

高等院校计算机专业教材

操作系统原理教程

黄珍生◎主 编

黄勇 杨世瀚 叶嘉◎副主编

- ◎操作系统概述
- ◎作业、进程与线程
- ◎进程同步与通信
- ◎调度与死锁
- ◎存储管理
- ◎设备管理
- ◎文件管理
- ◎系统安全性



知识产权出版社

全国百佳图书出版单位

高等院校计算机专业教材

操作系统原理教程

黄珍生◎主 编

黄勇 杨世瀚 叶嘉◎副主编

- ◎操作系统概述
- ◎作业、进程与线程
- ◎进程同步与通信
- ◎调度与死锁
- ◎存储管理
- ◎设备管理
- ◎文件管理
- ◎系统安全性



知识产权出版社

全国百佳图书出版单位

图书在版编目 (CIP) 数据

操作系统原理教程 / 黄珍生主编. —北京: 知识产权出版社, 2016. 9
ISBN 978-7-5130-4007-5

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

中国版本图书馆 CIP 数据核字 (2016) 第 005422 号

内容简介

本教材理论与实践相结合。全面介绍了现代操作系统的基本理论和知识。全书共分 8 章。第 1 章介绍操作系统概论; 第 2 章介绍作业、进程与线程管理; 第 3 章介绍进程同步与通信原理; 第 4 章介绍调度与死锁; 第 5 章介绍存储管理; 第 6 章介绍设备管理; 第 7 章介绍文件管理; 第 8 章介绍操作系统安全性。

本书可作为大学本科及专科计算机专业教材或考研参考书, 也可作为计算机工作者的自学用书。为了便于教师讲授好本课程和学生深入理解本课程, 本书还配套有《操作系统实训教程与习题解答》指导书, 并为教师配套提供电子教案, 提供全部调试通过的源程序下载。

责任编辑: 彭喜英 徐家春

操作系统原理教程 CAOZUO XITONG YUANLI JIAOCHENG

黄珍生 主编

黄勇 杨世瀚 叶嘉 副主编

出版发行: 知识产权出版社有限责任公司

电 话: 010-82004826

社 址: 北京市海淀区西外太平庄 55 号

责编电话: 010-82000860 转 8539/8573

发行电话: 010-82000860 转 8101/8102

印 刷: 三河市国英印务有限公司

开 本: 787mm×1092mm 1/16

版 次: 2016 年 9 月第 1 版

字 数: 456 千字

网 址: <http://www.ipph.cn>

<http://www.laichushu.com>

邮 编: 100081

责编邮箱: xujiachun625@163.com

发行传真: 010-82000893/82003279

经 销: 各大网上书店、新华书店及相关专业书店

印 张: 18.75

印 次: 2016 年 9 月第 1 次印刷

定 价: 49.00 元

ISBN 978-7-5130-4007-5

版权所有 侵权必究

如有印装质量问题, 本社负责调换。

前 言

操作系统是计算机系统中不可缺少的系统软件，在计算机专业的课程体系中占有重要的地位，是计算机专业及相关专业的一门必修课程，也是计算机专业工作者必须掌握的知识。因此，一本适用的教材对操作系统课程的学习显得特别重要。

笔者根据多年教学和科研的经验与体会，并汲取国内外操作系统方面优秀教材的精华，本着提高学生素质、培养创新意识的精神，遵循本科教学大纲的要求，兼顾考研的需要，力求做到概念清晰、表述准确、结构合理、取舍得当、由浅入深、循序渐进、通俗易懂、便于自学，以期达到较好的教学效果。

目前，用于普通高校本科生的操作系统教材较多，各有特色。本教材采用理论与实际相结合的方式，讲述传统和现代理论，以求达到学以致用目的。具体来说，本教材注意了以下4个方面。

(1) 传统经典理论与现代最新技术相结合，便于既在学习上循序渐进，不割断历史，又着眼于未来发展。考虑到与今后研究生课程的衔接及实际工作中的应用，既介绍作业、进程等方面的传统理论，又讲解线程、分布式操作系统、系统安全与保护等最新概念和技术。

(2) 理论讲解与例题讲解相结合，力求用简单浅显的例子或例题讲解抽象、深奥的理论，以期学生能更快、更容易理解操作系统的原理和管理思想。

(3) 体现“以学生和学习为中心”的新理念。内容表述上力求用最简洁的语言、最简单的语法、图文和表格相结合的方式，在讲授方法上注意由浅入深，由表及里，先引出问题，再给出概念、实现技术和典型算法，然后予以适当总结，以便学生自学。

(4) 理论与实践相结合。本书有配套的《操作系统实训教程与习题解答》教材，学生通过实验可以更直观地理解操作系统的地位、功能、管理思想与算法实现等内容，通过提供的详细习题解答可进一步巩固所学的知识。

考虑到本课程的课时数有限，我们对内容进行了精选。全书共分8章：第1章操作系统概论，介绍了操作系统的基本概念、操作系统的形成与发展、操作系统的功能和分类、操作系统的特征和作用；第2章介绍了作业、进程与线程的概念及其演进，引入作业、进程与线程的管理思想，现代操作系统引入线程的好处等；第3章介绍进程同步与通信原理，对进程控制、进程互斥与同步、进程通信等问题进行了分析和讨论，着重介绍了信号量机制；第4章介绍调度与死锁，着重讨论了批处理作业调度算法，探讨系统发生死锁的原因以及死锁预防、死锁避免、死锁检测与恢复的方法，着重介绍死锁避免算法的实现；第5章存储管理，介绍了存储管理的基本知识，讨论了存储管理的基本功能及各种存储管理技术和虚拟存储管

理技术；第6章设备管理，介绍了I/O系统结构、I/O控制方式和I/O缓冲技术，对设备的分配、磁盘存储的管理、驱动调度及算法、虚拟设备技术进行了讨论；第7章文件管理，介绍了文件及文件系统的概念，对文件目录、文件组织、文件存储、文件操作、文件保护与保密等问题进行了分析和讨论；第8章介绍操作系统安全技术，安全的威胁因素和安全策略等。

本书由黄珍生任主编，黄勇、杨世瀚、叶嘉任副主编。由于编者水平有限，书中难免会有错误和不当之处，恳请读者批评指正。为便于广大教师授课，本书将配套提供电子教案。

联系方式：zhenshengh@163.com。

编者

2016年8月于广西民族大学

目 录

第 1 章 操作系统概述	1
1.1 什么是操作系统	1
1.1.1 操作系统作为扩展机器	2
1.1.2 操作系统作为计算机系统资源的管理者	3
1.1.3 操作系统作为用户与计算机之间的接口	4
1.2 操作系统的目标和功能	4
1.2.1 操作系统的目标	4
1.2.2 操作系统的功能	5
1.3 操作系统的产生、完善与发展	8
1.3.1 操作系统的产生	8
1.3.2 操作系统的完善	12
1.3.3 操作系统的发展	21
1.4 操作系统的分类	22
1.4.1 嵌入式操作系统	22
1.4.2 单用户操作系统	23
1.4.3 网络操作系统	23
1.4.4 分布式操作系统	24
1.4.5 多处理器操作系统	24
1.5 操作系统的特性	24
1.5.1 并发性	25
1.5.2 共享性	25
1.5.3 异步性	25
1.5.4 虚拟性	25
1.6 操作系统的硬件环境	26
1.6.1 特权指令	26
1.6.2 系统调用	28
1.6.3 内存保护	28
1.6.4 中断机制	29
第 2 章 作业、进程与线程	32
2.1 作业与作业管理	32

2.1.1	作业的定义	32
2.1.2	作业的分类	33
2.1.3	批处理作业控制	34
2.1.4	交互式作业控制	37
2.2	进程与进程管理	39
2.2.1	进程的引入	39
2.2.2	进程状态模型	44
2.2.3	进程描述	48
2.2.4	进程控制块的组织	51
2.2.5	进程控制	53
2.3	线程与线程管理	58
2.3.1	线程的引入	58
2.3.2	线程的实现方式	62
第3章	进程同步与通信	69
3.1	进程同步与互斥	69
3.1.1	并发原理	69
3.1.2	互斥与同步的概念	71
3.1.3	临界资源与临界区	73
3.1.4	互斥实现的硬件方法	74
3.1.5	互斥实现的软件方法	77
3.1.6	信号量和PV操作	80
3.2	经典的进程同步与互斥问题	88
3.2.1	有限缓冲区的生产者-消费者问题	88
3.2.2	哲学家就餐问题	90
3.2.3	读者-写者问题	94
3.2.4	理发师睡觉问题	97
3.3	管程	100
3.3.1	管程的概念	100
3.3.2	管程解决生产者-消费者问题	103
3.3.3	管程解决哲学家进餐问题	104
3.4	进程通信	106
3.4.1	进程通信的方式	106
3.4.2	消息缓冲通信	107
3.4.3	信箱通信	108
3.4.4	共享文件通信	109
3.4.5	消息传递系统的若干问题	109

第 4 章 调度与死锁	114
4.1 调度类型与准则	114
4.1.1 调度类型	114
4.1.2 进程调度方式	116
4.1.3 进程调度时机	116
4.1.4 调度性能的评价准则	117
4.2 调度算法	118
4.2.1 先来先服务调度算法	118
4.2.2 短作业(进程)优先调度算法	120
4.2.3 最短剩余时间优先调度算法	121
4.2.4 响应比高者优先调度算法	122
4.2.5 时间片抢转调度算法	123
4.2.6 优先级调度算法	126
4.2.7 多级反馈队列调度算法	127
4.3 实时调度	128
4.3.1 实现实时调度的基本条件	128
4.3.2 实时调度算法的分类	130
4.3.3 实时调度算法	131
4.4 产生死锁的原因和必要条件	133
4.4.1 死锁定义	133
4.4.2 死锁产生的原因	134
4.4.3 死锁产生的必要条件	136
4.4.4 判断死锁的方法	137
4.4.5 死锁处理的基本方法	138
4.5 死锁的预防与避免	138
4.5.1 死锁预防	138
4.5.2 死锁避免	140
4.5.3 银行家算法及其实现	142
4.6 死锁的检测与解除	149
4.6.1 死锁检测	149
4.6.2 死锁解除	152
4.6.3 鸵鸟算法	154
第 5 章 存储管理	158
5.1 单道程序存储管理	159
5.2 分区存储管理	160

5.2.1	固定分区存储管理	161
5.2.2	可变分区存储管理	163
5.2.3	可变分区存储管理的实现	164
5.2.4	重定位和存储保护	173
5.2.5	伙伴系统	178
5.3	页式和段式存储管理	180
5.3.1	页式存储管理	180
5.3.2	段式存储管理	187
5.3.3	页式管理与段式管理的比较	191
5.3.4	段页式存储管理	191
5.4	覆盖技术与交换技术	193
5.4.1	覆盖技术	194
5.4.2	交换技术	195
5.5	虚拟存储技术	196
5.5.1	程序局部性原理	197
5.5.2	虚拟存储技术原理	197
5.5.3	虚拟页式存储管理	198
5.5.4	页面置换算法	202
5.5.5	缺页中断率	210
5.5.6	页表结构	213
5.5.7	虚拟段式存储管理	217
5.5.8	虚拟段页式存储管理	218
第6章	设备管理	222
6.1	I/O 硬件	222
6.1.1	I/O 设备	222
6.1.2	轮询和中断	224
6.1.3	直接内存访问	227
6.2	I/O 软件	229
6.2.1	I/O 软件结构的层次模型	229
6.2.2	I/O 应用程序接口	230
6.2.3	块设备与字符设备	231
6.2.4	网络设备	232
6.2.5	时钟与定时器	232
6.2.6	阻塞与非阻塞 I/O	233
6.2.7	I/O 向量	234

6.3 I/O 内核子系统	234
6.3.1 I/O 调度	234
6.3.2 缓冲	235
6.3.3 高速缓存	236
6.3.4 信息暂存与设备预留	237
6.3.5 错误处理	237
6.3.6 I/O 保护	237
6.3.7 内核数据结构	238
6.3.8 把 I/O 操作转换成硬件操作	239
6.4 I/O 性能	241
6.5 小结	243
第 7 章 文件管理	246
7.1 文件系统概述	246
7.1.1 文件及其分类	246
7.1.2 文件系统的功能	247
7.2 文件的逻辑组织	248
7.2.1 顺序文件	248
7.2.2 索引文件	248
7.2.3 索引顺序文件	249
7.2.4 哈希文件	249
7.3 文件的物理组织	250
7.3.1 连续分配	251
7.3.2 链接分配	251
7.3.3 索引文件	253
7.4 目录管理	255
7.4.1 基本概念	255
7.4.2 文件目录结构	257
7.4.3 目录查询	258
7.5 文件存储空间的管理	259
7.5.1 位示图法	259
7.5.2 空闲表法	259
7.5.3 空闲链表法	260
7.5.4 成组链接法	260
7.6 文件共享和保护	261
7.6.1 文件共享	261
7.6.2 文件保护	263

第 8 章 系统安全性	269
8.1 系统安全的基本概念	269
8.1.1 系统安全性质描述	269
8.1.2 系统安全的威胁	270
8.1.3 安全模型	272
8.1.4 信息技术安全评估标准	274
8.2 数据加密技术	275
8.2.1 密码学基本概念	275
8.2.2 对称加密算法和非对称加密算法	278
8.2.3 数字签名和数学证书	280
8.3 常见攻击手段	281
8.3.1 缓冲区溢出概念简介	281
8.3.2 缓冲区溢出常用攻击方式	282
8.3.3 Windows 堆栈溢出利用的七种方式	283
8.4 操作系统安全	284
8.4.1 操作系统安全的基本概念	284
8.4.2 Windows 操作系统安全机制	285
8.4.3 Windows7 安全框架	287
参考文献	290

第 1 章 操作系统概述

当今，计算机已走进千家万户，相信你对操作系统并不陌生。你或许组装过计算机，当你把计算机硬件安装完成后，计算机是不能直接使用的，随后首先安装且必须安装的是操作系统。你或许使用过计算机，当你按下电源按钮启动计算机时，计算机首先启动的是操作系统，启动完成后，在计算机屏幕上出现的是操作系统为我们提供的使用计算机的平台——桌面。目前大家所熟知的一些操作系统的名称，如 Windows 2000，Unix，Linux，Windows 8，以及微软最新推出的 Windows 10 操作系统，这些都是微型计算机上通常安装和使用的一些操作系统。那么，操作系统是干什么的？为什么要有操作系统？操作系统是如何工作的？本书将由浅入深地为你讲解操作系统的管理思想、原理、技术和方法等方面的知识。

操作系统在计算机科学的发展过程中发挥了极其重要的作用，没有它，就没有计算机科学的普及和发展。从计算机最初的产生到今天，操作系统经历了从无到有、从简单到复杂、从低级到高级的过程，经历从最初的监控程序逐渐演变成目前可以并发执行多用户多任务的高级系统软件的过程，同时也产生了许多与操作系统相关的基本理论和核心技术。操作系统之所以能够不断地推陈出新，其动力之一就是人们总是会发现其正在使用的操作系统中的一些问题与不足，这就需要人们不断地加以研究和改进。由此也可以说，操作系统的研究和发展过程也是计算机科学与技术不断创新的过程。

1.1 什么是操作系统

操作系统在其发展历史中曾被称作监控（督）程序（系统）（Monitor）、执行系统（程序）（Executive System）、控制系统（程序）（Control System）、管理程序（Supervisor/ Supervisory System）、任务管理程序（Task Manager）、核心（程序）（Kernel）等名称。这些名称目前基本已不再使用，其中有些甚至已有新的含义，现在只有一个统一的名称——操作系统（Operating System，OS）。由于计算机规模的不同，人们把计算机划分为巨型、大型、中型、小型和微型计算机，也产生了适应不同机型的操作系统，如微机上使用的操作系统有 Windows 2000、Windows 7、Windows 8 等；IBM OS/390 是 S/390 大型机上使用的操作系统；我国在研制巨型计算机银河系列的同时，也研制了相应的银河麒麟服务器操作系统。

如何看待一个操作系统，人们从不同的角度出发就有不同的观点，通常有如下三种观点。

1.1.1 操作系统作为扩展机器

计算机系统分为硬件系统和软件系统。硬件系统从顶层看，一台计算机主要由处理器（运算器的控制器）、存储器和输入/输出设备部件和系统总线组成，各个部件通过总线互联，以实现计算机执行程序的主要功能，如图 1-1 所示。硬件各部件安装完成的机器也称为裸机。

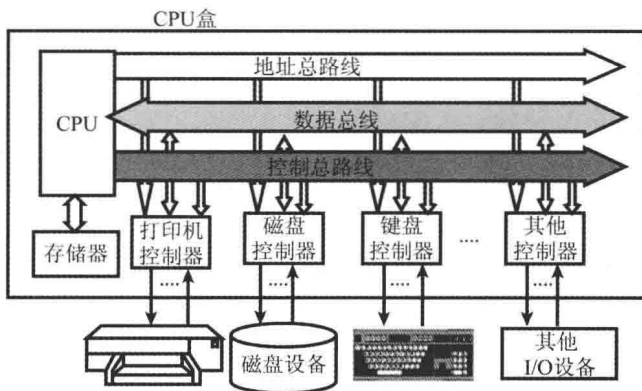


图 1-1 计算机硬件系统的体系结构

按照所起的作用和需要的运行环境不同，软件系统通常可分为三大类，即应用软件、支撑软件和系统软件。

应用软件是为解决某类应用需要或某个特定问题而设计的程序，如图形软件、财务软件、软件包等，这是范围很广的一类软件。

支撑软件是辅助技术人员从事软件开发和维护工作的软件，如各种开发工具、测试工具等，所以又称为工具软件，借以提高软件生产率，改善软件产品质量。

系统软件是指对计算机系统的资源进行控制、管理，并为用户使用和其他程序的运行提供服务的软件。包括操作系统、编译程序、汇编程序、连接装配程序、数据库管理系统、网络软件等。

计算机系统中硬件和软件是按层次结构组织的，如图 1-2 所示。操作系统是裸机之上的第 1 层软件，是最基本也最重要的系统软件。如果仅仅是裸机，则仅有机器语言可供用户利用，让用户直接针对内存、文件、外部设备等进行操作是相当麻烦的，这严重影响了工作效率和机器的利用率，甚至有些像现在常常应用的交互式命令操作都无法完成。

在裸机上加载操作系统，可使得一些硬件的细节及直接在硬件上的操作与用户或程序员无关。用户可利用由操作系统提供的各种命令或系统功能调用或进行程序设计来完成相应的任务。同时，加载了操作系统的机器又为软件开发人员创造开发、设计各种系统软件（或支撑软件）的环境，并在此基础上开发适合用户需求的各种应用软件，这进一步扩展了机器的功能，如在操作系统上安装 C 语言开发工具，程序员可以编制各种各样的解决不同实际问题的应用程序，让计算机为我们完成各种不同的工作；如安装办公软件，计算机可以协助完成日常的办公业务；安装财务管理软件，计算机可以协助完成财务管理工作，等等。这样使得原来简单的机器扩展为无所不能的、比真实的物理机器功能更强大、应用领域更广阔的机

器。因此，人们把经过软件扩充功能后的机器称为“虚拟机”。其中操作系统发挥了承上启下的作用。

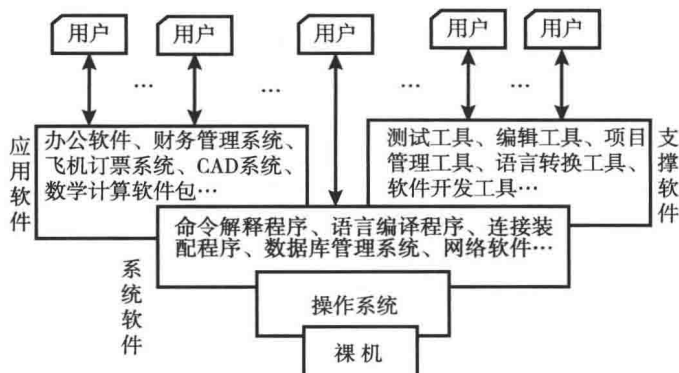


图 1-2 计算机系统层次，每层的扩充都为向上虚拟

1.1.2 操作系统作为计算机系统资源的管理者

计算机系统资源分为硬件资源和软件资源。这些资源按其性质来划分，大概可划分成四大类，即处理机、存储器、外围设备以及信息（程序和数据等）。这四类资源构成了操作系统本身和用户程序（或作业）赖以活动的物理基础和工作环境。如何管理这些资源，如何让这些资源得到合理、方便、高效、正确、安全的使用是操作系统必须解决的问题。它们的使用方法和策略决定了整个操作系统的规模、类型、功能和实现。因此，可以将操作系统看成一组资源管理子系统，它主要由处理机管理（调度）子系统、存储器管理子系统、外围设备管理子系统以及信息管理子系统（文件系统）这四个部分组成。随后的章节将对这些子系统加以分析，并研究其基本原理和相关实现技术。

现在，操作系统所管理的硬软资源越来越多，管理的范围也越来越宽，这不仅促进了操作系统的发展，而且也产生了对资源进行抽象研究的客观要求。尽管各种资源的性质各不相同，但是，从本质上考察，它们除了具有特性之外，又都具有共性，这就促使人们去研究资源的统一概念、资源的使用方法和策略，以便寻求一种管理资源的普遍原则和系统方法。总之，资源管理的目的在于为用户提供一种简单、有效的使用资源的方法，充分提高各种资源的利用率。为此，对每种资源管理来说，要研究如下几方面的内容。

①资源的使用状态，即使用、未使用、谁使用等问题。

②资源的分配原则和调度策略，即根据系统的设计目标，确定如何分配、分配给谁、何时分配、分配多少等问题。

③按具体要求执行分配，即根据②中的策略进行资源分配。

④资源回收，即在某些用户或程序不再需要某种资源时，系统应及时进行回收，以便进行再分配。

由此可见，计算机系统内的一切“活动”都需要一定的物理基础和执行的环境，这些都涉及系统内各种资源及其使用的状态，从这个意义上来说，资源管理的观点在众多观点中占有主导地位。

1.1.3 操作系统作为用户与计算机之间的接口

最初的计算机在没有操作系统的情况下，用户使用计算机时需要进行大量的手工操作。从程序和数据的装入，到程序的修改调试、结果的输出等计算机资源上的操作全部由人工指定好细节来完成。例如，选择何种输入设备，在什么地址上启动输入设备进行输入，在什么地址上启动相应的编译程序进行编译，编译的结果（目的代码）再重新在指定的输入设备上指定具体的启动地址启动运行等。

有了操作系统之后，原来由人工进行的许许多多烦琐而费时的操作就由操作系统来代替完成，这使得用户能够方便地使用计算机系统。现在的计算机系统在操作系统的支持下都配备了丰富的命令，给计算机发送操作命令可以通过键盘输入，也可以通过鼠标单击图标或菜单，这样方便了用户与计算机进行交互和操作。因此，也可以说现在的操作系统是人—机交互的界面与接口，用户与计算机进行的一切活动都要经过操作系统来实现。

人们还可以从其他角度来看待操作系统。如人们曾进行过各种形象化的描述和比喻：有的将操作系统描述为弥补用户与硬件双方的差距（Gap），有的将操作系统描述为硬件（裸机，机器语言）与用户之间的桥梁（Bridge）或中介（Intermediary），有的将操作系统描述为对系统资源使用进行管理的政府（Government）或环境（Environment），等等。这些比喻从各种不同角度强调了操作系统的方便服务性与统管权威性。这里就不再一一讨论了。

综上所述，以下几点有助于理解操作系统的定义。

①操作系统是硬件之上的第一层系统软件。操作系统是在计算机硬件之上首先加载的一组计算机程序。与其他计算机程序（包括用户程序、编辑程序等）类似，都需要在处理机上运行。主要区别在于程序的目的。操作系统用于控制和管理处理机使用其他系统资源，并控制处理机运行其他程序的时间。

②操作系统的基本职能是控制和管理计算机系统内各种资源，有效地组织作业在 CPU 上运行。例如，将用户以文件形式存在磁盘上的程序调入内存，为其分配内存资源，通过一定的策略分配 CPU 资源让其运行，然后根据选择分配输出设备资源，将运行结果显示在屏幕上或打印设备上，这些都由操作系统来完成。

③操作系统提供众多服务，方便用户使用，可扩充硬件功能。例如，Windows 操作系统通过提供图标、菜单命令展示各种各样可提供的服务，用户可以很方便地利用它们完成对文件输入/输出、程序运行、设备检测等多方面的控制、管理工作。

综上所述，我们把操作系统定义为：操作系统是能有效地控制和管理计算机系统内各种硬件和软件资源，合理地组织多道程序运行，并向用户提供各种服务功能，使用户能够灵活、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。

1.2 操作系统的目标和功能

1.2.1 操作系统的目标

在计算机的发展过程中，存在过许多类型的操作系统，有一些也是大家熟悉的，而不同

类型的操作系统的目标有所不同，有的追求通用性，面向多领域多用户，有的面向特殊领域，追求高、精、尖端技术的实现，追求产品精度、生产效率等，但总体来说，在计算机上加载操作系统的目标有以下几点。

1.2.1.1 方便性

加载了操作系统的计算机系统必须更容易操作、更方便使用。从利用计算机解决问题的角度来说，最简单的就是用户通过操作系统的某种接口形式（命令或鼠标等）只需输入几个命令，或单击几个图标便可完成。如人机界面方便，包括用户使用界面和程序设计接口要易用、易学、易维护；又如利用计算机上网查询各种信息，通过单击超链接很方便地链接到其他网址或媒体等，提供的这些方便也是通过操作系统的支持实现的。所有这些是目前的计算机用户都能体会到的。因此，具有操作系统的计算机系统对于用户是不可缺少和必要的。

1.2.1.2 有效性

加载了操作系统的计算机系统必须能提高系统运行效率、资源利用率，特别是处理器利用率。在没有加载操作系统的计算机系统中有许多资源，这些资源在很长的时间内由于没有一种机构来监管和调度而处于空闲状态，因而不能充分地得到利用，例如，CPU 和 I/O 设备（即输入/输出设备）的完全串行操作，都会出现 CPU 工作，而 I/O 闲置；或 I/O 工作，而 CPU 闲置的现象。这就是没有操作系统的计算机上单个用户上机常见的情形。另外的情形就是，内存与外存中所存放的数据没有统一的管理而带来的无序和空间的浪费。加载或安装了操作系统后，这些资源可以得到有效管理和充分利用。此外，通过操作系统合理地组织计算机的工作流程，也会提高资源的利用率和系统的吞吐率（单位时间内完成的工作数量），也相对缩短了用户“计算”的时间。

1.2.1.3 可扩展性

加载了操作系统的计算机系统必须容易地扩展计算机的功能和性能，如很容易地挂接各种应用软件、很容易地连接各种硬件设备。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得到了迅速发展，它们都对操作系统提出了更高的功能和性能上的要求。因此，在构造操作系统时，应该允许在不妨碍服务的前提下，有效地开发、测试和引进新的系统功能。由于硬件升级和新型硬件的出现带来了操作系统升级的例子就是如此，早期运行 Unix 和 IBM OS/2 的机器没有“分页”的硬件机制，因而也就没有使用页式存储管理机制，而在高性能芯片中出现了控制寄存器等新型硬件机制后，操作系统经过扩展就实现了页式存储管理机制。同样在 IBM PC 早期的 286 之前的机型由于没有相应的控制寄存器等硬件机制，因此那时的机器就没有页式管理机制。在新型的微处理器中出现了相应的控制寄存器等硬件机制后，也就有了新型的带有页式管理的操作系统。

操作系统不断升级和改造对其设计者提出了一定的要求。一个比较明确的观点就是，在设计和构造操作系统时应该采用模块化结构，并清楚地定义模块间的接口。

1.2.2 操作系统的功能

操作系统通过管理计算机资源来控制计算机的基本功能，因而从资源管理的角度来考察，操作系统要对计算机系统内的所有资源进行有效和自动的管理，要实现资源合理分配、高效使用。而用户通过提供简单的接口就可以使用计算机中的资源，因此，要求操作系统提

供各种各样的功能，以便用户使用方便。操作系统的基本功能主要分为如下五个方面。

1.2.2.1 存储器管理功能

存储器管理功能主要完成以下四个方面的任务。

①内存分配。内存分配的主要任务是为每道（用户）程序分配一定的内存空间。因此，操作系统必须记录整个内存的使用情况，处理用户提出的申请，按照某种策略实施分配，或回收用户释放的内存空间。

②地址映射。在多道程序系统中，所编写的程序在编译之后通常都是从 0 地址开始的，也称为逻辑地址或相对地址，而根据（1）中内存分配的地址通常不可能从 0 地址开始，内存中的 0 地址通常早被首先加载的操作系统占用。这样，就产生了一个问题——程序如何访问真正的内存物理单元，这就需要操作系统在硬件的支持下解决地址映射，即逻辑地址到内存物理地址的转换。

③内存保护。由于多道程序的问题带来了内存中各程序如何保证自己的空间不受“侵犯”的问题，特别是操作系统的内存空间更不能由用户程序进行“写”访问，因此需要操作系统在存储器管理上加以保护。具体如何保护的问题与系统所采用的存储管理机制有关。

④内存扩充。一台机器的物理空间通常是固定和有限的，而系统内各个用户是不关心（至少是不太关心）真正内存空间大小的，所关心的是系统工作的速度和时间，而要将若干用户程序全部同时放入内存中，内存可能无法容纳，因此就需要系统将物理空间通过虚拟存储器技术虚拟成比内存空间大得多的空间来满足实际运行的需要。

1.2.2.2 处理机管理功能

处理机管理功能的主要任务就是对处理机进行分配、运行控制和管理，包括以下几个方面。

①作业调度和进程调度。一个等待在后备队列上（外存空间）的用户程序通常要经过两级调度才得以在 CPU 上执行。首先是作业调度，它把选中的一批作业放入内存，并分配除 CPU 之外的其他必要的资源，为这些作业建立相应的进程，并让其进入就绪进程队列。此后由进程调度程序按一定的算法从就绪进程中选出一个合适进程，使之在 CPU 上运行。所谓的进程是系统中活动的实体，关于进程的进一步讨论将安排在进程管理章节。

②进程通信与控制。多个进程在活动过程中彼此会发生相互依赖或者相互制约的关系。相互合作的进程之间往往需要传递或交换信息，为保证系统中所有进程都能正常活动，就必须设置进程通信与控制机制，它分为同步方式和互斥方式。

1.2.2.3 设备管理功能

设备管理的主要任务是完成用户提出的 I/O 请求，包括以下几个方面。

①缓冲区管理。缓冲区管理的基本任务是管理各类 I/O 设备的数据缓冲区，解决 CPU 和外设速度不匹配的问题，从而使它们能充分并行工作，提高各自的利用率。

②设备分配。设备分配的基本任务是根据用户的 I/O 请求和相应的分配策略，为申请 I/O 操作的用户分配外围设备以及通道、控制器等。

③设备驱动。设备驱动的基本任务是实现用户提出的 I/O 操作请求，完成数据的输入/输出。这个过程是系统建立和维持的，它涉及相关的通道、控制器的控制。

④设备无关性。又称设备独立性，是指应用程序独立于实际使用的物理设备，所提出的