

新工科建设之路·计算机类专业规划教材



操作系统原理



胡元义 黑新宏 主编
金海燕 杨凯峰 鲁晓锋 副主编



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

新工科建设之路·计算机类专业规划教材

操作系统原理

胡元义 黑新宏 主编

金海燕 杨凯峰 鲁晓锋 副主编

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

操作系统主要涉及对计算机软、硬件资源的控制和管理。本书对操作系统的实现原理进行了详细和深入的分析,力求做到全面、系统、准确和通俗,以便读者透彻理解操作系统的设计思想,深化对基本概念的掌握。全书共分为 6 章,主要包括:引论、处理器管理、进程同步与通信、存储管理、设备管理和文件管理。

本书结构清晰、内容丰富、取材新颖,既强调知识的实用性,又注重理论的完整性,可作为高等院校计算机及相关专业的操作系统课程教材,也可作为从事计算机工作及报考研究生人员的参考资料。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

操作系统原理 / 胡元义, 黑新宏主编. —北京: 电子工业出版社, 2018.8

ISBN 978-7-121-34038-3

I. ①操… II. ①胡… ②黑… 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 (2018) 第 074475 号

策划编辑: 章海涛

责任编辑: 章海涛

文字编辑: 孟 宇

印 刷: 涿州市京南印刷厂

装 订: 涿州市京南印刷厂

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编: 100036

开 本: 787×1092 1/16 印张: 20 字数: 485 千字

版 次: 2018 年 8 月第 1 版

印 次: 2018 年 8 月第 1 次印刷

定 价: 48.00 元



凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888, 88258888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式: mengyu@phei.com.cn。

前 言

操作系统为计算机的使用提供了一个方便灵活、安全可靠的环境，特别是 Windows 操作系统的出现，使得使用计算机只需单击鼠标。

操作系统是计算机专业的一门核心课程，在计算机本科教学中占有十分重要的地位。操作系统主要涉及对计算机软、硬件资源的控制和管理，其理论性强，内容抽象。特别是进程管理，需要通过缜密、细致的逻辑思维来想象微观世界中处理器的调度与运行，这种抽象、复杂的内容不易理解，也难以掌握。本书对操作系统的实现原理进行了详细和深入的分析，力求做到全面、系统、准确和通俗，以便读者透彻理解操作系统的设计思想，深化对基本概念的掌握。

操作系统是现有软件系统中最复杂的软件系统之一，代码多达几亿条。著名计算机科学家 P. Denning 和他的助手及同事们认为，在操作系统方面取得了进程、内存管理、信息的保护与安全性、调度与资源管理以及系统结构等 5 项主要的成就。现今，信息的保护与安全性已作为一门独立的课程来讲授，即计算机信息安全；而系统结构的部分内容已经转化为硬件的内容，在计算机系统结构课程中讲授。本书围绕操作系统取得的主要成就（除上述已经独立设课的内容外），从原理出发，详细介绍了操作系统有关内容，注重操作系统理论的发展与传承，注意知识的连贯性与拓展性，并通过精选的示例和图例来帮助读者理解与掌握操作系统知识，同时还设计了一些不同于其他操作系统教材的算法，如睡眠理发师问题、吃水果问题、读者-写者问题等 P、V 操作和管程实现算法，对深入理解操作系统原理提供了更好的帮助。

全书共分 6 章：第 1 章引论，主要介绍操作系统的基本概念、操作系统的发展过程、操作系统的分类、操作系统运行的硬件环境以及操作系统与用户的接口；第 2 章处理器管理，主要介绍进程的引入和描述、进程状态及转换、进程调度及调度算法、进程控制和线程；第 3 章进程同步与通信，主要介绍进程同步与互斥、临界区的使用、实现进程同步与互斥的工具（P、V 操作和管程）、进程通信和进程死锁；第 4 章存储管理，主要介绍存储管理的基本概念和功能、各种存储管理技术、虚拟存储的思想及实现方法；第 5 章设备管理，主要介绍 I/O 设备的硬件结构和软件组成、I/O 设备控制方式、设备管理使用的有关技术及设备分配；第 6 章文件管理，主要介绍文件的概念、文件的逻辑结构和物理结构、文件目录、文件存储空间的组织和管理、文件的共享和保护。

本书结构清晰、内容丰富、取材新颖，既强调知识的实用性，又注重理论的完整性。本书是编者 30 年在操作系统方面取得的教学实践成果的总结，同时也汲取了国内外优秀操作系统教材的精华。此外，本书配套了辅助教材《操作系统原理习题解析与上机指导》

(ISBN 978-7-121-34037-6), 在学习中配合使用将会得到更好的效果。本书还配有教学用的电子教案, 以供教学与学习使用。本书可作为高等院校计算机及相关专业的操作系统课程教材, 也可作为从事计算机工作及报考研究生人员的参考资料。

限于编者水平, 书中难免有疏漏之处, 敬请读者赐教。

本书的相关教学资源, 可以从华信教育资源网站 (<http://www.hxedu.com.cn>) 下载。

编 者

2018 年 5 月

目 录

第 1 章 引论	1
1.1 操作系统的概念	1
1.1.1 什么是操作系统	1
1.1.2 操作系统的主要功能	4
1.1.3 操作系统的基本特征	6
1.2 操作系统的逻辑结构和运行模型	8
1.2.1 用户态和内核态的划分	8
1.2.2 操作系统的逻辑结构	9
1.2.3 操作系统的运行模型	11
1.3 操作系统的形成与发展	12
1.3.1 操作系统的形成时期	12
1.3.2 操作系统的成熟时期	15
1.3.3 操作系统的进一步发展时期	17
1.4 主要操作系统的类型	19
1.4.1 批处理操作系统	19
1.4.2 分时操作系统	20
1.4.3 实时操作系统	21
1.4.4 微机操作系统	23
1.4.5 网络操作系统	24
1.4.6 多 CPU 操作系统	25
1.4.7 分布式操作系统	27
1.4.8 嵌入式操作系统	29
1.5 操作系统安全性概述	30
1.5.1 操作系统安全的重要性	30
1.5.2 操作系统的安全观点	31
1.5.3 实现操作系统安全性的基本技术	31
1.6 操作系统运行基础	32
1.6.1 处理器及工作模式	32
1.6.2 中断技术	36
1.6.3 系统调用	40
习题 1	42
第 2 章 处理器管理	47
2.1 进程的概念	47

2.1.1	程序的顺序执行	47
2.1.2	程序的并发执行	48
2.1.3	进程	50
2.2	进程的状态及转换	53
2.2.1	两状态进程模型	53
2.2.2	进程的三态模型	53
2.2.3	进程的五态模型	55
2.2.4	进程的挂起	56
2.2.5	进程控制块	58
2.3	进程控制	60
2.3.1	进程切换	60
2.3.2	进程控制原语	62
2.4	处理器调度	65
2.4.1	作业与进程的关系	66
2.4.2	CPU 的三级调度	67
2.4.3	处理器调度队列模型	69
2.4.4	进程调度的方式和时机	71
2.5	单处理器调度算法	73
2.5.1	调度原则	73
2.5.2	常用调度算法	75
2.5.3	实时调度	80
2.6	线程	82
2.6.1	线程的引入	82
2.6.2	线程的概念	83
2.6.3	线程与传统进程的比较	84
2.6.4	线程实现原理	85
	习题 2	86
第 3 章	进程同步与通信	93
3.1	进程同步的基本概念	93
3.1.1	并发进程的关系	93
3.1.2	进程的互斥与同步	94
3.1.3	临界资源与临界区	95
3.2	进程互斥方法	97
3.2.1	实现进程互斥的硬件方法	97
3.2.2	实现进程互斥的软件方法	99
3.3	信号量机制	101
3.3.1	信号量	101
3.3.2	使用信号量实现进程互斥	103
3.3.3	使用信号量实现进程同步	105

3.4	经典互斥与同步问题	108
3.4.1	生产者-消费者问题	108
3.4.2	哲学家进餐问题	111
3.4.3	读者-写者问题	114
3.4.4	睡眠理发师问题	117
3.5	经典互斥与同步问题的应用	119
3.5.1	缓冲区数据传送问题	119
3.5.2	吃水果问题	120
3.5.3	汽车过桥问题	122
3.6	管程机制	123
3.6.1	条件变量与管程结构	123
3.6.2	生产者-消费者问题的管程解决方法	125
3.6.3	哲学家进餐问题的管程解决方法	127
3.6.4	读者-写者问题的管程解决方法	128
3.6.5	睡眠理发师问题的管程解决方法	129
3.6.6	吃水果问题的管程解决方法	131
3.7	进程通信	132
3.7.1	进程通信的概念	132
3.7.2	共享内存通信方式	133
3.7.3	消息缓冲通信方式	134
3.7.4	信箱通信方式	137
3.7.5	管道通信方式	139
3.8	死锁	140
3.8.1	产生死锁的原因和必要条件	141
3.8.2	死锁的预防	145
3.8.3	死锁的避免	146
3.8.4	死锁的检测与解除	152
	习题 3	155
第 4 章	存储管理	163
4.1	程序的链接和装入	163
4.1.1	逻辑地址和物理地址	164
4.1.2	程序链接	164
4.1.3	程序装入	165
4.2	存储器及存储管理的基本功能	168
4.2.1	多级存储器体系	169
4.2.2	内存的分配与回收以及地址转换	170
4.2.3	内存的共享、保护及扩充	171
4.3	分区式存储管理	173
4.3.1	单一连续分区存储管理	173

4.3.2	固定分区存储管理	175
4.3.3	可变分区存储管理	177
4.3.4	覆盖与交换技术	182
4.4	分页存储管理	184
4.4.1	分页存储管理的基本原理	184
4.4.2	分页存储管理的地址转换与存储保护	186
4.4.3	两级页表和多级页表	189
4.4.4	内存物理块的分配与回收	191
4.5	分段存储管理	192
4.5.1	分段存储管理的基本原理	192
4.5.2	分段存储管理的地址转换与存储保护	194
4.5.3	分段存储管理的优缺点	197
4.6	段页式存储管理	198
4.6.1	段页式存储管理的基本原理	198
4.6.2	段页式存储管理的地址转换与特点	199
4.7	虚拟存储管理	201
4.7.1	虚拟存储器的概念	201
4.7.2	请求分页存储管理	203
4.7.3	请求分段存储管理	212
4.7.4	请求段页式存储管理	216
习题 4		218
第 5 章	设备管理	225
5.1	设备管理概述	225
5.1.1	设备的分类	225
5.1.2	设备管理的目标和功能	226
5.2	I/O 设备管理系统的组成	228
5.2.1	I/O 系统的硬件组织	228
5.2.2	I/O 系统的软件组织	233
5.3	I/O 设备控制方式	238
5.3.1	程序直接 I/O 控制方式	238
5.3.2	程序中断 I/O 控制方式	239
5.3.3	直接存储器存取 I/O 控制方式	240
5.3.4	I/O 通道控制方式	242
5.3.5	4 种 I/O 控制方式的对比	245
5.4	缓冲技术与虚拟设备技术	246
5.4.1	缓冲技术	246
5.4.2	虚拟设备技术	250
5.5	设备的分配与回收	252
5.5.1	用于设备分配的数据结构	252

5.5.2	设备分配	254
5.5.3	设备回收	258
5.6	磁盘存储器管理	259
5.6.1	存储设备概述	259
5.6.2	磁盘调度	260
5.6.3	提高磁盘 I/O 速度的方法	264
5.6.4	磁盘阵列 (RAID)	266
习题 5		270
第 6 章	文件管理	274
6.1	文件系统基本概念	274
6.1.1	文件系统的引入	274
6.1.2	文件与文件系统	275
6.1.3	文件操作	278
6.2	文件的组织结构	280
6.2.1	文件的逻辑结构	280
6.2.2	文件的物理结构	282
6.3	文件目录	288
6.3.1	文件控制块与索引节点	289
6.3.2	目录结构	290
6.3.3	文件目录查找	292
6.4	文件存储空间管理	293
6.4.1	空闲分区表法	294
6.4.2	空闲块链法	294
6.4.3	位示图法	296
6.5	文件共享与文件安全	296
6.5.1	早期的文件共享方法	297
6.5.2	目前常用的文件共享方法	298
6.5.3	文件系统的安全	300
6.5.4	文件系统的可靠性	301
6.5.5	文件保护	303
习题 6		305
参考文献		309

引 论

计算机技术发展到今天,从个人计算机到巨型计算机,无一例外都配置了一种或多种操作系统(Operating System, OS)。计算机系统由硬件和软件两部分组成,操作系统是计算机系统最重要的系统软件,也是配置在计算机硬件上的第一层软件,是对硬件系统的首次扩充。它作为计算机硬件和计算机用户之间的中介,为应用程序提供使用基础,并成为整个计算机系统的控制中心。在现代计算机系统中,如果不安装操作系统,很难想象如何使用计算机。操作系统不仅将仅有硬件的裸机改造成为功能强、使用方便灵活、运行安全可靠的虚拟机,来为用户提供良好的使用环境,而且采用有效的方法组织多个用户共享计算机系统内的各种资源,最大限度地提高了系统资源的利用率。

1.1 操作系统的概念

1.1.1 什么是操作系统

1. 引子

计算机程序是如何运行的呢?首先,需要先进行编程,而编写程序是需要以计算机程序设计语言作为基础的。对大多数编写程序的人来说,使用的编程语言称为高级程序设计语言,如 C、C++、Java 等。但由于计算机并不识别用高级语言编写的程序,因此编写好的程序还需要将它编译成计算机能够识别的机器语言程序,而这需要编译程序或汇编程序的帮助才能完成。其次,编译好的机器语言程序需要加载(调入内存并将程序中的逻辑地址变成可以执行的内存物理地址)到内存中形成一个可执行的程序,即进程(见第 2 章),而这需要操作系统的帮助。进程需要在计算机芯片 CPU(中央处理器)上执行才是真正的执行,而将进程调度到 CPU 上运行也是由操作系统完成的。最后,在 CPU 上执行的机器语言指令需要变成能够在一个个时钟脉冲周期里执行的基本操作,这需要基本硬件指令系统及相关计算机硬件的支持,而整个程序的执行过程还需要操作系统提供服务和程序语言提供运行环境。这样,一个从程序到微指令执行的过程就完成了,图 1-1 给出了程序演变的全过程。

我们只是从一个线性角度来看待从编程开始到最终输出结果的演变,而没有考虑各种因素之间的穿插和交互过程。事实上,程序可以直接用机器语言或汇编语言这种称为“低级”的语言编写。用机器语言编写的程序无须经过编译程序的翻译就可以直接在 CPU 上运行,而用汇编语言编写的程序则还需经过汇编程序翻译后才能加载执行。

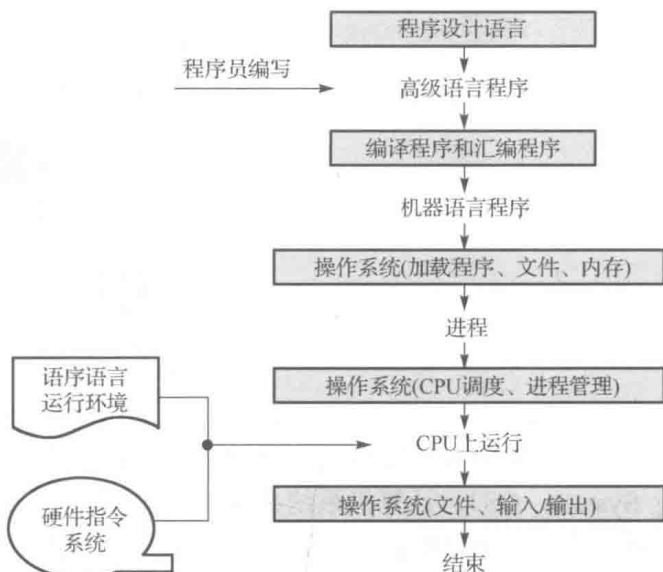


图 1-1 由程序到结果的演变

由图 1-1 的描述可以看出，程序的运行涉及四个方面：① 程序设计语言；② 编译程序；③ 操作系统；④ 硬件指令系统（计算机硬件系统）。而操作系统在程序的执行过程中具有关键作用。

2. 操作系统的定义

计算机系统由硬件系统和软件系统两大部分组成，硬件系统是计算机赖以工作的实体，软件系统则保证了计算机系统的硬件部分按用户指定的要求协调地工作。

计算机硬件系统由中央处理器（Central Processing Unit, CPU）、内存储器、外存储器和各种输入输出设备组成，它提供了基本的计算机资源。只有硬件的计算机称为裸机。

计算机硬件由软件来控制。按与硬件相关的密切程度，通常将计算机的软件分为系统软件和应用软件两类。用户直接使用的软件通常为应用软件，而应用软件一般需借助系统软件来指挥计算机的硬件完成其功能。

那么，操作系统是什么呢？英文中的 Operating System 意为掌控局势的一种系统，也就是说，计算机里的一切事情均由 Operating System 来掌控（管理）。现在我们面临两个问题：第一个问题，操作系统到底是什么东西？第二个问题，操作系统到底操控（管理）什么事情？

由图 1-1 可以大致得到第一个问题的答案：操作系统是介于计算机硬件和应用软件之间的一个软件系统，即操作系统的下面是硬件平台，而上面则是应用软件（如图 1-2 所示）。

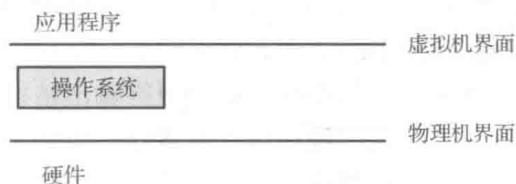


图 1-2 操作系统上下界面

下面分析第二个问题,操作系统掌控的事情当然是计算机上或计算机里发生的一切事情。最早的计算机并没有操作系统,而是直接由人来操控。随着计算机复杂性的增加,人们已经不能胜任直接掌控计算机了,于是编写出操作系统这个“软件”来掌控和管理计算机。这个掌控有着多层深远的意义。

(1) 由于计算机的功能不断趋向完善和复杂,操作系统所掌控的事情也就越来越多,越来越复杂,即操作系统必须掌控计算机的所有软硬件资源,并使计算机的工作变得有序。这是早期驱动操作系统不断改善的根本原因。

(2) 既然操作系统是专门掌控计算机的,那么计算机上发生的所有事情必须得到操作系统的允许,但由于所设计的操作系统不可能做到尽善尽美,因此给攻击者造成可乘之机,而操作系统设计人员和攻击者之间的博弈是当前驱动操作系统改善的一个重要动力。

(3) 为了更好地掌控计算机上发生的所有事情,同时也为了更好地满足人们对操作系统越来越苛刻的要求,操作系统必须不断地完善自己。

也就是说,从计算机管理的角度看,操作系统的引入是为了更加充分、有效地使用计算机系统资源,也就是合理地组织计算机的工作流程,有效地管理和分配计算机系统的硬件和软件资源,同时注意操作系统自身的安全与完善。从用户使用的角度看,操作系统的引入是为了给用户使用计算机提供一个良好的界面,使之既方便又安全。

因此,操作系统是掌控(管理)计算机上所有事情的系统软件,它需要完成以下5种功能。

- (1) 控制和管理计算机系统的所有硬件和软件资源。
- (2) 合理地组织计算机的工作流程,保证计算机资源的公平竞争和使用。
- (3) 方便用户使用计算机。
- (4) 防止对计算机资源的非法侵占和使用。
- (5) 保证操作系统自身的正常运转。

长期以来,(1)~(3)项一直是操作系统定义的内容,但随着越来越多对计算机系统恶意攻击的事件以及越来越多的计算机病毒出现,操作系统自身的安全越来越受到人们的重视。所以,现在的操作系统的定义又增加了(4)和(5)两项。

任何计算机,只有在安装了相应的操作系统后才构成一台可以使用的计算机系统,用户才能方便地使用计算机。只有在操作系统的支持下,计算机的各种资源才能安全、方便、合理地分配给用户使用,各种软件(编译程序、数据库程序、网络程序以及各种应用程序等)才能安全、高效、正常地运行。操作系统性能的高低直接决定了计算机整体硬件性能能否得到充分发挥。操作系统本身的安全性和可靠性在一定程度上决定了整个计算机系统的安全性和可靠性。操作系统在整个计算机系统中的地位如图1-3所示。

由图1-3可以看出,计算机系统具有层次结构,其中操作系统是在硬件基础上的第一层软件,是其他软件和硬件之间的接口。因此,操作系统是最重要的系统软件,它控制和协调各用户程序对硬件的使用。实质上,用户在使用计算机时直接面对的并不是计算机的硬件,而是应用软件,由应用软件在“幕后”与操作系统打交道,再由操作系统指挥计算机完成相应的工作。

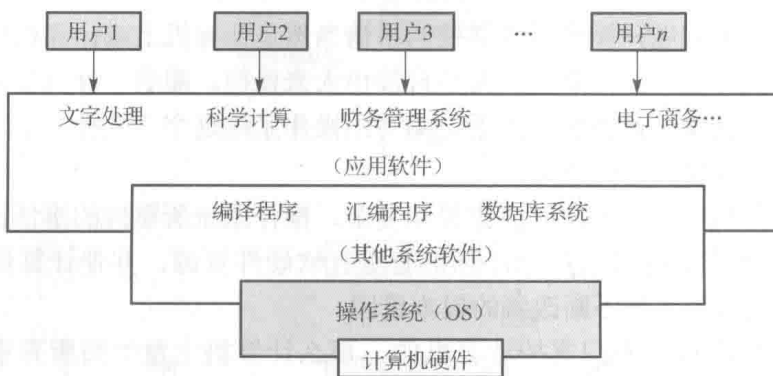


图 1-3 操作系统在计算机系统中的地位

3. 操作系统的设计目标

目前存在多种类型的操作系统，不同类型的操作系统其设计目标各有侧重。一般来说，设计的操作系统应达到以下 6 个目标。

(1) 方便性。操作系统应提供统一且界面友好的用户接口，以方便用户使用计算机。

(2) 有效性。操作系统应能有效分配和管理计算机的软、硬件资源，使系统资源得到充分利用。此外，操作系统应能合理地组织计算机的工作流程，改善系统性能，并提高系统运行效率。

(3) 可扩展性。操作系统应具备较好的可扩展性，以适应计算机硬件和计算机网络等发展的需要。可扩展性表现在能否很容易地增加新的模块。

(4) 开放性。操作系统应遵循国际标准进行设计，构造一个统一的开放环境，以便不同厂家生产的计算机和设备能通过网络集成，且能正确、有效地协调工作，实现应用程序的可移植性和互操作性。

(5) 可靠性。可靠性包括正确性、健壮性和安全性。操作系统除应满足正确性这一基本要求外，还应满足确保操作系统自身正常工作的健壮性，以及如何防止非法操作和入侵的安全性。

(6) 可移植性。可移植性是指将程序从一个计算机环境转移到另一个计算机环境中仍能正常运行的特性。由于操作系统开发是一项非常庞大的工程，为了避免重复工作及缩短软件研发周期，现在操作系统设计已将可移植性作为一个重要目标。

1.1.2 操作系统的主要功能

为了高效地使用计算机软、硬件资源，提高计算机系统资源的利用率和方便用户使用，在计算机系统都采用多道程序设计技术。多道程序设计技术是指内存中同时放入多道程序（进程）交替运行并共享系统资源，当一道程序（进程）由于某种原因（如输入/输出请求）而暂停执行时，CPU 立即转去执行另一道程序（进程）；这样，不仅使 CPU 得到充分利用，而且也提高了 I/O 设备和内存的利用率。在引入了多道程序设计技术后，使得操作系统具有多道程序同时（进程）运行且宏观上并行、微观上串行的特点，而操作系统也正是随着多道程序设计技术的出现而逐步发展起来的。要保证内存中多道程序的正常运行，在技术上需要解决以下 5 个问题。

(1) 在多道程序之间应如何分配 CPU, 使得 CPU 既能满足各程序运行的需要, 又能有较高的利用率? 此外, 一旦将 CPU 分配给某程序后, 应何时回收且如何回收?

(2) 如何为每道程序分配必要的内存空间, 使它们各得其所但又不会因相互重叠而丢失信息? 此外, 还要防止因某道程序出现异常而导致其他程序被破坏的情况出现。

(3) 系统中可能有多种类型的 I/O (输入/输出) 设备供多道程序共享, 应如何分配这些 I/O 设备? 如何做到既方便用户对设备的使用, 又能提高设备的利用率?

(4) 在现代计算机系统中通常都存放着大量的程序和数据, 应如何组织它们才便于用户使用并保证数据的安全性和一致性?

(5) 系统中的各种应用程序, 有的属于计算型, 有的属于 I/O 型, 有些既重要又紧迫, 有些又要求系统及时响应, 这时系统应如何组织这些程序(作业)的工作流程?

实际上, 这些问题就是操作系统的核心内容。因此, 操作系统应具有以下 4 个方面的管理功能: 处理器管理、存储管理、设备管理及文件管理。此外, 随着网络技术的不断发展, 操作系统还应具备相应的网络功能。同时, 为了方便用户使用计算机, 操作系统还应向用户提供方便、友好的用户接口。

1. 处理器管理

处理器管理主要是指对计算机系统中央处理器(CPU)的管理, 其主要任务是对 CPU 进行分配, 并对其运行进行有效地控制与管理。

为了提高计算机的利用率, 操作系统采用了多道程序技术。为了描述多道程序的并发执行引入了进程的概念, 进程可看作正在执行的程序, 通过进程管理来协调多道程序之间的关系, 以使 CPU 资源得到最充分的利用。在多道程序环境下, CPU 的分配与运行是以进程为基本单位的。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更细以降低并发执行的代价, 操作系统又引入了线程(Thread)的概念。对 CPU 的管理和调度最终归结为对进程和线程的管理和调度, 它的主要功能包括进程控制和管理、进程同步与互斥、进程通信、进程死锁、线程控制和管理以及 CPU 调度。

2. 存储管理

存储管理是指对内存空间的管理。程序要运行就必须由外存装入内存, 当多道程序被装入内存共享有限的内存资源时, 存储管理的主要任务就是为每道程序分配内存空间, 使它们彼此隔离互不干扰, 尤其是当内存不够用时, 要通过虚拟技术来扩充物理内存, 把当前不运行的程序和数据及时调出内存, 需要运行时再将其由外存调入内存。存储管理的主要功能包括内存分配、内存保护、地址变换和内存扩充。

3. 设备管理

设备管理是指计算机中除 CPU 和内存之外的所有输入输出设备(I/O 设备, 也称外部设备, 简称外设)的管理。其首要任务是为这些设备提供驱动程序或控制程序, 以便用户不必详细了解设备及接口的细节就可以方便地对设备进行操作。设备管理的另一个任务就是通过中断技术、通道技术和缓冲技术使外部设备尽可能与 CPU 并行工作, 以提高设备的使用效率。为了完成这些任务, 设备管理的主要功能包括外部设备的分配与释放、缓冲区管理、共享型外部设备的驱动调度、虚拟设备等。

4. 文件管理

文件是计算机系统中除 CPU、内存、外部设备等硬件设备之外的另一类资源，即软件资源。程序和数据以文件的形式存放在外存储器（如磁盘、光盘、磁带、优盘）上，需要时再把它们装入内存。文件管理系统的主要任务是有效地组织、存储和保护文件，使用户方便、安全地访问它们。文件管理的主要功能包括文件存储空间管理、文件目录管理、文件存取控制和文件操作等。

5. 用户接口

为了方便用户使用，操作系统向用户提供了使用接口。接口通常以命令、图形和系统调用等形式呈现给用户，前两种形式供用户通过键盘、鼠标或屏幕操作，后一种形式供用户在编程时使用。用户接口的主要功能包括命令接口管理、图形接口管理（图形实际上是命令的图形化表现形式）和程序接口管理。用户通过这些接口能方便地调用操作系统功能。

6. 网络与通信管理

随着计算机网络的迅速发展，网络功能已成为操作系统的重要组成部分。现代操作系统都注重为用户提供便捷、可靠的网络通信服务。为此，网络操作系统至少应具有以下 3 种管理功能。

（1）网络资源管理。计算机联网的主要目的之一是共享资源，网络操作系统应能够实现网上资源共享，管理用户程序对资源的访问，保证网络信息资源的安全性和完整性。

（2）数据通信管理。计算机联网后，站点之间可以互相传送数据。数据通信管理为网络应用提供必要的网络通信协议，处理网络信息传输过程中的物理细节，同时通过通信软件，按照网络通信协议完成网络上计算机之间的信息传输。

（3）网络管理。包括网络性能管理、网络安全管理、网络故障管理、网络配置管理和日志管理等。

1.1.3 操作系统的基本特征

目前存在多种类型的操作系统，不同类型的操作系统有各自的特征，但它们都具有并发性、共享性、虚拟性和不确定性等共同特征。在这些共同特征中，并发是操作系统中最重要的特征，而其他三个特征都是以并发为前提的。

1. 并发性

并发性（Concurrency）是操作系统最重要的特征。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（注意，不是同一时刻）。也就是说，在计算机系统中同时存在多个进程，从宏观上看，这些进程是同时运行并向前推进的；从微观上讲，任何时刻只能有一个进程执行，如果在单 CPU 条件下，那么这些进程就是在 CPU 上交替执行的。

操作系统的并发性能够有效地改善系统资源的利用率，提高系统的效率。例如，一个进程等待 I/O 时，就让出 CPU，并由系统调度另一个进程占用 CPU 运行。这样，在一个进程等待 I/O 时，CPU 就不会空闲，这就是并发技术。

操作系统的并发性会使操作系统的设计和实现变得更加复杂。例如,以何种策略选择下一个可执行的进程?怎样从正在执行的进程切换到另一个等待执行的进程?如何将内存中各交替执行的进程隔离开来,使之互不干扰?怎样让多个交替执行的进程互通消息并协作完成任务?如何协调多个交替执行的进程对资源的竞争?多个交替执行的进程共享文件数据时,如何保证数据的一致性?为了更好地解决这些问题,操作系统必须具有控制和管理进程并发执行的能力,必须提供某种机制和策略进行协调,从而使各并发进程能够顺利推进并获得正确的运行结果。此外,操作系统要充分发挥系统的并行性,就要合理地组织计算机的工作流程,协调各类软、硬件资源的工作,充分提高资源的利用率。

注意,并发性与并行性是两个不同的概念。并发性是指两个或多个程序在同一时间段内同时执行,即宏观上并行(同时执行),微观上串行(交替执行);而并行性则是指同时执行,如不同硬件(CPU与I/O设备)同时执行。

2. 共享性

共享性(Sharing)是操作系统的另一个重要特征。在内存中并发执行的多个进程可以共同使用系统中的资源(包括硬件资源和信息资源)。资源共享的方式可以分为以下两种。

(1) 互斥使用方式

互斥使用方式是指当一个进程正在使用某种资源时,其他欲使用该资源的进程必须等待,仅当这个进程使用完该资源并释放后,才允许另一个进程使用这个资源,即它们只能互斥地共享该资源,因此这类资源也称互斥资源。系统中的有些资源,如打印机、磁带机的使用就只允许互斥使用。

(2) 同时使用方式

系统中有些资源允许在同一段时间内被多个进程同时使用,这里的“同时”是宏观意义上的。典型的可供多个进程同时使用的资源是磁盘,可重入程序(可供多个用户同时运行的程序)也是可被同时使用的,如编译程序。

共享性和并发性是操作系统两个最基本的特征,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的问题。另一方面,如果系统不能对资源共享实施有效的管理,则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,导致整个系统的效率低下。

3. 虚拟性

虚拟性(Virtual)的本质含义是指将一个物理实体映射为多个逻辑实体。前者是实际存在的;后者是虚拟的,是一种感觉性的存在。例如,在单CPU系统中虽然只有一个CPU存在,且每一时刻只能执行一道程序,但操作系统采用了多道程序技术后,在一段时间间隔内,从宏观上看有多个程序在运行,给人的感觉好像是有多个CPU在支持每一道程序运行。这种情况就是将一个物理的CPU虚拟为多个逻辑的CPU。

4. 不确定性

在多道程序设计环境下,不确定性(Nondeterminacy)主要表现在以下三个方面。