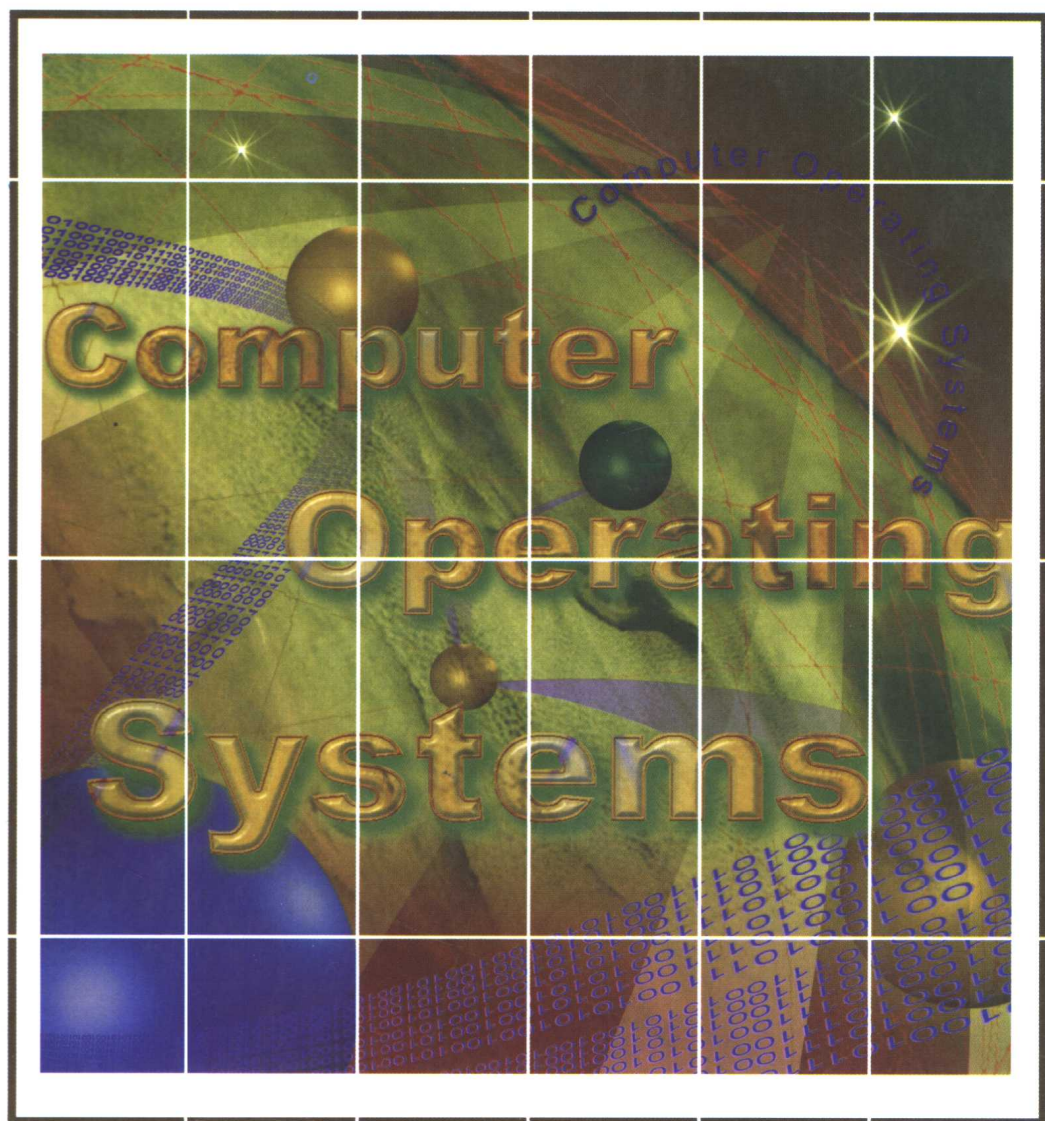


国家优秀教材汤子瀛主编《计算机操作系统》配套辅导书



# 《计算机操作系统》 学习指导与题解

汤子瀛 主审  
梁红兵 汤小丹 编著

西安电子科技大学出版社  
<http://www.xduph.com>

国家优秀教材汤子瀛主编《计算机操作系统》配套辅导书

# 《计算机操作系统》 学习指导与题解

汤子瀛 主审

梁红兵 汤小丹 编著

西安电子科技大学出版社

2003

## 内 容 简 介

本书全面介绍了计算机操作系统的基本原理。全书共分为 8 章，每章先扼要地阐述基本内容，再给出重点、难点的学习提示，典型问题的分析和解答，最后给出了大量的习题。附录部分中：附录 A 是操作系统实验指导，附录 B 给出了各章习题中选择题的答案，附录 C 提供了三套模拟试题，附录 D 给出了附录 C 中模拟试题的参考答案。

本书可作为计算机及相关专业的本科生和专科生学习操作系统的辅导教材，也可作为报考相关专业硕士研究生的复习用书，还可作为读者自学操作系统的参考书。

国家优秀教材汤子瀛主编《计算机操作系统》配套辅导书

《计算机操作系统》学习指导与题解

汤子瀛 主审

梁红兵 汤小丹 编著

策 划 李惠萍

责任编辑 龙 晖

出版发行 西安电子科技大学出版社（西安市太白南路 2 号）

电 话 (029)8227828 邮 编 710071

<http://www.xduph.com> E-mail: [xdupfbx@pub.xaonline.com](mailto:xdupfbx@pub.xaonline.com)

经 销 新华书店

印 刷 陕西画报社印刷厂

版 次 2003 年 3 月第 1 版 2003 年 3 月第 1 次印刷

开 本 787 毫米×1092 毫米 1/16 印张 14

字 数 328 千字

印 数 1~6 000 册

定 价 16.00 元

ISBN 7-5606-0496-X / TP·0232

XDUP 0766B01-1

\*\*\* 如有印装问题可调换 \*\*\*

# 前 言

本书是为汤子瀛主编的《计算机操作系统(修订版)》而编写的配套辅导教材,同时也可作为其他版本的操作系统教材辅导书、学习指导书。

全书共分为8章。前七章分别与《计算机操作系统(修订版)》的前七章对应,其中第一章简单地介绍了操作系统的作用、发展过程、特征和功能,第二章介绍了进程和线程的基本概念、进程控制、进程同步和进程通信,第三章介绍了处理机调度与死锁,第四、五、六、七章分别介绍了存储器管理、设备管理、文件管理和操作系统接口,第八章与《计算机操作系统(修订版)》的第九章对应,讲述了计算机系统的安全性。

每章均对操作系统的基本概念、基本原理和基本方法做了系统、扼要的阐述。为了帮助读者理解和掌握操作系统的基本内容,每章专门安排一节列出了其中的重点、难点内容,并给出了相应的学习提示;还在“典型问题分析和解答”一节中,给出了相当的典型问题,并对它们做了较详细和透彻的分析和解答。另外,各章还为读者提供了大量的习题,并在附录B中给出了其中选择题的答案。附录C提供了三套模拟试题供读者复习,附录D给出了附录C中模拟试题的参考答案。

在操作系统的教学中,实践环节同样是不容忽视的,为此,我们还在附录A中设计了7个有关操作系统的实验,它们可在Linux或UNIX环境下进行。

无论读者学习操作系统用的是哪本教材,相信都能从本书中得到一定的帮助。另外,本书亦可作为自学考试相关课程的辅导书。

本书的主审是汤子瀛教授。汤子瀛、哲凤屏教授认真细致、逐字逐句地审阅了全部书稿,并提出了许多宝贵的建议。本书的编写还得到了西安电子科技大学出版社,尤其是李惠萍同志的帮助和大力支持。在此谨向他们表示衷心的感谢。

限于编者水平,书中难免存在一些错误和不妥之处,恳请读者批评指正。

编 者

2003年1月

# 目 录

## 第一章 操作系统引论

1.1 基本内容 .....	1	1.3.1 OS的引入和发展过程中的典型问题 分析 .....	9
1.1.1 操作系统的目标和作用 .....	1	1.3.2 OS的基本特征和功能中的典型问题 分析 .....	10
1.1.2 操作系统的发展过程 .....	2	1.3.3 分层式结构和微内核结构中的典型 问题分析 .....	11
1.1.3 操作系统的基本特征和功能 .....	4		
1.1.4 操作系统的结构设计 .....	7		
1.2 重点难点学习提示 .....	7		
1.3 典型问题分析和解答 .....	9	1.4 习题 .....	11

## 第二章 进程管理

2.1 基本内容 .....	15	2.3.2 进程同步基本概念中的典型问题 分析 .....	35
2.1.1 进程的基本概念 .....	15	2.3.3 信号量机制及应用中的典型问题 分析 .....	38
2.1.2 进程控制 .....	18	2.3.4 经典进程同步问题中的典型问题 分析 .....	40
2.1.3 进程同步 .....	19	2.3.5 消息传递通信机制中的典型问题 分析 .....	46
2.1.4 经典进程的同步问题 .....	21	2.3.6 线程中的典型问题分析 .....	47
2.1.5 管程机制 .....	24	2.4 习题 .....	48
2.1.6 进程通信 .....	26		
2.1.7 线程 .....	28		
2.2 重点难点学习提示 .....	31		
2.3 典型问题分析和解答 .....	34		
2.3.1 进程基本概念中的典型问题分析 .....	34		

## 第三章 处理机调度与死锁

3.1 基本内容 .....	54	3.3 典型问题分析和解答 .....	63
3.1.1 处理机调度的基本概念 .....	54	3.3.1 进程调度及调度算法中的典型问题 分析 .....	63
3.1.2 调度算法 .....	55	3.3.2 实时调度中的典型问题分析 .....	66
3.1.3 实时调度 .....	56	3.3.3 多处理机调度中的典型问题分析 .....	67
3.1.4 多处理机系统中的调度 .....	57	3.3.4 死锁中的典型问题分析 .....	68
3.1.5 死锁的基本概念 .....	58	3.4 习题 .....	70
3.1.6 处理死锁的基本方法 .....	58		
3.2 重点难点学习提示 .....	61		

## 第四章 存储器管理

4.1 基本内容.....	73	4.3 典型问题分析和解答.....	90
4.1.1 存储器管理的基本概念.....	73	4.3.1 存储器基本概念中的典型问题 分析.....	90
4.1.2 连续分配方式.....	75	4.3.2 连续分配方式中的典型问题分析.....	91
4.1.3 基本分页存储管理方式.....	78	4.3.3 基本分页系统中的典型问题分析.....	93
4.1.4 分段式存储管理方式.....	80	4.3.4 基本分段系统中的典型问题分析.....	94
4.1.5 信息的共享与保护.....	83	4.3.5 虚拟存储器基本概念中的典型问题 分析.....	95
4.1.6 虚拟存储器的基本概念.....	84	4.3.6 请求分页/段系统中的典型问题 分析.....	95
4.1.7 请求分页存储管理方式.....	85	4.4 习题.....	101
4.1.8 置换算法.....	86		
4.1.9 请求分段存储管理方式.....	87		
4.2 重点难点学习提示.....	88		

## 第五章 设备管理

5.1 基本内容.....	107	5.3 典型问题分析和解答.....	120
5.1.1 I/O 系统的硬件组成.....	107	5.3.1 I/O控制方式中的典型问题分析.....	120
5.1.2 I/O 控制方式.....	109	5.3.2 缓冲管理中的典型问题分析.....	121
5.1.3 缓冲管理.....	110	5.3.3 设备分配中的典型问题分析.....	122
5.1.4 设备分配.....	111	5.3.4 虚拟设备中的典型问题分析.....	124
5.1.5 SPOOLing技术.....	113	5.3.5 设备处理中的典型问题分析.....	125
5.1.6 设备处理.....	114	5.3.6 磁盘存储器管理中的典型问题 分析.....	125
5.1.7 磁盘存储管理.....	115	5.4 习题.....	127
5.2 重点难点学习提示.....	118		

## 第六章 文件管理

6.1 基本内容.....	132	6.3 典型问题分析和解答.....	147
6.1.1 文件和文件系统.....	132	6.3.1 外存分配方式中的典型问题 分析.....	147
6.1.2 文件的逻辑结构.....	133	6.3.2 文件存储空间管理中的典型问题 分析.....	149
6.1.3 外存分配方式与文件的 物理结构.....	134	6.3.3 目录管理中的典型问题分析.....	151
6.1.4 文件存储空间的管理.....	137	6.3.4 文件共享和安全性中的典型问题 分析.....	152
6.1.5 目录管理.....	139	6.3.5 文件操作中的典型问题分析.....	153
6.1.6 文件共享.....	141	6.4 习题.....	154
6.1.7 文件系统的可靠性.....	142		
6.1.8 数据一致性控制.....	143		
6.2 重点难点学习提示.....	144		

## 第七章 操作系统接口

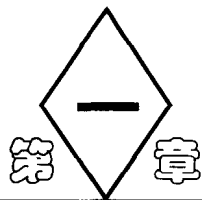
7.1 基本内容 .....	159	7.2 重点难点学习提示 .....	165
7.1.1 联机命令接口 .....	159	7.3 典型问题分析和解答 .....	166
7.1.2 脱机命令接口 .....	161	7.3.1 系统调用中的典型问题分析 .....	166
7.1.3 系统调用 .....	161	7.3.2 其他典型问题分析 .....	168
7.1.4 图形用户接口 .....	163	7.4 习题 .....	168

## 第八章 系统安全性

8.1 基本内容 .....	171	8.3 典型问题分析和解答 .....	182
8.1.1 安全威胁 .....	171	8.3.1 数据加密技术中的典型问题分析 .....	182
8.1.2 数据加密技术 .....	172	8.3.2 认证技术中的典型问题分析 .....	183
8.1.3 认证技术 .....	174	8.3.3 访问控制技术中的典型问题分析 .....	183
8.1.4 访问控制技术 .....	176	8.3.4 防火墙技术中的典型问题分析 .....	183
8.1.5 防火墙技术 .....	178	8.4 习题 .....	184
8.2 重点难点学习提示 .....	180		

## 附 录

附录A 操作系统实验指导 .....	187	附录D 模拟试题答案 .....	207
附录B 各章选择题答案 .....	193	参考文献 .....	216
附录C 模拟试题 .....	199		



# 操作系统引论

本章主要讲述操作系统的基本概念，具体包括操作系统的目标和作用、操作系统的发展过程、操作系统的基本特征和功能、操作系统的结构设计等内容。

## 1.1 基本内容

### 1.1.1 操作系统的目标和作用

操作系统(Operating System, OS)是一组控制和管理计算机硬件和软件资源、合理地调度各类作业，以及方便用户使用计算机的程序的集合。它是配置在计算机上的第一层软件，是对硬件功能的首次扩充。操作系统在计算机系统中占据着特别重要的地位，它是计算机中最重要的系统软件，是其他系统软件和应用软件运行的基础。

#### 1. 操作系统的目标

- (1) 方便性。操作系统使计算机系统更易于使用。
- (2) 有效性。操作系统使计算机资源的使用更有效，即使资源的利用率更高。
- (3) 可扩充性。操作系统必须能方便地开发、测试和引进新的系统功能，以适应计算机硬件和体系结构的迅速发展以及应用不断扩大的要求。
- (4) 开放性。操作系统必须能提供统一开放的环境，以使其应用在不同的系统中具有可移植性，并使不同的系统能够通过网络进行集成，从而能正确、有效地协同工作。

#### 2. 操作系统的作用

(1) 从一般用户的观点来看，操作系统是用户与计算机硬件系统之间的接口。用户并不直接与计算机硬件打交道，而是通过操作系统提供的命令、系统功能调用以及图形化接口来使用计算机。

(2) 从资源管理的观点来看，操作系统是计算机资源的管理者。处理机的分配和控制，内存的分配和回收，I/O设备的分配和处理，文件的存取、共享和保护工作都是由操作系统完成的。

(3) 从虚拟机的观点来看，操作系统是扩充裸机功能的软件。在裸机(即没配置任何软件的计算机系统)上覆盖了操作系统后，裸机将变成一台功能更强大、使用更方便的虚拟机。

(4) 从任务组织的观点来看，操作系统是计算机工作流程的组织者。它负责在众多作业间切换处理机，并协调它们的推进速度，从而进一步提高系统的性能。



## 1.1.2 操作系统的发展过程

### 1. 无操作系统的计算机系统

#### 1) 人工操作方式

在计算机发展的早期，由于还未出现操作系统，因而人们采用如下的人工操作方式使用计算机：由用户(即程序员)将已穿孔(对应于程序和数据)的纸带(或卡片)装入纸带(或卡片)输入机，再启动程序，将数据输入计算机，然后启动计算机运行程序；当程序运行完毕并由用户取走纸带和计算结果后，才允许下一个用户使用计算机。

人工操作方式具有用户独占全机和 CPU 等待人工操作的特点。由于人工操作的低速性和 CPU 运算的高速性，因而使计算机资源的利用率大大降低，此即人机矛盾。随着计算机的迅速发展及机器性能的不断提高，人机矛盾日益加剧。另外，高速的 CPU 和低速的 I/O 设备之间速度不匹配的矛盾也日趋严重。为了缓和这两对矛盾，便引入了脱机输入/输出方式和批处理技术。

#### 2) 脱机输入/输出方式

脱机输入方式是指在一台外围机(也叫卫星机，它是一台专门用来管理输入/输出的、功能较简单的计算机)的控制下，预先将程序和数据从低速输入设备输入到磁带，当 CPU 需要这些程序和数据时，再从磁带高速地读入内存。类似地，脱机输出方式是指当 CPU 需要输出时，先高速地将数据写入磁带，然后在—台外围机的控制下，通过低速输出设备进行输出。相反，在主机的直接控制下进行的输入/输出方式被称为联机输入/输出方式。

在脱机输入/输出方式下，是由外围机而不是主机的 CPU 等待人工操作，从而有效地减少了主机 CPU 的空闲时间，缓和了人机矛盾；另外，CPU 直接通过高速的磁带进行输入/输出，这又极大地提高了 I/O 的速度，进一步减少了 CPU 的空闲时间，从而较好地缓和了 CPU 与 I/O 设备之间速度不匹配的矛盾。

### 2. 单道批处理系统

批处理技术是指在系统中配置一个监督程序，并在该监督程序的控制下，能够对一批作业自动进行处理的一种技术。

早期采用批处理技术的系统，由于在内存中只能存放一道作业，因而称为单道批处理系统。而其中的监督程序就是操作系统的雏形。

单道批处理系统的处理过程如下：它将一批作业以脱机方式输入到磁带上，然后由配置在系统中的监督程序将磁带上的第一个作业装入内存，并把运行的控制权交给作业；当该作业处理完成或出现异常时，又把控制权交还给监督程序，再由监督程序调入磁带上的第二个作业……直至磁带上所有的作业全部完成。通过脱机输入和作业的自动过渡，单道批处理系统提高了机器资源的利用率。

单道批处理系统具有以下特征：

- (1) 自动性。磁带上的一批作业能自动地、逐个地依次运行，无须人工干预。
- (2) 顺序性。作业完成的顺序与它们进入内存的顺序以及作业在磁带上的顺序一致。
- (3) 单道性。内存中仅能存放一道作业。

### 3. 多道批处理系统

#### 1) 多道程序设计技术

在单道批处理系统中，内存中仅有一道作业，它无法充分利用系统中所有的资源，致使系统性能较差。为了进一步提高资源的利用率，又引入了多道程序设计技术。

多道程序设计技术是指在内存中同时存放若干个作业，使它们共享系统资源并同时运行的技术。在单处理机环境下，这些作业仅在宏观上同时运行，而在微观上交替执行。

由于在内存中可同时存放多个作业，因此当正在执行的作业因 I/O 等原因而暂停执行时，CPU 可立即调度给另一道作业，而系统中众多的 I/O 设备可与 CPU 并行地工作。多道程序设计技术可显著地提高内存、CPU 与 I/O 设备的利用率，增加系统的吞吐量(指系统在单位时间内完成的总工作量)。

#### 2) 多道批处理系统

采用多道程序设计技术的批处理系统被称作多道批处理系统。为了使系统中的多道程序能协调地运行，多道批处理系统中必须配置一组软件，来解决多道程序对系统资源的共享和争用问题，并对作业进行合理的组织和调度。这就形成了现代意义上的操作系统。

多道批处理系统具有以下特征：

(1) 多道性。内存中可同时存放多个作业。

(2) 调度性。需通过作业调度从外存中选取若干个作业装入内存，还需通过进程调度在内存的多个作业中分配 CPU。

(3) 无序性。通常，作业调度的次序与作业在外存中的次序无关，作业完成的次序与作业进入内存的次序也无关。

多道程序设计技术和批处理技术的采用，使多道批处理系统具有资源利用率高和系统吞吐量大的优点。但是，多道批处理系统将用户和计算机操作员分开，而且用户作业要排队，依次进行处理，故又具有用户无法直接与自己的作业进行交互和作业的平均周转时间(指作业从进入系统开始，直至作业完成并退出系统为止所经历的平均时间)较长的缺点。

### 4. 分时系统

为了解决批处理系统无法进行人机交互的问题，并使多个用户(包括远程用户)能同时使用昂贵的主机资源，又引入了分时系统。

分时系统是指在一台主机上连接有多个带显示器和键盘的终端，同时允许多个用户通过自己的终端以交互方式使用计算机，共享主机中的资源。

分时系统的关键问题是使用户能与自己的作业进行交互，或者说，它追求的主要目标是系统能及时响应用户的终端命令。为此，系统中采用了分时技术，即把处理机的时间划分成很短的时间片(如几百毫秒)，轮流地分配给各个终端作业使用。若在分配给它的时间片内，作业仍没执行完，它也必须将 CPU 交给下一个作业使用，并等下一轮得到 CPU 时再继续执行。这样，系统便能及时地响应每个用户的请求，从而使每个用户都能及时地与自己的作业交互。

分时系统具有以下特征：

(1) 多路性。一台主机上连有多个终端，能同时为多个用户服务。

- (2) 独立性。各个用户像独占主机一般，独立工作，互不干扰。
- (3) 及时性。系统按人们所能接受的等待时间(如 2~3 s)及时响应用户的请求。
- (4) 交互性。能进行广泛的人机交互。

## 5. 实时系统

### 1) 实时系统及其类型

实时系统是指系统能及时(或即时)响应外部事件的请求，在规定的时间内完成对该事件的处理，并控制所有实时任务协调一致地运行。它可分成两大类：

(1) 实时控制系统。通常是指以计算机为中心的生产过程控制系统和武器控制系统。这类系统要求实时采集现场数据，并对数据进行及时处理，进而自动地控制相应的执行机构，使某个(些)参数(如温度、压力、方位等)能按预定的规律变化，以保证产品质量和提高产量。

(2) 实时信息处理系统。通常是指对信息进行实时处理的系统。这类系统要求及时接收从终端(包括远程终端)发来的服务请求，按请求的内容对信息进行检索和处理，并在很短的时间内为用户做出正确的回答。典型的实时信息处理系统有证券交易系统、飞机订票系统、情报检索系统等。

### 2) 实时系统与分时系统特征的比较

(1) 多路性。实时信息处理系统与分时系统一样具有多路性，即系统能同时为多个终端用户服务；而实时控制系统也具有多路性，它主要表现在：系统经常对多路的现场信息进行采集，以及对多个对象或多个执行机构进行控制。

(2) 独立性。实时信息处理系统与分时系统一样具有独立性，每个终端用户可独立地向实时系统提出服务请求，彼此互不干扰；而实时控制系统对信息的采集和对象的控制也能独立进行、彼此互不干扰，因此也具有独立性。

(3) 及时性。实时信息系统对及时性的要求与分时系统类似，都是以人所能接受的等待时间来确定的；而实时控制系统的及时性通常高于分时系统，它是以控制对象所能接受的等待时间来确定的，一般为秒级、百毫秒级直至毫秒级，甚至有的要低于 100  $\mu$ s。

(4) 交互性。实时信息处理系统虽然也具有交互性，但其交互性通常不及分时系统，这里人与系统的交互，仅限于访问系统中某些特定的专用服务程序。它不像分时系统那样能向终端用户提供数据处理、资源共享等多方面的服务。

(5) 可靠性。分时系统也要求系统可靠，但相比之下，实时系统则要求系统高度可靠。因为任何差错都可能带来巨大的经济损失，甚至无法预料的灾难性后果。因此，在实时系统中，常采用多级容错措施来保障系统和数据的安全性。

批处理系统、分时系统和实时系统是三种基本的操作系统类型。一个实际的操作系统如果兼有三者或其中两者的功能，则被称为通用操作系统。

## 1.1.3 操作系统的基本特征和功能

### 1. 操作系统的基本特征

采用多道程序设计技术的现代操作系统都具有如下的基本特征：

### 1) 并发

并发性是指两个或多个事件在同一时间间隔内发生。在多道程序的环境下，并发性是指在一段时间内，宏观上有多个程序在同时运行。

应当指出的是，通常的程序是静态实体，它们是不能并发执行的。为使多个程序能并发地执行，系统必须分别为每个程序建立进程。进程是系统中能独立运行并能独立分配资源的基本单位，它由一组机器指令、数据、堆栈和进程控制块组成，是一个活动实体。多个进程之间可以并发执行和交换信息。现代 OS 中还引入了一个比进程更小的单位——线程，此时，一个进程中可包含若干个线程，资源分配的基本单位虽仍然是进程，但处理机分配和运行的基本单位却是线程。由于线程比进程更小，基本上不拥有资源，故它运行起来更为轻松，因此能更好地提高系统内多个程序间并发执行的程度。

与并发性相似的另一个概念是并行性，它是指两个或多个事件在同一时刻发生。可见，并行性具有并发的含义，但并发事件并不一定具有并行性。如单机系统中并发执行的程序，虽然宏观上是在同时运行，但微观上是在分时地交替执行，故不具有并行性的特征。

### 2) 共享

共享是指系统中的资源可供内存中多个并发执行的作业同时使用。根据资源性质的不同，可将资源共享方式分为以下两种：

(1) 互斥共享。系统中可供共享的某些资源，如打印机、变量、队列，一段时间内只能给一个作业使用，只有当这个作业使用完毕并释放这些资源后，其他作业才能使用它们。

(2) 同时访问。系统中的另一类资源，如磁盘、可重入代码，它们在同一段时间内可以被多个作业同时访问。虽然这种同时是指宏观上的同时，微观上可能是作业交替地访问该资源，但作业交替访问资源的顺序不会影响访问的结果。

### 3) 虚拟

虚拟是指通过某种技术，将一个物理实体变为若干个逻辑上的对应物。用来实现虚拟的技术，被称为虚拟技术。在现代 OS 中利用了多种虚拟技术，分别用来实现虚拟处理机、虚拟存储器和虚拟设备等。

### 4) 异步性

异步性是指在多道程序的环境下，每个程序不知何时执行、何时暂停，即它们以不可预知的速度向前推进。但同时操作系统应保证程序的执行结果是可再现的，即只要运行环境相同，一个作业的多次运行都会得到相同的结果。

## 2. 操作系统的功能

操作系统的主要功能包括处理机管理、存储器管理、设备管理、文件管理和提供友好的用户接口五个方面。

### 1) 处理机管理

处理机管理主要是对处理机的分配和运行进行管理。在传统的操作系统中，处理机的分配和运行都是以进程为基本单位，因此通常将处理机管理归结为对进程的管理；在引入线程的 OS 中，还必须包含对线程的管理。进程管理的主要功能包括：

(1) 进程控制。为作业创建进程、撤消进程，并控制进程在运行过程中的状态转换。在现代 OS 中，进程控制还应具有创建、撤消线程的功能。

(2) 进程同步。对进程(线程)的执行次序进行协调，使进程(线程)能有条不紊地运行。

(3) 进程通信。实现进程(线程)之间的信息交换，使进程(线程)能很好地相互合作。

(4) 进程调度。在多个就绪进程(线程)中分配处理机，并使分配到处理机的进程(线程)投入执行。

## 2) 存储器管理

存储器管理主要是为多道程序的运行提供良好的环境，它的主要功能包括：

(1) 内存分配。为每道程序分配内存空间，分配时要尽量提高内存的利用率。

(2) 内存保护。确保每道用户程序只在自己的内存空间中运行，从而不影响操作系统和其他程序的运行。

(3) 地址映射。将程序中的逻辑地址转换成内存中的物理地址，以使程序正确地执行。

(4) 内存扩充。在逻辑上扩充内存的容量，以方便大作业的运行和增加内存中并发作业的个数。

## 3) 设备管理

设备管理主要是完成用户的 I/O 请求，它的主要功能包括：

(1) 缓冲管理。利用缓冲来缓和 CPU 和 I/O 设备速度不匹配的矛盾，提高 CPU 和 I/O 设备的利用率和 I/O 的速度。

(2) 设备分配。为用户分配完成 I/O 所需的设备和设备控制器，在配置有通道的系统中，还需为用户分配通道。

(3) 设备处理。启动设备进行真正的 I/O 操作，响应并处理设备控制器发来的中断请求。

## 4) 文件管理

文件管理主要是使用户能方便、安全地使用各种信息资源，它的主要功能包括：

(1) 文件存储空间的管理。为文件分配必要的存储空间，并尽量提高文件存储空间的利用率和文件访问的效能。

(2) 目录管理。通过目录的方式来组织文件，以实现文件的按名存取，并提高文件的检索速度。

(3) 文件的读/写管理和保护。实现文件的读写操作，并提供有效的存取控制功能，保护文件的安全性。

## 5) 提供友好的用户接口

为方便用户使用计算机，操作系统还应提供友好的用户接口。该接口通常是以下列方式提供给用户的：

(1) 命令接口。提供一组联机命令和作业控制语言，供用户直接或间接地控制自己的作业。

(2) 程序接口。提供一组系统调用，供用户程序调用操作系统的功能。

(3) 图形用户接口。提供图标、窗口和菜单等元素，使用户可方便地通过指点设备(如鼠标)和少量的键盘操作，取得操作系统的服务。

## 1.1.4 操作系统的结构设计

操作系统是一个大型的系统软件，其内部的组织结构已经历了四代的变革。

### 1. 整体式系统

整体式系统中，整个操作系统是一堆过程的集合，每个过程都可以调用任意其他过程，操作系统内部不存在任何结构，因此也叫无结构操作系统。采用这种结构的操作系统不仅调试和维护不方便，而且其可读性和可扩充性都较差。

### 2. 模块化结构

模块化结构中采用了模块化程序设计技术，将操作系统按其功能划分成若干个具有一定独立性和大小的模块，并规定好各模块间的接口，使得它们之间能够交互。对较大的模块还可进一步细化为若干个子模块……采用这种结构可加速操作系统的研制过程。操作系统设计的正确性高、可适应性好，但模块的划分和接口的规定较困难，而且模块间还存在着复杂的依赖关系，使 OS 结构变得不够清晰。

### 3. 分层式结构

分层式结构是对模块化结构的一种改进，它将操作系统按其功能流图的调用次序以及其他一些原则划分为若干个层次，每一层代码只能使用较低层代码提供的功能和服务，并采用自底向上或自顶向下增添软件的方法来研制操作系统。由于它将模块之间的复杂依赖关系改为单向依赖关系，并消除了某些循环依赖关系，因此能使 OS 结构变得非常清晰，从而使系统的调试和验证更方便，正确性更高。

### 4. 微内核结构

微内核的主要思想是，在操作系统内核中只留下一些最基本的功能，而将其他服务尽可能地从中分离出去，用若干个运行在用户态下的进程(即服务器进程)来实现，形成所谓的“客户/服务器”模式。普通用户进程(即客户进程)可通过内核向服务器进程发送请求，以取得操作系统的服务。采用微内核结构，不仅提高了系统的灵活性和可扩充性，还增加了系统的可靠性。微内核结构的另一个优点是它比较适用于分布式系统。

在现代 OS 的设计中，常常还融入面向对象的程序设计技术。它利用被封装的数据结构和一组对数据结构进行的操作的过程来表示系统中的某个资源，这样，可使资源的管理因一致而简化。当前广泛使用的 Windows 2000 操作系统，就采用了微内核的结构，同时融入了面向对象的程序设计技术。

## 1.2 重点难点学习提示

学习本章的目的是使学生建立起 OS 的基本概念。为此应对以下几个重点、难点问题作认真的学习，切实掌握 OS 的一些基本概念。

### 1. OS 的引入和发展

由于 OS 随着计算机技术和应用需求的不断发展，由简单变为复杂，由低级变为高级，故在学习“OS 的引入和发展”时，应对下述几个问题有较清晰的认识：

(1) 早期无 OS 的计算机系统中,存在着所谓的“人机矛盾”和“CPU-I/O 设备速度不匹配的矛盾”。它们对计算机资源的利用率有何严重的影响?

(2) 单道批处理系统中引入了哪些技术,它们是如何解决上述两对矛盾的?

(3) 单道批处理系统还存在哪些不足之处,而多道批处理系统又是通过哪些技术措施来解决这些不足的?

(4) 多道批处理系统还有哪些地方不能满足用户的需求,或者说,是在什么样的需求推动力的作用下,由批处理系统发展为分时系统的?实现分时系统的关键技术是什么?

(5) 上述几种系统还有哪些地方不能满足用户的需求,或者说,是在什么样的需求推动力的作用下由分时系统发展为实时系统的?在学习时还应注意分析和比较分时系统与实时系统的特征。

## 2. OS 的基本特征和功能

在多道程序的环境下,OS 具有四大特征和五大功能。在学习该问题时,应对下述四个方面的内容有较深入的理解:

(1) OS 的特征。OS 具有并发、资源共享、虚拟和异步性四大特征,在学习时应应对每种特征的具体含义和形成原因有较清晰的认识。

(2) OS 四大特征之间的关系。这四大特征之中,最重要的是并发特征,其他三个特征都是以并发为前提的。在学习时必须弄清楚并发和资源共享之间的关系,并理解它们是如何导致虚拟和异步性特征的产生的。

(3) OS 的功能。OS 具有处理机管理、存储器管理、设备管理、文件管理和提供友好的用户接口等五大功能。在学习时应了解各个功能的主要任务,并必须深入了解处理机管理功能与 OS 的并发和共享特征之间的关系。

(4) OS 五大功能的必要性。为了保证多道程序能有条不紊地、高效地运行,并能方便用户对计算机系统的使用,OS 必须具备上述五大功能。在学习时请思考:如果缺少了其中的某个(些)功能(如处理机管理或内存管理功能),将会对系统的运行产生什么样的影响。

## 3. 分层式结构和微内核结构

在 OS 结构中,分层式结构是最为成熟的一种 OS 结构,被广泛应用了 20 多年。而 20 世纪 90 年代兴起的微内核结构是最具有发展前途的 OS 结构。在学习该问题时,应对下述内容有较深刻的了解:

(1) 分层式结构。该结构是指将 OS 按某些原则分为若干个层次,并规定了层次间的单向调用关系。在学习时应应对分层结构是为了解决什么样的问题,它又是如何解决该问题的等内容有较深入的了解。

(2) 分层的原则。在学习时应知道一个 OS 应设置哪些层次的主要依据,并清楚通常应把哪些功能放在最低层,哪些功能放在最高层。

(3) 客户/服务器技术。该技术是把 OS 分为两个部分:一部分是用于提供各种服务的服务器,另一部分是用于实现 OS 最基本功能(含通信功能)的内核。在学习时必须清楚,为什么要将 OS 一分为二,由此会带来什么好处。

(4) 面向对象技术。该技术是基于“抽象”和“隐蔽”原则来控制 OS 的复杂度的。它利用封装的数据结构和一组对它进行操作的过程,来表示系统中的某个对象,以达到隐蔽

系统内部数据结构和操作的实现细节的目的。在学习时应了解什么是对象、面向对象技术会给 OS 带来什么好处。

(5) 微内核结构。它是指将客户/服务器技术、面向对象技术用于基于微内核技术的 OS 中所形成的 OS 结构。在学习时应对该结构用于解决什么样的问题、又是如何解决该问题的, 以及该结构有何优点有较深刻的了解。

## 1.3 典型问题分析和解答

### 1.3.1 OS 的引入和发展过程中的典型问题分析

**【例 1】** 试说明操作系统与硬件、其他系统软件以及用户之间的关系。

答: 操作系统是覆盖在硬件上的第一层软件, 它管理计算机的硬件和软件资源, 并向用户提供良好的界面。操作系统与硬件紧密相关, 它直接管理着硬件资源, 为用户完成所有与硬件相关的操作, 从而极大地方便了用户对硬件资源的使用, 并提高了硬件资源的利用率。操作系统是一种特殊的系统软件, 其他系统软件运行在操作系统的基础之上, 可获得操作系统提供的大量服务, 也就是说, 操作系统是其他系统软件和硬件之间的接口。而一般用户使用计算机除了需要操作系统支持外, 还需要用到大量的其他系统软件和应用软件, 以使其工作更加方便和高效。可见, 硬件、操作系统、其他系统软件、应用程序和用户之间存在着图 1.1 所示的层次关系。

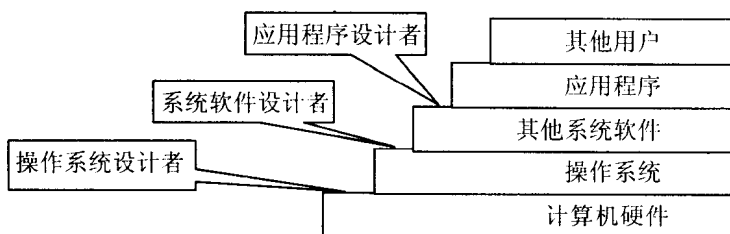


图 1.1 计算机系统的层次关系

**【例 2】** 推动批处理系统和分时系统形成和发展的主要动力是什么?

答: (1) 推动批处理系统形成和发展的主要动力是“不断提高系统资源利用率”和“提高系统吞吐量”。主要表现在: 脱机输入/输出技术的应用和作业的自动过渡大大地提高了 I/O 的速度及 I/O 设备与 CPU 并行工作的程度, 减少了主机 CPU 的空闲时间; 多道程序设计技术的应用更进一步提高了 CPU、内存和 I/O 设备的利用率和系统的吞吐量。

(2) 推动分时系统形成和发展的主要动力是“为了更好地满足用户的需要”。主要表现在: CPU 的分时使用缩短了作业的平均周转时间; 人机交互能力的提供使用户能方便地直接控制自己的作业; 主机的共享使多个用户(包括远程用户)能同时使用同一台计算机独立地、互不干扰地处理自己的作业。

**【例 3】** 实现分时系统的关键问题是什么? 应如何解决?

答: 实现分时系统的关键问题是使用户能与自己的作业交互作用, 即用户在自己的终



端上键入一命令以请求系统服务后，系统能及时地接收并处理该命令，并在用户能够接受的时延内将结果返回给用户。

及时地接收命令和返回输出结果是比较容易做到的，一般只要在系统中配置一多路卡，并为每个终端配置一个缓冲区用来暂存用户键入的命令和输出的结果便可以了。因此，要着重解决的问题是确保在较短的时间内，系统中所有的用户程序都能执行一次，从而使用户键入的命令能够得到及时处理。为此，系统不应让一个作业长期占用 CPU 直至它运行结束，而应设置一个较短的时间片，并规定每个程序只能运行一个时间片，然后，不管它是否运行完毕，都必须将 CPU 让给下一个作业，从而使所有的作业都得到及时的处理，使用户的请求得到及时的响应。

**【例 4】** 试从交互性、及时性以及可靠性三个方面，比较分时系统与实时系统。

答：(1) 从交互性方面来考虑。交互性问题是分时系统的一个关键问题，在分时系统中，用户可以通过终端与系统进行广泛的人机交互，如文件编辑、数据处理和资源共享。实时系统也具有交互性，但在实时系统中仅限于访问系统中某些特定的专用服务程序，也就是说它的交互性具有很大的局限性。

(2) 从及时性方面来考虑。分时系统的及时性是指用户能在很短的时间内获得系统的响应，此时间间隔是由人们能接受的等待时间决定的，一般为 2~3 s。对实时系统来说，及时性是它的关键问题之一，实时信息系统的及时性要求与分时系统相似，而实时控制系统的及时性要求则是由被控制对象所要求的开始截止时间和完成截止时间决定的，一般为秒级、百毫秒级直到毫秒级，甚至更低。

(3) 从可靠性方面来看。可靠性是实时系统的另一个关键问题，实时系统中的任何差错都可能带来巨大的经济损失，甚至无法预料的灾难性后果，所以实时系统往往采取多级容错措施来保证系统的高度可靠。分时系统虽然也要求可靠，但比实时系统的要求要低。

### 1.3.2 OS 的基本特征和功能中的典型问题分析

**【例 5】** 操作系统具有哪些特征？它们之间有何关系？

答：操作系统的特征有并发、资源共享、虚拟和异步性。它们的关系如下：

(1) 并发和共享是操作系统最基本的特征。为了提高计算机资源的利用率，OS 必然要采用多道程序设计技术，使多个程序共享系统的资源，并发地执行。

(2) 并发和共享互为存在的条件。一方面，资源的共享以程序(进程)的并发执行为条件，若系统不允许程序并发执行，自然不存在资源共享问题；另一方面，若系统不能对资源共享实施有效管理，协调好诸进程对共享资源的访问，也必将影响到程序的并发执行，甚至根本无法并发执行。

(3) 虚拟以并发和资源共享为前提。为了使并发进程能更方便、更有效地共享资源，操作系统常采用多种虚拟技术来在逻辑上增加 CPU 和设备的数量以及存储器的容量，从而解决众多并发进程对有限的系统资源的争用问题。

(4) 异步性是并发和共享的必然结果。操作系统允许多个并发进程共享资源、相互合作，使得每个进程的运行过程受到其他进程的制约，不再“一气呵成”，这必然导致异步性特征的产生。