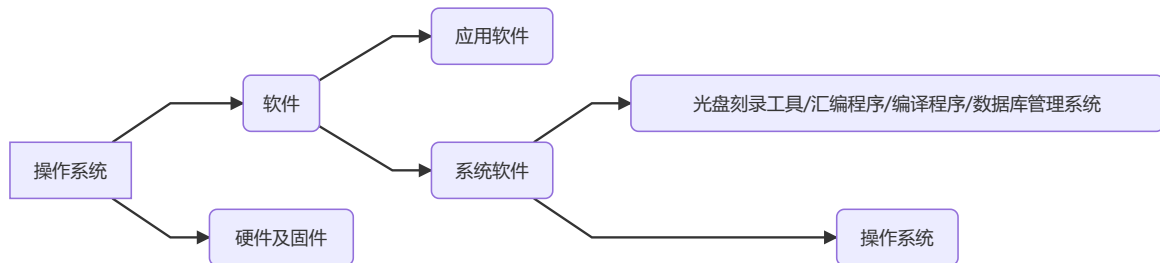


# 操作系统

## 什么是操作系统

### 操作系统的组成



### 操作系统由软件和硬件构成

- 计算机硬件作为基础
- 操作系统管理和调度硬件
- 系统工具例如iptables，可以修改内核的netfilter
- 应用软件 例如数据库管理软件
- 操作系统向上层提供系统调用
- 操作系统的作用就是将复杂的硬件调度封装起来，作为系统调用提供给其他软件

### 操作系统的设计目标

- 方便性  
提供用户接口(Interface)，使计算机便于使用
- 有效性  
合理分配资源，提高利用率，增加吞吐量(单位时间的工作量)
- 可扩充性  
适应将来的硬件体系和对应的更高功能和性能要求
- 开放性  
不针对特定的硬件，适应不同硬件，实现软件的可移植性和互操作性

## 操作系统的作用

### (1).用户与计算机硬件系统之间的接口

#### 外壳(Shell)

- 面向一般用户
- CLI, GUI, Script, Menu
- 使用户使用计算机更加便利

#### 系统调用(System Call)

- 面向开发人员
- 形式上类似于过程调用，编写程序的时候使用
- 软件和系统交互的基本机制

## (2).系统资源管理

### 管理对象

- 硬件对象和软件对象的控制

### 管理内容

- 资源分配，回收，访问操作
- 记录资源当前状态，相应管理策略

## (3).用作扩充机器

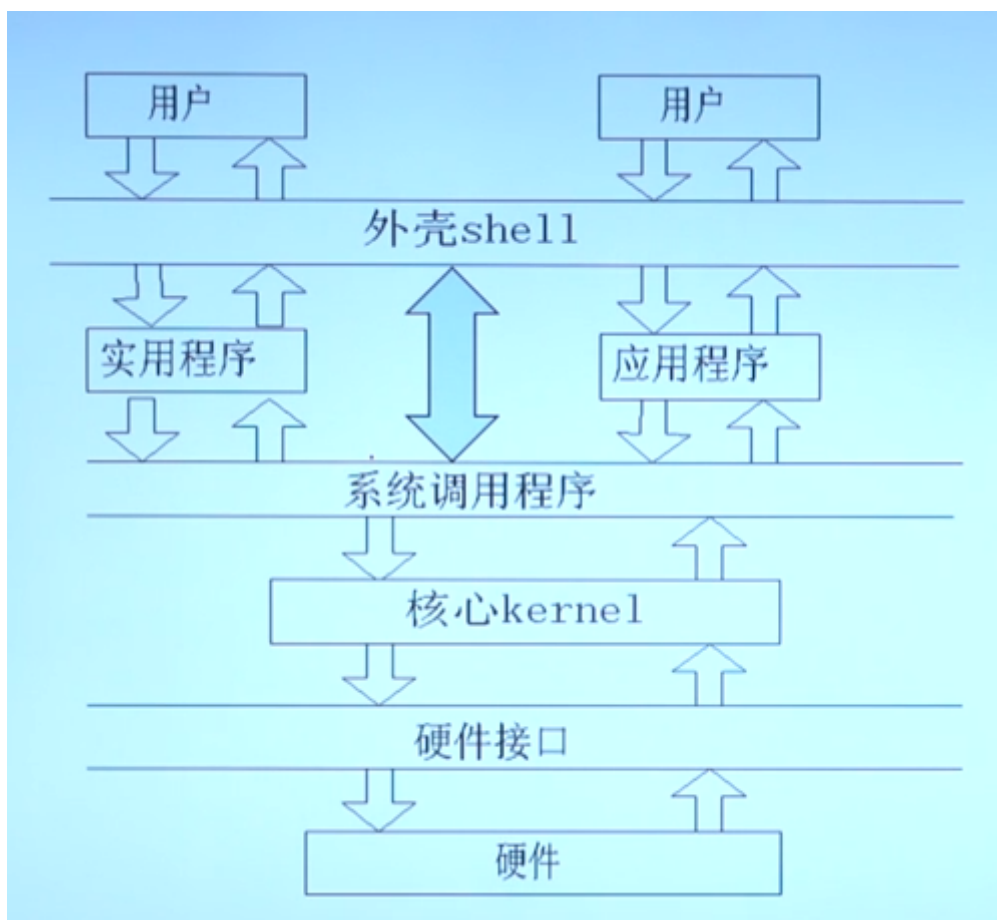
- 在硬件层上管理各种资源

## 操作系统组成及层次结构

### 三层结构

最上层为用户接口GUI,CLI等
中间层为对象管理软件
最底层是操作系统抽象硬件等作为对象

### 结构细化



## 操作系统定义

## 最基本的系统软件

- 有效管控硬件软件，对任务的合理调度，组织作业流程

## 操作系统发展的主要动力

---

### 硬件

- 提高硬件资源利用率
- 提高性能
- 支持新型硬件

### 人们的需求

- 改善用户体验

## 早期计算机

---

### 纸带计算机

- 1946-50年代中期
- 价格昂贵，只有计算功能
- 工作方式
  - 用户：管理员和使用者的双重身份
  - 输入输出：卡带或纸片
  - 编程语言：机器语言

#### 工作特点：

- 资源利用率低
- 计算前后，CPU空闲时间浪费算力

## 速度矛盾和缓和途径

---

- 通道技术，中断技术，缓冲技术
- 脱机输入输出
  - 主机不参与对输出结果的使用
- 专业操作员和批处理技术

### 单道批处理技术

- 提高系统资源利用率
- 在监督程序的控制下使作业连续进行
- 内存中始终存在着一道作业
- 吞吐量比纸带计算机提高很多

#### 特性：

- 自动性
  - 作业按次序执行，无需干预

- 顺序性
  - 按顺序进入内存执行
- 单道性
  - 只有一个程序可以运行，当其发生错误，调用监控程序管理作业，处理结束，退出内存，交出执行权

## 多道批处理技术

- 多道程序在同一处理器上交替运行
- 通过作业调度算法调度作业，处理异常
- 吞吐量比单道批处理提高很多

### 特性：

- 多道性
  - 内存中有多道程序，允许并发执行
- 无序性
  - 多个作业何时进入内存没有严格的限定
- 调度性
  - 作业调度
  - 进程调度

### 优点

- 资源利用率高
- 系统吞吐量大

### 缺点：

- 作业周转时间长，对短作业不公平
- 无交互，不利于调试和修改

### 需求分析：

- 处理机管理问题
  - 管理机分配
  - 提高利用率
  - 分配与回收
- 内存管理问题
  - 分配与保护
- IO设备管理问题
  - 设备共享，提高利用率
- 文件管理问题
  - 文件组织方便用户使用
  - 数据安全和一致性
- 作业管理问题
  - 作业调度
  - 工作流程

# 分时系统

---

- 人机交互
  - 程序修改调试
- 共享主机
  - 处理器昂贵
- 便于用户上机
  - 单处理器多终端
- 分时系统的概念
  - 单主机，多终端，多用户，可交互使用计算机

## 分时系统遇到的问题

- 及时接受输入
  - 多路卡，多路缓冲区
- 及时处理
  - 所有用户作业直接进入内存
  - 有限时间内就可以将每个程序都运行一次

## 分时系统的实现

### 作业直接进入内存

- 用户键入指令直接进入内存，不必等待

### 时间片轮转策略

- 时间片很短，每个程序都运行一些时间周期就切换到其他程序，给人一种并发的错觉

### 实现方式

- 单道分时系统
- 具有前台和后台的分时系统
- 多道分时系统

## 分时系统特征

- 多路性
  - 宏观上：多个用户同时工作，共享资源
  - 微观上：每个用户轮流执行时间片
- 独立性
  - 互不干扰
- 及时性
  - 响应时间短
- 交互性
  - 用户可以通过终端与系统进行交互

## 实时系统

---

- 时间约束

- 截止时间
- 可预测性
  - 对任务的执行时间进行预判
- 可靠性
- 交互作用

## 实时系统与分时系统比较

---

### 实时系统更追求于

- 及时性
- 高度可靠

### 区别

- 分时系统的目标是提供一种通用性很强的系统，有较强的交互能力；而实时系统则大都是具有特殊用途的专用系统，交互能力略差
- 分时系统对响应时间虽有要求，但一般来说，响应时间由人所能承受的等待时间来确定；而实时系统对响应时间要求很高，一般由控制系统或信息处理磁头所能接受的延迟时间来决定。