

● 高等学校 21 世纪计算机教材

计算机 操作系统

Operating System

王 清 李光明 李永勤 编著



Technoblog
Information

冶金工业出版社

计算机操作系统

王 清 李光明 李永勤 编著

北 京

冶金工业出版社

内 容 简 介

操作系统是计算机系统中一个重要的系统软件。本书从操作系统的基本概念和基本功能讲起,着重介绍了操作系统的基本工作原理、实现技术以及几种流行的操作系统。主要包括处理器管理、存储管理、文件管理、设备管理、作业管理、并发进程和死锁等内容,此外,为了加深大家对操作系统基本原理的理解,最后通过实例介绍了当今比较流行的操作系统在实际工作中的应用。

本书针对性强,结构清晰,通俗易懂,便于讲授和自学,同时力图反映操作系统技术的新发展和新成果。

本书既可作为高等院校相关专业的教材,也可供各类培训班和自学人士使用。

图书在版编目(CIP)数据

计算机操作系统 / 王清等编著. — 北京:冶金工业出版社, 2003. 1
ISBN 7-5024-3202-7

I. 计... II. 王... III. 操作系统 IV. TP316

中国版本图书馆 CIP 数据核字(2002)第 100193 号

出版人 曹胜利(北京沙滩嵩祝院北巷 39 号,邮编 100009)

责任编辑 程志宏

广东出版技校彩印厂印刷;冶金工业出版社发行;各地新华书店经销

2003 年 3 月第 1 版,2003 年 3 月第 1 次印刷

787mm×1092mm 1/16; 15.5 印张; 374 千字; 240 页; 1-2500 册

25.00 元

冶金工业出版社发行部 电话:(010) 64044283 传真:(010) 64027893

冶金书店 地址:北京东四西大街 46 号(100711) 电话:(010) 65289081

(本社图书如有印装质量问题,本社发行部负责退换)

前 言

一、关于操作系统

操作系统是配置在计算机硬件上的第一层软件，是对硬件系统的第一次扩充。它在计算机系统中占据了重要的地位，其他所有软件，如：汇编程序、编译程序、数据库管理系统等系统软件以及大量的应用软件，如Office办公软件等，都依赖于操作系统的支持，都需要它的服务。

操作系统已成为大型机、小型机直到微机、计算机网络都必须配置的软件。随着操作系统的运行平台——计算机体系结构的发展，计算机操作系统也得到了发展。操作系统从最初的监督程序，发展成我们熟知的DOS磁盘操作系统、Windows 3.X、Windows 95/Windows 98、Windows 2000、Windows XP，还有UNIX和Linux等。可以说，新的技术、新的概念层出不穷。

二、如何学习操作系统

操作系统的学习，到底要理解和掌握一些什么呢？又如何去学习这门课程呢？

对操作系统的学习，主要考查我们对操作系统基本概念、基本功能、基本思想和基本算法的掌握程度，以及利用这些知识去分析、解决一般的理论问题和实际问题的能力。

只有掌握了操作系统中基本的、共性的知识，才能基于这些共性的知识去理解和掌握各种操作系统的特性部分，才能深入到操作系统的内部，理解整个系统的工作流程是怎样控制的、系统资源是如何共享的、如何有效管理和分配资源以及用户又如何与系统交互作用等知识。归结为：处理器管理（第2章）、存储管理（第3章）、文件管理（第4章）、设备管理（第5章）及作业管理（第6章）等几大管理。有了这些知识，就为以后的学习和理解新型的操作系统，应用操作系统和在操作系统基础上开发自己的应用系统打下了基础。

三、本书的内容结构

本书是作者根据《操作系统大纲》的要求，结合多年的教学经验编写而成的。全书带着操作系统能做什么，怎么去做，为什么要这样做三个问题讲述。全书结构如下：

第1章：概论。介绍了操作系统的基本概念、特性以及硬件环境。

第2章：处理器管理。先引入进程概念，然后阐述了进程控制、进程调度以及调度算法等问题。

第3章：存储管理。介绍了存储管理的基本概念，讲述了分区管理、分页管理、分段管理、段页式管理等方法与特点，对覆盖与交换技术、虚拟存储技术等也做了描述。

第4章：文件管理。在介绍了文件基本概念的基础上，详细介绍了文件目录、文件结构、文件存取、文件的保护与保密及文件的使用等问题，重点介绍了磁盘存储空间管理的三种方法。

第5章：设备管理。主要介绍了I/O系统的结构、I/O缓冲区、设备的分配、磁盘的调度、设备的操作和虚拟设备等内容。

第6章：作业管理。介绍了批处理作业和交互式作业的管理。

第7章：并发进程。主要介绍了并发进程的基本概念、进程同步、信号量的应用、管程

的概念及进程间的通信等内容。

第8章：死锁。介绍了死锁的产生、死锁的特征、死锁的预防、死锁的避免与检测等内容，重点介绍了银行家算法和死锁定理。

第9章：操作系统实例。介绍了UNIX操作系统和MS-DOS操作系统的进程管理、存储管理、文件管理、设备管理及作业管理与用户接口等内容，然后对Windows 2000/XP操作系统和Linux操作系统作了介绍。

此外，本书的最后还附有参考答案，以方便读者对照课后的习题进行练习。

本书是按80学时编写的，第9章可以选学。

四、本书的特点

本书简明易懂，条理清晰，用简练的语言结合大量的图、表较详细的介绍了操作系统的基本原理和应用知识，并通过常用的操作系统实例总结了全书的知识要点。通过本书的学习，读者可以很快掌握操作系统的相关知识。

此外，本书每章后都提供了丰富的习题，读者可参照书后的参考答案进行练习，以巩固所学的知识。

五、本书的适用对象

本书既可作为高等院校相关专业的教材，也可供各类培训班和自学人士使用。

本书由王清担任主编，并编写第1章、第2章、第3章、第7章、第8章；李光明编写第4章、第5章和第6章，李永勤编写第9章。王清对全书进行统稿，曹琳教授担任主审。

读者如果有好的意见或建议，可以到相关网站进行探讨。网址：<http://www.cnbook.net>。

由于时间仓促，水平有限，书中不妥之处在所难免，敬请读者指正。

编 者
2003年1月

目 录

第 1 章 概论	1	2.2.5 阻塞进程 (block)	26
1.1 计算机系统	1	2.2.6 唤醒进程 (wakeup)	26
1.1.1 计算机硬件系统	1	2.3 进程调度	26
1.1.2 计算机软件系统	2	2.3.1 处理器调度级别	27
1.1.3 计算机系统层次结构	2	2.3.2 调度的功能与时机	28
1.2 操作系统的资源管理	3	2.3.3 调度方式	28
1.2.1 操作系统的功能	3	2.3.4 调度准则	29
1.2.2 操作系统的层次结构	4	2.4 调度算法	31
1.3 操作系统的服务与分类	5	2.4.1 先来先服务算法	31
1.3.1 操作系统的分类	6	2.4.2 最短者优先算法	32
1.3.2 操作系统的服务接口	9	2.4.3 最高响应比者优先算法	32
1.4 操作系统的特性	11	2.4.4 轮转法	34
1.5 操作系统的硬件环境	13	2.4.5 最高优先级算法	34
1.5.1 程序状态字	13	2.4.6 多级反馈队列算法	36
1.5.2 中断机构	13	2.5 线程概念	37
1.5.3 存储保护	16	2.5.1 线程的引入	37
小结	16	2.5.2 线程的描述	38
习题一	17	2.5.3 线程的模式	39
一、选择题	17	2.6 多处理器系统	40
二、填空题	17	小结	41
三、思考题	18	习题二	42
四、应用题	18	一、选择题	42
第 2 章 处理器管理	19	二、填空题	43
2.1 进程概念	19	三、思考题	43
2.1.1 多道程序设计	19	四、应用题	43
2.1.2 进程的定义	20	第 3 章 存储管理	45
2.1.3 进程的状态	21	3.1 概述	45
2.1.4 进程控制块	22	3.2 地址映射	46
2.1.5 进程结构	23	3.2.1 逻辑地址	46
2.2 进程控制	24	3.2.2 物理地址	47
2.2.1 创建进程 (create)	25	3.2.3 地址映射方式	47
2.2.2 停止进程 (halt)	25	3.3 分区管理	48
2.2.3 挂起进程 (suspend)	26	3.3.1 固定分区管理	48
2.2.4 激活进程 (activate)	26	3.3.2 可变分区管理	51

3.3.3 地址转换与存储保护	54	4.3.3 树形目录结构	80
3.4 覆盖与交换	55	4.3.4 无环图目录结构和 通用图目录结构	81
3.4.1 覆盖 (Overlay)	55	4.4 文件的组织结构	81
3.4.2 交换 (swapping)	56	4.4.1 文件的逻辑结构	82
3.5 分页管理	56	4.4.2 文件的物理结构	82
3.5.1 实现原理	56	4.4.3 记录的成组与分解	84
3.5.2 存储分配	57	4.5 磁盘存储空间的管理	84
3.5.3 碎片概念	58	4.5.1 位示图	84
3.6 分段管理	59	4.5.2 空闲块表	85
3.6.1 分段引入	59	4.5.3 空闲块链	85
3.6.2 实现原理	59	4.6 文件的保护与保密	86
3.6.3 分段与可变分区的区别	60	4.6.1 文件的共享	87
3.6.4 分段与分页的区别	60	4.6.2 文件的保护	87
3.7 段页式管理	60	4.6.3 文件的保密	88
3.7.1 实现原理	61	4.7 文件的操作与使用	88
3.7.2 数据结构	61	4.7.1 文件操作	88
3.7.3 地址映射	61	4.7.2 文件的使用	89
3.8 虚拟存储器管理	61	4.8 虚拟文件系统	89
3.8.1 基本概念	62	小结	90
3.8.2 请求分页管理	62	习题四	92
小结	68	一、选择题	92
习题三	70	二、填空题	93
一、选择题	70	三、思考题	93
二、填空题	71	四、应用题	93
三、思考题	72	第5章 设备管理	95
四、应用题	72	5.1 概述	95
第4章 文件管理	75	5.2 I/O 系统的结构	95
4.1 概述	75	5.2.1 I/O 中断处理程序	96
4.1.1 文件概念	75	5.2.2 设备驱动程序	96
4.1.2 文件系统的功能	75	5.2.3 逻辑 I/O 系统	96
4.1.3 文件分类	76	5.3 I/O 缓冲区	97
4.2 文件的存取方式	77	5.3.1 缓冲方式	97
4.2.1 文件的存储介质	77	5.3.2 缓冲区的设置	98
4.2.2 存取方式	78	5.4 外围设备的分类与分配	98
4.3 文件目录管理	78	5.4.1 外围设备的分类	99
4.3.1 一级目录结构	79	5.4.2 设备的绝对号与相对号	99
4.3.2 二级目录结构	79		

5.4.3 设备的分配	100	7.2 与时间有关的错误	130
5.5 存储设备简介	100	7.3 进程同步	131
5.6 磁盘的调度	102	7.3.1 同步概念	131
5.7 设备的操作	106	7.3.2 临界区	131
5.7.1 通道结构与通道程序	107	7.3.3 同步机构	132
5.7.2 CPU 与外围设备的操作	108	7.4 信号量的应用	133
5.7.3 I/O 中断处理	109	7.4.1 临界区的互斥	133
5.8 虚拟设备	109	7.4.2 合作进程的同步	134
5.8.1 概述	110	7.4.3 生产者与消费者关系	135
5.8.2 SPOOLing 技术	110	7.4.4 读者-写者关系	137
小结	111	7.5 管程概念	138
习题五	112	7.6 进程通信	141
一、选择题	112	小结	142
二、填空题	114	习题七	143
三、思考题	114	一、选择题	143
四、应用题	114	二、填空题	144
第 6 章 作业管理	116	三、思考题	144
6.1 作业概念	116	四、应用题	145
6.1.1 作业和作业步	116	第 8 章 死锁	146
6.1.2 作业控制方式	116	8.1 死锁的产生	146
6.2 批处理作业的管理	117	8.2 死锁的必要条件	148
6.2.1 批处理作业的组织	117	8.2.1 互斥条件	148
6.2.2 批处理作业的输入	117	8.2.2 占有并等待条件	148
6.2.3 批处理作业的调度	118	8.2.3 非剥夺条件	148
6.2.4 批处理作业的控制	122	8.2.4 循环等待条件	148
6.3 交互式作业的管理	123	8.3 资源分配图	149
6.3.1 交互式作业的概念	123	8.4 死锁定理	152
6.3.2 交互式作业的控制	123	8.5 死锁的预防	154
6.3.3 终端作业的管理	125	8.5.1 破坏互斥条件	154
小结	125	8.5.2 破坏占有并等待条件	154
习题六	126	8.5.3 破坏非剥夺条件	155
一、选择题	126	8.5.4 破坏循环等待条件	155
二、填空题	127	8.6 死锁的避免	156
三、思考题	128	8.6.1 安全状态	156
四、应用题	128	8.6.2 银行家算法	157
第 7 章 并发进程	130	8.7 死锁的检测与解除	161
7.1 并发进程概念	130	8.7.1 死锁的检测	161

8.7.2 死锁的解除	162	9.3.4 Windows 2000/XP 的 文件管理	202
小结	163	9.3.5 Windows 2000/XP 的 I/O 管理	206
习题八	164	9.4 Linux 操作系统简介	208
一、选择题	164	9.4.1 Linux 的内核	209
二、填空题	165	9.4.2 Linux 系统数据结构	210
三、思考题	165	9.4.3 Linux 的进程管理	210
四、应用题	165	9.4.4 Linux 的存储管理	213
第 9 章 操作系统实例	168	9.4.5 Linux 的文件系统	215
9.1 UNIX 操作系统	168	9.4.6 Linux 的设备管理	218
9.1.1 UNIX 的系统结构	168	小结	219
9.1.2 UNIX 的进程管理	168	习题九	221
9.1.3 UNIX 的存储管理	173	一、选择题	221
9.1.4 UNIX 的文件系统	174	二、填空题	222
9.1.5 UNIX 的设备管理	179	三、思考题	223
9.2 MS-DOS 操作系统	181	四、应用题	223
9.2.1 MS-DOS 的系统结构	182	习题参考答案	224
9.2.2 MS-DOS 的进程管理	183	习题一	224
9.2.3 MS-DOS 的存储管理	183	习题二	225
9.2.4 MS-DOS 的文件管理	185	习题三	226
9.2.5 MS-DOS 的设备管理	188	习题四	230
9.2.6 MS-DOS 的作业管理	189	习题五	231
9.3 Windows 2000/XP 操作系统简介 ..	190	习题六	232
9.3.1 Windows 2000/XP 的 体系结构	190	习题七	233
9.3.2 Windows 2000/XP 的 处理器调度	192	习题八	235
9.3.3 Windows 2000/XP 的 存储管理	196	习题九	238
		参考文献	240

第1章 概 论

科学技术，特别是信息技术的飞速发展，使得我们今天正步入信息时代。作为未来信息时代的核心和支撑点的“计算机”(Computer)，我们不再陌生，事实是现在越来越多的人已经与计算机交上了“朋友”。人与计算机之间如何交往？计算机如何根据人们的意愿去工作？计算机又如何管理和处理浩如烟海的信息和各种资源？这些问题都由一个复杂而庞大的计算机程序——操作系统来管理和实施。

操作系统(Operating System, 简称 OS)是计算机系统中一个重要的系统软件，其他系统软件和应用软件都需要操作系统的支撑。

1.1 计算机系统

我们知道，一个完整的计算机系统是由硬件系统和软件系统两大部分组成的。硬件(hardware)是由元器件构成的，是计算机系统赖以作用的实体。软件(software)是各种程序和技术资料的总和，指挥计算机各部件进行各种操作，以确保计算机系统协调、自动运行。可以说硬件是计算机系统的躯体，而软件则是计算机的头脑和灵魂，只有将这两者紧密地结合在一起，才能成为有生命、有活力的计算机系统。

1.1.1 计算机硬件系统

计算机硬件主要由以下 5 个部分组成：

- (1) 运算器。
- (2) 存贮器。
- (3) 控制器。
- (4) 输入设备。
- (5) 输出设备。

1. 运算器

直接完成各种算术运算和逻辑运算。在控制器的指挥下对数据进行加工处理，所需的运算数据从存贮器中取出，运算结果一般也送回存贮器单元中。

2. 存贮器

存放数据和程序。在计算机内不论是数据或指令都是一些用二进制表示的代码。存贮器的基本功能是能够把二进制代码，按需要存进去(称为“写入”)或者取出来(称为“读出”)，这种功能可以比喻为人的“记忆”。所以，存贮器也叫做记忆装置。

注意：向存贮器写入新的内容则覆盖了原来旧的内容，从存贮器中读出数据并不破坏原来存贮的内容，因此数据可以重复取出，多次利用。对存贮器进行读或写操作，叫做访问存贮器。

存贮器又可细分为内存贮器(也称内存或主存)和外存贮器(简称外存或辅存)两类。

内存根据其功能可分为只读存贮器 ROM(Read Only Memory)和随机存贮器 RAM(Random Access Memory)两种。

常用的外存有硬盘 (Hard Disk)、软盘 (Floppy Disk)、光盘 (Compact Disk) 等。

3. 控制器

计算机的指挥部, 其功能是根据存贮的程序, 指挥计算机各部件进行各种操作, 以确保计算机系统协调、自动运行。

4. 输入设备

向计算机输入数据和程序。它是人与计算机进行交往的入口。

常用的输入设备有键盘、鼠标器、光笔、扫描仪等。

5. 输出设备

输出计算机处理结果。它能把机内的信息以人们熟悉的形式 (如: 数字、字符、图形等) 输出。它是计算机与人交往的出口。常用的输出设备有显示器、打印机、绘图仪等。

除了上述 5 个部分外, 计算机还有电源及输入/输出接口部件, 如 I/O 插槽等。

一般地, 运算器和控制器组成计算机核心部件, 称为中央处理器 (Central Processing Unit, CPU)。通常把 CPU 和内存贮器一起称为计算机的主机, 而输入设备、输出设备及外存贮器统称为计算机的外部设备。

1.1.2 计算机软件系统

计算机软件通常分为系统软件和应用软件两大类。系统软件又可分为操作系统、语言处理系统、数据库管理系统和软件工具等。应用软件包括: 应用软件包和用户程序。

1. 系统软件

系统软件是计算机的基本软件, 一般由计算机厂家或专门的软件公司提供, 它们有的写入 ROM 芯片随机提供, 有的存入软盘或光盘供用户选购。为方便使用和管理计算机资源, 充分发挥计算机的功能的软件, 统称系统软件。

系统软件有两个特点: 一是通用性: 无论是哪个应用领域的计算机用户都要用到它们; 二是基础性: 应用软件要在系统软件支持下编写和运行。

系统软件又可分为 4 类: 操作系统、语言处理系统、数据库管理系统和软件工具。

2. 应用软件

为解决计算机各类应用问题而编写的程序称为应用软件。应用软件随着计算机应用领域的不断扩展和深入而与日俱增。应用软件又可分为应用软件包和用户程序。

1.1.3 计算机系统层次结构

计算机系统的最内层是硬件, 用户与计算机硬件之间的接口是计算机软件。如图 1-1 所示的计算机系统层次结构。

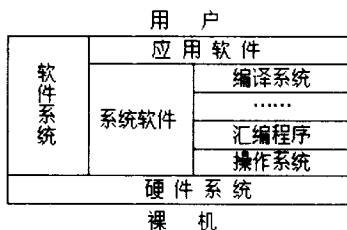


图 1-1 计算机系统层次结构

由计算机系统的层次结构,不难看出操作系统在整个计算机系统中的作用。

如果用户直接使用裸机,那么必须熟记机器指令系统,并且要了解各种外围设备的物理特性等等,这是非常麻烦的工作。

但在裸机上配置有系统软件之后,用户可以用高级语言编写程序,可以用各种命令提出控制程序执行的要求,而不必关心硬件的具体结构和内部特性。

分析计算机系统层次结构图,可以看出,处于最底层的硬件系统,只是“裸机”的形象,即没有附加任何软件的物理计算机;操作系统是加在裸机上的第一层软件,是对裸机的首次扩充,从而构成了一个比裸机功能更强、使用更方便的“虚拟计算机”;所有其他的系统软件以及更上层的应用软件都在操作系统虚拟机上运行,受操作系统的统一管理和控制,用户通过操作系统使用各种系统资源来完成特定的任务。

可见,操作系统是计算机系统中一个必不可少的关键组成部分,任何一个计算机系统可以不配置某些应用软件,但决不能没有操作系统——计算机系统的核心。

作为计算机系统的核心,操作系统是为提高计算机利用率,方便用户使用计算机而配备的一种最基本的系统软件。它协调控制和统一管理着整个计算机系统的软件资源和硬件资源,合理组织和统一安排计算机的工作流程,提供给用户方便使用计算机的接口,是用户与计算机的桥梁,它扮演着“全能管家”的角色,它在计算机系统中的地位 and 作用是举足轻重的。

下面我们从资源管理和服务来认识操作系统的作用。

1.2 操作系统的资源管理

从资源管理的观点来看,操作系统是管理计算机资源的软件。

计算机系统的资源包括两大类:硬件资源和软件资源,如图 1-2 所示。

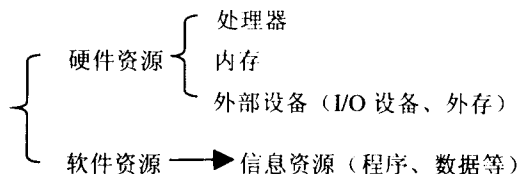


图 1-2 计算机系统资源

计算机系统资源细分成处理器、内存、外部设备和信息资源等 4 类资源,它们构成了操作系统本身和用户作业赖以活动的物质基础和工作环境。

一个配有操作系统的计算机,当用户在使用时,无需过问各类资源的分配使用状况,也没有必要编写各种 I/O 设备驱动程序。用户只需正确使用操作系统提供的各种命令和系统调用功能,指定的软件就会在操作系统的调度控制下自动而协调地运行。所以说一个优秀的操作系统能给计算机带来广阔的应用前景。

1.2.1 操作系统的功能

从资源管理的观点出发,与计算机系统的 4 类资源相对应,可将操作系统划分成处理器管理、存储管理、文件管理和设备管理 4 个基本部分;此外,操作系统要对程序任务使用资源的整个过程进行管理,即作业管理,它们构成了操作系统的 5 大功能。

操作系统这 5 大部分相互配合,协调工作,实现对计算机系统的资源管理,控制程序的

执行。

1. 处理器管理

处理器管理实质就是 CPU 调度，是指如何分配和控制处理器。

我们知道，所有程序都要在 CPU 上执行，宏观上，允许多个程序同时在主机上运行；微观上，CPU 在某一个时刻只能执行一个程序。当运行的程序数目大于 CPU 的个数时，这就存在 CPU 如何分配和控制才合理和公平的问题。

对请求 CPU 服务的多个程序任务如何排队？在某一时刻选择哪个程序使用 CPU？运行的程序任务占用多少时间？什么时候回收 CPU 的资源？这都是处理器管理的工作。

处理器管理的基本任务是：采用合理的调度策略进行 CPU 的分配与回收，使 CPU 充分发挥效率，并能满足各种程序的任务需求。

一个计算机系统可以配备一个或多个处理器，本书主要讨论的是单处理器管理。

2. 存储管理

存储管理主要指内存的管理。

我们知道，存储器分内存和外存两大类。内存与 CPU 直接打交道，程序必须进驻内存，方可被 CPU 执行；外存中的信息仅当需要时才装入内存，CPU 不能直接访问外存。

存储管理主要负责内存的分配与回收，进行内存中程序和数据的存储保护，以及进行内存扩充等，存储管理的目的是提高内存的利用率。

3. 文件管理

文件管理即软件资源的管理，负责文件的存取、共享与保护等。

它的主要任务是：负责文件物理存储空间的组织、分配及回收；实现按名存取；提供文件的保护与保密等。

4. 设备管理

设备管理负责各种外围设备的分配与操纵。

我们知道，外设品种多，各类外设之间以及主机与外设之间存在速度不匹配问题，各外设还可供多个运行程序共享等，因而设备管理是最操作系统中最复杂的部分。

设备管理的基本任务：提供方便的用户使用接口；进行设备的分配与回收；对磁盘设备优化调度以及实现 SPOOLing 技术等。

5. 作业管理

作业是系统的最大服务单位，是指用户要求计算机完成的一个独立的程序任务。

作业管理是指组织、控制与调度作业，一方面提供一个“作业控制语言”（Job Control Language，简称 JCL）供用户写作业说明书，另一方面还为操作员和终端用户提供与系统对话的“命令语言”，此外，作业管理的功能就是作业调度，根据不同的系统要求，将指定相应的调度策略。

作业管理的目的是实现作业的调度、控制作业的执行，使得各作业都能有效运行。

1.2.2 操作系统的层次结构

从用户角度来看，操作系统是一台虚拟机，它是在没有软件的裸机上运行。我们可以从模块分层的角度来认识操作系统。

通常把与硬件直接相关的处理器管理放在最内层，把与用户关系密切的作业管理放在最

外层,把存储管理、设备管理和文件管理放在中间层,如图 1-3 所示。

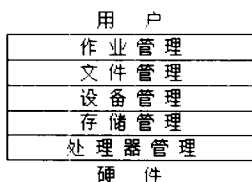


图 1-3 操作系统的层次结构

目前大型软件大都采用层次结构,即将软件分为若干层次来设计。从硬件开始,先实现最内层的功能,得到一个比硬件机器功能要强的第一级虚拟机;然后以第一级虚拟机为基础,实现中间层的功能,又得到一个功能更强的第二级虚拟机,如此逐层扩充,最后实现最外层的功能,从而得到一个功能最强的虚拟机,这就是用户眼中的操作系统了。这种分层方法叫“自底向上”设计方法,也可以反过来,采用“自顶向下”设计方法,在实际设计当中,我们往往把两者结合起来运用。

1.3 操作系统的服务与分类

操作系统是一种系统程序,除了要控制工作流程,要管理系统资源外,还要为用户服务。也就是说,操作系统的设计目标是:

- (1) 使系统资源得到高效利用。
- (2) 为用户创造良好的工作环境。

操作系统的服务观点是指:操作系统是系统服务软件,是用户与计算机系统之间的接口,用户通过操作系统使用计算机。操作系统为用户提供各类服务和支持,使得所有程序都能在系统内有效运行,从而使用户能方便、灵活、安全地完成各种应用任务。

操作系统的服务功能主要有:

- (1) 装入与执行程序。

任何用户程序都必须被装入内存,然后在 CPU 上执行;执行程序正常时,程序完成;异常时,终止程序执行。

- (2) I/O 操作。

通常,系统不允许用户去直接控制 I/O 设备,而是由操作系统统一实施管理,所以操作系统要提供 I/O 服务,使用户能通过统一简便的接口,方便使用各类外围设备,而无需了解用法不同的各类外设的硬件特性。

- (3) 文件使用。

各种程序和数据通常都是以文件的形式存储的,操作系统中的文件系统功能对用户显得特别重要。例如用户读写文件,按名存取文件,创建、删除、修改文件等,以及对文件进行保护与保密控制。

- (4) 控制作业的运行。

操作系统能够提供有效的通讯手段,使用户能够按照自己的控制意图来操纵作业的运行过程。

- (5) 检测与处理错误。

操作系统经常需要知道系统中所出现的差错,必要时应及时通知操作员或用户。无论是

硬件故障（CPU、存储器、I/O 设备、电源等发生的故障），还是软件故障（地址越界、非法存取、算术运算溢出、占用 CPU 时间过长等），对每一类差错，操作系统都应采取适当的措施，及时检测并修复，以保证用户不丢失数据和用户计算的一致性。

尽管不同的操作系统所提供的服务并不完全相同，但上述操作系统的服务功能却是最基本的。

1.3.1 操作系统的分类

早期的计算机是没有操作系统的，程序的装入与执行以及对外设的操作都是通过控制台进行的。直到计算机发展到第二代，才出现初级的操作系统，即监督程序。计算机发展到第三代，硬件和软件都有了很大发展，监督程序才发展成一个重要的软件分支——操作系统。

依操作系统的服务观点，大致将操作系统分为：

- （1）批处理操作系统（Batch Processing Operating System）。
- （2）分时操作系统（Time-sharing Operating System）。
- （3）实时操作系统（Real-time Operating System）。
- （4）单用户操作系统（Single User Interactive Operating System）。
- （5）网络操作系统（Network Operating System）。
- （6）分布式操作系统（Distributed Operating System）。

1. 批处理操作系统

所谓批处理是指系统对源源不断的作业流的连续处理。

用户把要计算的问题、数据和作业说明书一起交给操作员，操作员将一批算题输入到计算机，然后由操作员来控制执行。采用这种批量化处理作业技术的操作系统就称为批处理操作系统，简称批处理系统。常常分单道批处理系统和多道批处理系统两类。

1) 单道批处理系统

单道批处理系统是在单用户操作系统基础上发展的，目的是减少人工操作的时间，提高系统效率。

在单道批处理系统中，所有用户的作业均交给操作员，通过输入设备送入计算机，由操作员在控制台启动并控制计算机，监视作业运行，得到运行结果，交付用户。作业在整个运行过程中，不与用户发生任何关系，用户不能与作业交互作用，一旦作业出错则停止该作业，同时启动下一作业执行。而且在输入/输出操作过程中，CPU 处于等待状态。

单道批处理系统的特征是自动性（无需人工干预）；顺序性（各作业完成的顺序与它们进入内存的顺序完全相同）；单道性（内存中仅有一道程序在运行）。

20 世纪 50 年代 General Motors 研究室在 IBM 701 计算机上实现的第一个操作系统就是单道批处理系统。

2) 多道批处理系统

计算机发展到 20 世纪 60 年代，CPU 与外设，外设与外设之间均能并行工作，从而促进了操作系统的发展。多道批处理系统基于多道程序技术的应用，即在内存中可以同时存放多个用户作业，允许这些作业在系统中交替运行。

在多道批处理系统中，用户可以通过系统提供的各种功能，如作业控制语言、控制命令等，将用户程序与数据等一一提交给系统，在外存上形成作业的后备等待队列。

系统根据一定的调度原则,从这些后备等待队列中。选取若干作业调入内存。在内存中的这些作业按照多道方式组织它们的运行,某一道作业运行完毕或暂停运行,系统又将调入新的作业,内存中始终存放多个作业,它们交替地运行。如图 1-4 所示。

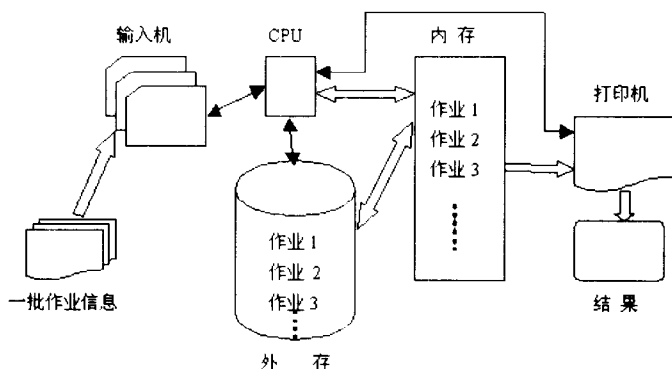


图 1-4 多道批处理系统

这样,作业不断进入系统,又不断退出系统,形成源源不断的作业流,从而大大提高了系统的资源利用率和吞吐量。

多道批处理系统的特征是多道性(内存中可以同时驻留多道程序);无序性(多个作业完成的先后顺序与它们进入内存的顺序之间,并无严格的对应关系);调度性(作业从提交给系统开始直到完成,需要经过两次调度:作业调度和进程调度,即从外存的后备作业队列中选取若干作业调入内存和从已在内存中的作业中选择一个作业,将 CPU 分配给它使之执行)。

为 IBM 360 机设计的 IBM DOS 是一个典型的多道批处理系统,后来扩展到 IBM 370 系统,以后又发展成 DOS/VS 和 DOS/VSE,运行于 IBM 4300 系列机上。

早在 20 世纪 60 年代就出现的多道批处理系统至今仍是三大基本操作系统类型之一,它具有资源利用率高和系统吞吐量大等优点;但也存在平均周转时间长、无交互能力等缺点。

2. 分时操作系统

如果说推动多道批处理系统形成和发展的主要动力是提高资源的利用率和系统的吞吐量,那么推动分时操作系统形成和发展的主要动力,则是用户的需要,即人机交互的需要,共享主机的需要等。

分时操作系统正是为了满足用户的这些需要而形成的。它是一种新型的操作系统,与多道批处理系统相比,在目标上(多道批处理系统是为了提高系统效率;分时操作系统是为了对用户请求做出快速响应)和提交给系统的作业性质上(对于短小作业,分时操作系统最有效;对于需要较长时间才能完成的作业则使用多道批处理系统较为有效)有着明显的差别。

分时操作系统简称分时系统,是指在一台主机(Mainframe)上连接了多个带有显示器和键盘的终端,同时允许多个用户共享主机中的资源,它使计算机为一组终端用户服务,虽然物理上只有一台计算机,但每个用户都可以得到及时的服务响应,都感到好像有一台支持自己请求服务的计算机在专门为他服务。在分时系统下,各个用户从终端直接输入各种命令请求系统服务,而系统采用时间片(Time Slice)的方式轮流地为各个终端上的用户服务,及时地对用户的请求予以响应,最后系统将处理的结果也通过终端报告给用户。

分时系统的特征是同时性(若干用户通过各自的终端同时使用一台计算机。从宏观上看,

所有用户在同时并行工作；从微观上看，各个用户轮流使用计算机）、独立性（虽然多个用户通过终端同时使用一台计算机，但用户之间相互独立，互不干扰，由分时系统保证各个用户程序执行的完整性）、及时性（系统保证对每个用户的请求做出及时的响应，使用户感到是他自己在使用和控制计算机）和交互性（用户能进行人机对话，交互操作和联机调试程序等）。

分时系统也称为交互式系统。

在 20 世纪 60 年代初期，美国麻省理工学院研制出了第一个分时系统 CTSS (Compatible Time-Sharing System)；60 年代末，麻省理工学院与 IBM 公司及 Bell 实验室合作开发的 MULTICS (Multiplexed Information & Computing Service) 是功能更强、更具代表性的分时系统；现今优秀的分时系统首推 UNIX。

3. 实时操作系统

早期的计算机基本上用于科学计算。20 世纪 50 年代后期，计算机开始用于生产过程的控制，形成了实时操作系统。实时操作系统简称实时系统，主要用于需要对外部事件进行及时响应并处理的领域，包括：实时控制系统（如生产过程的自动控制）和实时信息处理系统（如订票系统、情报检索系统、信息查询系统等）。对外部输入的信息，实时系统能够在规定的时间内处理完毕并做出反应。正因为实时系统对响应时间的要求很高，所以设计实时系统时，首先要考虑响应及时，其次才考虑资源的利用率。

实时系统的特征是：多路性（对多路的现场信息进行采集及对多个对象或多个执行机构进行控制）、及时性（对用户的响应时间要求秒级或毫秒级或微秒级）、交互性（实时系统允许用户使用终端，提供简单的联机服务，具有简单的交互能力，如限于查询系统中某些特定的专用服务程序而不能在系统运行中在终端输入或编写或调试程序等）和可靠性（实时系统要高度可靠，否则任何疏忽将导致灾难性后果，常常设置容错机制以保证系统的安全可靠或信息的完整性）。

4. 单用户操作系统

单用户操作系统是指在计算机系统内一次只能支持一个用户程序的运行，多数微机操作系统都属于单用户操作系统，如 Windows、MS-DOS、CP/M 等。

单用户操作系统是随着微型机的发展应运而生的，它不追求系统资源的利用率，不讲究资源的共享，也不要求系统支持实时响应等，而是强调个人的特点，注重使用的方便。

单用户操作系统功能比较简单，主要是文件管理和设备驱动，采用联机交互方式，除了提供键盘命令外，一些系统还提供菜单命令、窗口操作和鼠标操作等更为方便灵活的交互手段。

5. 网络操作系统

所谓计算机网络是指将地理上分散的具有独立功能的计算机系统和外围设备，通过各种媒体和通信手段，按不同的拓扑结构连接起来的集合体。计算机技术和通信技术的结合，产生了计算机网络，使资源共享成为可能。

提供网络通信和资源共享功能的操作系统就称为网络操作系统。网络操作系统把计算机网络中的各台计算机有机地联合起来，使用户可以突破地理条件的限制，方便地相互通信以及共享各种资源等。网络操作系统最突出的就是网络管理和网络服务功能。一个最著名的计算机网络就是 Internet 的前身即美国的 ARPANET。流行的网络操作系统有 Windows 2000、UNIX 和 Linux 等。