

高等院校计算机专业教育改革推荐教材

操作系统 简明教程

徐汀荣 徐云龙 孔 芳 卢晓东 编著

钱培德 审



机械工业出版社
CHINA MACHINE PRESS

16
5

高等院校计算机专业教育改革推荐教材

操作系统简明教程

徐汀荣 徐云龙 编著
孔 芳 卢晓东

钱培德 审



机械工业出版社

本书全面简要地介绍了计算机操作系统的原理及技术。首先概述了操作系统的形成、类型和功能等,介绍了几种目前常用的操作系统,尤其是Linux系统的功能,然后阐述了计算机操作系统进程、线程、管程的基本概念以及进程互斥、同步、通信、调度、死锁的原理,并用Linux系统的具体实现说明了这些原理。书中还介绍了存储器管理、设备管理、文件管理、分布式操作系统的各种原理及技术,并把目前流行的Linux操作系统贯穿在全书中,书后还附有Red Hat Linux 7.3的安装以及常用Linux命令的使用说明供读者参考。

本书注重基础知识的讲解和基本技能的培养,内容由浅入深、循序渐进地引导读者掌握现代操作系统的原理及其实现。注重简洁性、实用性、先进性、操作性。各章配备了相关的例题、习题。本书可作为各类大专院校计算机及信息工程类“操作系统”课程的教材,也可以作为各类计算机培训班的教材。

图书在版编目(CIP)数据

操作系统简明教程/徐汀荣等编著. —北京:机械工业出版社,2004.1

高等院校计算机专业教育改革推荐教材

ISBN 7-111-13393-5

I. 操... II. 徐 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(2003)第102244号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策 划:胡毓坚

责任编辑:蔡 岩

责任印制:闫 焱

北京京丰印刷厂印刷·新华书店北京发行所发行

2004年1月第1版·第1次印刷

787mm×1092mm¹/₁₆·11 75印张·285千字

0 001—5 000册

定价:17.00元

凡购本图书,如有缺页、倒页、脱页,由本社发行部调换

本社购书热线电话(010)68993821、88379646

封面无防伪标均为盗版

高等院校计算机专业教育改革推荐教材

编委会成员名单

主 编 刘大有

副主编 王元元

编 委 (按姓氏笔画排序)

刘晓明 李师贤 张桂芸 徐汀荣

耿亦兵 顾军华 黄国兴 薛永生

编者的话

计算机科学技术日新月异的飞速发展和计算机科学技术专业教育的相对滞后,已是不争的事实。

有两个发人深省的现象:一是,由于非计算机专业的学生既具有一门非计算机专业的专业知识,又具有越来越高的计算机应用技术水平,从而使计算机专业的学生感受到一种强烈的冲击和压力;二是,创建软件学院的工作已有近两年的历史,但软件学院的计算机专业教育的定位仍在探讨之中。

我们认为计算机科学与技术专业(以下简称计算机专业)教育的改革势在必行,正确认识和划分计算机专业教育的层次,对该专业的教育改革无疑是一个非常重要的问题。我国的计算机专业教育主要分三个层次。一般说来,这三个层次通常分布在以下三类高等院校:

第一层次主要以具有计算机一级学科博士学位授予权的教育部属重点高等院校为代表(包括具有两个博士点的大学)。这一类大学本科着重培养理论基础比较坚实、技术掌握熟练、有一定研究和开发能力的计算机专业学科型人才,其中部分学生(约本科生的10%)可攻读博士学位。

第二层次主要以具有一个计算机二级学科专业博士点的教育部属高等院校为代表。这一类高等院校本科着重培养有一定的理论基础、技术掌握比较熟练、有一定的研究或开发能力的计算机专业人才,其中一部分培养成学科型人才,另一部分培养成应用型人才,一小部分学生(约本科生的5%)可攻读博士学位。

第三层次主要以具有计算机二级学科专业硕士点的省属高等院校为代表。这一类高等院校本科面向企业应用,侧重培养对计算机技术或部分计算机技术掌握比较熟练,有一定的开发、应用能力的计算机专业应用型人才,其中很小一部分学生(约本科生的2.5%)可攻读博士学位。

国家教育部、计委批准的或省教育厅批准的示范性软件学院,就其培养目标和办学特色而言,分别与第二层次中应用型人才培养部分以及第三层次比较相近,但在如下方面有所不同:将软件工程课程作为专业教学重点;更加强调英语教学,更加重视实践能力培养,并对两者有更高的要求。

我们本着对高等院校的计算机专业状况的认识,主要面向与上述第二、第三两个层次对应的院校及与之相近的软件学院,总结多年的计算机专业的教改经验,在一定程度上溶入了ACM & IEEE CC2001和CCC2002(中国计算机科学与技术学科教程)的教改思路,组织我国一直投身于计算机教学和科研的教师,编写了这套“高等院校计算机专业教育改革推荐教材”(以下简称“推荐教材”)。自然,“推荐教材”中所贯穿的改革思路和做法,也是针对上述第二、第三两个层次对应院校的计算机专业学生。这些思路和做法可概括成以下三句话:

- 适度调整电子技术基础、计算机理论基础和系统软件的教学内容。
- 全面强化计算机工具软件、应用软件的教学要求。
- 以应用为目标大力展开软件工程的教学与实践。

电子技术基础、计算机理论基础、系统软件教学关系到学生的基本素质、发展潜力和日后

的应变能力。“推荐教材”在调整它们的教学内容时的做法是:适度压缩电子线路、数字电路和信号系统的教学内容,变三门课程为两门,并插入数字信号处理的基础内容;合并“计算机组成原理”、“微型计算机接口技术”和“汇编语言”为“计算机硬件技术基础”一门课程;注意适当放宽“离散数学”课程的知识面,使之与 CCC2002 的要求基本接轨,但适度降低其深度要求;更新系统软件课程的教学内容,以开放代码的 Linux 作为操作系统原理的讲授载体,更加关注系统软件的实践性和实用性。

为了提高计算机专业人才的计算机应用能力,全面强化计算机工具软件、实用软件的教学要求是十分重要的,这也是上述改革思路的核心。为此,“系列教材”的做法是:强化程序设计技术,强化人机接口技术,强化网络应用技术。

为强化程序设计技术,“推荐教材”支持在单片机环境、微机平台、网络平台的编程训练;支持运用程序设计语言、程序设计工具以及分布式对象技术的编程训练。大大加强面向对象程序设计课程的组合(设计了三门课程:面向对象的程序设计语言 C++,面向对象的程序设计语言 JAVA 和分布式对象技术),方便教师和读者的选择。

为强化人机接口技术,“推荐教材”设计了“人机交互教程”,“计算机图形学”和“多媒体应用技术”等可供选择的、有层次特色的课程组合。

为强化网络应用技术,“推荐教材”设计了“计算机网络技术”,“计算机网络程序设计”,“计算机网络实验教程”和“因特网技术及其应用”等可供选择的、新颖丰富的课程组合。

将软件工程课程作为专业教学重点,以应用为目标大力展开软件工程的教学与实践,是“推荐教材”改革思路的又一亮点。为改变以往软件工程课程纸上谈兵的老毛病,“推荐教材”从工程应用出发,理论联系实际,突出建模语言及其实现工具的运用,设计了“软件工程的方法与实践”,“统一建模语言 UML 导论”和“ROSE 对象建模方法与技术”等可供选择的、创新独特的软件工程课程组合。对于各类软件学院,“推荐教材”的这一特色无疑是很有吸引力的。

强调实践也是计算机学科永恒的主题,对计算机应用专业的学生来说更是如此。重应用和重实践是“推荐教材”的一个整体特点。这一特点,一方面有利于解决本文开始所指出的计算机专业学生较之非计算机专业学生,在应用开发工作中上手慢的问题;另一方面,使计算机专业的学生能在更大范围内、更高层面上掌握计算机应用技术。这一特点正是许多高等院校计算机专业教育改革追求的一个目标,也是国家教育部倡导软件学院的初衷之一。

“推荐教材”由基础知识、程序设计、应用技术、软件工程和实践环节等五个模块组成。各模块有其对应的培养目标与功能,从而构架出一个创新的、完整的计算机应用专业的课程体系。模块化的设计,使各学校可根据学生及学校的特点做自由的选择和组合,既能达到本专业的总体要求,又能体现具有特色的个性发展。整套教材的改革脉络清晰,结构特色鲜明,值得各高等院校在改革教学内容、编制教学计划、挑选教材书目时借鉴和参考。当然,很多书目也适合很多相关学科的计算机课程用作教材。

“推荐教材”的组成模块和书目详见封底。显然它不能说是完备的(实践环节模块更是如此),其改革的思路、改革的举措也可能有值得探讨的地方。我们衷心希望得到计算机教育界同仁和广大读者的批评指正。

高等院校计算机专业教育改革推荐教材

编委会

前 言

操作系统(OS)是计算机系统中的一个重要系统软件。本书遵循操作系统课程的教学大纲要求,以 Linux 作为应用实例,深入浅出地介绍了有关操作系统的基本原理。内容共分为 6 章。第 1 章概述了操作系统的形成、类型和功能等,介绍了几种目前常用的操作系统,尤其是 Linux 系统的功能;第 2 章为处理机管理,阐述了进程、线程、管程的基本概念以及进程互斥、同步、通信、调度、死锁的原理;第 3 章为存储器管理;第 4 章为设备管理;第 5 章为文件管理;第 6 章介绍了分布式操作系统。全书把目前流行的 Linux 操作系统贯穿在各个章节中,书后还附有 Red Hat Linux 7.3 的安装以及常用 Linux 命令供读者参考。

全书从内容上由浅入深、循序渐进地引导读者掌握现代操作系统的原理及其实现。注重简洁性、实用性、先进性、操作性。各章配备了相关的例子、习题。本书可作为大专学校和各类培训班的教材。

本书由徐汀荣教授主编,徐汀荣、徐云龙、孔芳、卢晓东编著,钱培德教授审阅。其中第 1 章由徐汀荣教授编写,第 2 章由徐云龙副教授编写,第 3,4 章及附录由孔芳老师编写,第 5,6 章由卢晓东老师编写,全书最后由徐云龙副教授润色。

在本书的编写过程中,还得到了中国人民解放军理工大学王元元教授、苏州大学计算机科学与技术学院、江苏省计算机信息处理技术重点实验室的大力支持,在此表示感谢。另外,苏州大学计算机科学与技术专业的金瑾、马利平等同学,也为本教材的编写提供了许多建议,在此也表示感谢。

本教材建议教学课时数为 56 学时,具体学时分配建议如下:

章	学 时 数
第 1 章 操作系统概述	4
第 2 章 处理机管理	12
第 3 章 存储器管理	12
第 4 章 设备管理	12
第 5 章 文件管理	8
第 6 章 分布式操作系统	8

建议实验环境软件配置要求:Red Hat Linux 7.3、C 语言。

对于书中存在的不妥和错误,敬请广大读者不吝指正。

编 者

目 录

编者的话

前 言

第 1 章 操作系统概述	1
1.1 操作系统的目标	1
1.2 操作系统作为用户/计算机接口	3
1.3 操作系统作为资源管理器	4
1.4 操作系统的发展	5
1.4.1 手工交互方式	5
1.4.2 手工批处理	6
1.4.3 第一代操作系统:单任务自动批处理操作系统	7
1.4.4 第二代操作系统:多任务和多用户操作系统	7
1.4.5 第三代操作系统:结构化和小型化操作系统	9
1.4.6 第四代操作系统:开放、并行、分布操作系统	10
1.5 操作系统实例简介	10
1.5.1 Linux 操作系统	11
1.5.2 DOS 和 Windows 操作系统	12
1.5.3 UNIX 操作系统	13
1.6 小结	14
1.7 习题	15
第 2 章 处理机管理	16
2.1 前趋图和程序的执行	16
2.1.1 前趋图	16
2.1.2 程序的顺序执行	16
2.1.3 程序的并发执行	17
2.1.4 程序并发执行的条件	18
2.2 进程的描述	19
2.2.1 进程的定义与特征	19
2.2.2 进程状态及其转换	20
2.2.3 进程控制块	21
2.3 进程控制	22
2.3.1 OS 内核	23
2.3.2 进程创建	24
2.3.3 进程撤消	24
2.3.4 进程阻塞与唤醒	25

2.4	线程	26
2.4.1	线程	26
2.4.2	线程与进程	26
2.5	进程互斥与同步	27
2.5.1	临界资源	27
2.5.2	临界区	29
2.6	信号量	30
2.6.1	整型信号量	30
2.6.2	记录型信号量	31
2.7	经典进程同步问题	32
2.7.1	生产者/消费者问题	32
2.7.2	读者/写者问题	34
2.7.3	哲学家进餐问题	35
2.8	进程通信	36
2.9	管程	37
2.9.1	管程的引入	37
2.9.2	基本概念	38
2.9.3	利用管程解决生产者/消费者问题	39
2.10	调度的类型和模型	40
2.10.1	调度类型	40
2.10.2	调度队列模型	42
2.10.3	选择调度方式和算法的若干准则	42
2.11	调度算法	44
2.12	死锁	48
2.12.1	产生死锁的原因	48
2.12.2	产生死锁的必要条件	49
2.12.3	处理死锁的基本方法	50
2.12.4	死锁的预防	51
2.12.5	死锁的检测	52
2.12.6	死锁的解除	53
2.13	Linux 的进程和线程管理	53
2.13.1	进程与线程的概念	54
2.13.2	进程的结构	54
2.13.3	进程的状态	55
2.13.4	Linux 的进程调度	55
2.13.5	Linux 的通信机制	58
2.14	小结	63
2.15	习题	63
第3章	存储器管理	65

3.1 存储管理的功能	65
3.1.1 存储空间的分配和回收	65
3.1.2 主存储器空间的保护	67
3.1.3 主存储器空间的共享	68
3.1.4 主存储器空间的扩充	68
3.2 单一连续存储管理	69
3.2.1 存储器的分配管理	69
3.2.2 存储保护	70
3.3 分区存储管理	70
3.3.1 固定分区存储管理	71
3.3.2 可变分区存储管理	72
3.4 分页存储管理	75
3.4.1 纯分页存储管理	75
3.4.2 请求式分页存储管理	78
3.5 分段式存储管理	85
3.5.1 分段式存储管理的基本原理	86
3.5.2 分段和分页的主要区别	89
3.5.3 分段式虚拟存储系统	89
3.5.4 分段存储管理的优缺点	90
3.6 段页式存储管理	91
3.6.1 实现原理	91
3.6.2 地址转换	91
3.6.3 段页式系统的优缺点	92
3.7 Linux 存储器管理	92
3.7.1 Linux 地址映射的过程	92
3.7.2 Linux 物理页面的分配	94
3.7.3 Linux 的页替换算法	95
3.7.4 Linux 内核存储器分配	95
3.8 小结	96
3.9 习题	96
第4章 设备管理	98
4.1 设备管理的基本功能	98
4.2 I/O 设备	98
4.2.1 I/O 设备分类	99
4.2.2 设备控制器	99
4.2.3 I/O 通道	100
4.2.4 I/O 控制方式	101
4.3 中断技术	103
4.3.1 中断的基本概念	103

4.3.2	中断的分类与优先级	104
4.3.3	中断处理过程	104
4.4	缓冲区技术	105
4.5	外围设备的分配	106
4.5.1	设备无关性	106
4.5.2	设备分配	107
4.6	外围设备驱动调度	111
4.6.1	设备驱动程序	111
4.6.2	设备驱动调度技术	112
4.7	Linux 设备管理	116
4.7.1	I/O 软件	116
4.7.2	设备驱动程序	117
4.7.3	Linux 设备驱动程序的几个通用函数	118
4.7.4	Linux 中的中断	118
4.8	小结	119
4.9	习题	119
第 5 章	文件管理	121
5.1	概述	121
5.1.1	文件的概念	121
5.1.2	文件的存储	122
5.1.3	文件系统的基本功能	124
5.2	文件的逻辑结构	125
5.3	文件的物理结构	126
5.3.1	磁带文件的组织	126
5.3.2	磁盘文件的组织	127
5.3.3	Linux 文件的结构	130
5.4	文件的目录结构	131
5.4.1	单级目录	131
5.4.2	二级目录	132
5.4.3	多级目录	133
5.4.4	Linux 目录分布	133
5.5	外存空间管理	134
5.5.1	空闲文件目录表	134
5.5.2	空闲块链表	135
5.5.3	空闲块索引表	135
5.5.4	空闲块成组链表	135
5.5.5	位示图	135
5.6	文件的使用	136
5.6.1	文件操作	136

5.6.2 文件的使用	137
5.7 文件的保护和保密	138
5.7.1 文件的共享	138
5.7.2 文件的保护	139
5.7.3 文件的保密	140
5.8 小结	141
5.9 习题	141
第 6 章 分布式操作系统	142
6.1 分布式系统的相关概念	142
6.1.1 集中式系统	142
6.1.2 网络系统	142
6.1.3 并行系统	143
6.1.4 分布式系统	143
6.1.5 常见应用模型	144
6.2 分布式进程管理	146
6.2.1 分布式系统中的处理器分配	146
6.2.2 分布式系统中的进程同步与通信	148
6.3 分布式文件系统	150
6.3.1 文件服务	150
6.3.2 目录服务	151
6.3.3 文件命名	152
6.3.4 文件共享	153
6.4 小结	154
6.5 习题	154
附录 A 安装 Red Hat Linux 7.3	155
附录 B 常用 Linux 命令	164
参考文献	174

第 1 章 操作系统概述

1.1 操作系统的目标

计算机系统是一个十分复杂的系统,要使其协调、高效地工作,必须有一套进行自动管理和便于用户操作的结构。操作系统(Operating System,简称 OS)就是用来管理计算机系统的资源、提高计算机系统资源利用效率、方便用户使用的程序集合。它是计算机系统进行自动管理的控制中心。

操作系统是计算机中最重要的系统软件,是计算机硬件(裸机)的直接外层,它是对计算机硬件进行的首次扩充。它的整体目标是负责完成所有与硬件因素相关的(硬件相关)和所有共需的(应用无关)基本工作,并解决这些基本工作中的效率和安全问题,为用户(操作和上层程序)能方便、高效、安全地使用计算机系统,而从最底层统一提供通用的帮助和管理。操作系统在计算机系统中的地位如图 1-1 所示。

那么,在用户操作和用户程序中,哪些工作内容是与硬件相关、与应用无关的呢?

概括地说,以下三方面的基础工作是与硬件相关、与应用无关的:

1) 负责启动执行每个用户程序,并负责完成用户程序的结束处理工作,使程序可以很方便、灵活地执行和中止。例如,通常启动一个程序只需在命令提示符后输入程序名即可,此外,操作系统还提供了批处理等多种启动方式。用户程序启动和结束过程涉及到对 I/O 设备、内外存的使用与处理、物理地址与物理接口寄存器的使用,这些都由操作系统加以处理。如果没有操作系统完成这些工作,用户操作和用户程序的结果将是所有程序都必须为片刻程序,即自包含引导装入代码,所有程序的启动操作都相当于一次关开机操作,因此麻烦、费时、不灵活。操作系统完成这些工作后,用户程序中不必包含启动自己执行的引导装入代码,不必包含与本程序具体功能无关的异常结束处理代码。所有结束出口处只需一条“返回系统”指令。用户不必考虑结束程序启动与结束,不必了解相关硬件的细节知识,不必频繁重复编写相应代码。

2) 每当用户程序使用到 I/O 设备和所存储的信息时,其中的硬件相关和应用无关工作都将通过调用操作系统来完成。用户程序运行过程涉及这些工作时,就通过一种特殊的调用方式(称为系统调用)或中断方式,调用操作系统来完成,从而为用户程序方便使用计算机系统,提供统一的帮助和管理。例如:在个人计算机上的用户程序需打印字符时,只需以一条 int 指令调用操作系统即可。若不编制操作系统,则打印一个字符需要设置和查看打印机接口寄存器的每一位,需要了解和使用打印机接口的 I/O 物理地址,导致用户程序对 I/O 设备的使用所涉及的代码经常占程序主体代码的大部分,这些代码中多处涉及接口寄存器,物理地址等复杂、繁琐的硬件细节,如果计算机系统换了一台不同型号的打印机,则程序中涉及打印输出的那些代码都要重新编写。因为任何用户程序的运行过程都需要使用内存、I/O 设备和外存信

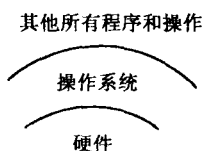


图 1-1 操作系统在计算机系统中的地位

息。只要使用 I/O 设备,就要与这些设备的接口寄存器打交道;只要使用内存,就要与物理地址打交道,这样所有程序中都包含硬件相关代码,程序员必须了解相应硬件细节知识,编写大量复杂代码。当硬件变化时,必须重新编写程序。引入操作系统后,用户程序中就不必包含硬件相关及应用无关代码,即不包含与接口寄存器、物理地址打交道的代码。用户不必了解相关硬件细节知识和编写相应的复杂、繁琐代码。

3) 操作系统对用户进行计算机基本操作,提供现成的使用程序和相应的管理,以使这些操作能很方便、有效地完成。这里所说的基本操作是指任何应用或开发下都通用的和普遍需要的,例如复制文件、删除文件、显示目录内容、显示文件内容、格式化磁盘等。例如: DOS 中为文件复制提供的现有程序有 COPY、DISKCOPY、XCOPY 等。用户需要复制时只需运行这些现有程序,不用自己编程。这些操作大量涉及对外存、I/O 设备的使用。任何一个应用程序,不管是财务应用或是人口统计应用,都需要进行这些文件复制工作,而复制一个财务总账文件和复制一个人口清单文件,其复制过程都是相同的,只不过被复制的数据不同,为避免用户为这些常用操作自行编制程序,系统提供了这些与硬件相关、应用无关的功能。因此,操作系统的第三个功能是为用户的常用基本操作提供现有程序,用户不必为这些常用操作亲自编程,从而有利于提高用户软件开发的效率。

4) 操作系统的第四个功能是改善上述三方面工作中的效率和安全问题,使计算机系统的各个部分和计算机系统得到高效、安全的使用。因为每个用户都希望提高使用效率并希望自己的数据有安全保障,每个机器主人都希望提高机器利用率,由用户自行解决计算机系统的所有安全和效率问题是不现实的,为此,要求操作系统提供一切可能的方法使计算机系统得到高效、安全的使用。

以上四方面工作似乎无关联,但他们都具有这样的共性:硬件相关和应用无关。

所谓一个(或一段)程序是硬件相关的,是指该程序代码中包含内外存及设备的物理地址,包含对设备接口寄存器和设备接口缓冲区的读写等操作。硬件相关的代码必然随硬件的变化而变化。这样的硬件变化包括内外存物理存储空间大小的变化,程序和数据在内外存物理存储空间中存放位置的变化,设备数量和类型的变化等等。这里所说的操作系统承担硬件相关工作,是指其上层的用户程序都是硬件无关的,即当被用户程序使用的硬件发生变化时,用户程序不必改变,人的操作更不必改变。

一个工作是应用无关的,是指不管用计算机来做什么,不管在计算机上运行什么程序,只要使用相应硬件或相应信息时就要涉及到的工作,它是用户共需的,且工作过程相同,有共性可循,却又与应用问题没有直接关系。

操作系统为用户的操作和程序完成所有硬件相关和应用无关的工作,目的和益处是显而易见的。硬件相关,必然意味着复杂繁琐、代码量很大、代码不通用和变化大,需要用户投入大量的精力来设计实现和维护修改,并了解相应大量硬件的细节知识,因此有必要统一管理,使用户摆脱负担。应用无关,就意味着更有必要统一管理(因为普遍和频繁涉及,与具体应用无直接关系)和能够统一管理(因为工作过程相同)。越是计算机使用中底层的、基本的工作,越具有硬件相关和应用无关的特点,对用户和系统的方便、效率、安全影响越大,越需要并可能由操作系统来完成,并解决其中的效率和安全问题,从而使用户程序成为硬件无关的,即独立于硬件,不包括硬件相关的代码,硬件改变时程序不必改变。使用户不必考虑这些底层基本使用过程和了解相关的硬件细节知识,避免频繁重复编写(或编译)与应用问题本身关系不大的大

量复杂繁琐的代码,专心于应用本身,更好地达到最终的具体应用要求和目标,总之,操作系统为保证用户的操作和程序最终使用的方便、效率、安全,承担了所有硬件相关和应用无关的工作,从最底层提供统一的帮助和管理。

操作系统为用户完成所有硬件相关和应用无关的工作,并不意味着操作系统本身的所有功能都是硬件相关的。操作系统分为实用程序层、命令解释层、核心层,其中只有核心层才是硬件相关的,甚至在核心层内部也做了进一步的隔离,即只有核心层中最底层的一些模块(例如 Windows NT 中的硬件抽象层、设备驱动程序等)才是硬件相关的。操作系统有狭义(核心)与广义(实用程序层,甚至更高层——产品捆绑)之分的说法,也缘于此。

1.2 操作系统作为用户/计算机接口

操作系统是计算机软件的核心。它在计算机系统中担负着管理系统资源、控制输入输出处理和实现用户和计算机系统间通信的重要任务。操作系统的功能可以从不同的角度来理解,例如可以从方便用户的角度或者从资源管理的角度理解。

计算机的界面是否友好,与操作系统人机交互功能的完善与否密切相关。人机交互功能主要靠输入输出的外部设备和相应的软件来完成。这些外部设备主要有键盘、显示器、鼠标等。驱动这些设备进行工作的软件,就是操作系统提供用户进行人机交互功能的工具。这些软件的主要作用是控制有关设备的运行,理解并执行通过人机交互界面传来的各种命令和要求。目前广泛使用的有键盘、鼠标和显示器等,近年来语音输入设备、文字读入设备、图形图像扫描输入设备等的使用也越来越普遍,使得计算机系统人机交互的功能越来越强。

下面,我们以程序启动为例,分析一下操作系统所做的工作。

用计算机做任何事,都需要先运行某个相应的程序。那么程序如何开始执行呢?从硬件知识可知,任一(可执行目标)程序的执行有两个前提条件:一是读程序(该程序已装入内存),二是 CPU 中的程序计数器 PC 被置为该程序在内存的起始执行地址(这就意味着 CPU 执行的下一条指令是该程序的指令)。这样,程序很自然就开始运行了。那么,一个程序是由谁来装入和启动的呢?关于“启动一个程序执行”的申请又是怎样提出的呢?

1. 程序的第一种启动方式——鼠标点击方式

在 Windows 系列操作系统下,通常通过双击一个程序的图标或程序名来运行该程序。这个图标或程序名通常出现在桌面、桌面上的开始菜单、资源管理器等处。程序图标或程序名的显示、双击动作的接收和翻译以及找到并启动执行相应的程序,都属于 Windows 操作系统的功能,通常由 Windows 操作系统中的资源管理器来完成。

在 Windows 中,一个正在运行的程序通常与一个窗口相联系,即启动了一个程序,就打开了一个相应的窗口(虽然这个窗口有时会被最小化或覆盖),关闭了这个窗口就关闭了这个程序。

这种鼠标操作方式和窗口界面形象生动,操作简单而有规律,用户不必记忆命令,是当今所有操作系统中最常有的程序启动方式。Linux 和 UNIX 操作系统中也提供了这种方式。

2. 程序的第二种启动方式——命令方式

在所有的操作系统下,都提供了命令方式来启动一个程序。

在 Windows 系列操作系统下,可以点击开始→运行→输入程序名(及其所有参数)来启

动一个程序,这称之为 Windows 程序名方式;也可以点击开始→程序→MSDOS 方式→输入程序名(及其所在路径与参数)来启动一个程序,这称之为 MS DOS 命令方式。MS DOS 是 Windows 98/95/3.x 的前身和基础,是本世纪 80 年代 IBM PC 机上最流行的操作系统。在 Linux 操作系统下,开机后通常默认进入的就是命令方式,标志是命令提示符(通常为%或#),在命令提示符后面输入程序名,就可以执行程序。命令提示符是操作系统向用户提供的一种提示标志,它的出现就表示“你可以在命令提示符后面输入程序名来启动该程序”。通常把命令提示符后面输入的程序名和参数称为命令行或命令。任一命令行(即任一命令)的本质都是申请某一程序执行。命令提示符有两重含义:一是标志着上一个程序的结束,也可以说上一条命令的结束;二是标志着可以启动运行下一个程序,即可以输入下一条命令。既然在开机通电状态的任一时刻都有程序在运行,或者说有指令在执行,那么在命令方式下,当命令提示符后为空(即光标紧挨在命令提示符后面闪烁)而操作者长久未归时,机器是在运行程序吗?如果是的话,是什么程序?答案是命令解释器程序正在运行,即正在等待用户输入命令(例如通过一个循环来等待,这个循环不停地检查是否有键盘中断,即是否有用户开始输入命令)。

通常把操作系统实现命令方式(以及批处理方式)的那部分程序称为命令解释程序或命令解释器。Windows 中 MS DOS 方式的命令解释器是 command.com,可以在 Windows 启动盘的根目录下看到这个文件。Linux 操作系统中的命令解释器称为 Shell。命令方式在 20 世纪 80 年代是启动程序运行最主要的方式,但如今它已经被鼠标点击方式代替而成为第二类常用的程序启动方式了。从作业管理角度说,鼠标点击方式与命令启动方式都属于交互式作业方式。

1.3 操作系统作为资源管理器

这里所谓“资源”是指计算机系统的硬件(含处理机、存储器、输入输出设备)和以文件形式存放在计算机中的信息。这些资源是操作系统管理的对象。也就是说,系统的硬件资源和软件资源都由操作系统根据用户需求按一定的策略分配和调度。因此,从资源管理的角度来看,操作系统的功能主要有以下四个方面:

1. 处理机管理

操作系统的重要任务是控制程序的执行。它负责对系统中各个处理机及其状态进行登记,管理各程序对处理机的要求,并按照一定的策略将系统中的各个处理机分配给申请的用户作业(进程)。目前计算机系统中所使用的大多数是“多任务”、“多线程”的操作系统。所谓“多任务”是指操作系统同时运行多个任务(多个程序)。例如:当一个报表正在打印时,CPU 可以同时运行另一个时钟程序和另一个大型科学计算任务。所谓“多线程”是指一个程序的多重执行能力,例如当计算机在写入一个大型文件时可以同时调用拼写检查程序。

2. 存储管理

操作系统按照一定的策略为用户作业分配内存存储器的存储空间,记录内存存储器的使用情况,并对内存存储器中的信息提供保护,在该作业执行结束后将它占用的内存单元收回以便其他程序使用。

3. 输入输出设备管理

由于输入输出设备管理的工作速度远远低于 CPU,所以操作系统应该对设备的输入输出

性能有很清晰的分类,以便当外部有输入输出要求时能及时地响应。操作系统记录系统中各个输入输出设备的状态,按照各个设备的不同特点采取不同的分配策略,并按照用户的要求控制外部设备的运行。当用户程序使用完毕后立即回收外部设备以便给另一个需要的用户使用。对于存储型外部设备,操作系统则提供存储空间给用户存放各类信息。

4. 文件管理

操作系统的文件管理功能是对存放在计算机中的信息进行逻辑组织和物理组织、维护文件目录的结构以及实现对文件的各种操作。例如,操作系统可以向用户提供创建文件、撤消文件、读写文件、打开文件和关闭文件等各种操作。有了操作系统的文件管理功能,用户就可以按照文件名存取数据而不必了解这些数据的确切物理位置。这不仅便于用户的操作使用而且还有利于用户间的共享数据。另外,操作系统的文件管理功能还可以允许用户在创建文件时规定文件的使用权限。这样,数据的安全性也可以得到保证。

1.4 操作系统的发展

操作系统最早产生于1955年,比计算机的出现晚了10年,经历了可粗略计算的以十年划分的四代,第一代单任务操作系统,出现在1955~1962年;第二代多任务和多用户操作系统,出现在20世纪60、70年代;第三代结构化与小型化操作系统,出现在20世纪70、80年代;第四代并行与分布式操作系统,出现在20世纪80年代末。

学习操作系统发展历史,尤其是发现和理解操作系统发展的原因、动机、契机,看到操作系统的各种组成成分和技术,认识操作系统是如何作为早期计算机系统中的问题的自然解决方案而提出与发展的,有助于加深对操作系统基本问题的理解。在学习时要注意操作系统与计算机使用方式二者之间密不可分的关系,因为操作系统决定用户对计算机的使用方式,二者是同步发展的。

上述四代操作系统是按照功能、规模、结构来划分的。其实,对操作系统的划分,角度是很多的,可以按功能、规模划分,也可以按外在表现划分,不同的划分并不完全相同甚至差异很大,但上述划分应能表现其发展过程的本质规律。

1.4.1 手工交互方式

1946年人类第一台电子计算机问世以后,操作系统并未随之出现,那时的机器非常原始,程序通常是在一排排的机械开关上一次一位地输入。手工交互阶段是计算机刚出现的5年左右,接着有了穿孔卡片或纸带,机器语言程序可以通过卡片或纸带输入。这一阶段对使用方式有影响的背景特点主要是没有系统软件,也没有像磁带这样的大容量快速设备。该过程具有如下特点:程序设计都是直接编制二进制目标程序,没有源程序的概念;输入输出设备主要是纸带和卡片,程序员在上机前必须将程序与数据通过纸带穿孔机或卡片穿孔机送到纸带或卡片上;程序员在上机前必须预约机时,可能要等待几天的时间;程序员自己上机操作,用户既是程序员也是操作员;程序的启动与结束处理都以手工方式进行,即手工装带卸带,通过操作控制台面板上的指示灯和按钮,操纵程序的装入、启动与结束(可能需要设置起始地址);程序员的操作以交互方式进行,程序员可以从控制台指示灯和打字机得到运行结果,并可以马上通过控制台按钮修改内存目标程序,再运行,再修改,直到程序运行正确,或到达预约时间,或因其