## 第1章 绪论

1 操作系统的概念

2 操作系统的发展历史与分类

单道批处理 多道批处理 分时 实时操作系统 通用操作系统

3 操作系统的功能

处理机管理 存储管理 文件管理 设备管理 用户接口

## 第2章 操作系统用户界面

用户与计算机之间的交互：是由用户界面来完成的

用户通过用户界面向计算机系统提交服务需求，计算机通过用户界面向用户提供所需要的服务。

计算机系统的用户有两类：

1、一类是普通用户（使用计算机的应用程序解决实际应用问题）和管理员用户（负责计算机和操作系统的正常与安全运行）

2、另一类是程序开发人员（需要使用操作系统提供的编程功能开发新的应用程序，完成用户所需求的服务）

计算机系统为不同用户所提供的不同界面：

**普通用户和管理员**：由一组以不同形式表示的操作命令组成——命令控制界面

用户利用这些操作命令来组织和控制作业的执行或管理计算机系统

**程序开发人员：**系统调用

操作系统为用户提供的一组系统功能

用户在源程序一级使用这些系统功能

## 第3章 进程管理

3.1 进程的概念

3.1.1 程序的顺序执行和并发执行

**顺序执行的特点：**

顺序性：一个程序开始执行必须要等到前一个程序已执行完成。

封闭性：程序一旦开始执行，程序独占全机资源，其计算结果不受外界因素影响。

可再现性：程序的结果与它的执行速度无关（即与时间无关），只要给定相同的输入，一定会得到相同的结果。

**并发执行：**

* 一组在逻辑上相互独立的程序或程序段在执行过程中，其执行时间在客观上互相重叠，即在计算机系统中同处于已开始执行且尚未结束的状态。
* 程序的并发执行会导致资源共享和资源竞争
* 程序将失去封闭性和可再现性

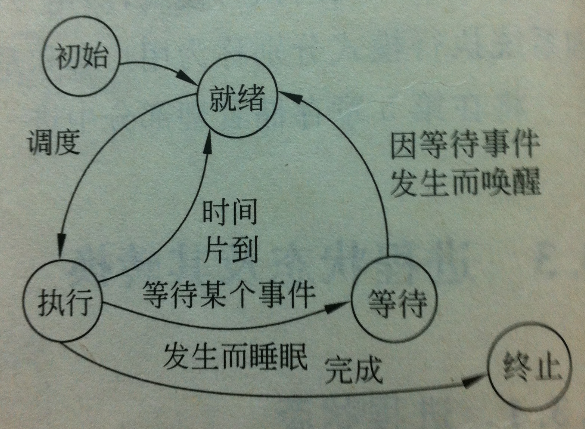
3.1.2 进程的定义

**进程（process）：是一个具有独立功能的程序对某个数据集在处理机上的执行过程和分配资源的基本单位。**

进程的静态描述：由三部分组成,进程控制块、有关程序段和该程序段对其进行操作的数据结构集。

3.3 进程的状态及其转换

进程在生命期内，至少具有5种基本状态：初始态、执行状态、等待状态、就绪状态、终止状态。



3.4 进程控制

创建一个新进程、终止一个已完成的进程，或终止一个因出现某事件而使其无法运行下去的进程，还负责进程运行中状态的转换。

操作系统通过原语进行进程控制：

原语的特点：

a、具有独立的系统功能；

b、在系统态运行；

c、不允许中断或不允许并发。

常用的进程控制原语：

创建原语Create

撤销原语Destroy

阻塞原语Block、

唤醒原语Wakeup

3.5 进程互斥

多道程序系统中，由于资源共享、进程合作，而产生进程之间的相互制约。

★因共享资源的方式不同，而导致进程间两种不同的制约关系

**（1）间接制约（进程互斥）**

由于共享资源而引起的临界区内不允许并发进程交叉执行的现象，称为由共享公有资源而造成的对并发进程执行速度的间接制约

临界资源(critical resource)：一次仅允许一个进程访问的资源。

硬件：输入设备，打印机，磁带机等

软件：公用变量，队列，表格等

临界区(critical section):不允许多个并发进程交叉执行的一段程序称为为临界部分或临界区。（访问公共数据的那段程序）

互斥：不允许两个以上的共享该资源的并发进程同时进入临界区称为互斥。

**（2） 直接制约（进程同步）**

由于并发进程互相共享对方的私有资源所引起的直接制约。相互协作，保证先后顺序。3.5.3 信号量与P,V原语

1. 信号量

一种特殊的变量。它的表面形式是：一个整型变量附加一个队列(等待该信号量的进程队列)。

value：>0时表示空闲临界资源总数，<0时表示等待临界资源的进程数。

信号量只能进行初始化，P，V操作

2. P, V 原语

P、V原语是不可中断的一段程序，它们在屏蔽中断的情况下连续执行。可用来实现并发进程对临界区的互斥访问。

sem是与临界区内所使用的公用资源有关的信号量，sem的值只能由P、V操作来改变，操作系统利用它的状态对进程和资源进行管理。

P 操作

阻塞原语，负责把当前进程由运行状态转换为阻塞状态

(1) sem减1；

(2) 若sem减1后仍大于或等于零，则进程继续执行;

(3) 若sem减l后小于零，

则该进程被阻塞在与 该信号相对应的队列中，然后转进程调度。

V 操作

(1) sem加1；

(2) 若sem加1后大于零，

则进程继续执行；

(3) 若sem减l后小于等于零，则从该信号的等待队列中唤醒一等待进程然后再返回原进程继续 执行或转进程调度

**信号量的物理含义：**

S>0：表示有S个资源可用

S=0：表示无资源可用

S<0：则| S |表示S等待队列中的进程个数

P(S): 表示申请一个资源

V(S)：表示释放一个资源。

信号量的初值应该大于等于0

**P，V操作必须成对出现，有一个P操作就一 定有一个V操作**

当为互斥操作时，它们同处于同一进程

当为同步操作时，则不在同一进程中出现

3.7 进程通信(communication)

进程通信：指进程间的信息交换。

主从式

会话式

消息或邮箱机制

共享存储区方式

**3.8 死锁问题**

死锁的定义：是指各并发进程彼此互相等待对方所拥有的资源，并且这些并发进程在得到对方的资源之前不会释放自己所拥有的资源，造成并发进程都无法向前推进，称这种现象为死锁现象。

★**死锁的原因：资源数少于并发进程要求的数量。**

产生死锁的必要条件

1)互斥条件：在一段时间内，一个资源只能由一个进程独占使用，若别的进程也要求该资源，则须等待直至其占用者释放；

2)不剥夺条件：进程所获得的资源在未使用完之前，不能被其它进程强行剥夺，而只能由其自身释放；

3)部分分配：允许进程在不释放其已分得资源的情况下请求并等待分配新的资源；

4)环路条件：存在一个等待进程集合，P0正在等待一个P1占用的资源，P1正在等待一个P2占用的资源，…，Pn正在等待一个由P0占用的资源。

3.8.2 死锁的排除方法

预防：破坏产生死锁的四个必要条件中的一个或多个，使系统绝不会进入死锁状态；

避免：产生死锁的四个必要条件有可能成立，但在资源动态分配的过程中使用某种办法防止系统进入不安全和死锁状态；

检测与恢复：允许系统产生死锁，然后使用检测算法及时地发现并解除它.

3.9 线程(Thread)

定义：线程是由进程派生出来的一组代码（指令组）的执行过程 ，是一个进程内的基本调度单位。

★特点——线程是由进程派生出来的，一个进程可以产生多个线程。它可以共享进程的内存空间，多线程之间可以并发、异步地执行。

★采用线程的优点——使同一个程序能有几个并发执行路径，提高了执行速度；线程需要的开销比进程小。

练习：

1. 三个进程P1,P2，P3互斥使用一个包含50个单元的缓冲区。P1每次用produce()生成一个正整数，并用put()送入缓冲区单元；P2每次用getOdd()从该缓冲区中取出一个奇数并用Countodd()统计奇数个数；P3每次用getEven()从该缓冲区中取出一个偶数并用CountEven()统计偶数个数。请用信号量机制实现这三个进程的同步与互斥活动，并说明所定义信号量的含义。

2.设系统中有3种类型的资源(A、B、C)和5个进程(P1、P2、P3、P4、P5) A资源的数量为17， B资源的数量为5， C资源的数量为20,T0时刻的系统状态如表所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Process | MAX | Allocation | Need | Available |
| A B C | A B C | A B C | A B C |
| P1 | 5 5 9 | 2 1 2 | 3 4 7 | 2 3 3 |
| P2 | 5 3 6 | 4 0 2 | 1 3 4 |  |
| P3 | 4 0 11 | 4 0 5 | 0 0 6 |  |
| P4 | 4 2 5 | 2 0 4 | 2 2 1 |  |
| P5 | 4 2 4 | 3 1 4 | 1 1 0 |  |

系统采用银行家算法实施死锁避免策略

(1)T0是否为安全状态，若是，请给出安全序列

(2)T0时刻若P2请求资源(0,3,4)能否分配？

(3)若此时P4请求资源(2,0,1)能否分配？为什么？

(4)在(3)的基础上P1请求资源(0,2,0)能否分配？为什么？

第4章 处理机调度  
处理机调度分四个级别

作业调度（高级）：按一定的原则从作业输入井中选择作业，为其创建进程、分配资源，当作业运行完毕后回收作业占用的资源。

交换调度（中级）：按某种策略将处于外存交换区的就绪进程调入内存、把内存中就绪状态或等待状态的进程调出内存。

进程调度（低级）：按某种策略选择一个就绪进程占用处理机，在确定了占用处理机的进程后，必须进行进程上下文切换，以便为运行进程准备好环境。

线程调度（微级）：负责各个线程的调度。

衡量指标：周转时间 平均周转时间

带权周转时间Wi = Ti / Tri

**调度算法：**

**先来先服务调度算法(FCFS)**

**时间片轮转算法(RR—Round Robin)**

**最短作业优先法（SJF,shottest job first）**

**高响应比优先调度算法(HRN，Highest Response\_ratio Next )**

**响应比R = 1 +（作业等待时间/ 作业执行时间）**

练习：

1.作业 提交时间 运行时间

1 8.00 2.50

2 8.20 1.20

3 8.30 0.30

4 9.00 0.50

用先来先服务和短作业优先调度算法进行调度，计算在每一种调度算法下的平均周转时间和平均带权周转时间。并说明哪一种算法的调度性能更好些。

2.使用轮转法实现处理机调度。现有4个进程：A、B、C、D,它们在系统中出现的时间相同，它们的运行时间分别为8s、6s、10s、4s,若现在规定时间片长度为2s,轮转的顺序按照进程名称的字母顺序执行。则4个进程工作结束后，它们的平均周转时间是多少？

## 第5章 存储管理

1. 基本概念：

什么是虚拟存储器？

什么是地址映射（地址重定位）？

2. 分区存储管理

整个进程作为一个整体进入内存连续空间。

* 固定分区存储管理
* 动态分区存储管理

最先适应法

最佳适应法

最差（坏）适应法

3. 页式存储管理

解决分区存储管理时的碎片问题

虚拟空间（进程空间）划分成页（page）

物理空间（内存空间）划分成块（页面，页框，帧frame）

* + 虚拟地址到物理地址的变换（注意十六进制地址转换的法则）
  + 静态页式存储管理
  + 动态页式存储管理

**页面置换（淘汰算法）**

**先进先出页面置换算法（FIFO）**

**最近最久未使用页面置换算法（LRU）**

**理想型淘汰算法（OPT）淘汰在访问序列中再也不出现或者在离当前位置最远的位置上出现的页**

3. 段式与段页式存储管理

分段存储管理的基本思路

分段存储管理内存的分配和释放-----类似分区存储管理

段页式存储管理的基本思路

补充：快表

多级页表

1. 在一个页存储管理系统中，页面大小为4096字节，逻辑地址为1B7C，

并且第0、1、2页分别对应的物理块为5，10，11，

试问该逻辑地址对应的物理地址是多少？

1. 在一个请求分页系统中，有一个长度为 5 页的进程，假如系统为它分配 3 个物理块 ，并且此进程的页面走向为 4，3，4，5，1，4，2，1，3，4，1，4。试用FIFO 和 LRU 算法计算出程序访问过程中所发生的缺页次数。

## 第8章 文件系统

1. 基本概念

什么是文件的逻辑结构？

什么是文件的物理结构？

2. 文件的逻辑结构

无结构的字符流式文件

记录式文件：记录式文件的结构

3. 文件的存取方法

顺序存取

随机存取

按关键字存取

4. 文件的物理结构

连续文件

串联文件

索引文件

5. 文件的存储空间管理

如何管理磁盘上的空闲块

空闲文件目录

空闲块链

位示图法

成组链法

6. 文件目录管理

如何实现的按名存取？

目录的构成？

一级目录 二级目录 多级目录

7. 文件的存取控制

实现文件的保护，共享和保密

存取控制矩阵

存取控制表

口令

密码术

## 第9章 设备管理

1. 设备管理的功能

2. 数据传送方式

控制设备和内存或CPU之间的数据传送

程序直接控制方式

中断方式

DMA方式

通道方式

3. 中断技术

基本概念：

中断源，中断请求，中断响应，中断禁止，中断屏蔽，中断寄存器，中断向量

中断的分类

外中断

内中断

4. 缓冲技术

单缓冲

双缓冲

多缓冲

缓冲池

5. 设备分配

设备分配用的数据结构

设备控制表DCT 系统设备表SDT 控制器表COCT 通道控制表CHCT

设备分配的原则与方式

独占分配

共享分配

虚拟分配：spooling技术