



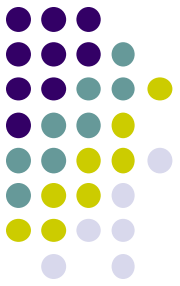
# 计算机组成与系统结构

## 第九章 操作系统支持

吕昕晨

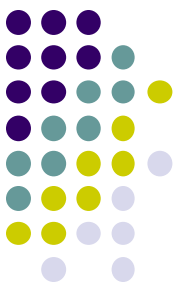
[lvxinchen@bupt.edu.cn](mailto:lvxinchen@bupt.edu.cn)

网络空间安全学院



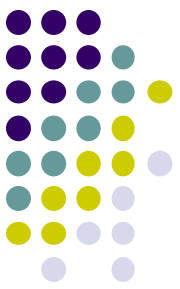
# 第九章 操作系统支持

- 操作系统概述
  - 操作系统概念
  - 操作系统功能
  - 操作系统特性与硬件环境
- 调度



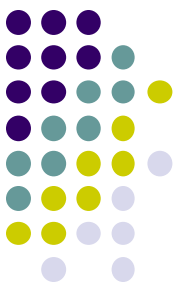
# 本章学习目的

- 操作系统整体作用
  - 管理计算机资源并为用户提供服务的系统软件
  - 是硬件与应用软件之间的接口
  - 操作系统起着承上启下的作用
- 本章学习目的
  - 了解操作系统与硬件之间的相互关系有助于理解计算机系统的整体工作过程
  - 本章仅就操作系统与硬件关系最密切的处理机调度问题



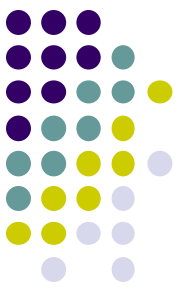
# 操作系统概念

- 计算机系统大体上可以分为三个部分
  - 硬件
  - 系统软件
  - 应用软件
- 操作系统（简称OS）是最重要的系统软件，是管理计算机系统资源、控制程序执行的系统软件
- 操作系统作为计算机用户与计算机硬件之间的接口程序，向用户和应用软件提供各种服务，合理组织计算机工作流程，并为用户使用计算机提供良好运行环境
- 操作系统是现代计算机系统不可分割的重要组成部分



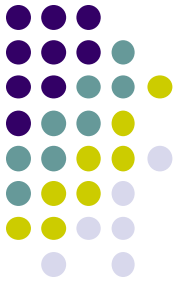
# 操作系统主要目标 (1)

- 操作系统依托计算机硬件并在其基础上提供许多新的服务和功能，从而使用户能够方便、可靠、安全、高效地操纵计算机硬件并运行应用程序
  - **管理系统资源**
    - 操作系统能有效管理系统中的所有硬件资源和软件资源，使资源得到充分利用
  - **提高系统效率**
    - 操作系统能合理地组织计算机的工作流程，改进系统性能，提高系统效率



# 操作系统主要目标 (2)

- 方便用户使用
  - 通过向用户提供友好的用户界面，操作系统能让用户更方便、更轻松地使用计算机系统。
- 增强机器功能
  - 操作系统能通过扩充改造硬件部件并提供新的服务来增强机器功能。
- 构筑开放环境
  - 操作系统通过遵循相关技术标准的方式支持体系结构的可伸缩性和可扩展性，支持应用程序在不同平台上的移植和互操作。



# 第九章 操作系统支持

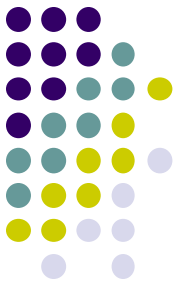
- 操作系统概述
  - 操作系统概念
  - 操作系统功能
  - 操作系统特性与硬件环境
- 调度



# 操作系统的核心任务

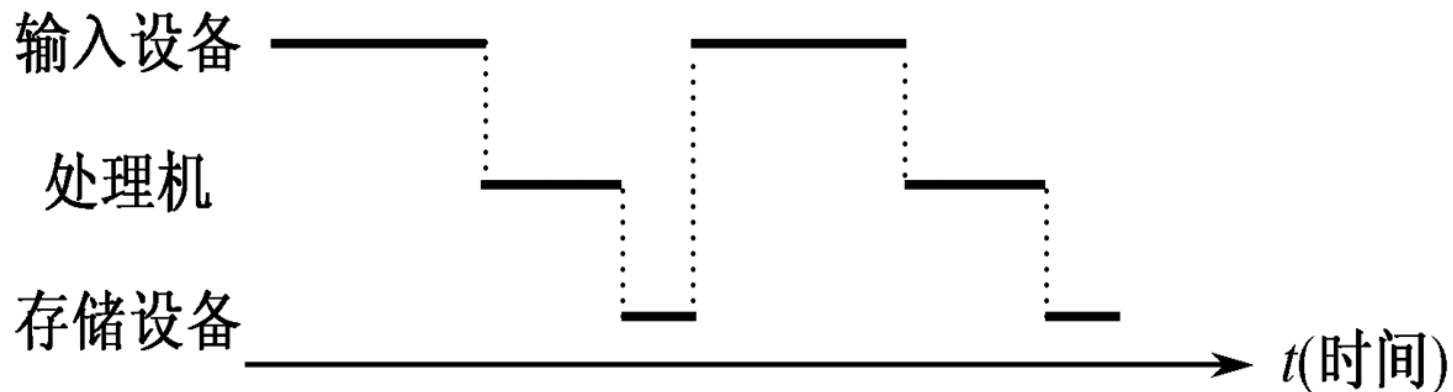
- 操作系统的核心任务
  - 管理计算机系统中的资源
- 从资源管理的角度来看，作为资源管理器的操作系统对计算机硬件资源的管理主要体现在以下三个方面：
  - 处理器管理
  - 存储器管理
  - 设备管理

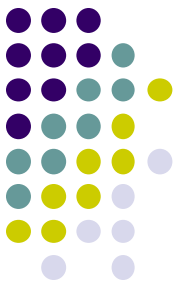




# 处理机管理——单任务系统

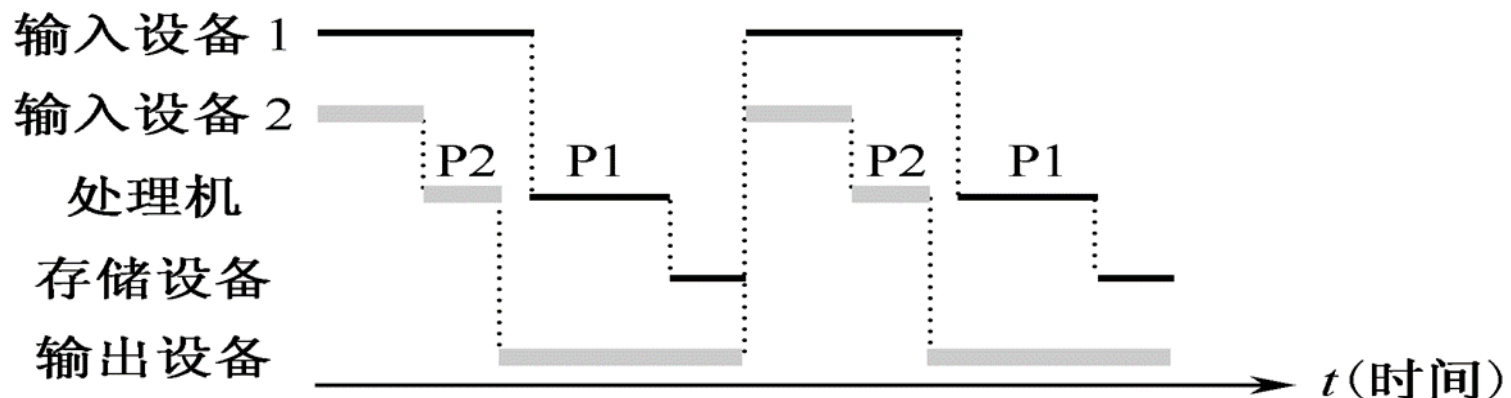
- 处理机是最重要的硬件资源，因为所有软件的执行和系统功能的实现都 依赖于处理机
- 能否充分发挥处理机的效能，是系统功能和性能的关键
- 如下图所示，早期的计算机系统是单用户、单任务系统，处理机仅为一个用户的一个任务服务
  - 特点：串行使用计算机各设备，效率较低

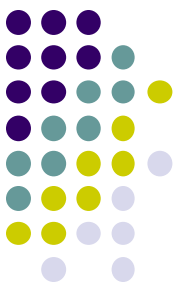




# 处理机管理——多任务系统

- 现代的计算机系统广泛采用多任务机制支持多个程序或多个用户并行使用计算机
- 在引入多任务机制后，当一个作业需等待I/O操作等外部条件满足时，处理机转去执行另一个作业，从而实现多任务的并行执行
- 操作系统负责组织多个任务的并行执行，并负责解决处理机调度、进程控制、进程同步等问题





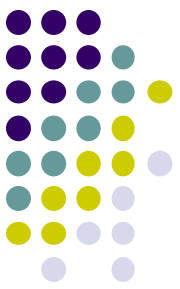
# 存储管理功能

- 存储分配

- 根据程序的需要为其分配存储器资源，在方便存储器使用的同时又要保证存储器的高利用率
- 例如，地址映射技术

- 存储共享

- 允许主存中的多个任务或多个用户程序共享存储器资源，这一方面可以提高存储器的利用率，另一方面又便于多任务间的数据交换
- 例如，线程之间数据共享与互锁



# 存储管理功能

- 存储保护

- 确保用户程序不会有意或无意地访问或破坏操作系统的关键代码和数据
- 各个用户程序之间也需要相互隔离、互不干扰
- 例如，进程之间各自独立一块内存

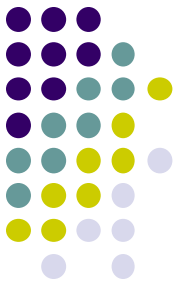
- 存储扩充

- 基于存储器的层次结构，存储管理需要为用户提供与实际物理内存空间不直接相关的逻辑编程空间，并在主存和辅助存储器的支持下实现逻辑地址空间与物理地址空间之间的映射与变换，方便用户的编程和使用
- 例如，虚拟存储器技术

# 设备管理

- 设备管理
  - 设备管理的主要任务是进行各类外围设备的调度与管理，协调各个用户提出的I/O请求，提高各I/O设备操作与处理机运行的并行性，提高处理机和I/O设备的利用率。
  - 设备管理还需提供每种设备的设备驱动程序，向用户屏蔽硬件使用细节。





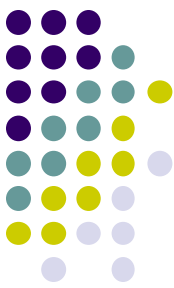
# 第九章 操作系统支持

- 操作系统概述
  - 操作系统概念
  - 操作系统功能
  - 操作系统特性与硬件环境
- 调度



# 操作系统特性

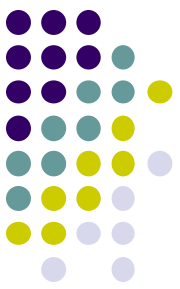
- 操作系统作为计算机系统的管理者，必须解决一系列复杂的管理问题，主要特性包括
- 并发性
  - 为了提高系统资源利用率，多任务系统采用并发技术消除计算机系统中部件和部件之间的相互等待
    - 两个或两个以上的程序可以在同一时间间隔内同时执行，设备的输入输出操作和处理机执行程序同时进行
- 共享性
  - 多个并发执行的程序需要共享系统中的硬件资源和信息资源
- 随机性
  - 用户操作是随机的，程序运行发生错误或异常的时刻是随机的，外部中断事件发生的时刻是随机的.....



# 操作系统硬件环境

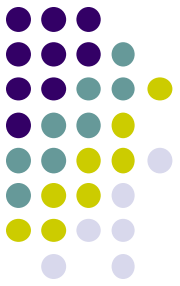
- 操作系统的管理功能只有在专门的硬件支持下才能充分保证系统工作的高效与安全。计算机底层硬件必须为操作系统提供支持
  - **处理机状态控制**
    - 为了支持操作系统，中央处理机需要知道当前执行的程序是操作系统代码还是一般用户程序代码。处理机中设置了状态标志。
  - **特权指令**
    - 特权指令是只能由操作系统核心程序执行的机器指令，用于系统资源管理与程序执行控制操作，如启动输入输出设备、设置系统时钟、控制中断屏蔽位、设置存储管理状态、加载程序状态字等。
  - **寄存器访问权限**
    - 暂存数据的通用数据寄存器和用于存放处理器的控制和状态信息的控制寄存器。





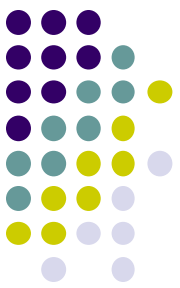
# 操作系统硬件环境

- **程序状态字和程序执行现场**
  - 为了记录计算机系统当前的工作状态，需要专门设置程序状态字（PSW）用于控制指令的执行并存储与程序有关的系统状态。
  - 例如，算术运算指令运行结果是否为零、溢出等
- **中断机制**
  - 通过响应硬件定时器中断，操作系统可以执行周期性的例行管理任务，例如进程调度。这可以确保某个进程不会独占系统资源。以中断方式实现处理机与外界进行信息交换的握手联络，能保证CPU与外设的并行工作。
- **存储管理**
  - 系统硬件通过存储管理部件支持操作系统实现多级存储体系和存储保护功能。



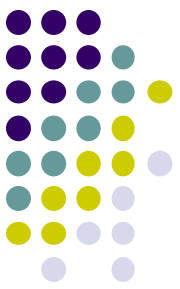
# 第九章 操作系统支持

- 操作系统概述
- 调度
  - 进程
  - 调度的层次
  - 处理机调度的实现



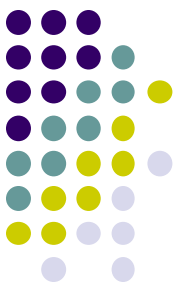
# 进程的概念

- 在操作系统中，通常使用**进程(process)**这一概念描述程序的动态执行过程。
  - **程序是静态实体**；
  - **进程是动态实体，是执行中的程序。**
    - 进程不仅仅包含程序代码，也包含了当前的状态（这由程序计数器和处理机中的相关寄存器表示）和资源。因此，如果两个用户用同样一段代码分别执行相同功能的程序，那么其中的每一个都是一个独立的进程。虽然其代码是相同的，但是数据却未必相同。



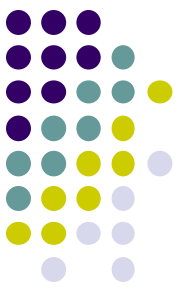
# 进程的状态

- 进程在不同阶段会处于不同状态。进程的状态会随着进程当前进行的活动而改变，包括
  - 创建：正在被创建，并未准备好运行
  - 就绪：已经准备好运行，并正等待分配处理器时间
  - 运行：正占用处理机执行
  - 阻塞：等待I/O操作完成或某些事件出现而被系统挂起
  - 终止：完成操作结束运行
- 在不同的系统中，进程的状态种类和名称不尽相同
- 系统中可能有多个进程处于创建、就绪、阻塞和终止状态



# 进程控制块

- 为了管理和控制进程，操作系统为每个进程设置一个进程控制块PCB（Process Control Block）。记录所需要的用于描述进程情况及控制进程运行所需的全部信息：
  - 进程标识符：用于唯一标识当前进程
  - 状态：记录进程当前状态
  - 程序计数器：将要执行的下一条指令的地址
  - 进程上下文：进程执行时CPU内部寄存器的取值，是进程执行的现场数据
  - 存储管理信息：指明进程占用的内存空间的范围
  - I/O状态信息：例如打开的文件、未完成的I/O请求、占用的I/O设备等



# 进程调度——非抢占模式

- 通常将进程调度分为抢占式和非抢占式两种模式
- 非抢占模式
  - 指一旦将处理机时间分配给某个进程后，便让该进程一直运行，直到该进程因运行完毕或因发生某事件而被阻塞，才把处理机时间重新分配给其它进程。
  - 先来先服务（FIFO）策略就是一种非抢占调度模式，先进入就绪队列的进程首先运行，直到运行结束或被阻塞为止。
  - 优点：实现简单、系统开销小
  - 缺点：无法满足实时系统对紧急事件处理的时间要求



# 进程调度——抢占模式

- 抢占模式

- 允许进程调度程序根据某种策略，暂停某个正在运行的进程，将处理机时间重新分配给另一个进程

- 时间片策略

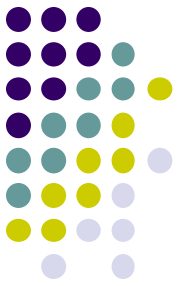
- 处理机被分割为等长的时间单位（时间片）。当该时间片超时，重新进行进程调度，将处理机交给另一个进程

- 优先权策略

- 操作系统为某些重要或紧急的进程指定高优先级，暂停正在运行的低优先级进程，将处理机分配给优先级高的进程

- 短进程优先策略

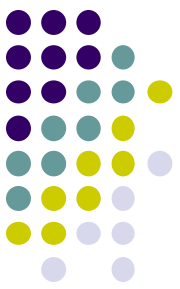
- 当就绪队列中的某个进程比正在运行的进程的运行时间明显短，操作系统将剥夺长进程的执行



# 第九章 操作系统支持

- 操作系统概述
- 调度
  - 进程
  - 调度的层次
  - 处理机调度的实现

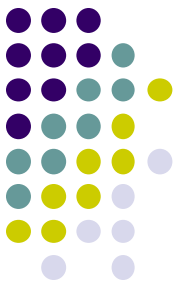




# 高级调度

调度的层次 {  
    高级调度：作业调度  
    中级调度：内存调度  
    低级调度：进程调度

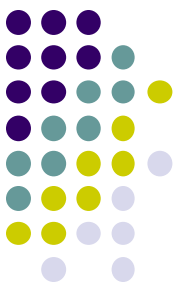
- **高级调度**（管理是否创建进程）
  - 按一定原则把辅存上处于后备队列中的作业调入内存，并为它们创建进程、分配必要的资源，再将新创建的进程排在就绪队列上准备执行
  - 高级调度决定哪些作业可以进入系统竞争系统资源
  - **在较大的粒度上**决定对处理机时间的使用权分配，故又称**作业调度、收容调度或长程调度**



# 中级调度

调度的层次 { 高级调度：作业调度  
中级调度：内存调度  
低级调度：进程调度

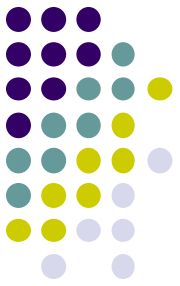
- 中级调度（管理进程是否就绪/调入内存）
  - 为了提高内存的利用率和系统吞吐量。让那些因为某些原因暂时不能运行的进程不再占用宝贵的内存资源，操作系统通过中级调度将这些进程调出至辅存等待
  - 当这些处于挂起状态的进程再次准备好运行，由中级调度决定将辅存上的哪些处于就绪驻外存状态的进程重新调入内存（又称内存调度、中时间粒度）



# 中级调度

调度的层次 { 高级调度：作业调度  
中级调度：内存调度  
低级调度：进程调度

- 低级调度（管理进程竞争处理机资源的分配）
  - 决定当存在多个就绪进程时，哪一个就绪进程将分配到中央处理机的运行时间
  - 在较小的粒度上决定对处理机时间的使用权分配，又称为短程调度。低级调度执行非常频繁，进程调度程序每秒钟可执行多次进程调度，故进程调度程序必须常驻内存。



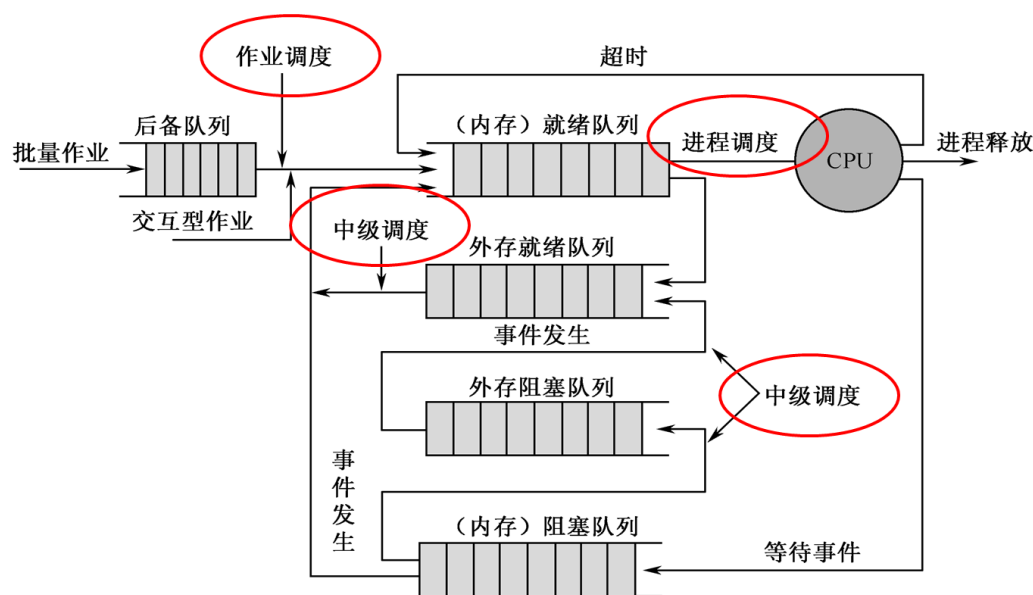
# 第九章 操作系统支持

- 操作系统概述
- 调度
  - 进程
  - 调度的层次
  - 处理机调度的实现

# 处理机调度队列



- 为了调度的方便，操作系统会建立并维护若干个进程队列。每个队列均用于维护一个等待某些资源的进程的列表。
- 在批处理系统中，作业进入系统后，先驻留在辅存的后备队列中，作业调度从辅存的后备队列中选择作业调入内存，并为之创建进程，然后送入内存就绪队列，并等待进程调度。



# 处理机调度队列



- 每个进程运行时都可能在分配给其运行的时间出现以下三种情况：
  - 该进程在分配给其运行的时间内运行完成，**进程释放**
  - 该进程在运行期间因等待某些事件（如资源）而无法继续运行，操作系统将该进程**送入内存阻塞队列**
  - 该进程在分配给其运行的时间超时时尚未运行完成，将该进程**重新排入内存就绪队列**

