

面向21世纪

高职高专系列教材

Linux 操作系统

◎陈建辉 主编

◎周岳山 审

6.89

 机械工业出版社
China Machine Press



面向 21 世纪高职高专系列教材

Linux 操作系统

陈建辉 主编
周岳山 审



机械工业出版社

本教材主要介绍计算机操作系统的基本工作原理和网络操作系统 Linux 的结构和资源管理、Linux 系统的安装和管理等。

全书共 18 章。按教材内容自然形成四大部分，彼此相对独立，可以根据教学对象的实际情况，选择性搭配使用。本教材内容具有较强灵活性和较宽的适用性的特点。为了方便教学和课外复习，每章结束均安排内容小结和习题，并根据教学进程安排适当数量的实验。

本教材可作为高职高专计算机专业“操作系统”课程教学用书，也可作为中等技术专业学校计算机专业学生深入掌握“操作系统”或学习 Linux 的参考书。

图书在版编目 (CIP) 数据

Linux 操作系统/陈建辉主编. —北京: 机械工业出版社, 2002.9

面向 21 世纪高职高专系列教材

ISBN 7-111-10809-4

I .L... II .陈... III .Linux 操作系统—高等学校: 技术学校—教材
IV .TP316.89

中国版本图书馆 CIP 数据核字 (2002) 第 062857 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: 周艳娟

责任印制: 闫 焱

北京第二外国语学院印刷厂印刷·新华书店北京发行所发行

2002 年 9 月第 1 版·第 1 次印刷

1000mm × 1400mm B5 · 11.125 印张 · 457 千字

0 001—5 000 册

定价: 27.00 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、68326677 - 2527

封面无防伪标均为盗版

面向 21 世纪高职高专 计算机专业系列教材编委会成员名单

- 顾问** 曾玉崑 王文斌 陈瑞藻 李 奇 凌林海
林 东
- 主任委员** 周智文
- 副主任委员** 周岳山(常务副主任) 詹红军 陈付贵
穆天保 赵佩华 黄甘洲 武文侠 吕何新
- 委员** 郭曙光 王德年 刘瑞新 陈丽敏 孔令瑜
李 玲 鲁 辉 陶书中 赵增敏 马 伟
孙心义 翟社平 廖常武 于恩普 王春红
王娟萍 屈 圭 汤新广 谢 川 姜国忠
汪赵强 董 勇 梁国浚 张晓婷
- 秘书长** 胡毓坚
- 副秘书长** 陈丽敏(兼)

出版说明

积极发展高职高专教育，完善职业教育体系，是我国职业教育改革和发展的一项重要任务。为了深化职业教育的改革，推进高职高专教育的发展，培养 21 世纪与我国现代化建设要求相适应的，并在生产、管理、服务第一线从事技术应用、经营管理、高新技术设备运作的高级职业技术应用型人才，尽快组织一批适应高职高专教学特色的教材，已成为各高职高专院校的迫切要求。为此，机械工业出版社与高职高专计算机专业、电子技术专业和机电专业教材编委会联合组织了全国 40 多所院校的骨干教师，共同研究开发了一批计算机专业、电子技术专业和机电专业的高职高专系列教材。

各编委会确立了“根据高职高专学生的培养目标，强化实践能力和创新意识的培养，反映现代职业教育思想、教育方法和教育手段，造就技术实用型人才为立足点”的编写原则。力求使教材体现“定位准确、注重能力、内容创新、结构合理和叙述通俗”的编写特色。

本套系列教材是由高职高专计算机专业、电子技术专业、机电专业教材编委会分别会同各院校第一线专业教师针对高职高专计算机、电子技术和机电各专业的教学现状和教材存在的问题开展研讨，尤其针对目前高职高专教学改革的新情况，分别拟定各专业的课程设置计划和教材选题计划。在教材的编制中，将教学改革力度比较大、内容新颖、有创新精神、比较适合教学、需要修编的教材以及院校急需、适合社会经济发展的新选题优先列入选题规划。在广泛征集意见及充分讨论的基础上，由各编委会确定每个选题的编写大纲和编审人员，实行主编负责制，编委会通过责任编委和主审对教材进行质量监控。

担任本套教材编写的老师们都是来自各高职高专院校教育第一线的教师，他们以高度的责任感和使命感，经过近一年的努力，终于将本套教材呈现在广大读者面前。由于高职高专教育还处于起步阶段，加上我们的水平和经验有限，在教材的选题和编审中可能出现这样那样的问题，希望使用这套教材的教师和学生提出宝贵的意见和建议，以利我们今后不断改进，为我国的高职高专教育事业的繁荣而共同努力。

高职高专系列教材编委会
机械工业出版社

前 言

根据高职高专学生的培养目标,本教材参考教学时数为 80 课时,主要介绍操作系统的基本工作原理和网络操作系统 Linux 的结构、安装及管理、功能及维护等。

全书分四部分,共 18 章。第一部分(第 1~7 章)主要介绍了操作系统的工作原理和计算机系统的作业管理、进程管理、存储管理、文件管理以及设备管理。第二部分(第 8~13 章)以 Linux 操作系统为例,介绍操作系统资源管理在 Linux 中的具体实现。第三、四部分(第 14~18 章)以 Linux 实际应用和网络架构为主,介绍 Linux 的安装;RPM、GNOPRM 软件包管理和 Linux 系统管理与维护技术等。第一部分内容是全书的基础,其余三部分内容相对独立,可以根据教学实际情况对第一部分内容和第二、三、四部分内容选择搭配使用。本教材具有实用性较强、内容新颖(Linux 操作系统)、使用面宽等特点。为了便于学生课外复习,每章均附上小结与习题。

本教材根据高职高专教学的特点,贯彻理论教学和实际教学相结合的原则编写各章节内容。教学实施中可根据章节内容和实际条件安排实验、演示或实习,以加深对相关内容的理解,培养学生使用系统和管理系统的能力。

本书由陈建辉主编,周岳山担任主审。第 1~7 章由陈建辉编写,第 8~13 章由章瑞平编写,第 14~18 章由孙志强编写,全书由陈建辉统稿。在本书的编写过程中得到杨建南老师的协助,林东老师提出了许多宝贵意见,在此表示诚挚的感谢。

由于编者水平有限,书中难免有疏漏之处,敬请读者批评、指正。

编 者

目 录

出版说明	2.4.3 键盘操作命令	20
前言	2.5 作业控制	21
第 1 章 计算机操作系统	2.5.1 脱机控制方式	21
1.1 操作系统的概述	2.5.2 联机控制方式	22
1.1.1 什么是操作系统	2.6 小结	22
1.1.2 计算机系统操作系统	2.7 习题	23
1.2 操作系统的功能和特征	第 3 章 进程管理	24
1.2.1 操作系统的功能	3.1 并发程序	24
1.2.2 操作系统的特征	3.1.1 程序的顺序执行及特点	24
1.3 操作系统的基本类型	3.1.2 程序的并发执行及特点	24
1.3.1 批处理系统	3.2 进程模型	27
1.3.2 分时系统	3.2.1 进程的定义和特征	27
1.3.3 实时系统	3.2.2 进程结构	29
1.3.4 通用操作系统	3.3 进程同步与通信	32
1.4 小结	3.3.1 临界资源与临界区问题	33
1.5 习题	3.3.2 进程之间同步机构	34
第 2 章 作业管理	3.4 进程控制与调度	39
2.1 作业、作业类别和组织	3.4.1 进程控制	39
2.1.1 作业、作业步和作业流	3.4.2 进程调度	41
2.1.2 作业类别和组织	3.5 线程	44
2.2 批量型作业的管理和调度	3.5.1 线程的概念	44
2.2.1 作业的组成	3.5.2 线程和进程的关系	44
2.2.2 作业的进入	3.5.3 线程的执行状态及其转换	44
2.2.3 作业的建立	3.5.4 引入线程的优点	45
2.2.4 作业的调度	3.6 小结	45
2.3 终端型作业的管理	3.7 习题	46
2.3.1 作业的建立和进入	第 4 章 存储管理	47
2.3.2 终端型作业与批量型作业 在管理方面的其他区别	4.1 概述	47
2.4 用户与操作系统的接口	4.2 实存管理技术	48
2.4.1 系统调用	4.2.1 用户连续存储管理	49
2.4.2 作业控制语言	4.2.2 分区存储管理	50
	4.2.3 覆盖与交换技术	53

4.2.4 简单页式存储管理	55	6.2.5 文件目录检索	90
4.3 虚拟存储管理技术	57	6.3 文件的共享和安全	92
4.3.1 概述	57	6.3.1 文件共享	92
4.3.2 请求页式存储管理	59	6.3.2 文件的保护和保密	93
4.3.3 分段式存储管理	61	6.4 小结	93
4.3.4 段页式存储管理	64	6.5 习题	94
4.4 小结	65	第7章 网络操作系统	95
4.5 习题	66	7.1 计算机网络概念	95
第5章 设备管理	67	7.2 网络协议和网络操作系统	95
5.1 设备的分类和设备管理		7.2.1 OSI 参考模型	95
功能	67	7.2.2 网络操作系统的定义	
5.1.1 设备的分类	67	和特性	97
5.1.2 设备管理的目标和功能	69	7.2.3 网络操作系统的分类	98
5.2 I/O 设备控制与驱动软件	69	7.2.4 网络操作系统的结构模式	100
5.2.1 设备控制器	70	7.2.5 网络操作系统的功能	101
5.2.2 I/O 控制方式	71	7.2.6 网络操作系统介绍	102
5.2.3 I/O 驱动软件	72	7.3 小结	104
5.3 设备管理实现	73	7.4 习题	104
5.3.1 操作系统与中断处理	73	第8章 Linux 操作系统概述	105
5.3.2 设备驱动程序	75	8.1 Linux 简介	105
5.3.3 与硬件无关的 I/O 软件	76	8.1.1 什么是 Linux	105
5.4 设备的调度和分配	77	8.1.2 计算机系统以 Linux 为	
5.4.1 设备的调度	77	平台的好处	105
5.4.2 设备的分配	78	8.1.3 Linux 的发行版本	106
5.5 小结	78	8.2 Linux 的特性	107
5.6 习题	79	8.2.1 多任务	107
第6章 文件管理	80	8.2.2 多用户	107
6.1 文件和文件系统	80	8.2.3 可编程 Shell	107
6.1.1 文件的概念	80	8.2.4 多文件系统	108
6.1.2 文件操作	82	8.2.5 支持的硬件多	108
6.1.3 文件系统功能	83	8.2.6 丰富的软件	108
6.2 文件系统的实现	84	8.3 Linux 使用基础	108
6.2.1 文件的结构	84	8.3.1 Linux 的开机程序和正常	
6.2.2 文件的存取方法	87	关机步骤	108
6.2.3 文件存储空间管理	87	8.3.2 使用者的管理	114
6.2.4 文件目录	89	8.3.3 文件系统的结构	114

8.3.4 Linux 的基本命令	119	10.2.4 进程的结束(exit)	176
8.4 小结	137	10.3 Linux 进程调度	177
8.5 习题	137	10.3.1 调度的时机	177
第 9 章 Linux 的作业管理	138	10.3.2 调度的策略	178
9.1 Linux 的 Shell	138	10.4 Linux 进程通信	180
9.1.1 什么是 Shell	138	10.4.1 信号	180
9.1.2 Shell 的主要版本	139	10.4.2 管道	182
9.2 Shell 命令的语法分析	140	10.4.3 System V 的进程通信	183
9.2.1 Shell 的命令格式	140	10.5 小结	188
9.2.2 通配符	142	10.6 习题	188
9.2.3 重定向	144	第 11 章 Linux 中的存储管理	190
9.2.4 管道	147	11.1 Linux 的虚拟存储管理	190
9.2.5 命令别名	147	11.1.1 分页机制	191
9.2.6 命令替换	148	11.1.2 共享和保护	192
9.2.7 自动补全命令行	149	11.2 Linux 内存映射	194
9.2.8 命令行编辑	149	11.3 Linux 存储管理的实现	
9.2.9 命令历史	150	技术	196
9.2.10 引用	151	11.3.1 内存的分配和回收	196
9.2.11 作业管理	152	11.3.2 需求分页机制	198
9.3 Shell 编程	155	11.3.3 交换机制	199
9.3.1 变量	156	11.3.4 高速缓存	203
9.3.2 Shell 程序设计的流程控制	159	11.4 小结	205
9.3.3 Shell 脚本的建立和执行	162	11.5 习题	205
9.3.4 调试 Shell 脚本	164	第 12 章 Linux 中的设备管理	207
9.3.5 实例	165	12.1 概述	207
9.4 小结	166	12.1.1 Linux 中的总线使用	207
9.5 习题	167	12.1.2 Linux 中的数据传送方式	211
第 10 章 Linux 中的进程管理	168	12.2 Linux 设备驱动程序与	
10.1 Linux 中的进程	168	内核的接口	215
10.1.1 Linux 中的进程控制块	168	12.2.1 字符设备的管理	216
10.1.2 Linux 的进程标识与状态	172	12.2.2 块设备的管理	217
10.2 Linux 的进程控制	173	12.3 Linux 设备驱动程序框架	219
10.2.1 进程的创建(fork 和		12.4 Linux 设备的 I/O 调用	220
vfork)	173	12.5 小结	222
10.2.2 进程的执行(exec)	175	12.6 习题	222
10.2.3 进程的等待(wait)	176	第 13 章 Linux 文件系统	223

13.1 概述	223	14.4.4 配置 Linux	262
13.1.1 文件的用户接口	223	14.5 小结	271
13.1.2 Linux 文件系统框架	224	14.6 习题	272
13.2 Ext2 文件系统	225	第 15 章 软件包的管理	273
13.2.1 Ext2 文件系统的特特点	225	15.1 RPM 的设计目的	273
13.2.2 Ext2 文件的逻辑结构 与物理结构	228	15.2 使用 RPM	274
13.2.3 Ext2 文件系统存储 空间管理	230	15.2.1 安装软件包	274
13.2.4 Ext2 文件系统目录文件的 实现	232	15.2.2 卸载	276
13.3 虚拟文件系统(VFS)	234	15.2.3 升级	276
13.3.1 VFS 对文件系统的描述	235	15.2.4 查询	277
13.3.2 VFS 对文件系统的管理	237	15.2.5 验证	278
13.3.3 文件系统的接口	240	15.3 RPM 使用实例	279
13.3.4 VFS 的缓冲机制	241	15.4 小结	281
13.4 小结	244	15.5 习题	281
13.5 习题	245	第 16 章 GnoRPM	282
第 14 章 Linux 系统的安装和管理	246	16.1 启动 GnoRPM	282
14.1 安装 Linux 的硬件需求	246	16.2 显示软件包	283
14.2 硬盘分区	246	16.3 选择软件包	283
14.2.1 为 Red Hat Linux 的安装 准备空间	247	16.4 安装新的软件包	284
14.2.2 分区命名方案	248	16.5 设置 GnoRPM	285
14.2.3 磁盘分区和其他操作系统	248	16.6 软件包操作	287
14.2.4 分区数	249	16.6.1 查询软件包	287
14.2.5 LILO 的介绍和使用	249	16.6.2 验证软件包	288
14.3 Linux 的安装方法与 启动盘的制作	249	16.6.3 卸载软件包	288
14.3.1 Linux 的安装方法	249	16.6.4 升级软件包	290
14.3.2 制作系统安装启动盘	250	16.6.5 查找软件包	290
14.4 Red Hat Linux 的安装 过程	251	16.7 小结	290
14.4.1 关于安装的几点说明	251	16.8 习题	291
14.4.2 启动安装程序	251	第 17 章 系统设置和维护	292
14.4.3 开始安装	252	17.1 用户、组和私有用户组	292
		17.1.1 用户和组管理的概念	292
		17.1.2 私有用户组	293
		17.2 设置控制台的访问权	295
		17.2.1 取消控制台程序访问 权限	295
		17.2.2 禁止所有控制台访问	295

17.2.3	定义控制台	295	18.1	使用 Linuxconf 设置系统 ...	312
17.2.4	使文件可被控制台访问 ...	296	18.1.1	运行 Linuxconf	312
17.2.5	其他可通过控制台访问的 应用程序	296	18.1.2	组	317
17.3	软盘组	297	18.1.3	文件系统	318
17.4	带 PAM 的用户认证	297	18.1.4	配置网络	321
17.4.1	PAM 的体系结构	297	18.1.5	日期和时间的设置	324
17.4.2	PAM 的功能与模块	298	18.2	使用控制面板设置系统 ...	325
17.4.3	PAM 的配置	300	18.2.1	设置打印机	325
17.5	Shadow 工具	302	18.2.2	内核守护进程的配置	330
17.6	构造自定义内核	302	18.2.3	网络配置	332
17.6.1	创建模块化内核的 流程	303	18.3	小结	336
17.6.2	构造 initrd 镜像	305	18.4	习题	336
17.7	匿名 FTP	306	附录	实验	337
17.8	设置 NFS	307	实验一	操作系统与用户交互 界面实验	337
17.8.1	安装 NFS	307	实验二	进程管理实验	339
17.8.2	配置 NFS 服务器	307	实验三	存储器管理实验	340
17.8.3	配置 NFS 客户端	310	实验四	文件管理实验	342
17.9	小结	311	实验五	Linux 系统的用户交互 界面实验	343
17.10	习题	311	参考文献	345
第 18 章	系统管理	312			

第 1 章 计算机操作系统

现代计算机系统是由硬件系统和软件系统两大部分构成的。计算机系统中的所有硬件和软件,统称为计算机的资源。现代计算机系统如何全面将资源管理起来,如何让计算机中资源发挥出作用,这是我们所关心和感兴趣的问题。计算机如果离开了软件将成为一堆废铜烂铁。有了软件,计算机可以对信息进行存储、处理和检索,可以显示多媒体文档、搜索 Internet 网并完成其他工作。计算机系统的硬件是计算机系统的各种物理设备,包含一个或多个处理器(CPU)、若干内存、磁盘、打印机、网络接口及其输入/输出设备。计算机系统的软件是各种各样的程序和数据及各种文档资料。计算机系统的硬件和软件之间具有密切关系,硬件是软件运行的基础,软件的运行需要硬件的支持。

如何才能方便、有效地管理和使用计算机系统是人们必然提出的问题。用户通过直接编程使用硬件资源是很困难的,原因在于使用者必须熟悉硬件中各器件所提供接口的参数,才能正确编程,但要做到这一点必须是计算机专业人员才能胜任的,这是由硬件的复杂所引起的,称为硬件的复杂性。因此人们就设想用一个最基础的系统软件,称为“操作系统”来管理计算机系统,同时为用户提供一个更容易理解的程序设计接口。这样用户就可以方便地通过操作系统使用计算机系统的硬件和软件。实际上,计算机被操作系统改造了,构成和原来计算机系统完全不同的一台计算机了,这被称为虚拟机。

操作系统也不是直接管理底层硬件的,而是通过主机板中的只读存储器(ROM或 EPROM)中固化的微程序来控制最低一层的物理器件,包括集成电路芯片、连线、电源、监视器等。微程序实际是一个解释器,它先取得机器指令,如 ADD(加法运算)、MOVE(移动)和 JUMP(转移)等指令,由微程序解释这一套指令,如图 1-1 所示。

有了操作系统就可以将硬件的复杂性隐藏起来,同时为程序员提供一套更加方便的命令(人们多把这些命令称为“系统调用命令”)使用计算机的硬件和软件,这种类型的命令就是上面所讲的接口。

在操作系统之上是其他系统软件,包括命令解释器(shell)、窗口系统、编译器、编辑器及类似的独立于应用的程序,它们并不是操作系统的组成部分。操作系统专指在核心态(kernel mode),或称管态(supervisor mode)下运行的软件,装入内存的空间受硬件保护,用户程序不能进入此空间,保证了操作系统在内存中的安全。而编译器和编辑器运行在用户态(user mode),编译器可以根据用户的需要进行不同选择。系统软件之上是应用软件,应用软件可以由用户根据使用的需要采用某种特定的语言

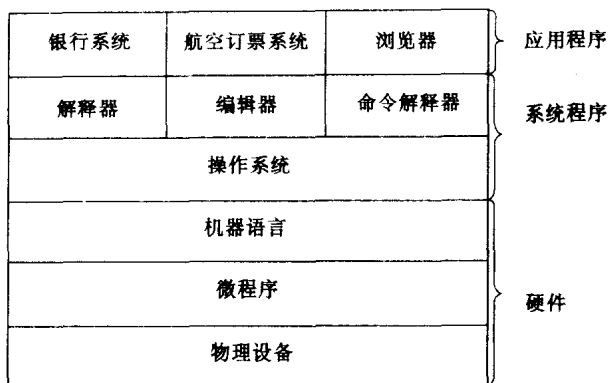


图 1-1 计算机系统的组成

自行开发,也可以向软件公司购买已经开发成熟的软件。

1.1 操作系统的概述

1.1.1 什么是操作系统

操作系统 (Operating System) 简称 OS,它是现代计算机系统运行和工作必不可少的软件。作为特殊的系统软件,操作系统的主要目标就是对计算机系统的资源进行高效的管理,并向用户提供一个方便、易用的计算机操作环境。

无论是巨型机、大型机,还是中、小型机,也无论是台式个人计算机、便携式微型机,及多台计算机联网的计算机网络,都离不开操作系统。和任何事物发展一样,操作系统的形成和发展也是逐步演变的。

1946 年第一台计算机问世时,并没有操作系统,甚至没有任何软件。

早期计算机的工作,基本采用手工操作方式,即程序员将自己编写的程序和数据先用凿孔机送到纸带或卡片上,然后由操作员将纸带(或卡片)装入纸带输入机(或卡片输入机)等输入设备上,经手工启动输入设备,把程序和数据输入计算机内存,再由程序员启动程序。在程序运行的过程中出现问题时,由操作员借助扳键和指示灯查找问题,再利用扳键进行修改,再次启动程序,当程序运行完毕后,由用户取走纸带和计算结果,此后才能让下一个用户上机操作。这种手工操作方式有如下特点:

(1) 资源独占。即一台计算机的全部硬件资源(如处理机、存储器、外设等)均由一个用户独占使用,系统中的全部资源由一个用户支配,不会出现资源被其他用户占用而引起的等待现象。

(2) 串行工作。用户和用户之间,程序和程序之间,操作与计算机的运行,以及计算机各部件之间都是串行工作,这使得系统资源的利用率非常低。

(3) 人工干预。即计算机在人的直接联机干预下进行工作,人工操作的时候,

CPU 空闲等待。

随着半导体技术的飞速发展,计算机的运算速度大幅度提高,计算机各个部件和设备逐渐增多,手工操作方式的特点成为制约提高计算机发展和效率的严重障碍。人们开始设想通过专门编写程序实现计算机的使用、管理和操作,该程序开始时称为监督程序(monitor),只是担当计算机的输入输出、运行控制、出错处理等工作。开始时监督程序与用户程序一起装入内存,通过运行监督程序管理计算机的资源和控制用户程序的运行。

后来,随着计算机硬件运行速度的提高,监督程序的程序结构和功能不断完善,监督程序扩展成为计算机的操作系统。从资源管理观点来认识操作系统,它是管理计算机系统硬、软件资源,成为最大可能地提高计算机资源利用率、方便用户使用计算机的大型系统软件。

1.1.2 计算机系统中的操作系统

对于操作系统的认识可以有两个方面:一是从外部认识操作系统;二是从内部认识操作系统。

1. 从外部认识操作系统

从操作系统的外部来认识操作系统,长期以来有两种主要观点,它们表现了操作系统的两种基本功能,一种是虚拟机的观点,另一种是用户环境的观点。

(1) 操作系统作为虚拟机。对多数计算机而言,在机器语言一级的体系结构(机器指令集、存储组织、I/O 和总线结构)上编程是很困难的,尤其是输入输出操作。例如进行软盘的 I/O 操作,最基本的命令是读数据和写数据。通过指令参数的设置指定的信息有:欲读取的磁盘块地址、每条磁道的扇区数、物理的数据记录格式、扇区间隔以及对删除数据地址标识的处理方法等。显然,程序员不想涉及硬件的这些具体细节(对于其他的外设也同样如此,只是参数的设定因外设的不同而异),而需要一种简单的高度抽象的设备。

有一种典型的方法是把设备抽象成一组命名的文件,通过对该文件的打开、读写,到最后的关闭实现对设备的操作。设备一经抽象,设备操作细节就对用户隐藏。

这种将硬件的细节与程序员隔离开来,同时提供一种简洁的命名文件方式的程序,就是操作系统。它还隐藏了其他低层的硬件的特性,包括中断、时钟、存储器等。

(2) 用户环境的观点。从用户来看,操作系统是一个工作环境,是用户与计算机系统打交道的界面。用户使用操作系统为他们提供的各种服务、功能接口、系统调用等来使用计算机系统及其各类资源。用户不必知道各种系统资源的细节和控制过程。

用户当然关心用什么方法与操作系统打交道,这就是所谓操作系统界面问题。目前,这个问题已成为系统的重要组成部分。随着硬件不断发展,操作系统在不断更新换代,界面已经从字符界面、菜单界面、窗口界面发展到目前的图形界面和多媒体

界面。如今的操作系统设计者不再是只考虑如何提高硬件资源的利用率,同时还必须考虑如何有效地为用户服务,这样界面在系统中的位置显得越发重要。

2. 从内部认识操作系统

从内部来认识操作系统主要看它的实体结构,即操作系统是如何组成的,如何工作,如何对外提供服务。有一种称为资源管理的观点,这种观点是自底向上的观点,操作系统则用来管理一个复杂系统的各部分。操作系统的任务是在相互竞争的程序间有序地控制设备的分配。

设想在一台计算机上运行的三个程序同时试图在一台打印机上输出计算结果,如果按照头几行是程序 1 的输出,接下来几行是程序 2 的输出,然后又是程序 3 的输出等,最终造成打印结果的混乱。操作系统采用将打印输出到磁盘上的方法可以避免这种混乱。当一个程序结束后,操作系统将暂存磁盘文件上的输出结果送到打印机,同时其他程序可以继续运行,产生新的输出结果也是送到磁盘文件暂存,而这些程序并不知道这些输出没有立即送至打印机。

当一台计算机(或网络)有多个用户时,因为用户间可能相互影响,所以管理和保护存储器、I/O 设备以及其他设备的需求随之增加。而且用户往往不仅需要共享硬件,还要共享信息(文件、数据库等)。总之,此时操作系统的首要任务是跟踪资源的使用状况、满足资源的请求、提高资源利用率,以及协调各程序和用户对资源的使用冲突。

(1) 资源管理的观点。资源管理中,操作系统的任务是如何使系统资源得到充分合理地使用、解决用户作业争夺资源产生的矛盾,主要功能是:

1) 监视资源。保持资源分配的全局信息,了解系统资源的总数,已分配和未分配的资源,资源的增减和变动情况,通过系统中各类数据结构和表格记录资源的使用情况和环境状态。

2) 分配资源。处理对资源的使用请示,调解请求中的冲突,确定资源分配策略,当多个用户竞争某个资源时,进行仲裁。同时,根据资源分配的条件、原则和环境,确定是否立即分配。对可以分配的资源,记录相应的分配情况,更新记录分配信息的数据结构和表格的内容。

3) 回收资源。用户使用资源结束,提出释放请求,系统按照与分配过程相反的操作回收资源,更新相应的数据结构和表格内容。

4) 保护资源。对系统共享资源要进行保护,防止它们被有意无意地破坏。在多用户系统中,对每一个用户及程序运行进行保护,防止彼此间的干扰和冲突。对系统或者网络中重要的、敏感的(保密的)信息也进行保护。同时防止非授权的操作,防止来自系统内部或者外部的人侵。

(2) 任务组织观点。计算机系统要完成用户的各种请求,这些请求通过作业或者任务的方式提交给系统,它们各自都是一次独立计算过程,操作系统则是在众多的计算过程之间切换处理机,它是计算机系统工作流程的组织者,负责协调在系统中运

行的各个任务的推进速度,提高系统的吞吐率。

大多数操作系统引进了进程的概念,将其作为工作基本单位。随着操作系统的发展,根据系统工作的层次和并发控制程度引入了作业、任务、进程、线程,以及对象等概念,对它们进行组织和管理是操作系统的重点。

操作系统的内部运行过程,也是进程状态不断变迁的过程。一个用户作业从外部提交给系统,就等待调度,等待被装入内存。一旦作业被装入,根据情况会创建若干任务或者进程,并等待处理机的分配。一旦进程得到处理机就开始运行,运行中的进程要么正常运行结束而退出,要么被某个事件中断暂时停止,而转去运行另一进程,直至所有进程全部终止。操作系统也就主导这些进程管理和调度。

综上所述,我们可以说操作系统是对计算机系统中所有资源进行高效管理的一种系统软件,是所有软件运行的基础,是用户使用计算机的接口。

1.2 操作系统的功能和特征

1.2.1 操作系统的功能

操作系统的目标之一是提高系统资源的利用率,为了提高资源的利用率,应尽可能地提高设备和程序并行执行程度,为此引入了多道程序技术。所谓多道程序技术是指,内存中同时放几道程序,允许它们交替执行,共享系统中的各种硬、软件资源。当一道程序因 I/O 请求而暂停原来程序的执行(程序在 I/O 请求的指令处停止向前推进),而通过 I/O 指令启动设备做输入、输出工作,这时 CPU 已经空闲,如果能让 CPU 立即转去执行另一道程序,不仅使 CPU 能够得到充分利用,同时还可以提高 I/O 设备和内存的利用率,允许多道程序运行的系统称为多道程序系统。

在多道程序环境下,用户提交给系统作业所需资源的总和,远远多于系统所拥有的资源,致使系统不能同时满足所有作业对资源的要求,各作业为使自己投入运行,必将引起对资源的争夺,为使多道程序能有条不紊地运行,操作系统必须具备以下五个方面的功能。

1. 存储管理——系统存储空间的有效利用

(1) 存储的分配和回收。在计算机系统中,多道程序并发执行的首要条件是各道程序有自己的内存空间,因此,为多道程序分配内存是存储管理的首要功能。当多道程序运行结束后,必须释放它所占用的存储空间,由存储管理程序回收以便再分配。常用的分配方式有静态分配(程序运行前一次性分配)和动态分配(运行过程中随时分配)。

(2) 信息保护。由于硬件的限制,存储器的容量是有限的。如果由多个用户程序共享存储器,它们彼此之间不能相互冲突和干扰,尤其是不能让用户程序破坏操作系统常驻内存程序,存储管理必须提供信息保护措施。即每道程序在执行时随时检

查对内存的访问权限是否合法,若越界则拒绝执行,发出越界信号。

(3) 地址变换机构和地址变换功能。由目标程序所限定的地址范围称为该程序的地址空间,地址空间的地址是逻辑地址。内存空间是内存物理地址的集合。在多道程序系统中,操作系统必须把程序地址空间中的逻辑地址转换为内存空间的物理地址,需要内存管理中有地址变换机构和地址映射功能。

内存的大小会直接影响大型作业或多个作业的并发执行,为了满足用户对内存容量的要求,并改善系统性能,必须对内存加以扩充,但由于价格、技术上的原因,我们无法去无限地增大内存空间,而只好采用虚拟存储技术解决这个问题,详见第4章。

2. 处理机管理

处理机管理也称进程管理。因为在多道环境下,处理机的分配和运行都是以进程为基本单位进行(进程是指程序的一次执行)。例如,进程因某事件而无法继续执行时,应引起对处理机的重新分配。因此对处理机的管理就是通过引入进程的概念,提高 CPU 的利用率。

进程管理应具备的功能包括:进程控制、进程同步、进程通信、进程调度。

3. 设备管理

随着信息社会的发展,计算机的外部设备得到了迅速发展,这些设备的多样化大大改变了传统计算机外部设备的观念,如何有效地分配和使用设备,如何处理与设备操作间的时间差异,提高系统总体性能,就是操作系统设备管理模块的主要功能。

设备管理主要包括:缓冲区管理、设备分配、设备处理、实现虚拟设备功能。

4. 文件管理

在现代计算机系统中,总是把大量信息(程序和数据)以文件的形式存放在外存储器中,供用户使用。系统也允许用户把处理结果保存在系统中,供以后使用。这样,操作系统中配置了文件管理系统,作为信息管理机构。

文件管理系统的主要功能包括:文件存储空间的管理、目录管理、对文件的读和写、文件保护、提供接口。

5. 作业管理

作业管理的主要任务是根据系统和用户的要求,对作业的运行进行合理的组织及相应的控制。作业管理具备以下功能:

(1) 作业调度。作业调度是指根据系统的能力和当前作业运行情况,按一定策略,从后备作业队列中选出一批作业,为它们分配所需要的 I/O 设备和存储空间,将它们调入内存并为之建立相应的进程,使之成为获得处理机资格的备选进程。

(2) 作业控制。作业控制是指作业从进入系统开始,直到运行完成的整个过程中,用户可以通过某种形式发出各种命令,以对自己的作业进行控制和管理。在实际中,作业有两种类型:一是批量型作业;二是终端型作业。这两种作业有不同的输入和运行方式,这就形成两种作业的控制方式,分别称为脱机控制方式和联机控制方