

操作系统原理
与
Linux
系统

马季兰 冯秀芳等 编著

人民邮电出版社

操作系统原理与 Linux 系统

马季兰 冯秀芳等 编著

人民邮电出版社

内 容 提 要

本书全面讲述了操作系统的一般原理，并在此基础上详细讲述 Linux 操作系统的基础知识和实现原理，Linux 系统的安装、启动、系统管理和网络功能等内容。本书将一般操作系统的基础知识与 Linux 系统的相关知识分章对照讲解，便于读者学习和掌握 Linux 操作系统原理和使用方法。本书既可作为计算机专业本、专科学生教材，也可作为计算机工程技术人员的参考资料。

操作系统原理与 Linux 系统

- ◆ 编 著 马季兰 冯秀芳 等
责任编辑 张瑞喜
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
北京鸿佳印刷厂印刷
新华书店总店北京发行所经销
 - ◆ 开本:787×1092 1/16
印张:17.25
字数:419 千字 1999 年 10 月第 1 版
印数:5 001 - 9 000 册 2000 年 3 月北京第 2 次印刷
- ISBN 7-115-08048-8/TP·1268
-

定价:26.00 元

前 言

操作系统是现代计算机系统中最主要的系统软件，是计算机专业的必修课程，也是从事计算机工作的技术人员必不可少的知识。而 Linux 是符合 POSIX 标准的、具有全部 UNIX 功能的操作系统。Linux 是真正的多用户、多任务操作系统。在个人计算机和工作站上使用 Linux，能更有效地发挥硬件的功能，使个人计算机可以作为工作站和服务器使用。Linux 的源程序是开放的，它是允许任何人获取、修改和重新发布的自由软件。Linux 的基本系统对硬件要求很低。可以从网上下载 Linux 或用很低的价格买到 Linux 的光盘。Linux 继承了 UNIX 的主要特征。它在 Internet 的应用中占有明显的优势，在教学科研领域具有广阔的应用前景。

本书一、三、五、七、九、十一章讲述了操作系统的基础知识，二、四、六、八、十、十二章讲述了 Linux 系统的实现原理。十三、十四、十五章讲述了 Linux 系统的使用，包括安装、启动、系统管理和网络功能等。本书既讲述了操作系统的基础知识，又讲述了 Linux 的实现原理和使用方法，便于用户学习和掌握 Linux 操作系统原理和使用。

在本书编写过程中，得到余雪丽教授的悉心指导和多方面的热情帮助，在此表示衷心感谢。

本书第一、三、五、六、七、八、九、十、十一、十二章由马季兰编写，第二、四章由冯秀芳编写，第十三、十四章由戴敬周编写，第十五章由李维民编写。全书由马季兰、冯秀芳审定。程涛、纪文平对本书的编写做了不少工作，在此一并表示感谢。

由于水平有限，时间比较仓促，难免有错误和不妥之处。恳请读者批评指正。

目 录

第 1 章 操作系统引论	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 操作系统和计算机系统的关系	2
1.2 操作系统的功能和特征	4
1.2.1 操作系统的功能	4
1.2.2 操作系统的特征	7
1.3 操作系统的分类	8
1.3.1 多道批处理系统	8
1.3.2 分时系统	10
1.3.3 实时系统	12
1.3.4 通用操作系统	13
1.4 分析研究操作系统的几种观点	14
1.4.1 操作系统的结构	14
1.4.2 操作系统的主要性能指标	14
1.4.3 分析研究操作系统的几种观点	15
第 2 章 Linux 操作系统概述	16
2.1 什么是 Linux	16
2.2 Linux 简史	16
2.3 Linux 特性综述	17
2.3.1 多任务	17
2.3.2 多用户	18
2.3.3 可编程 shell	18
2.3.4 多文件系统	18
2.3.5 支持的硬件多	18
2.3.6 丰富的软件	19
2.4 Linux 的发行版本	21
2.5 Linux 的发展前景	22
2.6 Linux 使用入门	24
2.6.1 Linux 的使用介绍	24
2.6.2 Linux 基本命令	26

第 3 章 作业管理	31
3.1 作业、作业步、作业流	31
3.2 作业类别和组织	31
3.3 批量型作业的管理和调度	32
3.3.1 作业的组成	32
3.3.2 作业的进入	32
3.3.3 作业的建立	33
3.3.4 作业调度	34
3.3.5 从作业到进程	34
3.3.6 作业的运行	35
3.3.7 作业的终止和完成	35
3.4 终端型作业的管理	36
3.4.1 作业的建立和进入	36
3.4.2 终端型作业与批量作业管理方面的其他区别	36
3.5 用户与操作系统的接口	36
3.5.1 系统调用	37
3.5.2 作业控制语言	40
3.5.3 键盘操作命令	41
3.5.4 键盘命令的执行过程	43
3.6 作业控制	44
3.6.1 脱机控制方式	44
3.6.2 联机控制方式	45
第 4 章 Linux 的用户接口	46
4.1 了解 Linux 的 Shell	46
4.2 Shell 命令的语法分析	47
4.2.1 Shell 变量	47
4.2.2 通配符	48
4.2.3 shell 脚本	48
4.3 shell 编程	49
4.3.1 使用 echo	49
4.3.2 使用 shell 程序中的变量	49
4.3.3 Shell 编程结构	50
4.4 shell 程序举例	53
4.5 X Windows 系统简介	56
4.6 X Windows 系统的安装和配置	56
4.7 X Windows 的使用	58
4.7.1 Xinitrc	58
4.7.2 窗口管理程序	59

第 5 章 并发程序和进程	61
5.1 并发程序	61
5.1.1 程序的顺序执行及其特点	61
5.1.2 程序的并发执行及其特点	62
5.2 进程的概念	64
5.2.1 进程的定义和特征	64
5.2.2 进程的状态及其转换	66
5.2.3 进程控制块	67
5.3 进程同步	68
5.3.1 临界资源和临界区问题	69
5.3.2 实现进程同步的机构	70
5.4 进程控制	76
5.4.1 建立进程原语	77
5.4.2 挂起进程原语	78
5.4.3 撤消进程原语	78
5.5 进程调度	79
5.5.1 进程调度的概念	79
5.5.2 进程调度程序	79
5.5.3 进程调度算法	80
5.6 死锁	82
5.6.1 死锁的概念	82
5.6.2 资源的概念	83
5.6.3 形成死锁的原因和必要条件	84
5.7 线程	85
5.7.1 线程的概念	85
5.7.2 线程和进程的关系	85
5.7.3 线程的执行状态及其转换	85
5.7.4 引入线程的优点	85
5.7.5 Windows NT 的线程特点	86
第 6 章 Linux 中的进程管理	87
6.1 Linux 中的进程	87
6.1.1 Linux 的进程控制块	87
6.1.2 Linux 中的进程标识	88
6.2 Linux 中的进程调度	88
6.3 Linux 中的进程控制	90
6.3.1 进程的建立和 fork () 系统调用	90
6.3.2 程序执行和系统调用 exec ()	91
6.3.3 等待子进程结束系统调用 wait ()	91

6.3.4	结束子进程系统调用 <code>exit()</code>	91
6.3.5	shell 进程工作过程	91
6.4	Linux 进程通信	92
6.4.1	信号	92
6.4.2	“管道”	93
6.4.3	Systemv 的进程通信机制	96
第 7 章	存储管理	98
7.1	引言	98
7.1.1	存储管理的研究内容	98
7.1.2	存储组织	99
7.1.3	存储管理及其策略	99
7.1.4	分级存储体系	99
7.2	实存管理技术	100
7.2.1	用户程序的处理步骤	100
7.2.2	地址重定位	100
7.2.3	固定分区多道程序设计	102
7.2.4	可变分区多道程序设计	103
7.2.5	覆盖技术	108
7.2.6	对换技术	109
7.2.7	简单页式存储管理	109
7.3	虚拟存储管理技术	112
7.3.1	概述	112
7.3.2	请求页式虚拟存储管理技术	113
7.3.3	分段式存储管理技术	114
7.3.4	段页式存储管理技术	116
第 8 章	Linux 中的存储管理	117
8.1	Linux 存储管理的特点	117
8.1.1	段页式虚存技术	117
8.1.2	新建进程只占两页内存	117
8.1.3	新建子进程从父进程处仅复制页表	117
8.1.4	写时复制	117
8.1.5	利用高速缓存 cache	118
8.2	Linux 内存管理	118
8.2.1	基本思想	118
8.2.2	实现技术	119
第 9 章	设备管理	123
9.1	概述	123

9.1.1	设备的分类	123
9.1.2	设备管理的目标和功能	125
9.2	外围设备的分配	126
9.3	I/O 控制方式	127
9.3.1	设备控制器	127
9.3.2	I/O 控制方式	127
9.4	设备驱动程序	130
9.5	缓冲技术	130
9.6	中断技术	131
9.6.1	基本概念	131
9.6.2	中断类型	131
9.6.3	中断响应和处理过程	131
第 10 章	Linux 中的设备管理	134
10.1	概述	134
10.2	Linux 中的设备控制方式	135
10.2.1	测试和中断	135
10.2.2	设备驱动器与内核的接口	135
10.2.3	字符设备的管理	136
10.3	块设备的管理	136
10.4	磁盘	136
10.5	终端	137
10.5.1	终端分类	137
10.5.2	终端软件	137
10.5.3	终端输入输出过程	138
10.6	I/O 软件	139
10.6.1	I/O 软件的目标	139
10.6.2	与设备无关的 I/O 软件	140
10.6.3	用户空间的 I/O 软件	140
10.6.4	设备驱动程序	146
第 11 章	文件管理	147
11.1	概述	147
11.2	文件结构和存取方法	148
11.2.1	文件的逻辑结构	148
11.2.2	文件的物理结构	148
11.3	文件的存取方法	150
11.4	文件存储设备简介	150
11.5	文件结构、存储设备、存取方法之间的关系	152
11.6	文件存储空间的管理	152

11.6.1	位示图法	152
11.6.2	成组链接法	153
11.7	文件目录	154
11.7.1	什么是文件目录	154
11.7.2	二级目录和多级目录	155
11.7.3	对文件目录的管理	157
11.8	文件的保护和保密	157
11.8.1	文件的共享	157
11.8.2	文件的保护和保密	158
11.9	文件的使用	158
第 12 章	Linux 文件系统	159
12.1	Linux 文件系统概述	159
12.2	Linux 文件系统的实现	160
12.3	虚拟文件系统转换 (VFS)	163
12.3.1	VFS 超级块	164
12.3.2	VFS 的 inode 节点	164
12.3.3	登记文件系统	165
12.3.4	加载文件系统	166
第 13 章	LINUX 的安装	167
13.1	Linux 的版本介绍	167
13.2	Linux 的获取	168
13.3	硬件的支持	169
13.4	RedHat Linux 安装指南	172
13.4.1	准备工作	172
13.4.2	启动机器	173
13.4.3	安装系统	174
13.5	常见问题解答	181
第 14 章	系统管理	183
14.1	Root 帐号与系统安全	183
14.2	引导系统	187
14.2.1	从软盘引导	187
14.2.2	从硬盘引导	188
14.2.3	用 Loadin 程序引导	190
14.3	系统初始化程序和文件	191
14.3.1	初始化信息	191
14.3.2	初始化程序和文件	193
14.4	关闭系统	204

14.5	用户管理	205
14.5.1	用户帐号的信息	205
14.5.2	创建帐号	206
14.5.3	修改帐号的属性	208
14.5.4	删除和查封帐号	209
14.5.5	组帐号的管理	209
14.6	管理文件系统	211
14.6.1	文件系统的基本概念和类型	211
14.6.2	创建文件系统	213
14.6.3	安装文件系统	214
14.6.4	检查, 维护文件系统	214
14.7	交换空间	215
14.8	系统升级	215
14.8.1	内核的升级	216
14.8.2	库的升级	216
14.8.3	GCC 的升级	217
14.8.4	其他软件的升级	217
第 15 章 网络管理与服务		218
15.1	网络管理	218
15.1.1	TCP/IP 简介	218
15.1.2	获得 Linux 的网络软件	222
15.1.3	配置 Linux 的 TCP/IP 网络	222
15.1.4	配置 Linux 的域名服务	227
15.1.5	Linux 的 SLIP 配置	232
15.1.6	Linux 的 PPP 配置	245
15.2	建立 Linux 的普通站点	247
15.2.1	安装远程主目录	247
15.2.2	设置网络其他设备	248
15.2.3	创建帐户	249
15.2.4	处理连接	250
15.2.5	使用虚拟控制台	254
15.2.6	关闭系统	254
15.3	Internet	254
15.3.1	Internet 简介	254
15.3.2	使用电子邮件	254
15.3.3	在 Linux 下安装 Web/Gopher/FTP/UseNt 服务器软件简介	256
15.3.4	使用 FTP/gopher/Usenet/telnet 简介	258
15.3.5	安全体系	260

第1章 操作系统引论

1.1 操作系统的概念

1.1.1 什么是操作系统

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：即程序员将自己编好的程序或数据先经穿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显示灯查找问题所在，并利用扳键进行修改，再次启动程序运行；当程序运行完毕后，由用户取走纸带和计算结果，带回去进行研究；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。用户与用户之间，程序与程序之间，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的，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，计算机的各种部件和设备逐渐增多，规模日益庞大，上述特点成了发展探索和提高计算机效率的严重障碍。比如：一个程序在速度为 1 千次/秒的计算机上运行，假设它所需要的运行时间为 1 小时，而建立作业和全部人工操作时间为 5 分钟，则运行时间和操作时间之比为 12：1；若该作业改为在 6 万次/秒的计算机上运行，则运行时间由 1 小时变为 15 分钟，而手工操作时间仍为 5 分钟，两者时间之比由 12：1 变为 3：1，说明有效机时浪费惊人。人们不得不提出这样的问题：既然计算机可用于各行各业，既然各种问题都可以通过编程由计算机去解决，为何计算机的使用、管理和操作就不能通过编程，由计算机自己运行程序去完成呢？有了这种思想以后，人们就把对计算机的输入输出，运行控制，出错处理等编为程序，最初称之为监督程序 (monitor)，并连同用户程序一起送入计算机，通过执行监督程序去管理计算机资源和指挥用户程序的运行，以摆脱人工干预。这个监督程序便是操作系统的雏型。

后来，随着计算机硬件的不断发展，运行速度的提高，监督程序的功能逐步扩展成为管理整个计算机系统硬、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，这便是现在的操作系统（Operating System, OS）。操作系统通过有效地控制和管理计算机系统中各种硬件和软件资源，合理地组织计算机系统的工作流程，以改善系统性能(如响应时间，吞吐量等)。此外，操作系统提供了一个计算机用户和计算机硬件之间的接口，以使用户更方便地使用计算机。

1.1.2 操作系统和计算机系统的关系

一、计算机系统的组成

我们平时见到的计算机仅是计算机硬件，而一个完整的计算机系统，不论是大型机、小型机、甚至微型机和微处理机，都由两大部分组成：即计算机的硬件部分和软件部分。

计算机硬件部分通常是指计算机物理装置本身的实物，它可以是电的、磁的、机械的、光的元件或装置，或者是由它们组成的计算机部件和计算机本体。总之，是指计算机系统中所有那些“硬”的设施。也就是指计算机的各种处理机（如中央处理机，输入输出处理机和包含在该计算机系统的其他处理机）、存储器、输入输出设备和通讯装置等。

软件部分是指计算机系统的所有软件，包括系统软件和应用软件。术语“软件”是相对于硬件而言的，它是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。计算机软件又分为系统软件和应用软件。系统软件用于计算机的管理、维护和运行，以及对运行的程序进行翻译、装入等服务工作。系统软件常指操作系统，包括各语言的编译和解释程序、汇编程序、装入程序、连接编辑程序、编辑程序、检查程序、调试程序、数据库管理系统、数据通讯等。应用软件常指那些为了某一方面的应用而设计的程序或用户为解决某个特殊问题而编写的程序。

那么计算机系统的各硬件是如何连接，构成一个完整的计算机呢？各软件之间是什么关系？硬件软件之间是什么关系？整个计算机系统和操作系统又是什么样的关系？图 1.1 给出计算机系统中各硬件之间的连接关系。

随着计算机技术的发展，计算机硬件的功能越来越强，软件资源也日趋丰富。从功能上，可把整个计算机系统划分为四个层次:硬件、操作系统、实用程序和应用程序，它们之间的层次关系如图 1.2 所示。

上述四个层次表现为一种单项服务关系，即外层的软件必须事先以约定好的方式使用内层软件或硬件提供的服务。通常把这种约定称为界面（Interface），各层次特点如下：

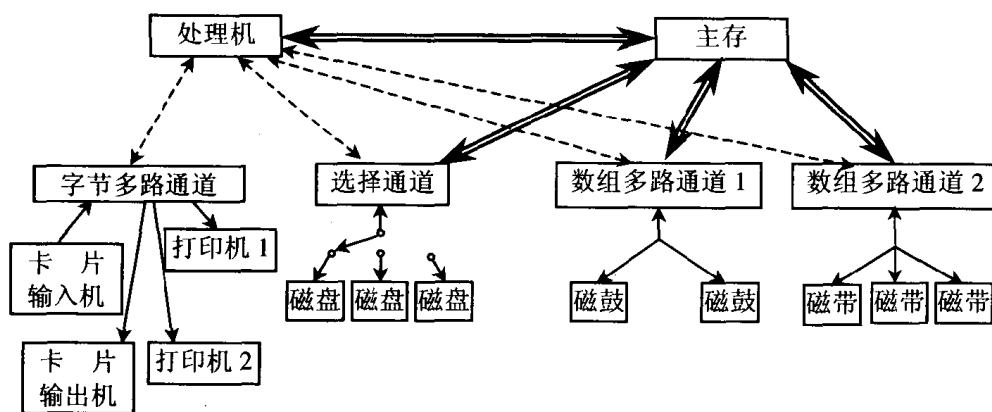


图 1.1 IBM370 硬件结构

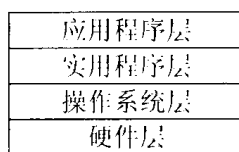


图 1.2 计算机系统的四个层次及关系图

1. 硬件

硬件层的对外界面由机器的指令系统组成。操作系统及其外层软件通过执行各种机器指令访问和控制各种硬件资源。机器指令系统与硬件系统的组织结构密切相关。为了能使操作系统高效地运行，硬件系统的组织结构不断地得到改进，与此同时，机器指令系统也变得复杂和庞大。图 1.1 是目前大、中、小型计算机系统经常采用的组织方式。

传统的计算机硬件一般都采用以 CPU 为中心的组织结构，但现代的计算机硬件则大都采用以主存为中心的结构。这种结构可使 CPU 和 I/O 设备并行工作，如图 1.1 所示。CPU 和各种 I/O 通道(例如图 1.1 中的几种通道)通过一个双端口的内存相互通讯，交换信息。I/O 通道实际上是一台专用的 I/O 处理机，它能执行由自己的通道指令编写的通道程序，从而控制系统中的高速、中速和低速外围设备并行工作。这样就使 CPU 摆脱了对各种外设的繁杂控制，从而充分地利用其高速度的特点集中进行计算。由于 I/O 通道还可以使各种外设之间并行地工作，因而整个计算机系统的处理能力，以及各种资源的利用率大大提高。

计算机硬件通常称为裸机，一个裸机的功能即使很强，也不方便用户使用，因为在裸机上运行程序必须用机器语言编写程序，这显然使用户感到十分不便，而且将严重降低工作效率和机器的利用率。

2. 操作系统

操作系统正是为了填补人与机器之间的鸿沟，即建立用户与计算机之间的友好界面，而为裸机配置的一种系统软件。操作系统是裸机上的第一层软件，也是最基本的系统软件，是对硬件系统功能的首次扩充。它密切地依赖于计算机硬件，直接管理着系统中各种硬件和软件资源。由于操作系统的主要部分驻留在主存中，所以，把它称为操作系统内核或核心。操作系统核心由五个部分组成：进程管理、存储管理、设备管理、文件管理和作业管理。

操作系统的对外界面是系统调用。实用层软件及各种应用软件均是通过系统调用访问计算机系统的软、硬件资源的。所谓系统调用，实际上是由操作系统软件提供的，能访问系统核心的接口程序。

3. 系统实用程序

实用程序层的界面是由一组在操作系统控制下运行的实用程序构成。与操作系统核心程序不同，这些程序通常驻留在磁盘上，仅当需要运行这些程序时，才把它们装入内存。实用程序的功能是为应用层软件以及最终用户加工自己的程序或数据提供服务。此外，计算机系统的管理员还要利用实用程序对系统进行日常维护。实用程序是计算机系统的基本组成部分，它们通常由计算机系统的销售者提供，并随机器的硬件及操作系统一起出售。实用程序主要有以下几类：

(1) 文本编辑程序。其功能是将用户编写的程序或用户提供的数据送到计算机外存储器中，从而形成可以长期保存的文件。通常有两类文本编辑程序，一是行编辑程序，以行为单位进行编辑和修改，如 EDLIN (PCDOS) 和 ed (UNIX)。另一种是全屏编辑程序，它可

以在全屏幕范围内进行编辑和修改，如 WPS (DOS) 和 vi (UNIX)。在有条件的情况下，建议使用全屏幕编辑程序。

(2) 装配程序。装配程序的逻辑功能是将浮动地址的目标程序装配成绝对地址的可执行程序，即从地址空间的逻辑地址向存储空间的物理地址的映射。在不同的系统中，常用的装配命令有 Load、Link、ld、ln 等。

(3) 查错调试程序。该程序主要用于检查系统或用户有何错误，错在何处。常用的调试程序有 DEBUG 和 ODT (联机调试技术) 等。

(4) 语言处理程序。软件系统中的语言处理程序一般由编译 (汇编解释) 程序和汇编程序两部分组成。编译程序是将高级语言的源程序翻译成为汇编语言程序或目标程序。汇编程序是将汇编语言的源程序翻译成目标程序。

4. 应用程序

计算机系统的最外层软件是应用程序。这些程序通常由计算机用户或专门软件公司编制。例如数据库管理系统、办公自动化系统、事务处理系统等。这些应用软件常作为计算机系统的选件，用户根据自己的需要选择购买。

二、操作系统和计算机系统的关系

操作系统是计算机系统的重要组成部分，用于管理、指挥、控制整个计算机系统的工作，这就是操作系统和计算机系统的辩证关系。操作系统在计算机系统中占据特殊重要的地位，所有其他软件，都是建立在操作系统基础上的，并得到它的支持和服务的；用户则是利用操作系统所提供的命令和服务去操纵和使用计算机的。对一个计算机系统来说，若没有操作系统，犹如一个人没有大脑思维，它将一事无成。

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的并行执行程度，为此引入了多道程序技术。所谓多道程序技术是指：内存中同时存放几道程序，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 请求而暂停执行时，CPU 立即转去执行另一道程序。这样不仅使 CPU 能得到充分利用，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，用户提交给系统的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，从资源管理的角度来看，操作系统应具有以下五方面的功能。

一、存储器管理

存储器管理具备以下功能：

(1) 内存的分配和回收。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。常用的存储分配方式有静态分配法（一次性分配）和动态分配法（运行过程中随时分配）。

(2) 信息保护。为保证各道程序能在自己已分到的内存空间中正确运行，互不干扰，互不破坏，尤其是不能让用户程序破坏操作系统常驻内存的程序，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，提出越界错误信号。

(3) 地址变换机构和地址变换功能（地址映射）。由目标程序所限定的地址范围称为该程序的地址空间，地址空间中的地址是逻辑地址。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中的物理地址。当然要有地址变换的机构和地址映射功能。

内存的大小会直接影响到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，必须对内存容量加以扩充。但由于价格、技术等方面的客观原因，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来解决此类问题。

二、处理机管理

处理机管理也叫进程管理。因为在多道程序环境下，处理机的分配和运行都是以进程为基本单位进行的（进程是指程序的一次执行）。假如某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。进程管理应实现下述主要功能：

(1) 进程控制。当用户作业要运行时，应为之建立一个或多个进程，并为它分配除处理机以外的所需资源，将它们放入进程就绪队列。当进程运行完成时，立刻撤消进程，以便及时释放其所占用的资源。进程控制的基本功能就是创建和撤消进程以及控制进程的状态转换。

(2) 进程同步。所谓进程同步是指系统对并发执行的进程进行协调。最基本的进程同步方式是诸进程以互斥方式访问临界资源（即一次只允许一个进程使用的资源）。此外，对于彼此相互合作，去共同完成任务的诸进程，则应由系统对它们的运行速度加以协调。

(3) 进程通信。对于相互合作的进程，在它们运行时，相互之间往往要交换一定的信息。例如，进程 A 需要把它运行的结果数据传送给另一相关进程 B，而进程 B 仅在收到该数据后方能开始执行。必要时，进程 B 还应向进程 A 发送确认信息。我们把这种进程之间所进行的信息交换称为进程通信。

(4) 进程调度。当一个正在执行的进程已经完成或因某事件而无法继续执行时，系统应进行调度，重新分配处理机。进程调度是指按一定的算法，从进程就绪队列中选出一进程（就绪队列中的进程处于一种“万事俱备，只欠处理机”的状态），把处理机分配给它，为该进程设置运行现场，并使之投入运行。

三、设备管理

设备管理的主要任务是：

- ① 为用户程序分配 I/O 设备。
- ② 完成用户程序请求的 I/O 操作。
- ③ 提高 CPU 和 I/O 设备的利用率。
- ④ 改善人机界面。

为实现上述任务，设备管理程序应具有下述功能：

(1) 缓冲区管理。几乎所有的外围设备与处理机交换信息时,都要利用缓冲区来缓冲 CPU 和 I/O 设备间速度不匹配的矛盾,提高 CPU 与外设、外设与外设之间操作的并行程度,以提高 CPU 和 I/O 设备的利用率。因此,在系统中设置了许多类型的缓冲区,系统必须对它们进行有效管理。

(2) 设备分配。系统根据用户所请求的设备类型和所采用的分配算法对设备进行分配,并将未获得所需设备的进程列入相应设备的等待队列。

(3) 设备处理。启动指定的 I/O 设备,完成用户规定的 I/O 操作,并及时响应由设备发来的中断请求,根据中断类型进行相应的处理。

(4) 实现虚拟设备功能。通常,把一次仅允许一个进程使用的设备(如打印机)称为独享设备。系统可通过某种技术(如 Spooling 技术),使该设备成为能被多个用户共享的设备,既可提高利用率又能加速程序的执行过程。可使每个用户都感觉到自己在独占该设备。如前所述,这种只为用户所感觉到而实际上并不存在的设备,称为逻辑设备或虚拟设备。

四、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,操作系统中又配置了文件管理系统,作为信息管理机构。文件管理系统应具有下述五个功能:

(1) 文件存储空间的管理。通常,系统中的文件都存放在外存中,供多个用户共享。若由用户自己对文件的存储进行管理,不仅极其困难,而且效率也很低,这就要求文件系统对文件存储空间进行分配与回收。对于磁盘空间的管理,通常是以盘块为单位进行的,典型的盘块大小为 512 字节或 1024 个字节。

(2) 目录管理。为使用户方便地存取所需文件,通常由系统为每个文件建立一个文件控制块。它包含文件名、文件属性、文件所在的物理位置及有关其他信息。而若干个文件控制块的集合构成文件目录。对于大型文件系统,又把文件目录也组织成文件,叫目录文件,文件系统应妥善地对它们加以组织,以便于对文件目录的查找,获得令人满意的文件检索速度。

(3) 对文件的读或写。文件读写管理系统在对某文件进行读操作时,根据用户给定的目标地址和要传送的字节数,将文件信息从外存储器读入缓冲区后,再复制到指定的用户区或系统区。文件的写操作是根据用户指定的源地址和要传送的字节数,把信息从指定的用户区写到磁盘上。

(4) 文件保护。为防止文件被其他用户有意无意地破坏或盗用,应对文件进行保护和加密,保护的措施有口令、存取权限、加密等。

(5) 提供接口。为了方便用户,操作系统应向用户提供一个使用方便的接口,用户通过该接口,取得文件系统的服务。例如,使用建立、打开、关闭、撤消文件,读或写文件等系统调用,供编程时使用。

五、作业管理

作业管理的主要任务是根据系统和用户要求,对作业的运行进行合理的组织及相应的控制。作业管理应具备两方面的功能:

(1) 作业调度。作业调度是指根据系统的能力和当前作业的运行情况,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,使之成为具有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