访问矩阵 .....	335
7.8 小结 .....	314	8.4.3 访问矩阵的实现 .....	336
7.9 思考练习 .....	315	8.5 小结 .....	337
7.10 实验室项目六 软盘驱动程序 .....	316	8.6 思考练习 .....	338
<b>第 8 章 操作系统的安全和保护 .....</b>	<b>323</b>	<b>参考文献 .....</b>	<b>339</b>
8.1 操作系统和计算机 系统安全 .....	323		

# 第1章 操作系统概论

## 教学目标和要求

- 了解操作系统在整个计算机系统中的作用。
- 掌握操作系统的概念和现代操作系统的特征、功能。
- 了解操作系统的发展历程以及发展趋势，掌握批处理系统、分时系统和实时系统的特点。
- 了解操作系统的设计目标，熟悉操作系统的设计结构。

## 学习建议

本章内容是操作系统的总论，涉及的内容比较多，也比较抽象和枯燥，因此，在学习时，应加强对概念的理解，结合 Windows 和 Linux 等商用操作系统的发展历程来理解整个操作系统的发展，进而掌握现代操作系统的特点和功能。

## 本章导读

操作系统是计算机系统中最重要系统软件，它管理整个计算机系统的软件和硬件资源，是其他软件和程序的运行基础，是沟通用户与计算机硬件的桥梁。由于各种操作系统的应用领域不同，所以它们具有不同的设计目标和要求，但这些操作系统仍然具有一些共性。

本章叙述了操作系统的发展历史，介绍了现代操作系统的特征和功能，并对现代操作系统的设计问题作了简单介绍，为进一步学习操作系统奠定了基础。

## 1.1 操作系统的概念

### 1.1.1 计算机系统的层次结构

计算机系统是一种能够按照用户的要求接收和存储信息、自动进行数据处理，并输出结果的系统。一个完整的计算机系统由硬件和软件两大部分组成。硬件包括中央处理器(CPU)、存储器、输入和输出设备等，它们向用户提供了基本的计算资源，用户可以借助这些资源来完成计算任务。显然，每类硬件资源具有自己的物理特性和操作方式，其功能非常有限，因此，为了方便用户的使用和扩展计算机系统的功能，现代计算机系统均向用户提供了相应的软件来实现计算任务，并规范用户使用硬件资源的方式。因此，

现代计算机系统是硬件和软件的有机统一体，硬件是基础，软件是工具。

根据计算机系统中硬件和软件所起的作用不同，可以将一个完整的计算机系统粗略分成：硬件、操作系统、系统软件和应用程序 4 个层次，如图 1-1 所示。硬件层提供了基本的可计算性资源，包括：具有一组指令的处理器、可被访问的寄存器和存储器，可被使用的各种 I/O 设施和设备。这些设施和设备是操作系统和软件赖以工作的基础，也是操作系统设计者可以使用的功能和资源。操作系统层是硬件层上的第一层软件，是对硬件的首次扩充和改造，主要用于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工作。操作系统是上层软件运行的基础，它为编译程序、编辑程序、数据库系统等设计者提供了有力的支撑。系统程序层建立在经过操作系统改造和扩充过的机器上，提供了扩展指令集，实现各种语言处理程序、数据库管理系统和其他系统程序。此外，还向用户提供了种类繁多的实用程序，如链接装配程序、库管理程序、诊断排错程序、分类/合并程序等。应用程序层用于解决用户的不同应用要求，如娱乐、办公等。通过使用操作系统提供的支撑环境，应用程序开发者可以借助各种程序设计语言来快捷、方便地开发出各种应用程序，以满足用户的应用要求，而不需要考虑计算机系统硬件之间的差异。因此，完全可以认为，在整个计算系统中，操作系统和硬件组成了一个运行平台，其他软件都运行在这个平台之上。

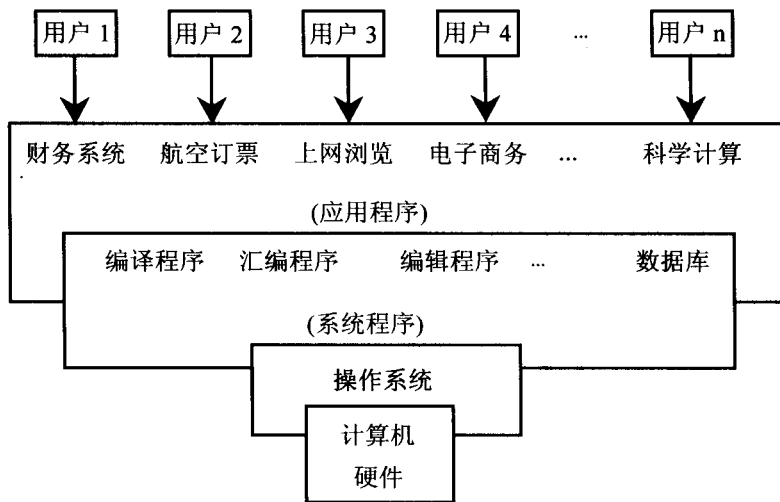


图 1-1 计算机系统组成逻辑图

### 1.1.2 操作系统的作用

操作系统是计算机硬件上的第一层软件，可以从不同的角度来观察操作系统的作用。一般来说，操作系统在计算机系统中的作用可以从以下 3 个方面来理解：

(1) 操作系统是用户与计算机硬件之间的接口

不同的用户使用计算机的目的是不同的，例如：有的是为了娱乐，有的是为了完成某

项工程等,因此,每个用户使用的操作系统界面应该也是不完全相同的,但是计算机系统的基本工作方式大致相同,这就要求操作系统能够屏蔽计算机硬件的物理特性差异,向用户提供一个良好的工作环境,从而方便用户使用计算机来完成计算任务。操作系统作为计算机硬件上的第一层软件,其主要作用就是向用户提供良好的工作环境,使用户可以直接调用操作系统提供的各种功能,而无需了解软、硬件本身的细节。从用户的角度来看,操作系统便成为了用户与计算机硬件之间的一个接口。通过这个接口,用户可以通过使用不同的界面(如:Windows的图形界面和控制台方式)方便地使用计算机完成自己的计算任务。

### (2) 操作系统为用户提供虚拟计算机

在机器语言一级上,计算机的体系结构是原始的且编程是很困难的,尤其是输入和输出操作。例如,当用户使用磁盘来进行I/O操作时,必须得了解磁盘的各种参数(如:磁盘的扇区数、物理介质的记录格式等)。很显然,这对程序员的编程造成了很大的困难,而对于一般的程序员来说,他们并不想涉足磁盘编程的详细细节,他们需要的是—种简单的、高度抽象的、可以与之打交道的设备。这就需要采用软件技术将硬件的复杂性和用户隔离,向用户提供一个更好的使用计算机设备的接口,这种软件就是操作系统。操作系统隐藏了计算机硬件的底层特性,向用户提供了一个虚拟的计算机系统,使用户能够实现处理器的管理、存储空间的管理、输入和输出设备的控制和管理等。也就是说,当在计算机上安装了操作系统后,计算机系统的功能显著增强,使用更加方便、安全。从某种意义上讲,操作系统为用户提供了一台扩展的机器,称为虚拟机(Virtual Machine),它比底层硬件的功能更强,更易于编程和使用。这种虚拟功能还可以重叠,在裸机上覆盖了第一层软件之后,得到第一层扩展,然后,在此基础上再加上一层软件,得到第二层扩展,以此类推。

### (3) 操作系统是计算机系统的资源管理者

用户程序的运行需要相应的资源,典型的资源有处理器、内存储器、输入和输出设备等。在现代操作系统中,计算机系统的这些资源是被其他程序共享的,因而程序的性能受到资源和其他同时执行程序的影响。这就要求操作系统必须监视系统资源的使用状态,优化各种计算资源,合理分配资源以保证各个程序的正确执行,并能够满足某些特殊的要求。

现代操作系统都承担着计算机系统资源管理者的角色,是保证各种程序运行性能的基础。现代操作系统的重要任务之一就是—对资源进行抽象研究,有序管理计算机系统的资源,保证用户对资源的需求,协调各程序对资源的使用冲突,提高资源的利用率。

总之,操作系统是计算机系统中最重要—的系统软件,是其他程序运行的基础。但是,到底什么是操作系统?它应该具有哪些功能?目前还没有一个完整的定义,结合上面的介绍,本书把操作系统定义为:操作系统是控制和管理计算机硬件和软件资源,并合理地对各种资源进行分配和调度,规范计算机工作流程,以方便用户使用的程序的集合。

操作系统是计算机系统的基础软件,它常驻内存,向用户程序提供支撑环境,所以,操作系统由哪些成分组成变得非常重要。—个比较普遍的定义是:操作系统是—直运行在计算机上的系统程序(通常称为内核),其他程序则为应用程序,它们运行在操作系统提供的良好环境中。因此,操作系统类似于政府,它本身并不能实现任何有用的功能,只不过是向其他程序提供了一个工作环境。1998年,美国司法部控告微软公司的垄断经营,就是

因为微软公司将过多的功能加到操作系统中，妨碍了其他应用程序开发商的公平竞争。

### 1.1.3 操作系统的主要特征

尽管现代操作系统种类繁多，功能差别很大，但是它们仍然具有一些共同的特征。现代操作系统具有并发性、共享性、虚拟性和不确定性。

#### 1. 并发性(Concurrence)

并发性指的是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的系统，并发性是它最重要的特性。操作系统的并发性指的是计算机系统中同时存在若干个运行的程序，这些程序在执行时间上重叠。并发性能够消除计算系统中各个部件之间的相互等待，有效地改善了系统资源的利用率，提高了系统的吞吐量和系统效率。例如，一个程序等待 I/O 时，就会让出 CPU，操作系统调度另一个程序来占有 CPU 并运行，因此，在程序等待 I/O 时，CPU 便不会空闲，使得多个 I/O 设备可同时进行输入和输出，也使得 I/O 设备和 CPU 计算同时进行，这就是并发技术。

尽管操作系统的并发性能有效地改善资源的利用率，但是也会引发一系列的问题，使操作系统的设计和实现变得复杂化，如程序之间如何切换、协调等问题，所以，操作系统必须具有控制和管理各种并发活动的的能力，保证各程序的正确执行。在计算机系统中，并发实际上是一个物理 CPU 在若干个程序之间的多路复用，它与并行性(Parallelism)不同。并行性是指两个或两个以上的事件或活动在同一时刻发生。可见，并行的事件或活动一定是并发的，但是，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。实现并发性的关键技术之一是如何对系统内的多个程序进行切换，这将涉及到进程调度的问题。

#### 2. 共享性(Sharing)

共享性是现代操作系统的另一个重要特征。共享指的是系统中的硬件和软件资源不再为某个程序所独占，而是供多个用户共同使用。资源共享的方式有两种：一种是互斥访问，系统中的某些资源，如打印机、磁带机等，它们虽然可以提供给多个程序使用，但在同一时间段内却只允许一个程序访问这些资源，即要求互相排斥地使用这些资源；另一种是同时访问，计算机系统中有些资源，允许多个程序在同一时间内对它们进行访问。典型的可同时访问的设备是磁盘，各种可重入程序也可被同时访问。

并发和共享是现代操作系统的两个最基本的特征，这两者之间又是互为存在条件的。资源共享以程序的并发为前提，若系统不允许程序并发执行，自然不存在资源共享的问题。若系统不能对资源共享实施有效的管理，也必将影响到程序的并发执行，甚至根本无法并发执行。

#### 3. 虚拟性(Virtual)

虚拟性是操作系统采用的一种管理技术，它把一个物理上的实体变为若干个逻辑上的对应物，或者把物理上的多个实体变为逻辑上的一个对应物的技术。显然，物理实体(前者)

是实的，而后者是虚拟的。采用虚拟技术的目的是为了向用户提供一个易于使用的、高效的操作环境。现代操作系统使用了多种虚拟技术，分别用来实现虚拟处理器、虚拟内存、虚拟外部设备和虚拟信道等。

#### 4. 异步性(Asynchronism)

在多道程序环境下，允许多个进程并发执行，但是，由于竞争资源等因素的限制，使进程的执行不是“一气呵成”，而是以“走走停停”的方式运行。也就是说，在多道程序环境下，程序的执行是以异步方式进行的。每个程序在何时执行，多个程序间的执行顺序以及完成每道程序所需的时间都是不确定和不可预知的。异步性(也称为随机性)是现代操作系统的重要特征。操作系统运行在一个随机的环境中，但这并不能说明操作系统不能很好地控制资源的使用和程序的运行，而是强调操作系统的设计和实现需要考虑各种可能性，以便稳定、高效、可靠、安全地达到程序并发和资源共享的目的。

### 1.1.4 操作系统的功能

操作系统的主要任务是为多道程序的运行提供良好的运行环境，保证多道程序的高效运行，提高资源的利用率和方便用户的使用。为实现上述任务，现代操作系统应具有以下功能。

#### 1. 处理器管理

处理器是计算系统中最重要的资源，各种程序最终都要在处理器上执行，因此，必须尽可能地提高处理器的利用率。为了提高处理器的利用率，现代操作系统采用了多道程序设计技术。如果一个程序因等待某一条件而不能继续运行时，就把处理权交给另一个可以运行的程序。或者当一个比当前运行程序更重要的程序到达时，它可以抢占当前程序所占用的 CPU。为了描述多道程序的并发执行，操作系统引入进程或线程的概念来描述程序的动态执行过程。处理器的分配和调度都是以进程或线程为基本单位，因此，处理器的管理可归结为对进程或线程的管理。操作系统负责下列进程管理的活动：

- 创建或删除用户进程和系统进程
- 暂停或重启进程
- 提供进程同步机制
- 提供进程通信机制
- 提供死锁处理机制

#### 2. 存储管理

内存是现代计算机系统的中心，是可以被 CPU 和 I/O 设备共同访问的数据仓库。内存通常是 CPU 直接寻址和访问的唯一的容量存储器。例如，如果 CPU 需要处理磁盘中的数据，那么，这些数据必须首先通过 CPU 产生的 I/O 调用传送到内存中。同样，如果 CPU 需要执行指令，这些指令必须在内存中。如果需要执行一个程序，必须先将该程序映射成

绝对地址并装入内存。随着程序的执行，进程可以通过产生绝对地址来访问内存中的程序指令和数据。最后，程序终止，其内存空间得以释放，接着下一个程序可以被装入并执行。

为了改善 CPU 的利用率和计算机对用户的响应速度，必须在内存中保留多个程序。内存管理方法很多，不同算法的效能和特定环境有关。对于某一特定系统的内存管理方法的选择取决于多种因素，尤其是系统的硬件设计。每个算法都要求有特定的硬件的支持。操作系统负责下列内存管理的活动：

- 记录内存的哪些部分正在被使用和被谁使用。
- 当内存空间可用时，决定哪些进程可以装入内存。
- 根据需要分配和释放内存空间。
- 确保多道程序环境下，各个程序的运行只在自己的内存空间中运行，互不干扰。
- 当内存空间不足时，采取何种策略去扩展逻辑内存。

### 3. 设备管理

计算机系统及设备有着不同的物理特性，为了方便用户使用设备，操作系统提供设备管理功能来隐藏特定设备的硬件特质。操作系统的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各种设备的驱动程序和中断程序。因此，设备管理应具有以下功能：提供外围设备的控制与处理；提供缓冲区的管理；提供设备独立性；外围设备的分配和驱动调度；实现虚拟设备。

### 4. 文件管理

文件管理是操作系统最常见的功能。计算机可以在多种物理媒介上存储信息。磁带、磁盘和光盘是最常用的媒介，这些媒介有着不同的特点和物理组织。为了方便用户使用计算机系统，操作系统对存储设备的物理属性进行了抽象，定义了逻辑存储单元，即文件。操作系统将文件映射到这些媒介上，并通过这些存储设备访问文件。文件是由创建者定义的一组相关信息的集合。文件通常可以用来表示程序和数据，可以是有严格格式的，也可以是没有格式的。为了实现文件的管理，操作系统必须提供文件的存储、检索和修改等功能，解决文件的共享、保密和保护等问题，以便用户能方便、高效、安全地访问文件。现代操作系统一般都提供了很强的文件系统。

通常，操作系统负责下列与文件管理相关的功能：

- 创建或删除文件。
- 创建或删除目录。
- 提供操作文件和目录的原语。
- 将文件映射到辅存上。
- 在稳定的存储媒介上备份文件。

### 5. 用户接口管理

为了方便用户灵活、方便地使用计算机系统，操作系统提供了一组友好的用户接口。

通过这些接口,用户能够方便地调用操作系统提供的功能,有效地组织任务及其工作和处理流程,并使整个系统能够高效的运行。操作系统提供的接口有两大类:命令接口和程序接口。程序接口是为了在用户程序执行中访问系统资源而设置的,是用户程序取得操作系统资源的唯一途径。它由一组系统调用组成。命令接口可分为基于文本的接口(通称称为 shell)和基于图形的接口(通常称为 GUI)两种。用户通过命令接口可以实现与操作系统的交互。

## 6. 其他功能

随着计算机技术和网络的广泛应用,操作系统的功能得到了进一步的加强。除了传统的操作系统的功能之外,操作系统还必须提供其他新的功能来满足计算机发展的需要,这些新功能主要包括系统安全和网络通信。

### (1) 系统安全

操作系统的安全直接关系到信息自身的安全。安全已成为衡量操作系统性能的重要的方面。通常情况下,系统安全有两个方面的含义:一方面指的是安全,即保存在系统中的数据或信息不会被未经授权的任何单位和个人进行操作、复制或修改;另一方面指的是防护机制,即操作系统应该为用户提供一套信息防护手段。这套防护系统能够按照不同信息系统的防护要求,采取相应的机制来防止一定强度的入侵攻击。

### (2) 网络通信

在计算机网络时代里,网络上任何计算机的应用都离不开网络通信,所以,现代操作系统都注重为用户提供可靠、快捷的网络通信功能。这里所说的网络通信功能主要是指操作系统为应用提供必要的网络协议栈,即一组网络通信所必需的通信程序,常用的网络协议是 TCP/IP 协议(Transmission Control Protocol/Internet Protocol)。通信协议的实现涉及到操作系统内部的多种功能,将网络通信协议作为操作系统的一项功能一并设计,有利于提高网络通信的效率和可靠性。

通过前面的介绍,读者可了解操作系统的特点、功能和作用。为了更好地理解操作系统的概念和识别操作系统的共性,以下将简单介绍操作系统的发展历史。

## 1.2 操作系统的形成和发展

操作系统的形成和发展是随着计算机体系结构的变化而来的。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,因为与操作系统相关的许多基本概念都是在操作系统的发展过程中出现并逐步发展和成熟的。了解操作系统的发展历史,有助于更深刻地认识操作系统基本概念的内在含义。通过对操作系统发展历程的回顾,我们将会发现,在操作系统中,新的需求将导致引入新的硬件功能。



## 1.2.1 人工操作阶段

从计算机诞生(1946年)到20世纪50年代中期的计算机属于第一代计算机。这时的计算机体积庞大、速度慢、没有操作系统,由用户(程序员)采用手工方式直接控制和使用计算机硬件,即由用户(程序员)将事先准备好的程序和数据穿孔在纸带或卡片上,然后将这些纸带或卡片装入纸带输入机或卡片输入机,接着启动它们将程序和数据输入到计算机中,随后启动计算机并运行程序。当程序运行结束并取走结果后,才让另一个用户使用计算机。这种人工操作方式有以下缺陷:用户上机时独占全机资源,造成资源利用率不高,系统效率低;手工操作多,浪费处理器时间,也极易发生差错;数据的输入、程序的执行、结果的输出均联机进行,从上机到下机的时间很长。

这种人工操作方式在慢速的计算机上还能容忍,但是,随着计算机速度的提高,其缺点就暴露出来了。例如,一个作业在每秒1万次的计算机上执行时需要运行1个小时,作业的建立和人工干预花了3分钟,那么,手工操作的时间占总运行时间的5%。当计算机的速度提高到每秒钟10万次时,作业运行仅需要6分钟,而手工操作的时间仍为3分钟,这时手工操作时间占了总运行时间的50%。由此可见,缩短手工操作和人工干预的时间是十分必要的。另外,随着CPU速度的迅速提高,而I/O设备速度却难以提高,导致了CPU和I/O设备之间的矛盾越来越突出,操作系统必须妥善解决这些问题。

## 1.2.2 管理程序阶段——操作系统的雏形

20世纪50年代,随着晶体管计算机的广泛应用和计算机高级语言(FORTRAN、ALGOL、COBOL等)的出现,此时的用户可以采用高级语言编写程序来控制计算机的执行。用户需要完成某一任务,首先将程序写到纸上,然后将纸穿成卡片,再将卡片带到计算中心交给操作员。计算机运行完当前的任务之后,其结果由打印机输出。接着操作员从卡片盒中选择另一个任务(作业)交给计算机执行。在这个阶段,用户提交的任务是在操作员的干预下成批执行的。

由于处理机的速度与手工操作设备的输入和输出的速度不相匹配,人们设计了监督程序(或管理程序)来实现任务的自动转换处理。这期间,每个任务由程序员提供一组在某种介质(如纸、磁盘)上的任务信息(文件),包括任务说明书以及相关的程序和数据。任务说明书由程序员提交给系统操作员,操作员集中一批用户提交的作业,由管理程序将这批作业从纸带或卡片机输入到磁带上,每当一批作业输入完成后,管理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,作业又把控制权交回管理程序,管理程序再调入磁带上的第二个作业到内存中执行,以此类推,直到所有的作业完成。这种处理方式称为“批处理方式”。由于是串行操作,所以又称为单道批处理。