

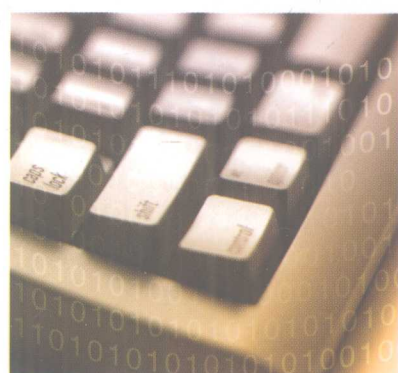
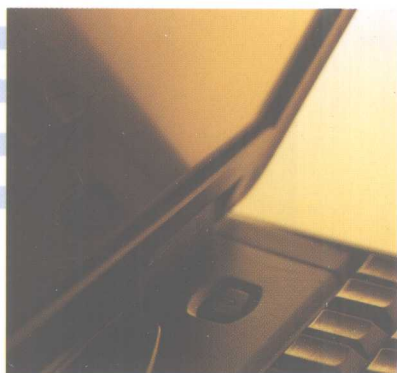


21 世纪高等学校应用型教材

操作系统原理 与实训教程

□ 范 辉 主 编

□ 华 臻 谢青松 陆 虹 副主编



高等教育出版社
Higher Education Press

21 世纪高等学校应用型教材

主要内容

操作系统原理与实训教程

范 辉 主 编
华 臻 谢青松 陆 虹 副主编



高等教育出版社

88
8020-810-0288
http://www.hep.com.cn
http://www.hep.edu.cn

北京西城区德胜门内大街
100011
010-25028899

中国农业大学出版社
北京海淀区学院路

2004年9月第3次印刷
2003年11月第1版
19.00元

350 000
13.2
787×1092 1/16

ISBN 7-03-013282-0

内容提要

本书根据职业教育“理论够用、注重实践”的特点和要求编写,对操作系统设计原理作了简要生动的介绍,并结合主流操作系统平台组织了比较丰富的实训内容。

本书共 11 章,分为原理篇和实训篇。原理篇主要内容包括:操作系统处理机管理、存储管理、设备管理、文件管理、作业管理和用户接口。每章后都有题型丰富的综合练习题,有助于读者及时消化所学知识。实训篇包括进程管理实训、存储管理实训、设备管理实训和文件管理实训。

本书语言通俗,图文并茂,理论与实践紧密结合,可作为高等学校应用型本科、成人高校以及高职高专院校计算机相关专业教材,也可作为其他院校非计算机专业教学参考用书。配套电子教案可从高等教育出版社网站下载,网址为 <http://www.hep.edu.cn> 或 <http://cs.hep.com.cn>。

图书在版编目(CIP)数据

操作系统原理与实训教程/范辉主编. —北京:高等教育出版社, 2003.11 (2004 重印)

ISBN 7-04-013289-3

I. 操... II. 范... III. 操作系统-高等学校:技术学校-教材 IV. TP316

中国版本图书馆 CIP 数据核字 (2003) 第 099583 号

策划编辑 雷顺加 责任编辑 萧潇 封面设计 王凌波 责任印制 杨明

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100011
总 机 010-82028899

购书热线 010-64054588
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>

经 销 新华书店北京发行所
印 刷 中国农业出版社印刷厂

开 本 787×1092 1/16
印 张 13.5
字 数 320 000

版 次 2003 年 11 月第 1 版
印 次 2004 年 9 月第 3 次印刷
定 价 19.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号: 13289-00

郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话：(010) 58581897/58581896/58581879

传 真：(010) 82086060

E - mail: dd@hep.com.cn

通信地址：北京市西城区德外大街4号

高等教育出版社打击盗版办公室

邮 编：100011

购书请拨打电话：(010)64014089 64054601 64054588

急需用章新示表我一预奇。替安代大随书编出有选等高职研官编的
系类出书编。玉流刊出香第味家步分各新息。成数网有对种目而预中并, 别官平本香书任由

前 言

jsn. fdc @ wmlnet. lsm-3 式 5 式

操作系统是现代计算机系统中必不可少的系统软件,控制和管理计算机的所有资源,为用户提供种种方便,是用户开发和使用应用软件不可缺少的支撑环境。操作系统课程是计算机专业的一门专业主干课程,主要研究操作系统的基本原理和实现方法,是计算机专业学生的知识结构中重要的组成部分。

已出版的操作系统教材大多注重理论,对实际应用比较忽视。这无疑增加了读者自学的难度,并降低了教材的实用性。因为操作系统原理中有很多概念比较抽象,而大多数读者今后并不参与操作系统的研究工作,他们需要的是对理论的理解和对实用系统的灵活应用。

根据高职高专教育的特点(理论够用、注重实践),结合作者多年来讲授操作系统课程的经验,并汲取现有操作系统研究的理论成果和相关教材的营养,本书除了对操作系统的设计原理作了简明扼要的介绍外,还结合主流操作系统平台,组织了比较丰富的实训内容。全书分两大部分——原理篇和实训篇。通俗易懂、图文并茂、注重理论与实践相结合是本书的特点。特别是书中多处引用了源于生活的生动有趣的例子,对于读者理解操作系统原理中的一些重要而抽象的概念以及掌握进程同步互斥问题等学习难点将大有裨益。书中各章相关的内容前后呼应,这种内容上的对比和关联可以帮助读者融会贯通,从整体上深入理解操作系统原理。另外,本书每一章开头部分都给出了“本章导读”和“本章主要知识点”,对读者学习和理解该章内容起到一定的指导作用;每章后都安排了大量的理论和实训题目,读者能通过练习巩固所学的理论知识并提高自己的动手能力。原理篇简要介绍了操作系统的产生、发展、类型和结构,操作系统对处理机、存储器、I/O 设备、文件和作业的管理功能及其实现原理;实训篇则从使用级、观察级、系统管理级、源码阅读级和实现级等不同深度,结合 UNIX、Linux 和 Windows 操作系统平台,安排了许多实训题目,并对实训内容作了比较详细的指导性介绍。本书实训内容比较多,使用者可根据实验条件和自己的需要等酌情选择。

本书第 1 章为引言,概述操作系统的内容。第 2 章为处理机管理,主要介绍进程管理和处理机调度,其中进程的互斥与同步是全书的学习难点。第 3 章为存储管理,介绍实模式和虚模式的内存管理策略,主要以实用的分页方式为主。第 4 章为文件管理,主要介绍文件的结构、基本操作、共享与保护、目录检索和外存空间的管理。第 5 章为设备管理,主要介绍设备的分类、驱动、调度和缓冲的管理。第 6 章为作业管理,主要介绍作业的概念、分类、控制方式以及系统对批处理和分时作业的处理。第 7 章~第 11 章为实训内容,包括进程管理实训、存储管理实训、设备管理实训和文件管理实训。

本书原理篇由华臻、谢青松、冯烟利编写,实训篇由陆虹、李晋江编写。全书由范辉统稿。参加本书排版、录入、程序调试等工作的还有邹海林、张奎平、朱智林、刘惊雷、范宝德、张玉林等同志。

在本书编写大纲的讨论会上,深圳职业技术学院、上海第二工业大学、洛阳大学、沈阳电力高等专科学校、宁波高等专科学校等 20 多所院校的一线老师提出了许多宝贵的意见和建议。本书

的编写得到高等教育出版社的大力支持。在此一并表示诚挚的谢意。

由于作者水平有限,书中疏漏与错误在所难免,恳请各位专家和读者批评指正。编者的联系方式为 E-mail:fanlinw@263.net。

前 言

编 者

2003年9月

本书在编写过程中,得到了许多领导和专家的关心、支持与帮助,在此一并表示诚挚的谢意。

由于作者水平有限,书中疏漏与错误在所难免,恳请各位专家和读者批评指正。编者的联系方式为 E-mail:fanlinw@263.net。

本书在编写过程中,得到了许多领导和专家的关心、支持与帮助,在此一并表示诚挚的谢意。

由于作者水平有限,书中疏漏与错误在所难免,恳请各位专家和读者批评指正。编者的联系方式为 E-mail:fanlinw@263.net。

本书在编写过程中,得到了许多领导和专家的关心、支持与帮助,在此一并表示诚挚的谢意。

由于作者水平有限,书中疏漏与错误在所难免,恳请各位专家和读者批评指正。编者的联系方式为 E-mail:fanlinw@263.net。

目 录

原 理 篇

第 1 章 引言	(3)	1.7.3 管态与目态	(15)
1.1 操作系统的概念	(3)	1.7.4 存储结构	(15)
1.1.1 操作系统的地位	(3)	1.7.5 存储保护	(15)
1.1.2 操作系统的管理目标和主要功能	(4)	综合练习题一	(16)
1.1.3 操作系统的定义	(5)	第 2 章 处理机管理	(18)
1.2 操作系统的历史回顾	(5)	2.1 多道程序设计	(18)
1.2.1 操作系统的产生	(5)	2.1.1 顺序程序的执行	(19)
1.2.2 操作系统的完善	(6)	2.1.2 并发程序的执行	(19)
1.2.3 操作系统的发展	(6)	2.1.3 并发程序执行的条件	(20)
1.3 操作系统的类型	(6)	2.2 进程的描述	(21)
1.3.1 批处理操作系统	(6)	2.2.1 进程的定义	(21)
1.3.2 分时操作系统	(7)	2.2.2 进程的特性及其与程序的区别	(22)
1.3.3 实时操作系统	(8)	2.2.3 进程的基本状态及其转换	(23)
1.3.4 单用户操作系统	(9)	2.2.4 进程控制块	(24)
1.3.5 网络操作系统	(9)	2.2.5 进程的队列	(25)
1.3.6 分布式操作系统	(9)	2.3 进程的控制	(26)
1.4 操作系统的特征	(10)	2.3.1 进程控制机构	(26)
1.4.1 并发性	(11)	2.3.2 进程控制原语	(27)
1.4.2 共享性	(11)	2.4 进程的互斥	(28)
1.4.3 虚拟性	(11)	2.4.1 互斥的定义	(28)
1.4.4 异步性	(11)	2.4.2 上锁和开锁原语	(29)
1.5 操作系统与用户的接口	(12)	2.4.3 用上锁和开锁原语实现进程的互斥	(29)
1.5.1 命令接口	(12)	2.5 信号量机制	(29)
1.5.2 程序接口	(12)	2.5.1 信号量的概念	(30)
1.5.3 图形用户接口	(12)	2.5.2 P、V 操作原语	(30)
1.6 操作系统的结构	(13)	2.5.3 用 P、V 操作原语实现进程的互斥	(31)
1.6.1 整体式系统	(13)	2.6 进程的同步	(33)
1.6.2 层次式系统	(13)	2.6.1 同步的定义	(34)
1.6.3 虚拟机系统	(13)	2.6.2 用 P、V 操作原语实现进程的同步	(34)
1.6.4 客户-服务器系统	(14)	2.7 进程的通信	(41)
1.7 操作系统的硬件环境	(14)	2.7.1 进程通信的定义	(41)
1.7.1 CPU 与外设并行工作	(14)		
1.7.2 I/O 中断的作用	(15)		

2.7.2 发送和接收原语——实现进程通信的基本原语	(42)	4.1.3 设备管理的目标	(94)
2.7.3 消息缓冲通信方式	(42)	4.1.4 设备管理结构	(94)
2.7.4 信箱通信方式	(43)	4.1.5 设备控制器和 I/O 通道	(94)
2.8 死锁问题	(45)	4.2 I/O 控制方式	(95)
2.8.1 死锁的定义	(45)	4.2.1 程序直接控制方式	(95)
2.8.2 产生死锁的原因	(47)	4.2.2 中断控制方式	(96)
2.8.3 产生死锁的必要条件	(47)	4.2.3 DMA 控制方式	(96)
2.8.4 死锁的防止	(47)	4.2.4 通道控制方式	(97)
2.8.5 死锁的避免	(48)	4.3 中断技术	(99)
2.8.6 死锁的检测与解除	(52)	4.3.1 中断的基本概念	(99)
2.9 处理机调度	(52)	4.3.2 中断分类与优先级	(100)
2.9.1 调度算法	(53)	4.3.3 中断处理过程	(100)
2.9.2 调度时机	(55)	4.4 缓冲技术	(101)
2.9.3 调度过程	(55)	4.4.1 缓冲技术的基本思想	(101)
2.10 线程的概念	(55)	4.4.2 缓冲技术的分类	(101)
2.10.1 线程的定义	(56)	4.5 设备分配及设备的处理程序	(103)
2.10.2 线程与进程的比较	(56)	4.5.1 设备分配的数据结构	(103)
综合练习题二	(57)	4.5.2 设备分配策略	(104)
第3章 存储管理	(62)	4.5.3 设备处理	(105)
3.1 存储管理的概念	(62)	4.6 SPOOLing 系统	(106)
3.1.1 存储系统的分类	(62)	4.7 磁盘设备管理	(108)
3.1.2 物理地址和逻辑地址	(63)	4.7.1 磁盘结构	(108)
3.1.3 静态重定位和动态重定位	(63)	4.7.2 磁盘的访问时间	(109)
3.1.4 存储管理的功能	(64)	4.7.3 磁盘调度算法	(110)
3.1.5 内存扩充技术	(65)	4.7.4 磁盘缓存置换算法	(113)
3.1.6 存储管理的分类	(66)	综合练习题四	(114)
3.2 分区存储管理	(67)	第5章 文件管理	(117)
3.2.1 固定分区存储管理	(67)	5.1 文件系统的概述	(117)
3.2.2 可变分区存储管理	(69)	5.1.1 基本概念	(118)
3.3 页式存储管理	(73)	5.1.2 文件分类	(119)
3.3.1 分页式存储管理	(73)	5.2 文件结构与存取方法	(120)
3.3.2 虚拟页式存储管理	(78)	5.2.1 文件的组织	(120)
3.4 段式存储管理	(83)	5.2.2 文件的逻辑结构及存取方式	(121)
3.4.1 分段式存储管理	(83)	5.2.3 文件的物理结构及存取设备	(122)
3.4.2 虚拟段式存储管理	(86)	5.2.4 存储空间管理	(124)
3.5 段页式存储管理	(89)	5.3 文件目录	(126)
综合练习题三	(89)	5.3.1 目录内容	(127)
第4章 设备管理	(92)	5.3.2 目录结构	(127)
4.1 设备管理概述	(92)	5.3.3 文件别名的实现	(129)
4.1.1 设备管理的分类	(92)	5.4 文件使用	(131)
4.1.2 设备管理的功能	(93)	5.4.1 文件访问	(131)
		5.4.2 文件控制	(131)

5.4.3 目录管理	(131)	6.1.1 启动程序执行的方式	(139)
5.4.4 伪文件	(132)	6.1.2 用户与操作系统之间的接口	(139)
5.5 安全与完整性	(132)	6.1.3 作业的基本概念	(142)
5.5.1 文件的访问权限	(133)	6.2 批处理系统的作业管理	(142)
5.5.2 文件的并发访问	(133)	6.2.1 作业的状态及转换	(142)
5.5.3 文件的存取控制	(133)	6.2.2 作业的建立	(143)
5.5.4 文件的完整性	(135)	6.2.3 作业调度的评估	(144)
综合练习题五	(135)	6.2.4 作业调度算法	(145)
第6章 作业管理和用户接口	(138)	综合练习题六	(148)
6.1 概述	(139)		
实 训 篇			
第7章 实训基础	(153)	8.5.2 进程同步和通信	(178)
7.1 概述	(153)	8.5.3 模拟进程调度	(178)
7.1.1 UNIX系统调用	(154)	第9章 存储管理实训	(179)
7.1.2 简单命令	(156)	9.1 UNIX/Linux存储管理	(179)
7.1.3 重定向与管道命令	(157)	9.1.1 请求调页管理的数据结构	(179)
7.1.4 后台命令	(158)	9.1.2 交换进程	(180)
7.1.5 通信命令	(158)	9.1.3 请求调页	(180)
7.2 Windows 2000/NT实训基础	(159)	9.2 UNIX/Linux存储管理实训指导	(181)
7.2.1 结构部件	(159)	9.2.1 观察内存工作情况	(181)
7.2.2 Windows NT对象	(160)	9.2.2 有关系统调用	(181)
7.3 实训	(161)	9.2.3 有关系统文件/proc	(181)
7.3.1 实训目的	(161)	9.2.4 动态存储分配	(182)
7.3.2 实训内容	(161)	9.3 Windows 2000/NT存储管理实训指导	(183)
第8章 进程管理实训	(162)	9.4 存储管理设计模拟	(184)
8.1 UNIX/Linux进程管理	(162)	9.4.1 目的和要求	(184)
8.1.1 UNIX进程的组成	(162)	9.4.2 示例说明	(184)
8.1.2 进程状态	(164)	9.4.3 示例程序	(184)
8.1.3 进程控制	(165)	9.5 实训	(186)
8.2 UNIX/Linux进程管理实训指导	(168)	9.5.1 实训目的	(186)
8.2.1 启动进程	(168)	9.5.2 实训内容	(186)
8.2.2 查看进程	(169)	第10章 设备管理实训	(187)
8.2.3 进程调度	(170)	10.1 UNIX/Linux设备管理	(187)
8.3 Windows 2000/NT进程管理实训指导	(171)	10.1.1 缓冲区管理	(187)
8.4 进程调度设计模拟	(172)	10.1.2 块设备的管理	(188)
8.4.1 目的和要求	(172)	10.1.3 字符设备的管理	(189)
8.4.2 示例说明	(173)	10.2 UNIX/Linux设备管理实训指导	(190)
8.4.3 示例程序	(173)	10.2.1 观察/控制设备工作情况	(190)
8.5 实训	(177)	10.2.2 磁盘管理	(190)
8.5.1 进程并发执行	(177)	10.3 Windows 2000/NT设备管理实训	

4 目 录

10.3 实训指导	(191)	11.2.3 文件查找命令	(200)
10.4 实训	(192)	11.2.4 文本处理命令	(200)
10.4.1 实训目的	(192)	11.2.5 统计文件大小命令	(200)
10.4.2 实训内容	(192)	11.2.6 文件比较命令	(201)
第11章 文件管理实训	(193)	11.2.7 文件复制、删除和移动命令	(201)
11.1 UNIX/Linux 文件管理	(193)	11.2.8 文件链接命令	(201)
11.1.1 文件管理系统的数据结构	(193)	11.2.9 目录的创建与删除命令	(202)
11.1.2 索引结点的管理	(194)	11.2.10 改变工作目录、显示目录内容命令	(202)
11.1.3 文件存储空间管理	(195)	11.3 Windows 2000/NT 文件管理实训	(202)
11.1.4 文件目录	(195)	11.3.1 实训指导	(202)
11.1.5 文件的分配	(196)	11.4 实训	(204)
11.1.6 文件卷的安装与拆卸	(196)	11.4.1 实训目的	(204)
11.2 UNIX/Linux 文件管理实训指导	(197)	11.4.2 实训内容	(204)
11.2.1 文件和目录操作相关命令	(197)	参考文献	(205)
11.2.2 文件内容查询命令	(199)		
11.2.3 文件查找命令	(199)		
11.2.4 文本处理命令	(200)		
11.2.5 统计文件大小命令	(200)		
11.2.6 文件比较命令	(201)		
11.2.7 文件复制、删除和移动命令	(201)		
11.2.8 文件链接命令	(201)		
11.2.9 目录的创建与删除命令	(202)		
11.2.10 改变工作目录、显示目录内容命令	(202)		
11.3 Windows 2000/NT 文件管理实训	(202)		
11.3.1 实训指导	(202)		
11.4 实训	(204)		
11.4.1 实训目的	(204)		
11.4.2 实训内容	(204)		
参考文献	(205)		

原 理 篇

- 第 1 章 引 言
- 第 2 章 处理机管理
- 第 3 章 存储管理
- 第 4 章 设备管理
- 第 5 章 文件管理
- 第 6 章 作业管理和用户接口

第 1 章 引 言

计算机系统由硬件和软件两部分组成。硬件部分指计算机物理装置本身，主要包括处理器（如中央处理器、I/O 处理器）、存储器、I/O 控制器和各种外部设备等。软件部分指操作系统、应用程序等。操作系统是计算机系统的核心，它管理计算机系统的资源，为用户提供操作界面。应用程序是用户为解决特定问题而编写的程序。计算机系统的发展经历了从简单到复杂、从低级到高级的过程。随着科学技术的进步，计算机系统的性能不断提高，应用范围不断扩大。本章主要介绍操作系统的概念、类型、特征、接口、结构和硬件环境。



本章导读

大量复杂计算的需求催生了计算机，早期计算机的复杂操作引起了使用者的不满，这种不满又引发出新的发明……这样的过程形成了计算机的发展链。计算机的发展链是更大范围发展链的组成部分，而它本身又孕育着新的发展链。操作系统(Operating System, OS)——在今天的计算机系统中举足轻重的系统软件，就是在计算机发展链上从无到有、从简单到复杂地逐步发展起来的。它的发展与计算机用户的需求和计算机制造技术的进步以及计算机理论科学的发展密切相关。它是人机之间的“桥梁”，是硬件最亲密的“伙伴”，是其他软件的“基石”。

本章主要知识点

- 操作系统的概念
- 操作系统的类型
- 操作系统的特征
- 操作系统与用户的接口
- 操作系统的结构
- 操作系统的硬件环境

1.1 操作系统的概念

每个用过操作系统的人都得到过它的服务和帮助，然而要给操作系统下一个精确定义却并非易事。几十年来，人们从不同角度对操作系统作过多种解释，但至今尚无统一的定义。以下通过分析操作系统的地位和配置它的目的，给出操作系统概念的一个非形式化的描述。

1.1.1 操作系统的地位

一个完整的计算机系统由两部分组成：计算机硬件和计算机软件。硬件部分指计算机物理装置本身，主要包括处理器（如中央处理器、I/O 处理器）、存储器、I/O 控制器和各种外部设备等。

中央处理器是对信息进行高速运算和控制处理的部件;存储器用于存放程序和数据;I/O 处理器和控制单元用于控制和管理各种外设与主存间的信息传送。软件部分指由计算机硬件执行以完成一定任务的所有程序及其数据,主要包括系统软件和应用软件两大类。操作系统、编译器、编辑器和数据库管理系统等是常见的系统软件,其中操作系统是当今计算机必不可少的,只有装上它,别的软件才能运行,用户才能够使用计算机。财务管理程序、电脑摇奖程序、火车订票系统和浏览器等都属于应用软件。

没有配置软件的计算机称为裸机,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安装过若干层软件后的计算机,如图 1-1 所示。尽管人们最早使用的电子计算机是裸机(当时的使用者也只是极少数的专家),但今天如果给我们的读者每人发一台裸机的话,哪怕这些裸机的配置非常高,大概谁也不会用。

由图 1-1 可以看出,操作系统是现代计算机系统中最基本和最重要的系统软件,是裸机的第一层软件,是对硬件功能的首次扩充。所以说,它是人机之间的“桥梁”,是硬件最亲密的“伙伴”,是其他软件的“基石”。

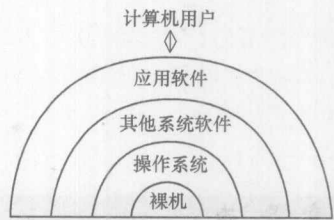


图 1-1 操作系统是用户与裸机之间的接口

1.1.2 操作系统的管理目标和主要功能

操作系统作为计算机系统的“管家”,必须明确管理目标,或者说必须明白计算机配置操作系统的目的。为了充分理解这些管理目标,我们不妨借鉴对人类社会中成功管理的经验。例如,一个公司的管理部门,要提高经济效益,至少需要实现三个管理目标:开拓市场,搞好生产,用好资源,即:

- (1) 为客户提供种种方便,以争取接到尽量多的订单;
- (2) 制定生产计划,组织加工流程,提高生产效率,保证产品质量;
- (3) 及时获取并管理好所需各种资源,充分发挥资源作用,尽量消除浪费资源的现象。

类似地,操作系统的三大目标是:为用户使用计算机提供方便,合理地组织计算机的工作流程,有效控制和管理计算机系统的各类资源。

再比如,一个公司为了实现自己的管理目标,主管人员至少要设置几个主要部门协同做好以下主要管理工作:处理订单、管理仓库、组织生产、保管物资和运输调度等。与此相似,作为计算机系统的“管家”,操作系统中由几部分模块协作完成其五大管理功能:处理机管理、存储器管理、设备管理、文件管理和作业管理。

操作系统的处理机管理与公司组织生产相似,它主要负责协调系统中的多个程序,使它们正确、有效地占用处理机。存储器管理与公司管理仓库相似,它主要负责内存的分配与回收、程序的相对地址到物理地址的转换、内存的共享与保护以及内存的逻辑扩充等。设备管理与公司运输调度相似,它主要负责输入/输出设备的分配、共享、驱动调度和方便使用等。文件管理与公司保管物资相似,它主要负责文件存储介质的管理(主要是外存空间的分配与释放)、文件的按名存取(即文件目录的组织与查询等)、文件的保护与共享、文件的基本操作和文件系统的安全性等。作业管理与公司处理订单相似,它主要负责命令处理和作业调度。另外,操作系统还有用户接口、中断处理、时钟管理、出错处理甚至网络管理等功能。

本书主要介绍单处理机多道程序环境下操作系统的五大主要管理功能的实现原理,这五大管理功能实际上也是多数操作系统内核(指操作系统中运行在核心态的部分)的主要功能。

1.1.3 操作系统的定义

根据前面关于操作系统的地位以及操作系统的目标的描述,可以给出关于操作系统的描述性定义如下:

操作系统是合理组织计算机的工作流程,有效控制和管理计算机系统的各类资源,并方便用户使用计算机的程序集合。

1.2 操作系统的历史回顾

操作系统的发展与计算机用户的需求和计算机制造技术的进步以及计算机理论科学的发展是密切相关的,它是从无到有、由简单到复杂、逐步发展起来的,至今已有近 50 年的历史。

1.2.1 操作系统的产生

在 1946 年到 1955 年的第一代电子管计算机中是没有配置操作系统的。早期的每台机器都有一个小组专门来设计、制造、编程、操作和维护。编程全部采用机器语言,通过在一些插板上的硬连线来控制其基本功能。机器的使用方式是程序员到机房按照要运行的程序的指令要求将插板上的线缆插好,再开始长达几个小时的计算。要改变程序的运行,就要重新插拔这些导线,变更电路的连接,这都要靠手工来完成。这时的计算机很不可靠,只要一个真空管发生故障,计算机就无法运行。这个阶段基本上所有的应用都是数值计算问题。运行程序所需的过多的人工干预,包括运行前的手工准备、运行中的手工控制、运行后的手工整理,以及随时需要的人工维修,使得计算机难以发挥高速运算的才能。到 20 世纪 50 年代早期,出现了穿孔卡片,就不用插板了,而是将程序写在卡片上然后读入计算机,但其他过程则依然如故。随着计算机速度的迅速提高,人工干预与计算机计算所花时间之比也迅速增大,这种人机矛盾日益尖锐化。

解决人机矛盾的努力与计算机硬件的发展有着密切的关系。1955 年到 1965 年的第二代晶体管计算机已经很可靠了,其运算和存储能力有了很大提高,外部设备中也出现了相对高速的磁带机。为了解决人机矛盾,人们首先想到的是如何摆脱从一个用户程序过渡到另一个用户程序时的人工干预,使其转换能自动进行。这就产生了由计算机对一批用户程序进行自动处理的所谓(单道)批量处理技术,相应的操作系统就产生了。其中,典型的有后来被称作现代操作系统雏形的 FMS(FORTRAN Monitor System, FORTRAN 监控系统)和 IBSYS(IBM 为其 7094 机配备的操作系统)。这时的系统属于脱机 I/O 系统,机器的使用方式是操作员使用廉价的外围计算机从读卡机上依次把一批作业读到磁带上,再把这个输入磁带卸下,移到较昂贵的主机上,然后启动主机上特殊的监控程序(现代操作系统的前身),由它从磁带上将第一个作业读入内存并运行,将输出写到第一盘输出磁带上。每个作业结束后,监控程序(操作系统)自动地读入下一个作业运行。当这一批作业结束后,操作员输入和输出磁带,将输入磁带换成下一批作业,然后把输出磁带拿

到外围机上进行脱机打印。由此可见,这时的人机矛盾仍然存在。

1.2.2 操作系统的完善

从20世纪50年代末到20世纪60年代初,通道(指专用的I/O处理机)、中断和缓冲技术的出现和使用,使得多道程序并发执行成为可能。1965年到1980年的第三代集成电路计算机非常可靠,其运算和存储能力有了更大提高,外部设备中也出现了带键盘的终端和硬盘等更高速的设备。内存中可以同时存放多道程序,使它们同时(在单机系统中只能交替)运行,于是,共享系统资源的多道程序系统出现了,相应的旨在提高系统资源利用率和方便用户使用计算机的三种基本的操作系统也相继产生,它们是多道批处理系统、分时系统和实时系统。这时作业运行过程中几乎没有人工干预,系统资源利用率和吞吐量(指系统在单位时间内所完成的总工作量)很高,用户作业周转时间(指作业从进入系统开始,直到其完成并退出系统为止所经历的时间)也比较令人满意。操作系统终于代替人工操作成为计算机系统的“管家”,其发展进入了成熟期,UNIX是这个时期的典型代表。

1.2.3 操作系统的发展

从1980年至今,建立在大规模集成电路基础上的第四代计算机得到蓬勃发展。从个人计算机到并行机,再到网络,计算机体系结构也不断发展变化。相应的微机操作系统、并行操作系统、分布式操作系统、网络操作系统和嵌入式操作系统等也相继产生。操作系统的使用界面也从多年一贯的字符界面变成了图形界面。操作系统的结构除了有序分层的模块化结构外,还出现了虚拟机结构和客户/服务器加微内核结构等。DOS、OS/2、Windows和Linux等是这一时期的典型代表。操作系统已经进入了一个新的发展时期。

1.3 操作系统的类型

对操作系统进行分类的方法很多,从不同的角度可以得到不同的划分。例如,按照机器硬件的结构与规模可分为大型机操作系统、中型机操作系统、小型机操作系统、微型机操作系统、网络操作系统和嵌入式操作系统等。按照系统所能同时响应的用户和任务数可分为单用户单任务操作系统、单用户多任务操作系统和多用户多任务操作系统等。而过去使用更广泛的一种分法是,按照系统处理任务的方式分为三种基本类型:多道批处理操作系统、分时操作系统和实时操作系统。现在则又增加了一种基本类型,即分布式操作系统。下面简单介绍几种主要的操作系统。

1.3.1 批处理操作系统

多道批处理系统实际上是把批处理技术和多道程序技术相组合的产物,是成批处理的或顺序共享式的系统,允许多个用户以高速、非人工干预的方式进行成组作业工作和程序执行。它出现于20世纪60年代初期,目前仍在使用。

1. 多道批处理系统的运行方式

这时的系统采用了假脱机工作方式,机器的使用方式是操作员把用户提交的作业卡片放到读卡机上,系统通过 SPOOLing(Simultaneous Peripheral Operation On-Line,联机外设同时操作)输入程序,及时把这些作业送入外存的磁盘输入井,形成后备作业队列。然后,系统的作业调度程序根据系统的当前情况和后备作业的特点,按照一定的调度原则选择一个或几个搭配合理的作业,将其装入内存并为之创建进程准备运行。接着,系统的进程调度程序再按照一定策略调度各个进程依次占有处理机运行用户作业程序。当某个作业完成时,系统把它的计算结果交给 SPOOLing 输出程序准备输出,并回收该作业的全部资源。重复上述过程,使得作业一个接一个地流入系统,经过处理后又一个接一个地退出系统,形成一个源源不断的作业流。可见,这时的操作员再也不必搬运磁带了,系统资源得到充分的利用,但用户在提交作业的同时必须提交用系统提供的作业控制语言写的作业说明书,以提前向系统说明各个作业步的脱机控制需求,而且在作业运行整个过程中,用户不能与作业进行交互,这对程序员修改和调试程序是极不方便的。

2. 多道批处理系统的优缺点

多道批处理系统的主要优点如下:

(1) 资源利用率高。由于在内存中装入了多道程序,它们共享资源,使得系统的各类资源得到更充分的利用。

(2) 系统吞吐量。合理的调度算法会使 CPU 和其他资源尽可能地保持忙碌。

(3) 系统开销小。仅当作业完成或无法运行时才进行切换,系统的时空开销小。

多道批处理系统的主要缺点如下:

(1) 用户没有交互能力。用户一旦把作业提交给系统后就失去了对自己作业的控制,系统只能根据用户提供的作业说明书来控制作业的执行,这对程序的修改和调试是非常不便的。

(2) 作业平均周转时间(指一批作业中所有作业周转时间的平均值)长。由于作业要排队,需依次进行处理,因而周转时间较长。特别对于排在队尾的作业,其运行请求会被长期推迟响应。

1.3.2 分时操作系统

配置了分时操作系统的计算机系统是这样的:一台主机连接着多个带有显示器和键盘以及控制器的本地或远程终端,每个用户都可以在自己的终端以交互方式使用主机,共享系统的资源,如图 1-2 所示。

这样的系统现在仍然很普遍,特别在金融、电信等行业网点。如果用工作站代替终端,那么今天的网络也可看成是这种系统进化的结果。分时操作系统的典型代表是 UNIX,它对以后的操作系统的设计产生了重大影响。

1. 分时的概念与实现

分时是计算机系统中一个普遍的概念,它可以理解为:两个或多个事件按时间划分轮流使用计算机系统中的某一资源。例如,CPU 和通道同时使用内存,多台设备同时使用

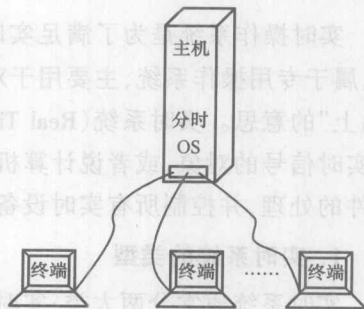


图 1-2 分时系统示意图