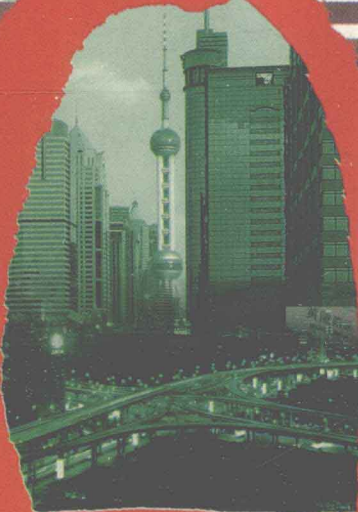


高等职业技术院校教学用书

计算机操作系统

(Win95/Win98/Linux)

上海第二工业大学



夸

越

世

纪

之

门

中国民航出版社

高等职业技术院校教学用书

计算机操作系统

上海第二工业大学

编者：周敬贤、陈敏超、陆虹

策划：何学仪

中国民航出版社

图书在版编目(CIP)数据

计算机系统/周敬贤等编著

——北京:中国民航出版社,1999.10

ISBN 7-80110-371-8

I.计… II.周… III.操作系统(软件)-高等教育:职业教育-教材
IV.TP316

中国版本图书馆 CIP 数据核字(1999)第 66871 号

计算机系统

周敬贤等编著

*

中国民航出版社出版发行

(北京市朝阳区光熙门北里甲 31 号楼 5 层)

上海市长城绘图印刷厂

开本:787×1092 1/16 印张:22.5 字数:540 千字

1999 年 10 月第 1 版 1999 年 10 月第 1 次印刷 印数:1~5,000 册

ISBN 7-80110-371-8/G·173 定价:26.00 元

(发行电话:(021)63052990 本书如有印装错误,我社负责调换)

前 言

随着信息时代的到来,计算机系统特别是微型计算机系统的应用已深入到社会的各个领域。近年来,微机操作系统的更新速度很快。

操作系统是计算机系统中重要的系统软件,操作系统课程是计算机专业的重要专业的主干课程。让这门课程的教学内容跟上技术更新的潮流,是高等学校计算机专业的急需,更是培养应用型人才的高等职业教育的急需。我们在多年应用实践和教学实践的基础上,针对高等职业教育的培养目标编写了这本教材。

本教材兼顾基础理论和实际应用,在理论教学的基础上,再以目前流行的、有发展潜力的微机操作系统为实例,既能加深对基本原理的理解,又能提高实际应用的能力。

本教材由三篇组成。

第一篇介绍操作系统的基本原理,包括操作系统的形成、特征和类型,进程和线程的概念及进程的同步机制,资源管理,虚拟存储器,文件系统,磁盘存储器管理等。

第二篇介绍 Windows 操作系统。从微机多任务操作系统的概念、Windows 的发展介绍了 Windows 的基本知识;从 Windows 的体系结构、注册表、内存管理、文件和磁盘管理介绍了 Windows 95/98 的内部机制。其中穿插着一些“自己动手”的内容,可以让学生通过实际操作来加深对概念的理解。

第三篇介绍 Linux 操作系统。从版本知识、系统安装、使用方法,到运行机制,较详细地作了系统的介绍,其间穿插着与理论要点相联系的解释和分析,同时介绍了实际操作和观察的实例。

以上内容在教学中可以根据课时和学生情况裁剪使用。

本书的第一篇由周敬贤编写,第二篇由陈敏超编写,第三篇由陆虹编写。

编写这样的教材,可供借鉴的资料不多;对深浅难易、理论实际等矛盾的处理分寸见仁见智,尚无充分的实践检验结论;加之作者水平有限,处理不当甚至差错之处在所难免,敬请读者批评指正。

目 录

第一篇 基本理论

第一章 操作系统引论

1.1 什么是操作系统	1
1.1.1 操作系统是硬件的延伸	1
1.1.2 操作系统是系统资源的管理者	2
1.1.3 操作系统的定义	2
1.2 操作系统的形成	2
1.2.1 手工操作阶段	3
1.2.2 监督程序(早期批处理)阶段	3
1.2.3 操作系统的形成	3
1.3 操作系统的特征	4
1.4 操作系统的类型	4
1.4.1 传统操作系统的基本类型	5
1.4.2 微机操作系统	6
1.4.3 网络操作系统	8
1.4.4 分布式操作系统	9

第二章 进程

2.1 进程概念	11
2.1.1 进程的提出	11
2.1.2 进程的定义	11
2.1.3 进程的特征	12
2.2 进程的描述	12
2.2.1 进程的表示	13
2.2.2 进程的基本调度状态	15
2.2.3 进程状态间的转换	15
2.2.4 进程的挂起状态	16
2.2.5 UNIX 的进程状态	16
2.3 进程控制	17
2.3.1 进程控制机构	17
2.3.2 进程控制原语	17
2.4 线程概念	18
2.4.1 线程概念	18
2.4.2 为什么要引入线程	19
2.4.3 线程和进程的比较	20
2.5 并发程序设计	21
2.5.1 前趋图和程序的执行	22
2.5.2 进程间的同步与互斥	25

2.5.3 同步机制	26
2.5.4 进程互斥的实现	27
2.5.5 简单同步问题的实现	29
2.5.6 经典同步问题	30
2.6 进程通信	33
2.6.1 共享存储器系统	34
2.6.2 管道通信	35
2.6.3 消息系统	36

第三章 系统的资源管理

3.1 概述	45
3.1.1 资源管理的目的和任务	45
3.1.2 资源共享和资源分配	46
3.1.3 资源管理的结构和策略	49
3.1.4 系统中资源管理的分类	51
3.2 处理机管理	52
3.2.1 处理机的多级调度	52
3.2.2 处理机调度模型	53
3.2.3 调度算法	55
3.2.4 多处理机调度	60
3.3 存储器管理	62
3.3.1 基本概念	62
3.3.2 地址空间管理	66
3.3.3 存储空间管理	71
3.3.4 几种存储管理方法	75
3.3.5 高速缓冲存储器	86
3.4 设备管理	89
3.4.1 设备的分类	89
3.4.2 设备管理的功能和任务	90
3.4.3 数据传送控制方法	91
3.4.4 缓冲管理	94
3.4.5 设备分配	96
3.4.6 设备驱动	98
3.5 死锁	99
3.5.1 产生死锁的原因	99
3.5.2 死锁例举	100
3.5.3 死锁的描述	102

3.5.4 产生死锁的必要条件和死锁的预防 …	102	5.3.1 文件的逻辑结构 ……………	138
3.5.5 死锁的避免 ……………	104	5.3.2 文件和物理结构 ……………	138
第四章 虚拟存储器管理		5.4 文件共享 ……………	142
4.1 虚拟存储器的基本概念 ……………	112	5.4.1 “当前目录”法 ……………	142
4.1.1 程序运行的局部性原理 ……………	112	5.4.2 连访文件法 ……………	143
4.1.2 虚拟存储器概念 ……………	113	5.4.3 基本文件目录法 ……………	143
4.1.3 虚拟存储器的实现 ……………	114	5.4.4 符号链接法共享文件 ……………	144
4.2 请求页式存储管理 ……………	114	5.5 文件保护 ……………	145
4.2.1 请求页式的页表机制 ……………	115	5.5.1 影响文件安全性的因素 ……………	145
4.2.2 缺页中断机构 ……………	115	5.5.2 文件的存取控制 ……………	145
4.2.3 页面淘汰算法 ……………	115	5.5.3 文件的分级安全管理 ……………	148
4.2.4 性能分析 ……………	119	第六章 磁盘存储器管理	
4.2.5 请求页式系统中的页面分配 ……	120	6.1 磁盘存储器概述 ……………	151
4.3 请求段式存储管理 ……………	120	6.1.1 磁盘的结构 ……………	151
4.3.1 请求段式的段表机制 ……………	120	6.1.2 磁盘上的信息组织 ……………	152
4.3.2 缺段中断机构 ……………	121	6.1.3 物理扇号和逻辑扇号 ……………	153
4.3.3 段的共享 ……………	122	6.1.4 磁盘的访问时间 ……………	153
4.4 32位微机的虚拟存储机构 ……………	122	6.2 驱动调度算法 ……………	154
4.4.1 保护虚拟地址模式 ……………	122	6.2.1 先来先服务(FIFS)算法 ……………	154
4.4.2 32位芯片的地址变换机制 ……	122	6.2.2 最短寻道时间优先(SSTF)算法 ……	155
4.4.3 段机制 ……………	123	6.2.3 扫描(SCAN)算法 ……………	155
4.4.4 页机制 ……………	125	6.2.4 循环扫描(C-SCAN)算法 ……	156
第五章 文件系统		6.2.5 其它改进调度算法 ……………	157
5.1 文件和文件系统 ……………	130	6.3 磁盘空间的管理 ……………	157
5.1.1 文件(File) ……………	130	6.3.1 硬盘空间的分区 ……………	157
5.1.2 文件的分类 ……………	131	6.3.2 空闲存储空间的管理 ……………	158
5.1.3 文件系统 ……………	131	6.3.3 磁盘空间的分配方法 ……………	161
5.2 文件目录 ……………	132	6.4 磁盘容错技术 ……………	163
5.2.1 目录组成和管理的基本要求 ……	132	6.4.1 低级磁盘容错技术 ……………	163
5.2.2 简单一级目录结构 ……………	134	6.4.2 中级磁盘容错技术 ……………	164
5.2.3 二级目录结构 ……………	134	6.4.3 后援备份 ……………	165
5.2.4 多级目录结构 ……………	135	6.5 磁盘性能的改善 ……………	167
5.2.5 目录管理 ……………	137	6.5.1 磁盘高速缓存 ……………	167
5.3 文件的逻辑结构和物理结构 ……	138	6.5.2 数据分布优化 ……………	169
		6.5.3 提高磁盘 I/O 速度的其它措施 ……	170

第二篇 Windows 操作系统

第一章 概述

1.1 微机多任务操作系统概述 ……………	172	1.1.3 多任务和多线程 ……………	173
1.1.1 内在式与贴附式 ……………	172	1.2 Windows 概述 ……………	174
1.1.2 先决式、协作式和抢占式 ……	172	1.2.1 Windows 的发展 ……………	174
		1.2.2 Windows 的运行模式 ……………	177

1.2.3 几种 Windows 的比较 178

第二章 Windows 的启动

2.1 Windows 98 的启动文件 180

2.2 Windows 98 的启动顺序 183

2.3 Windows 98 的启动菜单 184

第三章 Windows 的体系结构

3.1 操作系统组件 186

3.2 Intel 保护环体系结构 187

3.3 虚拟机 187

3.3.1 系统虚拟机 188

3.3.2 MS - DOS 虚拟机 188

3.4 消息系统 189

3.5 多任务的调度 191

3.6 设备驱动程序 191

3.6.1 驱动程序的类型 192

3.6.2 虚拟设备驱动程序 (VxD) 193

3.6.3 通用驱动程序/微型驱动程序 194

3.6.4 WIN32 驱动程序模型 (Win32 Driver Model, WDM) 194

3.7 Windows API 层 195

第四章 配置管理

4.1 注册表概述 197

4.2 注册表的历史 198

4.2.1 Windows 3.1 198

4.2.2 Windows NT 199

4.2.3 Windows 95 199

4.2.4 Windows 98 199

4.3 Windows 95/98 中的 INI 文件 199

4.4 Windows 95/98 中的注册表 200

4.4.1 启动注册表编辑器 200

4.4.2 注册表的组成文件 201

4.4.3 注册表的键和键值 202

4.4.4 注册表的主键及内容 203

4.4.5 使用注册表编辑器 208

第五章 内存管理

5.1 不同类型的内存 211

5.1.1 常规内存 211

5.1.2 上位内存 211

5.1.3 扩展内存 212

5.1.4 扩页内存 212

5.1.5 虚拟内存 212

5.2 Windows 95/98 中 32 位内存模式 212

5.2.1 系统的地址空间 212

5.2.2 虚拟内存管理机制 213

第六章 磁盘和文件管理

6.1 磁盘和文件系统结构 219

6.1.1 磁盘存储空间的划分 219

6.1.2 主引导记录和引导记录 219

6.1.3 文件的存储结构 220

6.1.4 文件分配表 221

6.1.5 文件目录 222

6.1.6 长文件名 223

6.1.7 归纳与举例 225

6.1.8 优化分区和簇的大小 226

6.1.9 建立硬盘分区 227

6.2 磁盘管理 230

6.2.1 用 SCANDISK 分析并修复磁盘 230

6.2.2 用 DEFRAG 去除磁盘碎片 234

第三篇 Linux 操作系统

第一章 Linux 系统概述

1.1 Linux 的历史、现状和发展 236

1.2 Linux 的特点 237

1.3 Linux 的软件支持 238

1.4 Linux 与程序设计语言 238

第二章 安装 Linux

2.1 Linux 的环境需求 240

2.2 Linux 的安装界面 242

2.3 Linux 的安装 242

2.3.1 安装程序启动 242

2.3.2 选择一种语言 243

2.3.3 选择一种键盘类型 243

2.3.4 选择安装或升级 244

2.3.5 从 CD - ROM 安装 244

2.3.6 选择安装或升级 245

2.3.7 选择安装类别 246

2.3.8 SCSI 的配置 246

2.3.9 为 Red Hat Linux 创建分区 246

2.3.10 格式化交换 (Linux Swap) 空间 248

2.3.11 格式化分区	249	3.3.11 Shell 编程	287
2.3.12 选择需要安装的包	250	第四章 Linux 系统内核	
2.3.13 选择鼠标	251	4.1 Linux 系统的体系结构	296
2.3.14 Y Windows 的基本安装	251	4.2 进程管理	296
2.3.15 选择显示卡的时钟芯片	254	4.2.1 Linux 进程及线程	297
2.3.16 网络配置	255	4.2.2 Linux 的进程调度	300
2.3.17 设定时区	255	4.2.3 Linux 的通讯机制	302
2.3.18 选择启动时的服务程序	256	4.2.4 Linux 系统的进程管理操作	305
2.3.19 选择与配置打印	256	4.3 内存管理	311
2.3.20 设置 root 口令	259	4.3.1 Linux 的虚拟内存	312
2.3.21 创建启动盘	260	4.3.2 内存页表	313
2.3.22 LILO 的安装	260	4.3.3 分配与释放	313
2.3.23 为 LILO 启动命令行增加选项	261	4.3.4 缺页中断与页面置换	314
第三章 Linux 的基本使用		4.4 设备驱动	317
3.1 Linux 的基本命令	263	4.4.1 中断及中断处理	317
3.2 Linux 的 Vi 编辑器	276	4.4.2 设备驱动程序	318
3.2.1 Vi 功能简介	276	4.4.3 常用硬件设备的使用	321
3.2.2 光标的移动	277	4.5 Linux 文件系统	324
3.2.3 命令列表	278	4.5.1 EXT2 文件系统	325
3.3 Linux 的 Shell 反 Shell 编程	278	4.5.2 虚拟文件系统	336
3.3.1 Shell 的命令补全和通配符	279	4.5.3 缓冲区高速缓存	341
3.3.2 Shell 的命令历史记录	280	第五章 Linux 的引导与关闭机制	
3.3.3 Shell 的别名	281	5.1 Linux 系统的引导	343
3.3.4 标准输入与标准输出	282	5.1.1 BIOS 启动主引导程序	343
3.3.5 输入重定向	282	5.1.2 Linux 内核和装载	343
3.3.6 输出重定向	283	5.1.3 内核初始化系统硬件	343
3.3.7 使用管道	284	5.1.4 挂装 root 文件系统	344
3.3.8 作业控制	284	5.1.5 init 过程	344
3.3.9 提示符	285	5.2 Linux 系统的关闭	347
3.3.10 配置自己的 bash 及 bash 命令、变量 一览表	285	参考文献	347

第一章

操作系统引论

电子计算机发展到今天,不论是简单的个人计算机系统,还是十分复杂的计算机系统,毫无例外地都配置一种或几种操作系统(Operating Systems)。如果让用户离开操作系统直接去使用计算机,那是十分困难的,也很难想象的。

操作系统是系统硬件平台上设置的第一层软件,它在整个计算机系统中具有特殊重要的地位,它不仅是硬件和用户间的一个接口(界面),使用户利用操作系统所提供的命令和服务去控制和使用计算机,摆脱直接使用硬件的繁琐和困难,而且,操作系统也是硬件平台和所有其它软件(包括系统软件和应用软件)的接口,如编辑程序、汇编程序、语言编译程序、字处理程序、数据库管理系统、以及大量其它应用软件都是建立在操作系统基础上的,并得到它的支持和取得它的服务。

总之,操作系统是现代计算机系统中必不可少的最重要的系统软件。

1.1 什么是操作系统

尽管很多计算机用户已具有使用操作系统的经历或经验,但什么是操作系统也许知之甚少,这里可从两个不同的方面来看操作系统究竟是什么。

1.1.1 操作系统是硬件的延伸

可以把一个计算机系统看成是由硬件和软件按层次构成的系统,如图 1-1 所示计算机在机器语言一级的体系结构,大多数是很原始的,对于程序设计,尤其是 I/O 程序设计,变得十分困难。例如,微机系统普遍使用的软盘驱动器,一般有十几个命令,如读和写数据,移动

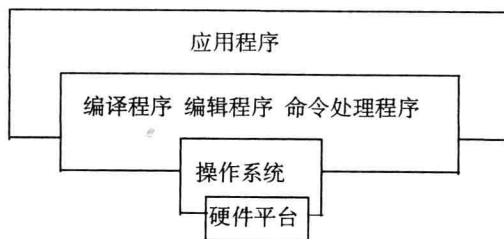


图 1-1 计算机系统硬件和软件层次结构

读/写头,格式化磁道,以及控制器和驱动器的初始化,检测,复位等。其中最基本的读和写命令,一般需要十几个参数,用来指定磁盘存储块的地址、在物理介质上的记录方式、每一磁道扇区数、以及扇区间的空隙等。对于一般程序设计人员来说,要求掌握这些细节是不现实的。相反,他们希望有一个简单且有效的高层次的抽象。例如,把磁盘抽象为一个文件卷,其中包含一批命名的文件,对文件可以打开供存取之用,然后可以读写,最后关闭之。而无

需关心磁盘上数据存放的物理位置、各磁道的扇区号、扇区之间的间隙、控制器返回的状态和错误字段、甚至驱动器电机是否已经启动、启动时间长短等硬件细节。

能向程序员隐蔽裸机硬件细节和特征的系统软件就是操作系统。操作系统为用户提供了一个十分友好的界面,它不仅能让用户以文件形式来方便地使用磁盘,而它们所提供的高级命令,可对程序进行装入、控制、打印和执行等操作,而无须了解在硬件层次上的复杂细节。所以,建立在硬件层上的操作系统实际上已把硬件平台改造为一台功能更强,使用更方便的扩展机器(Extended Machine)或称为虚拟机(Virtual Machine)

此外,大多数操作系统还提供一个多用户环境,但它给每一个用户的感受,好象只有该用户在独占使用系统。

1.1.2 操作系统是系统资源的管理者

从“自上向下”观点看,操作系统是向用户提供一个方便的界面,即是一台功能更强的虚拟机,但从另一个“自下向上”的观点来看,操作系统却是一个系统资源的组织和管理者。

一个计算机系统包含的硬件、软件资源可分为以下几部分:处理机(CPU),存储器(内存和外存或称为主存和辅存)、外部设备和信息(文件)。现代计算机系统都支持多用户多任务运行环境,面对众多的程序(作业)争夺处理机、存贮器、各种 I/O 设备,以及共享软件资源,操作系统必须进行统一协调,并有条不紊地进行资源分配。操作系统必须记录谁在使用什么样的资源,系统中还有哪些资源空闲以及当前响应谁对资源的请求,以及回收哪些不再使用的资源等。操作系统还要提供一定的机制去协调程序间的竞争与同步,提供机制对资源进行合理使用,施加保护,以及采取虚拟技术来“扩充”资源等等。

总之,操作系统是一个资源管理者,从这一角度看,操作系统是四大资源管理程序的集成。

1.1.3 操作系统的定义

虽然操作系统存在已几十年了,且任何计算机系统都配置一种或多种操作系统,但至今世界上并没有一个被普遍接受的对操作系统的定义。对操作系统的定义有各种说法,不同说法反映了人们从不同视角去揭示操作系统的本质特征。

当前,世界上广为流行的观点是把操作系统看成为计算机系统的资源管理者,也就是说操作系统的主要任务是管理并调度计算机系统资源的使用。

但从用户视角看操作系统,则它应能为用户提供比裸机功能更强,服务更完善,使用户更觉方便的灵活的虚拟机器。对于程序设计人员,操作系统应提供一个开发其它系统软件和应用软件的良好环境,即提供一组功能强大的,方便使用的广义指令(系统调用)。对一般用户,操作系统应成为一个友好的软件运行环境,即命令界面(shell)。用户通过终端或其它方法使用这些命令来控制自己作业的运行。

综上所述,我们可以非形式地把操作系统定义为:操作系统是计算机系统中的一个系统软件,它是管理和控制计算机系统硬件和软件资源,合理地组织计算机工作流程,以及方便用户使用的程序的集合。

1.2 操作系统的形成

操作系统是在人们不断地改善计算机系统性能和提高资源利用率的过程中,逐步地形成和发展起来的。

1.2.1 手工操作阶段

早期计算机的工作,基本上采用人工控制方式:由操作员(也就是用户)通过手工操作方式去启动设备和主机运行。每次只能一个用户使用计算机,一切资源全部由该用户所占用。并且在一个作业运行过程中,以及在作业完成后转换到另一作业都是由人工干预的。显然手工操作方式,资源利用率低下,但对早期的计算机来说,此矛盾尚不突出,因为计算机本身运算速度低,拥有的资源很少。随着计算机速度的提高,规模的扩大,效率的矛盾变得严重起来,甚至可能使资源利用率下降为百分之几,甚至更低。

1.2.2 监督程序(早期批处理)阶段

随着 CPU 速度的大幅度提高,使得手工操作设备与计算机计算速度不匹配的矛盾日益突出。例如:第一代计算机用一小时运行一个用户程序,其中人工干预时间为三分钟。如果计算机速度仅提高十倍,再运行这一程序也许需要六分钟,而手工干预的时间不变,仍为三分钟,这相当于手工干预占据了运行时间的 50%。而且,事实上 CPU 速度的提高远远大于十倍。因此,人们设计了监督程序(管理程序),来实现作业的自动交接处理。首先,每道作业由程序员提供一组记录在某种介质上的作业信息(文件),它们包括相应程序、程序所需的数据文件,以及用作业控制语言(JCL)编写的,反映程序员对作业控制意图的作业说明书,并提交给系统操作员。而操作员将作业“成批”地输入到计算机中,由监督程序识别一个作业,进行处理后再取下一个作业。这种自动定序的处理方式称为“批处理”方式。而且,由于是一个接一个地串行执行作业,因此称为“单道批处理”。由此可见,这一阶段实现了作业和作业的自动交接,缩短了作业建立时间和手工操作时间,从而也提高了系统资源的利用率。

1.2.3 操作系统的形成

到了 60 年代,硬件技术取得了两个方面重大进展:一是通道技术的引进,二是中断技术的发展,使得通道具有中断主机工作的能力。通道是专门用来控制输入输出设备的处理机,在中断技术的配合下,它能与主机并行工作,这样可把原来由主机直接控制的输入输出工作转交给通道去完成,使得宝贵的中央处理机(主机)的时间都用来进行主要的数据处理工作。这时,原来的监督程序的功能扩大了,它不仅负责作业运行的自动调度,而且还要提供输入输出控制功能。这个发展了的监督程序常驻内存,称为执行系统(Executive System)。

但是,这时计算机系统运行的基本特征仍是单道顺序执行。那么可能会出现两种情况:对于以计算为主的作业,输入输出量少,外围设备空闲;然而对于以输入输出为主的作业,CPU 会造成空闲。总之,单道系统计算机资源利用率是不高的。为了改善 CPU 和其它系统资源的利用率,而引入了多程序设计技术。

所谓多道程序设计是指“把一个以上的程序存放在主存中,并且同时处于运行状态。这些程序共享处理机时间和其它系统资源”。当一道程序因 I/O 请求而暂停执行时,CPU 便立即转去执行另一道程序。这样不仅使 CPU 时间能得到充分利用,同时也提高了其它资源的利用率。对于一个传统的单处理机系统来说,“多个程序同时处于运行状态”这显然只是一个宏观上的概念,其意义只是指每个程序都已开始运行,但尚未完成。就微观上来说,在任一特定时刻,在处理机上真正运行(执行)的程序只有一个。

由于在主存同时存放多个正在运行的作业,就给系统带来了一系列复杂的问题(这是本书以后各章将详细阐述的问题),这就导致了操作系统的逐步完善。然后,分时系统也相继

出现,多道程序和分时系的出现,标志操作系统的正式形成。

1.3 操作系统的特征

多道程序(设计)系统具有很大的优点,它可以使 CPU 和输入输出设备以及其它系统资源得到最充分的利用。例如,从理论上来说,如果主存中的程序数量不断增加,CPU 的利用率最终应达到百分之百。但是,多道程序设计也带来了不少新的复杂问题,使操作系统具有如下一些特征:

一、并发(Concurrence)

并行性和并发性是既相似又有区别的两个概念。并行性是指两个或多个事件在同一时刻发生;而并发性是指两个或多个事件在同一时间间隔内发生。显然,在单 CPU 系统中,所谓多道程序同时处于运行状态,是一个并发的概念,这是从宏观上看,在一段时间内有多道程序在同时运行,而每一时刻仅仅只是其中一道程序在执行。能并发执行的程序称为并发程序,相应的系统称为并发系统。

二、共享(Sharing)

共享是指系统中的所有资源(如 CPU 时间,主存空间,I/O 设备,以及各种软件资源)不再为一个程序所独占,而是供同时存在于系统中的多道程序所共同使用。根据资源属性的不同,可分为两种不同的共享方式:

1. 互斥共享。是指在一段时间内仅允许一个作业(程序)进入访问(使用),而多个作业只能依次顺序进入访问,所以也称此为顺序共享。系统中极大多数资源的属性都要求互斥共享,如各种字符设备、磁带机、以及变量、表格等,这类资源也称为临界资源。

2. 同时共享。是指在一段时间内可由多个作业(程序)同时进入访问(使用),也称为交叉共享。典型可供同时共享的资源是磁盘和一些可重入的程序等。

三、虚拟(Virtual)

所谓虚拟,是指把一个物理上的实体,变为若干个逻辑上的对应物。在单 CPU 系统中,通过分时技术,可把 CPU 时间分配给多个程序,从宏观上看这一 CPU 能同时运行多道程序。这种情况,也可看成是一台物理的 CPU,经过扩充和改造,得到速度较慢但数量较多的逻辑的 CPU,使每道程序都有一个 CPU 在为它服务。亦即,多道程序技术可以把一台物理上的 CPU 虚拟为多台逻辑上的 CPU。同样,我们通过操作系统的功能,把主存扩充并改造为虚拟存储器,把 I/O 设备扩充并改造为使用更方便的虚拟设备。此外也可以把一条物理信道虚拟为多条逻辑信道。

四、不确定性

在操作系统中,不确定性有两种含义:

1. 程序执行的速度不确定 多道程序环境下,各程序的执行是以异步方式进行的。换言之,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的,因而也是不可预知的。

2. 程序执行结果的不确定 即对同一程序,给定相同的初始条件,在相同的环境下进行多次执行,却可能获得完成不同的结果。这也就是程序并发执行的不可再现性特征。执行结果的不确定是绝对不允许的,因而也成为操作系统必须解决的主要问题。关于这一问题将在第二章中作详细阐述。

1.4 操作系统的类型

操作系统作为独立的系统软件,已有 30 多年历史了,它经过 60 和 70 年代的大发展,已日趋完善和成熟。但这一时期,操作系统的发展主要集中于传统的操作系统。传统操作系统按其功能特征的不同,可分为批处理系统、分时系统和实时系统。它们共同的特点是单一处理机,即所构成的整个系统只有唯一的一台 CPU。所以在传统操作系统中,所谓多道程序同时执行,只能是并发的概念,而不是真正的并行。到了 70 年代后期和 80 年代初期,由于微电子技术的迅速发展,和计算机系统体系结构的不断改进,作为系统重要组成部分的操作系统也出现了许多新的类型。总体上沿着两个方向发展,一是随着微机和微机应用的发展而形成了微机操作系统;另一个则是随着计算机体系结构的发展,形成了多机操作系统、网络操作系统和分布式操作系统。

1.4.1 传统操作系统的基本类型

一、批处理操作系统

这类系统突出的特征是“批量”。它把提高系统的处理能力,即作业的吞吐量作为主要设计目标,同时也兼顾作业的周转时间。所谓周转时间就是从作业提交给计算机中心到用户作业完成取得计算结果的这段时间。批处理系统可分为单道批处理系统和多道批处理系统两种。单道批处理系统比较简单,类似于单用户操作系统。多道批处理系统引入了多道程序设计技术。这类系统把用户提交的作业成批送入计算机,然后由作业调度程序按一定的调度原则选择一个或几个作业进入主存。主存中几个作业交替运行,直到某作业完成计算任务,输出其结果,收回该作业占用的全部资源,然后再调入其它作业。

多道批处理系统一般用于计算中心的较大的计算机系统中。由于它的硬件设备比较全,价格较高,所以这类系统十分注意 CPU 与其它设备的充分利用,追求高的吞吐量。故多道批处理系统的特点是对资源的分配策略和分配机制,以及对作业和处理机的调度等功能均进行精心设计。

二、分时系统

分时系统既是操作系统的一种类型,又是对配有分时操作系统的计算机系统的一种称呼。所谓分时是指多个用户按时间轮转的方法来共享同一台计算机,也就是说把计算机的系统资源(尤其是 CPU 时间)进行时间上的分割,即将整个工作时间分成一个个的时间段,称为时间片,多个用户依次轮转地被分配获得 CPU 的时间片,从而使一台计算机同时为多个用户服务。

在分时系统中,每个用户都有其独立的终端设备,用户通过各自的终端,以交互方式使用计算机,得到系统的服务。分时系统具有如下特性:

- 多路性 系统将若干个用户终端通过多路卡连接到一台主机上。宏观上,多个用户同时工作,共享系统资源;微观上,各终端作业轮流在 CPU 上运行一个时间片。多路性提高了资源利用率,促进了计算机的广泛应用。

- 独立性 每个用户有其独立的终端,彼此独立操作互不干扰。从用户角度说,每一用户并不感觉到其它用户的存在,就象整个系统被他所独占。

- 及时性 终端用户请求能在很短时间间隔内获得系统响应,这一段时间间隔称为

响应时间。响应时间一般为 2~3 秒钟。

- **交互性** 用户能与系统进行较广泛的人机对话,即用户从键盘键入命令,请求系统服务和控制程序运行。系统能及时响应用户的命令,并在终端上显示响应结果。交互性可使用户请求系统提供多方面的服务,如文件编辑、数据处理和资源共享等。

由于分时系统的主要目的是及时地响应和服务于联机用户,因此分时系统设计的主要目标是一个合理的响应时间。

三、实时系统

实时系统是另一类联机的操作系统,它主要随着计算机应用于实时控制和实时信息处理领域而发展起来的。

实时系统的主要特点是提供即时响应和高可靠性。而且系统本身要安全可靠。实时系统主要有两种典型应用形式:

- (1)实时控制系统。用于工业生产的自动控制,系统实时采集现场数据,如温度、压力、方位等。进而对采集的数据进行及时分析处理,最终产生控制信息返回生产现场,对相应的执行机构进行控制调度,以控制保持设定的正常(或最佳)工况。类似地,实时控制系统也可用于武器系统,如火炮的自动控制系统、导弹的制导系统等。

- (2)实时信息处理系统。这类系统要求对大量信息进行实时处理。它又可分为信息查询系统,这类系统配有大型文件系统或数据库系统,并具有向用户提供简便、快速地查询的能力,如仓库管理系统、医疗信息系统等。另一类是事务处理系统,如订票系统、银行储蓄系统等。实时系统主要为联机的实时任务服务的,其主要特点如下:

- **实时性** 系统对对象的实时要求必须能及时响应。以控制对象不同,一般可以在秒级、毫秒级,甚至可低于 100 微秒的。

- **高可靠性** 系统的任何差错都可能带来严重后果,甚至酿成无法预料的灾难性后果。因此在实时系统中,往往采用多级容错措施,来保证系统的安全可靠。相对,系统的效率则可放在次要地位。

- **过载保护** 在实时系统中,进入系统的实时任务时间和数目有很大的随机性,因而在某一时刻有可能超出系统的处理能力,这就是所谓瞬时过载。系统应具有过载保护机制,如对多个实时任务按一定的策略排入缓冲区,等待调度处理,最急迫的任务最优先得到处理。对于持续性过载,则需拒绝某些较低级任务的进入。

- **交互能力** 这类系统“人一机”之间必须具有一定的交互能力。但是,不能象分时系统那样,用户可随意地输入各种命令,或在终端上调试和修改程序。为了保证系统的可靠性,人与系统的交互,仅只限于访问系统中某些特定的专用服务程序。

传统操作系统虽可分为上述三种基本类型,但是一个实际的操作系统,可能兼有三者或其中两者的功能,例如,在 VAX-11 系列机上所配置的 VMS 操作系统来说,便是一个兼有分时、实时和批处理功能的操作系统

1.4.2 微机操作系统

随着 60 年代后期开始的集成电路的发展和性能/价格比的提高,微型机和个人计算机得到迅猛的发展。微机已成为世界上应用最广的机种,1995 年的产量达到 6000 万台,使用中的积累总数已经超过 2 亿台。配置在微机上的操作系统称为微机操作系统。最早出现在微机上的操作系统,是配置在 8 位微机上的 CP/M。后来出现了 16 位微机操作系统,最典型

的是 MS - DOS。

随着 32 位微机的出现;相应的 32 位微机操作系统也相继产生,最典型例子是 UNIX、Linux 和 Windows NT 等。

一、单用户微机操作系统

由于早期微机主要用于单用户情况,而且微机的价格十分便宜,所以其操作系统的设计不象中、大型系统那样刻意追求 CPU 及其它资源的利用率。通常不考虑多用户多道程序共享系统资源,所以,第一代微机操作系统的功能是十分有限的,实际上只有信息管理,即文件系统。单用户微机操作系统最有代表性的是 CP/M 和 MS - DOS。

CP/M 是带有软盘系统的 8 位微机操作系统,它主要配置在以 Intel.8080.8085 和 Z80 芯片为基础的微机上,后来又推出了增加硬盘管理功能的 CP/M2.2 版本。由于 CP/M 具有较好的层次结构,以及适应性、可移植性好,使之成为事实上的 8 位微机操作系统的标准。1981 年 IBM 公司首次推出了 IBM - PC 个人计算机,并在其上配置了由 Microsoft 公司开发的 MS - DOS 操作系统。MS - DOS 是在 CP/M 基础上开发的,它增加许多内部和外部命令,使该操作系统具有较强的功能及性能优良的文件系统。又因为它是配置在 IBM - PC 机及其兼容机上,随该机种的不断增强和改进,MS - DOS 几乎每年都有新的版本问世。新版本不仅保证与原版本兼容,且加入了许多创新和增强的功能。MS - DOS 操作系统也就成了事实上的 16 位微机单用户单任务操作系统的标准。随着 32 位微机的出现和发展,也就产生了单用户多任务操作系统。单用户多任务的含义是:只允许一个用户上机,但允许将一个用户程序分为若干个任务,使它们并发执行,从而有效地改善系统的性能。目前,在 32 位微机上所配置的 32 位微机操作系统,大多数是单用户多任务操作系统,其中最具有代表性的是 OS/2 和 Windows。OS/2 最初版本 OS/21.x 是针对 80286 开发的,OS/22.x 版本则是针对 80386 和 80486 开发的。是真正的 32 位微机操作系统。

Windows 是由 Microsoft 公司开发,并以其友好的图形用户界面、并能支持多任务的优点,得以很快地推广,占领了微机市场。本书有专门篇幅介绍 Windows。

二、多用户微机操作系统

随着超级微机 32 位机出现,以及微机应用的日益广泛,并广泛应用在多机系统中,许多微机已具有以前的小型机甚至大型机系统中所具有的性能。这些微机系统采用多用户多任务的多道程序设计技术,存储管理上采用虚拟存储管理技术。为此,在 32 位微机上,也需要配置多任务多用户操作系统。其中,最具有代表性的是 UNIX 操作系统。

UNIX 是由美国 AT&T 公司的 Bell 实验室开发的,至今已有 20 多年的历史。最初,UNIX 是配置在 DEC 公司的小型机 PDP 上,后来它被移植到微机上。UNIX 是唯一能在个人计算机、工作站、小型机到大型机上都能运行的操作系统,它已成为事实上的操作系统标准。

在 1980 年前后,UNIX 第七版本首先移植到基于 Motorola 公司的 MC68000 芯片的微机上。与此同时,Microsoft 公司也同样基于 UNIX 第七版本推出了相当简洁的、用于 Intel8088 微机上的 UNIX 版本,称为 Xenix。1986 年 Microsoft 又发表了 Xenix 系统 V 版本。后来 SCO 公司也公布了 SCO Xenix 系统 V 版本,使 UNIX 可以在 386 微机上运行。SCO 公司又把 SCO Xenix 系统 V 和 AT&T 的 UNIX 系统 V 结合起来,形成了 SCO UNIX 系统 V/386。

另一被世界所关注并被迅速推广应用的微机操作系统是 Microsoft 公司推出的 Windows-NT。它是一个 32 位可移植的操作系统,具有工作站和小型机上的操作系统所具有的强大功能,其性能在许多方面已经和 UNIX 相近。

事实上,很多微机用户因费用的问题,在其微机上安装 UNIX 有一定困难。80 年代末产生了目前已被越来越多用户所接受的 Linux 操作系统。Linux 是一个 UNIX 在个人计算机上的完整实现,可以在个人机上运行各种 UNIX 命令,使用各种 UNIX 软件,并可从 Linux 获得免费的各种为 UNIX 编写的软件、各种工具以及各种游戏程序。Linux 对机器系统的要求很低,只要有 386SX 以上的 CPU 或兼容机,2MB 以上的主存容量,就可以安装运行。对于教学, Linux 有更进一步的意义,它的内核和上层的所有软件的源代码都能方便的获得。

1.4.3 网络操作系统

计算机网络可以定义为一些互连的自主计算机系统的集合。所谓自主计算机是指计算机具有独立处理问题的能力;而互连则是表示计算机之间能够实现通信和相互合作。可见,计算机网络是在计算机技术和通信技术高度发展的基础上,相互结合的产物。配置在网络系统上的操作系统,称为网络操作系统。

在网络上的计算机由于各机器的硬件特性不同、数据表示格式及其它方面要求不同,在互相通信时为正确进行并相互理解通信内容,相互间应有许多约定。这些约定称为协议或规程。因此通常网络操作系统定义为:网络操作系统是使网络上各计算机能方便而有效地共享网络资源,为网络用户提供所需求的各种网络服务的软件和有关规程的集合。

网络操作系统除了应具有通常操作系统应具有的处理机管理、存储管理、设备管理和文件管理的基本功能外,还应具有如下有关网络的功能:

1. 网络通信

这是网络最基本的功能。除了能支持终端和主机之间的通信外,还要支持主机和主机的无差错通信,以及多个“用户对”之间同时通信的能力,即将一条物理链路虚拟为多条虚(逻辑)链路。在出现异常事件时,还应具有及时处理的能力。

2. 网络服务

为了方便用户,应能直接向用户提供多种网络服务功能,主要包括:

(1) 电子邮件服务。为源用户把电子邮件传送给目标用户。
(2) 文件传输、存取和管理服务。把源用户存放在站上的文件传送到指定目标站显示或存盘,或从目标点索取文件

(3) 共享硬盘服务。

(4) 共享打印服务。

3. 网络管理

网络管理最基本的任务是安全管理。这是通过“存取控制”来确保存取数据的安全性;通过“容错技术”来保证系统故障时数据的安全性。

4. 互操作能力

在 90 年代推出的网络操作系统,又提供了一定范围的互操作能力。所谓互操作,在客户/服务器模式的 LAN 环境下,是指连接在服务器上的多种客户机和主机,不仅能与服务器通信,而且还能以透明的方式访问服务器上的文件系统;而在互连网络环境下的互操作,是指不同网络间的客户机不仅能通信,而且也能以透明的方式,访问其它网络中的文件服务器。

通常的网络操作系统是在传统操作系统之上附加上网络软件,如 MS-DOS 3.1 就是可增加网络软件,使其成为网络操作系统。现代操作系统已把网络功能包含到操作系统的内

核中,作为操作系统核心功能的一个组成部分。例如:BSD UNIX 4.2/4.3 已把 TCP/IP 网络功能包含在内核中。Linux 也具有内置的 TCP/IP 网络功能。WindowsNT 也把网络功能做到操作系统之中,而且是该操作系执行体的输入输出系统的一部分。

1.4.4 分布式操作系统

80 年代以来,计算机网络发展非常迅速。局域网(LAN)可以连接小范围内的数十台甚至上百台计算机,信息传输速度达 100Mbps 到 1000Mbps。而广域网(WAN)则可以将全世界的计算机连接起来。网络技术的发展使一些计算机系统从集中式走向分布式。

分布式计算机系统(Distributed Computer Systems)是由多个分散的计算机经互连网络连接而成的计算机系统,其中各个资源单元(物理的或是逻辑的)既相互协同又高度自治,能在全系统范围实现资源管理,动态地进行任务分配或功能分配,并能并行地运行分布式程序。

在分布式系统上配置的操作系统,称为分布式操作系统。虽然从系统结构上看,网络系统和分布式系统有很多相似之处,但从操作系统的角度来看,分布式操作系统和网络操作系统是有很大区别的。

网络中的各台计算机配置各自的操作系统,而网络操作系统把它们有机的联系起来。因此。它除了具有一般操作系统所具有功能外,还应具有高效可靠的网络通信能力和向用户提供多种网络服务的能力。分布式操作系统除了最低级的 I/O 设备资源外,所有的系统任务都可以在系统中任何别的处理机上运行,并提供高度的并行性和有效的同步算法和通信机制,自动实现全系统范围的任务分配,并自动调度各处理机的工作负载,为用户提供一个方便、友善的使用环境。分布式操作系统的主要特点是:

- 进程通信不能借助公共存储器,因而常采用信息传递方式;
- 系统中资源分布于多个场点(Site),因而进程调度、资源分配及系统管理等必须满足分布处理要求,并采用保证一致性的分散式管理方式和具有强健性的分布算法;
- 不失时机地协调各种场点的负载,使其达到基本平衡,以充分发挥各场点的作用;
- 故障检测与恢复及系统重构和可靠性问题的处理和实施都比较复杂。

一般认为,高级操作系统即多机操作系统,而分布式操作系统则是多机操作系统中的典型和卓越代表。它与单机操作系统和网络操作系统都有不同程度的区别,其复杂程度也明显高于它们。

习 题

- 1.1 在计算机系统中配置操作系统的目的是什么?
- 1.2 说明操作系统在整个计算机系统中的作用和地位?
- 1.3 从资源管理的角度看操作系统应具有哪些功能?
- 1.4 给出操作系统的定义? 操作系统形成的标志是什么?
- 1.5 什么是多道程序设计? 多道程序设计为操作系统带来了哪些特征,请详细说明之?
- 1.6 实现多道程序设计应解决哪些问题?
- 1.7 操作系统形成和发展经历了哪几个阶段? 当前操作系统发展的趋势如何?
- 1.8 操作系统有哪三个基本类型? 试比较其特点,各适用于哪些场合?
- 1.9 试比较单道与多道批处理系统的特点?
- 1.10 为什么要引入分时系统? 分时系统的特点是什么?