



# 操作系统

(一)

# 功能各异、品种繁多的操作系统



# 这门课程要讲什么内容？

内容：操作系统的工作原理

注意：只讲原理

# 推荐一本理论联系实践的书

附带实例代码





# 操作系统是什么？

- 1、它是一个软件
- 2、它是一个“管理员”，帮助我们管理和使用计算机
- 3、其他软件都依靠它来运行，使用各种资源时都要向它提出申请

## ■ 为什么要使用操作系统

- ① 直接使用计算机，很不方便
- ② 大部分人无需知道计算机系统的工作原理
- ③ 设备昂贵、使用效率却很低
- ④ 技术层出不穷，资源管理麻烦
- ⑤ 不同厂家生产的设备要有有效的协同工作



# 第一章 操作系统引论

1.1 操作系统的目标、作用

1.2 操作系统的发展过程

1.3 操作系统的基本特性

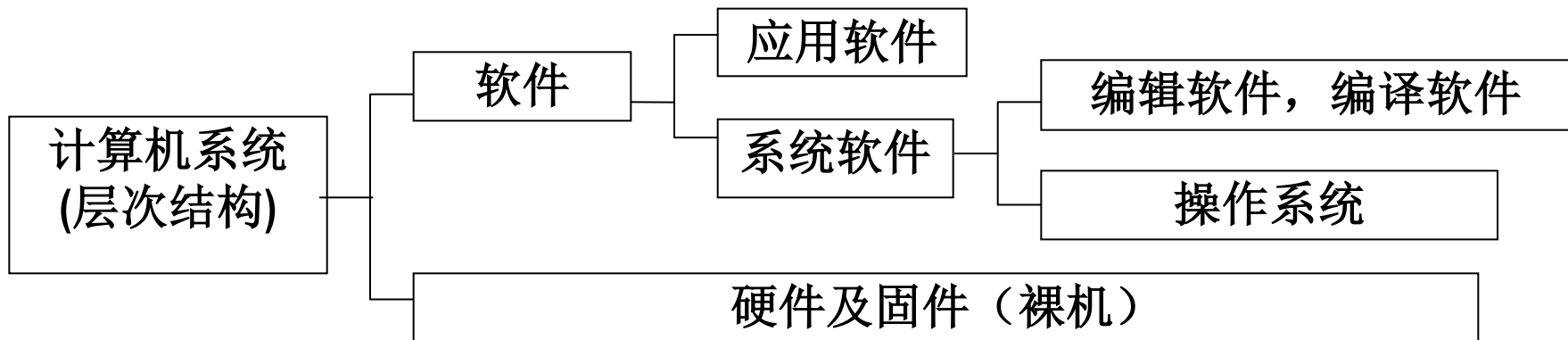
1.4 操作系统的主要功能

1.5 操作系统的结构设计

# 1.1 操作系统的目标和作用

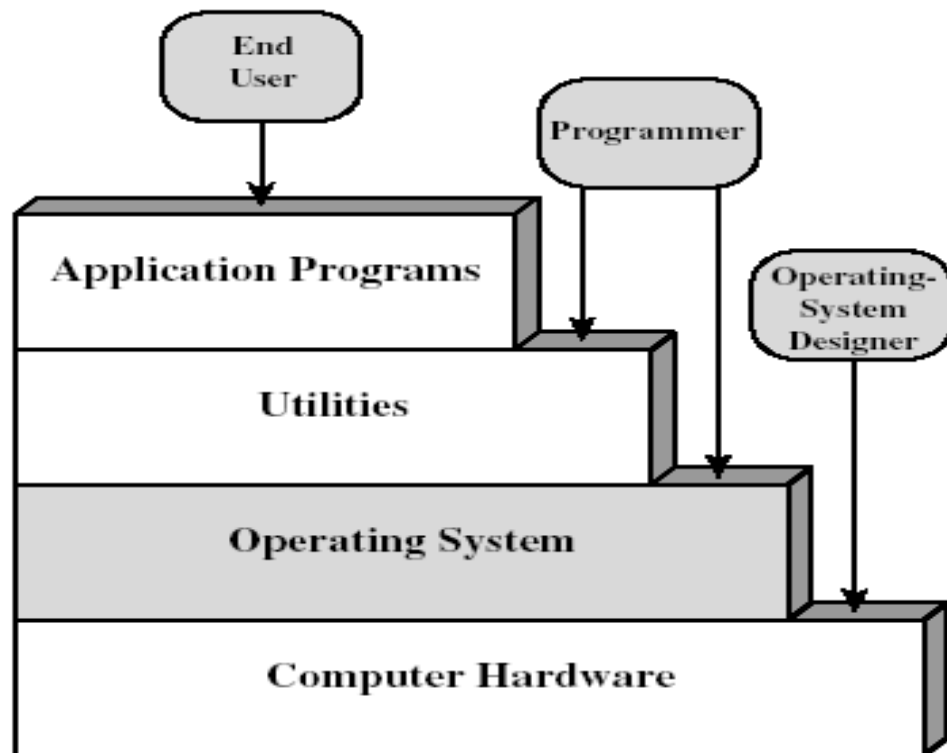
## ■ 计算机系统的组成

- 由硬件系统、软件系统两部分组成





## ■ 计算机系统的层次结构图



- OS的位置：紧贴硬件之上，其他软件之下（支撑其他软件的运行）
- **OS是配置在硬件上的第一层软件，是对硬件系统的首次扩充。**

# ■ 操作系统的设计目标

- **方便性**：使用户方便使用计算机
- **有效性**：提高资源利用率，提高系统吞吐量
- **可扩充性**：为适应计算机硬件、体系结构以及计算机应用发展的要求；推动OS结构的不断发展
- **开放性**：不同软硬件之间的兼容，遵循世界标准规范

方便性、有效性是设计**OS**时最重要的目标

开放性是**20世纪90年代**以后计算机技术的核心问题

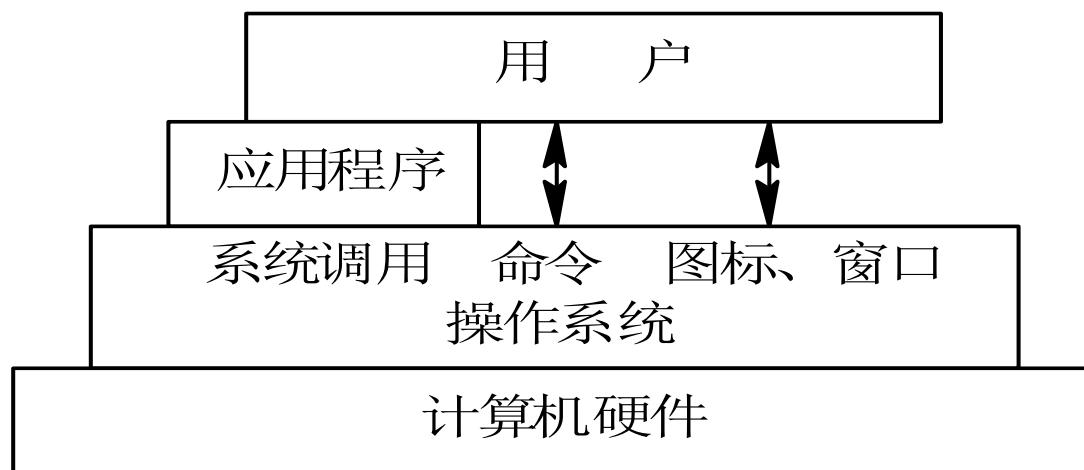
## ■ 操作系统的作用

### □ 从不同的角度来看

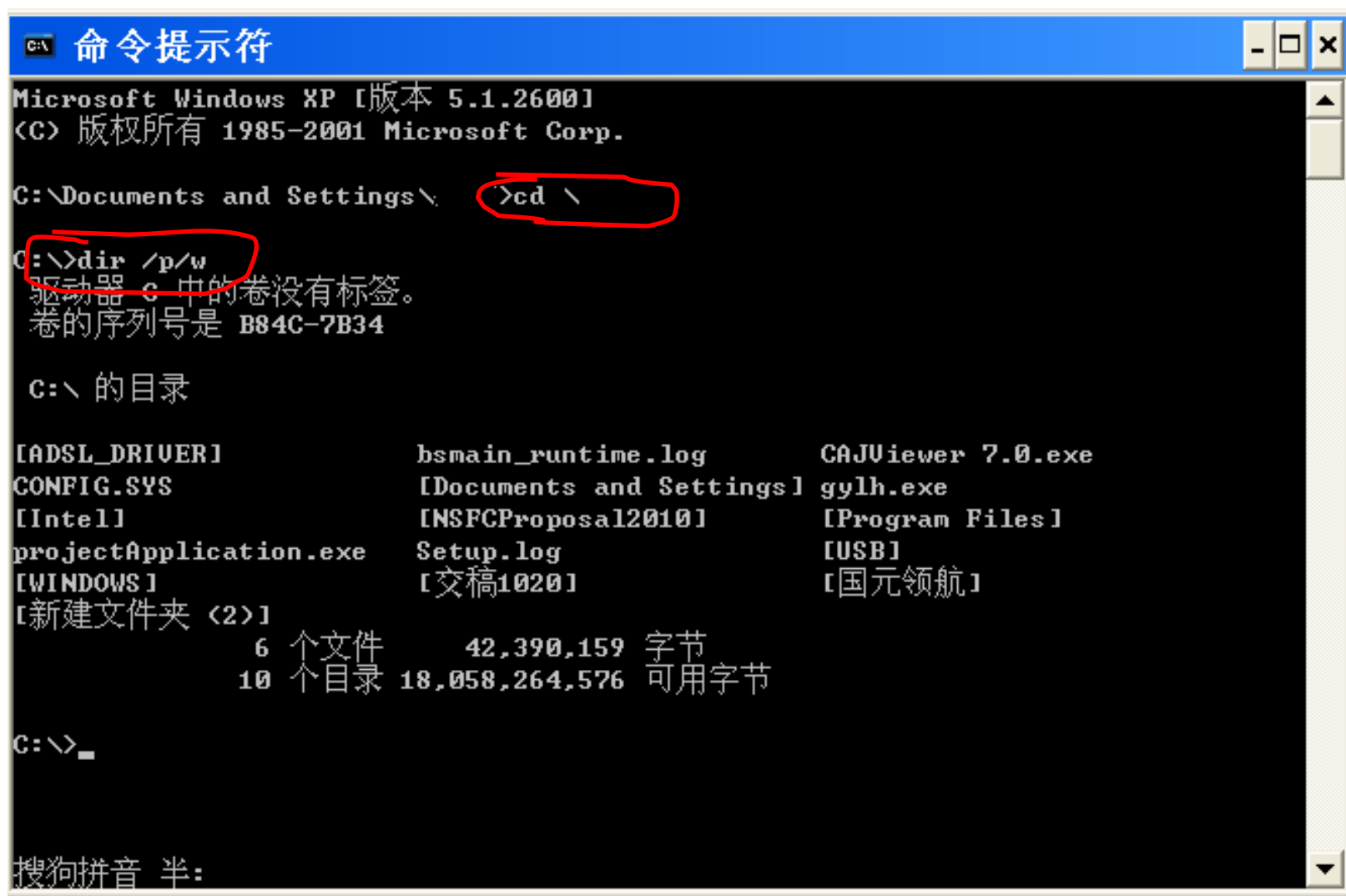
- ① OS是用户使用计算机硬件系统的接口
- ② OS是计算机系统资源的管理者
- ③ OS实现了对计算机资源的抽象

## ① OS是用户使用计算机硬件系统的接口

- **命令方式**（命令行、菜单式、命令脚本式）；
- **系统调用**（形式上类似于过程调用，在应用编程中使用）
- **图形方式**



## □ 命令方式



The screenshot shows a Windows XP Command Prompt window titled "命令提示符" (Command Prompt). The window has a blue title bar with standard Windows XP window controls. The command history is as follows:

```
Microsoft Windows XP [版本 5.1.2600]
(C) 版权所有 1985-2001 Microsoft Corp.

C:\Documents and Settings\ >cd \

C:\>dir /p/w
驱动器 C 中的卷没有标签。
卷的序列号是 B84C-7B34

C:\> 的目录

[ADSL_DRIVER]          bsmain_runtime.log      CAJViewer 7.0.exe
CONFIG.SYS             [Documents and Settings] gylh.exe
[Intel]                [NSFCProposal2010]      [Program Files]
projectApplication.exe Setup.log                        [USB]
[WINDOWS]              [交稿1020]              [国元领航]
[新建文件夹 (2)]

                6 个文件          42,390,159 字节
                10 个目录 18,058,264,576 可用字节

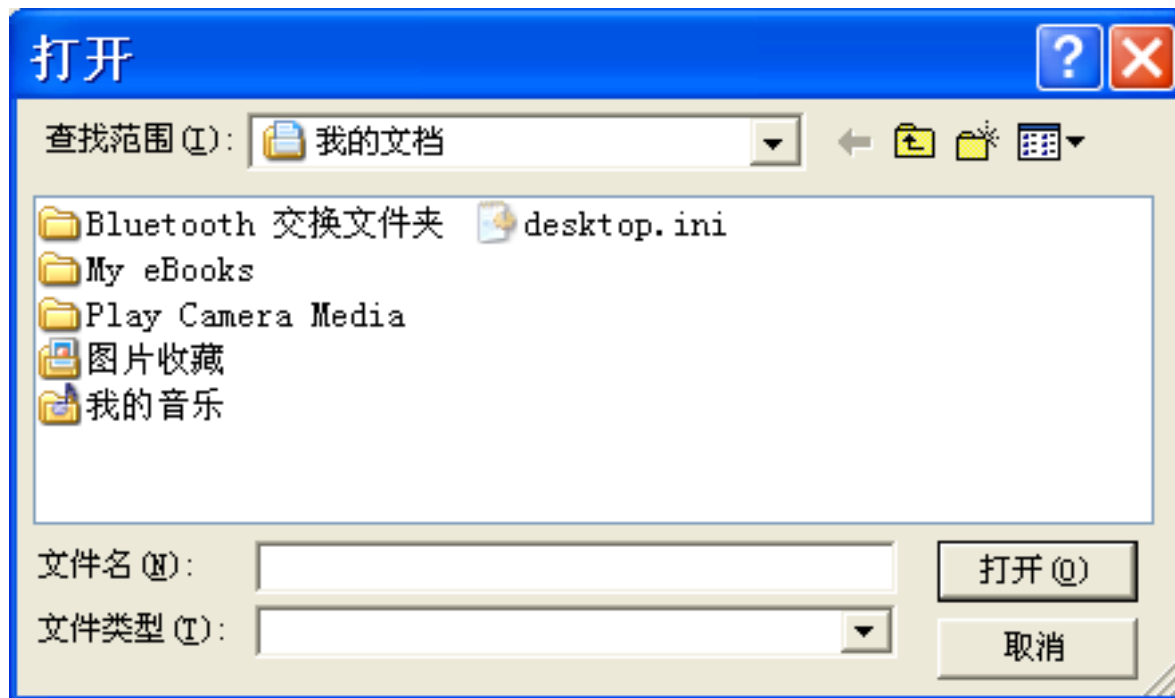
C:\>_

搜狗拼音 半:
```

Red circles highlight the command prompt title bar, the `>cd \` command, and the `dir /p/w` command.

## □ 系统调用方式 (以VC6为例)

```
CFileDialog dfile(TRUE); //定义文件打开对话框  
dfile.DoModal ();        //显示对话框
```



## □ 图形窗口方式



## ② OS是计算机系统资源的管理者

### □ 管理的对象：

- 硬件——CPU、存储器、外部设备
- 软件——程序、数据

### □ 管理的内容：

- 资源的当前状态（数量和使用情况）、资源的分配/回收和访问操作、相应管理策略（包括用户权限）。



### ③ OS实现了对计算机资源的抽象

□ 在裸机上添加：

- 设备管理
- 文件管理
- 存储管理（针对内存和外存）
- 处理机管理（针对CPU）

每当在计算机系统中覆盖一层软件后，计算机系统的功能就增强一级。掩盖了资源的使用细节，使用更方便。

## ■ 推动操作系统发展的动力

- ① 不断提高计算机资源利用率
- ② 方便用户
- ③ 器件的不断更新换代
- ④ 计算机体系结构的不断发展(多处理机OS、网络OS)
- ⑤ 不断提出新的应用需求

## 1.2 操作系统的发展过程

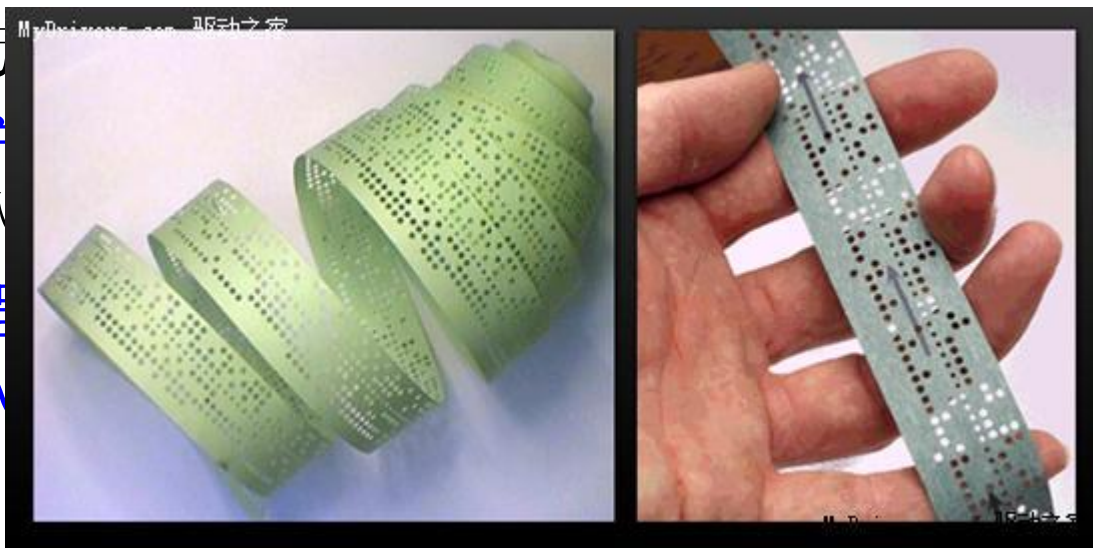
- ① 未配置OS
- ② 单道批处理系统
- ③ 多道批处理系统
- ④ 分时系统
- ⑤ 实时系统

# ① 未配置OS

- 1946 ~ 50年代(电子管), 计算机造价昂贵;

- 工作方式

- 用户
- 业人
- 编程
- 输入



是计算机专

- 计算机的工作特点

- **用户独占全机**: 资源不与其他用户共享, 资源利用率低;
- **CPU等待人工操作**: 计算前, 手工装入纸带或卡片; 计算完成后, 手工卸取纸带或卡片; **CPU**利用率低;

## ■ 手工操作方式存在的问题：人机矛盾

- 计算机处理能力高，但手工操作效率低。手工操作方式严重降低了计算机资源的利用率，造成浪费。

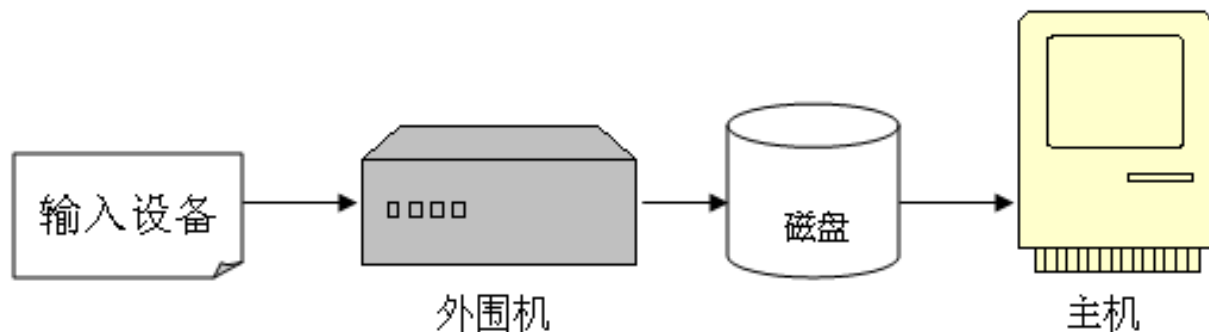
## ■ 脱机输入/输出方式(50年代末)

### □ 工作方式

- 将纸带装入纸带机，在外围机的控制下，将数据装入磁带上。CPU从磁带上高速调入内存。

### □ 优点

- ① 减少CPU的空闲时间
- ② 提高I/O速度



## ② 单道批处理系统

### ■ 50年代末 ~ 60年代中期（晶体管）

- **引入的目的：** 为充分利用计算机系统，应尽量让系统连续运行，以减少空闲时间。
- **工作方式：** 一批作业以脱机方式输入到磁带上，用监督程序控制使它们一个接一个的连续处理
- **什么叫单道：** 内存中始终只保持一道作业运行
- **什么叫批处理：** 成批处理作业

## ■ 单道批处理系统的特征

- 单道批处理系统是最早出现的一种OS。严格地说，它只算是OS的前身，并非现在人们所理解的OS。
- 尽管如此，该系统比手工操作方式已有很大进步。



## ■ 单道批处理的主要问题

- 在单道批处理系统中，内存中仅有一道作业，无法充分利用资源（比如内存），致使系统性能较差。
- 为进一步提高效率，在60年代中期又引入多道程序设计技术，由此形成了多道批处理系统。

### ③ 多道批处理系统

- 60年代中期 ~ 70年代中期（集成电路），利用多道批处理提高资源的利用率。
  - 用户所提交的作业都先存放在外存上并排成一个队列，称为“**后备队列**”
  - 由**作业调度程序**按一定的**算法**从后备队列中**选择若干个作业**调入内存，使它们**共享CPU**和系统中的各种资源。

## ■ 单道和多道批处理的比较

	单道	多道
内存使用	每次一道作业	每次多道作业 (充分利用资源)
作业次序	先进先出	无确定次序

- **多道程序系统**和**多处理系统**的区别：前者指多个程序同时在内存中并发运行，后者指多个处理器。

# ■ 多道批处理系统的优缺点

## □ 优点

- 资源利用率高

- 系统吞吐量大 (**系统吞吐量**是指从作业进入系统开始所完成的总量)

## □ 缺点

- 平均周转时间长。 (**周转时间**: 从作业进入系统开始, 直至完成并退出系统为止所经历的时间)
- 无交互能力。

## ■ 多道批处理系统需要解决的问题

- 多道批处理系统是一种有效但复杂的系统，为了使多道程序能协调运行，须解决几个问题：
  - ① 处理机管理问题
  - ② 内存管理问题
  - ③ I/O设备管理问题
  - ④ 文件管理问题
  - ⑤ 作业管理问题
  - ⑥ 用户与系统的接口问题
- 所以应增加一组软件，以解决上述问题。正是这样一组软件构成了现代意义上的操作系统。



## ■ 操作系统定义

一组能有效地组织和管理计算机硬件和软件资源，合理地对各类任务进行调度，以及方便用户使用的程序的集合。

## ④ 分时系统

### ■ 分时系统的产生

#### □ 产生原因

- 人机交互、共享主机、便于用户上机

#### □ 实现的关键

- 如何使用户与自己的作业交互？

#### □ 基本实现思路

- 及时接收、及时处理
- 所以，作业要直接进入内存，且不允许一个作业长时间占据CPU

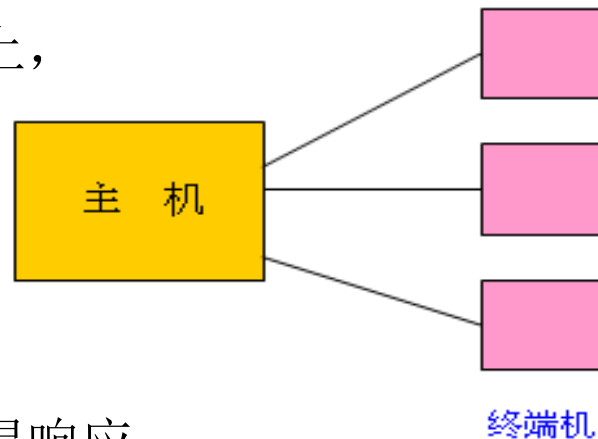
#### □ 解决方法

- 把计算机的系统资源（尤其是CPU时间）进行时间上的分割，每个时间段称为一个时间片，每个用户依次轮流使用时间片。

## ■ 分时系统的特征

- **多路性**：多台终端同时连接到一台主机上，

并按分时原则为每个用户服务。



- **独立性**：各用户独立操作，互不干扰。
- **及时性**：用户的请求能在很短时间内获得响应。
- **交互性**：用户可通过终端与系统进行人机交互。



## ⑤ 实时系统

### ■ 实时系统的引入

- **实时系统**：指系统能**及时响应**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。
- **运行的正确性**：**不仅由运行结果决定，还由产生结果的时间决定。**
- **主要作用**：用于过程控制、事务处理等有实时要求的领域，其主要特征是**实时性**和**可靠性**。



## □ 应用场合

- 工业（武器）控制系统
- 信息查询系统
  - 订票系统
- 多媒体系统
- 嵌入式系统

## ■ 实时任务的类型

### □ 周期性实时任务

外部设备周期性发出信号，要求任务按指定周期循环执行，以便周期性地控制某外部设备。

### □ 非周期性实时任务

无明显周期性，但必须联系一个截止时间：开始截止时间 或 完成截止时间。

## ■ 实时任务的类型

根据对截止时间的要求来划分：

### □ 硬实时任务

系统必须满足任务对截止时间的要求，否则可能出现难以预测的结果。

### □ 软实时任务

它也联系着一个截止时间，但并不严格，若偶尔错过了任务的截止时间，对系统产生的影响也不会太大。

## ■ 实时系统与分时系统的区别

### □ 多路性

- 实时系统也采用分时的思想，但系统**经常性**对多路现场信息进行采集、对多个对象进行控制

### □ 独立性

- 与分时系统相同。对信息的采集、对对象的控制都是彼此互不干扰的

### □ 及时性

- 分时系统：人能接受的等待时间
- 实时系统：以控制对象所要求的开始截止时间或完成截止时间来确定

### □ 交互性

- 仅限于访问系统中某些特定的专用服务程序

### □ 可靠性

- 要求高度可靠性，否则，后果很严重

## ■ 微机OS的发展

### □ 单用户单任务OS

MS-DOS

### □ 单用户多任务OS

Windows

### □ 多用户多任务OS

UNIX、Linux

# 1.3 操作系统的基本特征

## ■ 操作系统的基本特征

- ① 并发
- ② 共享
- ③ 虚拟
- ④ 异步

## ① 并发

- 并发：多个事件在**同一时间段内**发生(在多道程序中，宏观上并发，微观上交替执行)。
- 并行：多个事件**在同一时刻**发生。



- OS是一个并发系统，各进程间并发，操作系统与应用程序间并发。OS要完成这些并发进程的管理。
- 程序的静态实体是可执行文件，而动态实体是进程（或称作任务），并发指的是进程间的并发。
- “并发”和“进程”是现代OS最重要的基本概念，也是OS的运行基础

## ② 共享

- 多个进程共享有限的计算机系统资源，OS要对系统资源进行合理分配和使用。
- 目前主要有以下两种资源共享方式：
  - **互斥共享**：资源分配后到释放前，不能被其他进程使用，如打印机
  - **“同时”访问**：如磁盘文件

- **临界资源**(独占资源): 在一段时间内只允许一个进程访问的资源。
  - 如: 打印机
  - 对临界资源的访问, 应使用 “**互斥共享方式**”。

- “并发”和“共享”是多用户（多任务）OS的两个最基本的特征。

### ③ 虚拟

- 虚拟，是指通过某种技术把一个物理实体变为若干个逻辑上的对应物。
  - 物理实体是实际存在的，而后者是用户感觉上的东西。
- 虚拟是OS管理系统资源的重要手段，可提高资源利用率。
- 如何虚拟：分时或分空间。
  - CPU —— 每个用户（进程）的“虚处理机”（多道程序设计技术）
  - 内存 —— 虚拟内存
  - 设备—— 共享打印机等

## ④ 异步性

- 异步性，也称不确定性，指进程的执行顺序和执行时间的不确定性
  - 多道程序环境下，允许进程并发执行。因此，运行时走走停停。程序完成时间、推进速度等不可预知。
  - **异步性**：进程是以人们不可预知的速度向前推进。
  - 但运行环境相同，运行结果是完全相同的。

# 1.4 操作系统的功能

- ① 处理机管理功能
- ② 存储器管理功能
- ③ 设备管理功能
- ④ 文件管理功能
- ⑤ 人机接口

.....

## ■ 引入OS的主要目的

- 为多道程序的运行提供良好的运行环境；
- 保证多道程序有条不紊、高效地运行；
- 最大程度地提高系统中各种资源的利用率；
- 方便用户使用。



# 1.5 操作系统的结构设计

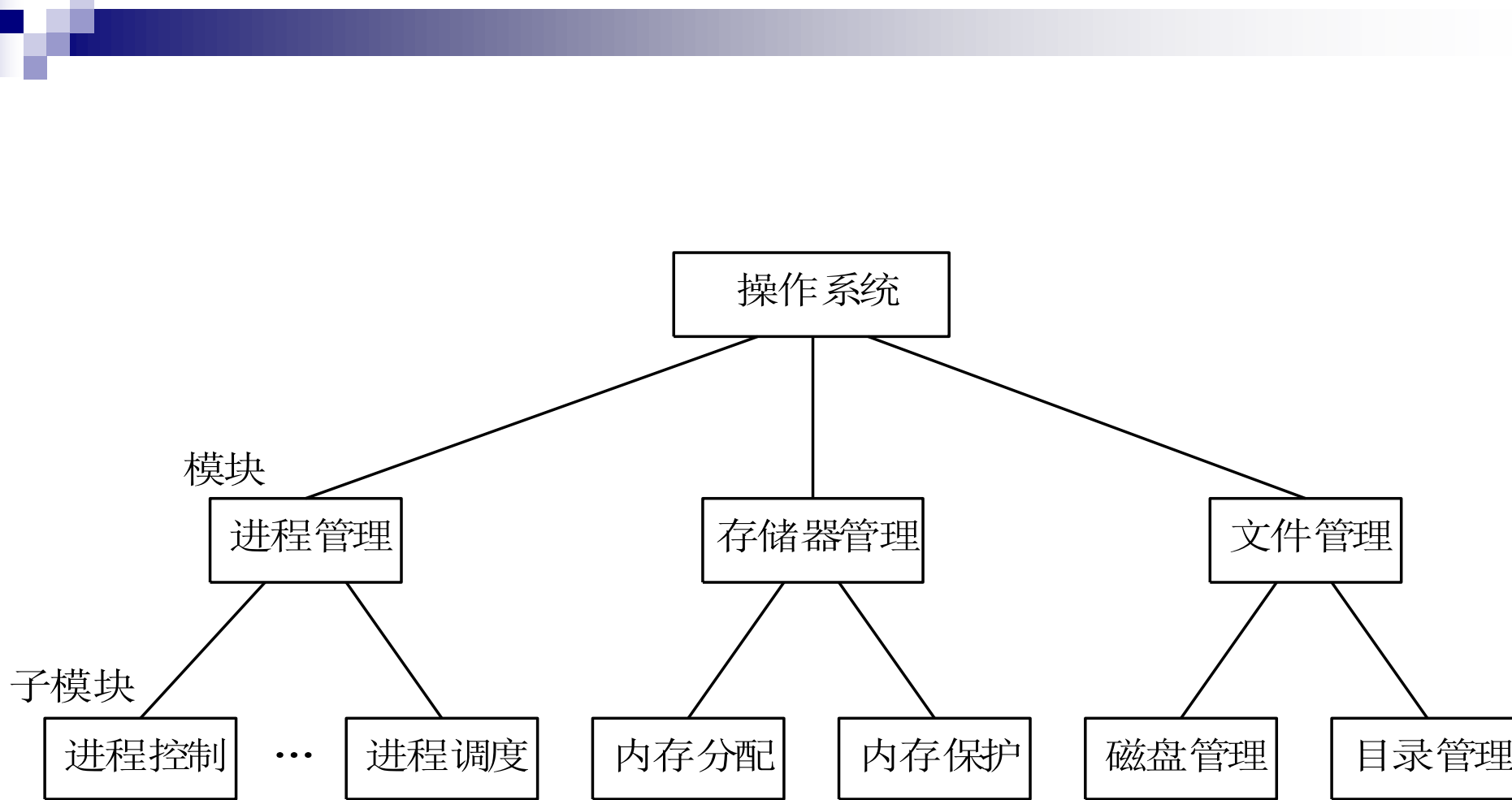
- OS是一个大型系统软件，其结构已经历4代的变革
  - ① 无结构OS
  - ② 模块化结构OS
  - ③ 分层式结构OS
  - ④ 微内核结构OS
- 后来，人们把工程学的基本原理和方法，引入到软件开发过程中，形成软件工程学。

## ■ 传统的操作系统结构：

- OS是一个十分复杂的大型软件。为了控制其复杂性，在开发OS时，先后引入了分解、模块化、抽象和隐蔽等方法。
- 开发方法的不断发展，促进了OS结构的更新换代。我们把第1代至第3代的结构，称为传统结构，而把微内核结构称为现代结构。

## ■ 模块化OS结构

- 模块化程序设计技术，是最早(20世纪60年代)出现的一种程序设计技术。该技术是基于“分解”和“模块化”原则来控制大型软件的复杂度的。
- OS不再是由众多的过程直接构成，而是将OS按功能划分为若干具有独立性和大小的模块。
- 每个模块具有某方面的管理功能，并规定好各模块间的接口。
- 可将各模块再细分为更小的模块，同样也要规定各子模块之间的接口。



## ■ 模块化OS的优缺点

### □ 优点

- 提高了OS设计的正确性、可理解性和可维护性。
- 增强了OS的可适应性。
- 加速了OS的开发过程。

### □ 缺点

- 模块接口的规定难以满足实际需求
- 决定顺序“无序性”

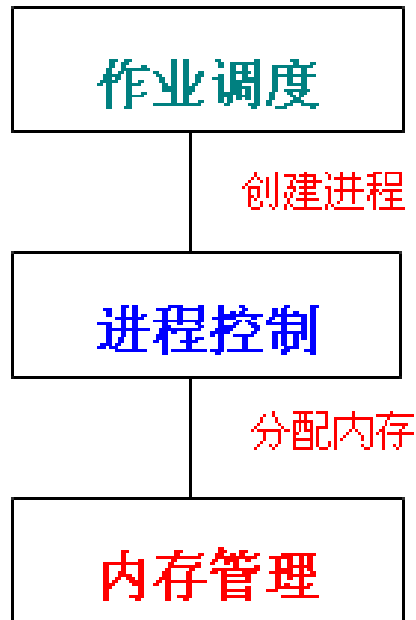
# ■ 分层式结构OS

## □ 基本原理

- 每一层都使用底层提供的功能和服务，使系统的调试和验证变得容易

## □ 设计关键

- OS分哪几层，应怎样确定各层间的顺序。



## ■ 微内核结构

- 支持多处理机，适用于分布式系统环境

- 微内核技术的引入

- 微内核是精心设计的、实现现代OS核心功能的小型内核 (更小、更精炼)。
- 并非是一个完整的OS， 只是为构建通用OS提供一个重要基础。
- 在微内核结构中，通常采用客户/服务器模式(C/S)， OS的大部分功能和服务，都是由若干服务器来提供的， 如文件服务器、作业服务器和网络服务器等。

## ■ 微内核的基本功能

- 进程(线程)管理
- 低级存储器管理
- 中断处理 等

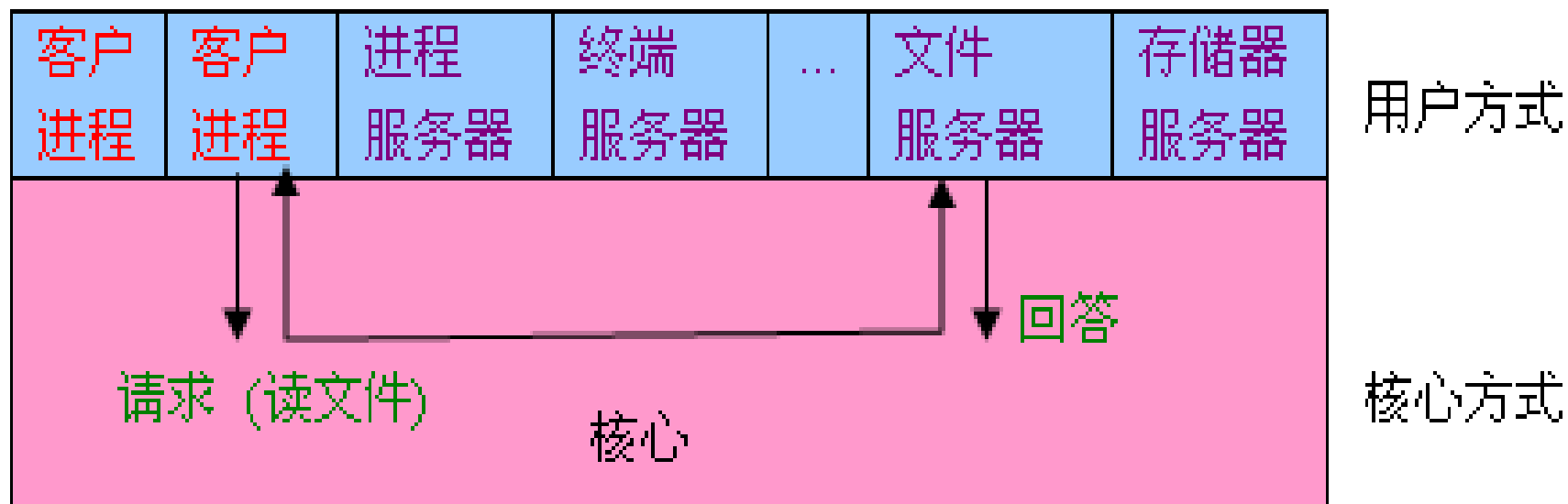
## ■ 微内核结构的特征

- 以微内核为OS的核心，以C/S为基础，采用面向对象的程序设计方法。



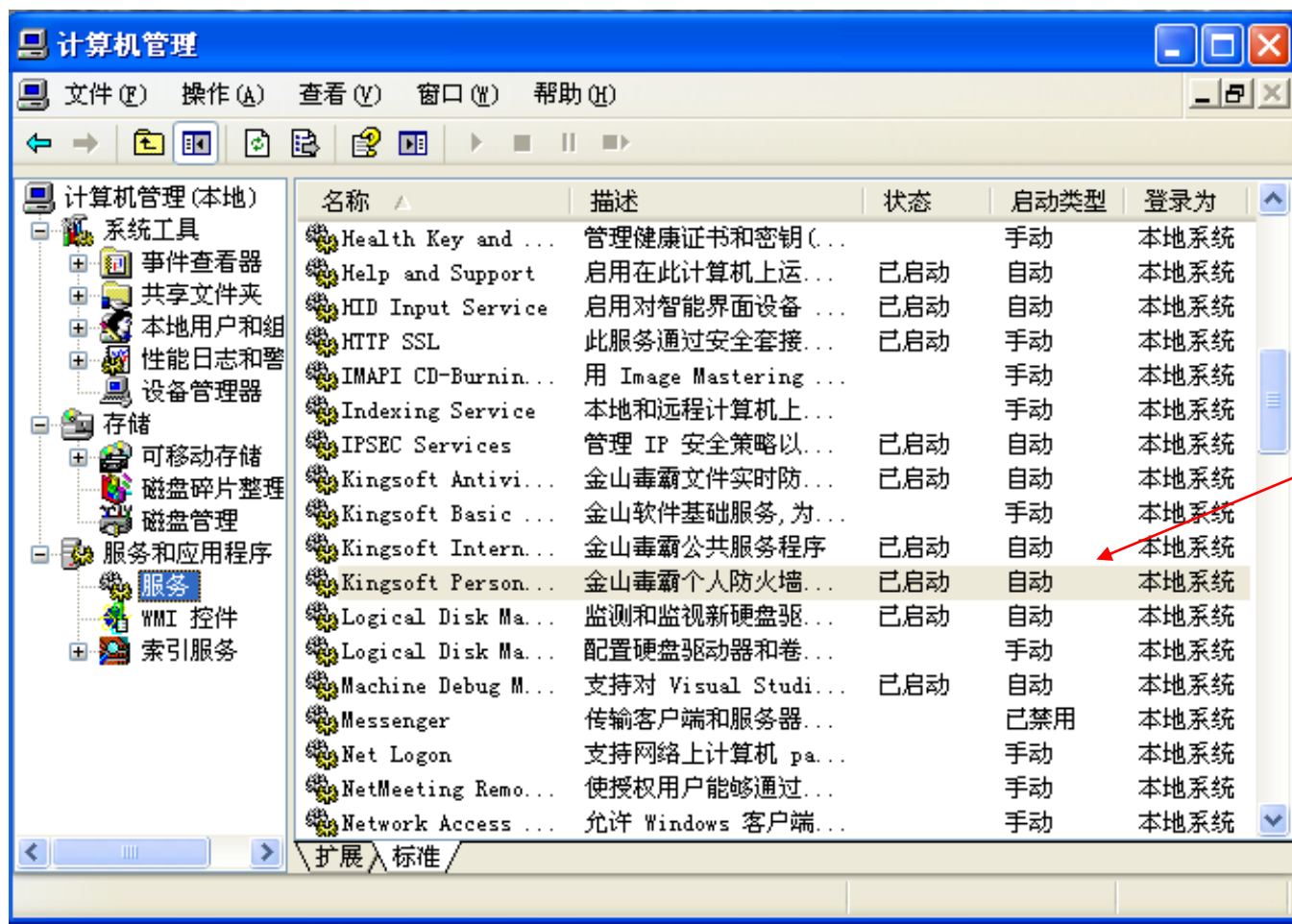
## ■ 客户/服务器模式 (C/S)

- OS划分为两部分：服务器、内核
- 进程分为两部分：服务器进程、客户进程




## ■ 优点:

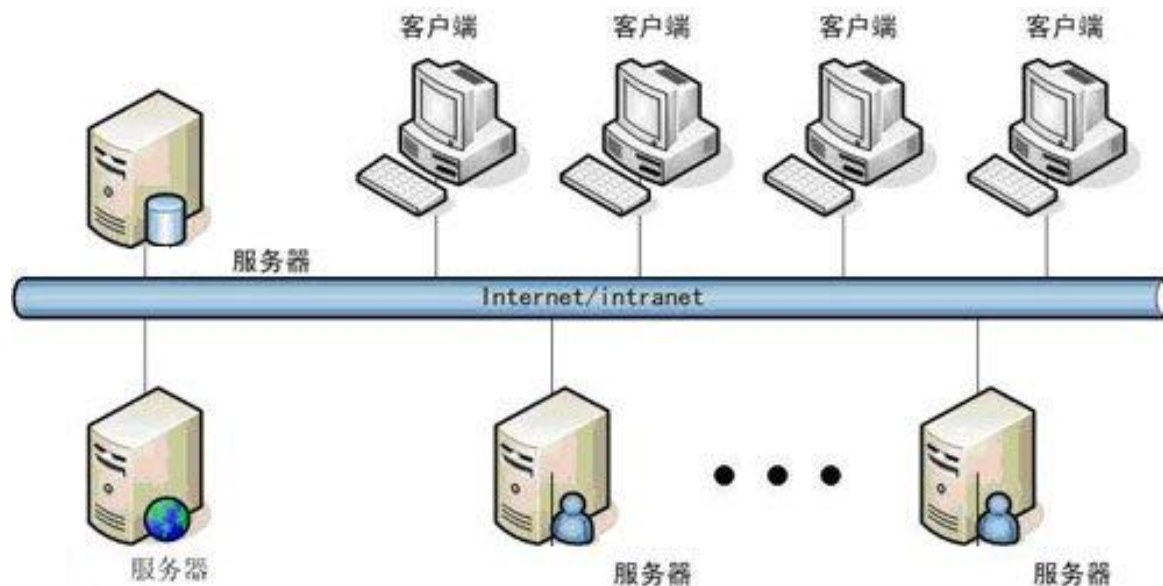
(1) 提高系统的灵活性、可扩充性: 根据需要选配部分或全部服务器, 根据计算机的发展更新或增加服务器



金山毒霸  
(扩充的)

- 
- (2) 提高OS的可靠性：服务器进程出错时仅影响它自己，不影响内核/硬件。
- (3) 可移植性强：因为与底层硬件相关的部分都放在内核中，因此进行移植时，只需做较小的修改。

(4) 可运行于分布式系统中：服务器和客户可以部署在不同的机器上



(5) 融入面向对象技术：明显减少开发系统所付出的开销。

# 小结

- 理解OS的目标和作用
- 掌握OS的各种类型(单道和多道批处理、分时、实时)
- 掌握OS的基本特征(并发、共享、虚拟和异步)
- 理解OS的主要功能(各种管理功能)
- 了解OS的结构设计(特别是微内核技术)

# 练习

- 1. 操作系统是一种\_\_ (2) \_\_，在操作系统中采用多道程序设计方式能提高CPU和外部设备的\_\_ (1) \_\_。

A: (1)通用软件； (2)系统软件； (3)应用软件； (4) 软件包

B: (1)利用效率； (2)可靠性； (3)稳定性； (4)兼容性

■ 2. 分时系统中，为使多个用户能够同时与系统交互，最关键的问题是\_\_\_ **C** \_\_\_

A 计算机具有足够的运行速度

B 内存容量应足够大

C 系统能及时地接收多个用户输入

D 能在一段的时间内，使所有用户程序都能运行

■ 3. 在设计分时操作系统时，首先要考虑的是\_\_\_ **B** \_\_\_

A 灵活性和可适应性      B 交互性和响应时间

C 周转时间和系统吞吐量      D 实时性和可靠性

■ 4. 在设计实时操作系统时，首先要考虑的是\_\_\_ **D** \_\_\_

A 灵活性和可适应性      B 交互性和响应时间

C 周转时间和系统吞吐量      D 实时性和可靠性。



■ 5.在设计批处理系统时，首先要考虑的是\_\_ **C** \_\_

A 灵活性和可适应性      B 交互性和响应时间

C 周转时间和系统吞吐量   D 实时性和可靠性