计算机考研系列书课包

# 玩转操作系统

| 主讲人 | 刘财政

# 第六讲 IO系统

# 本讲内容

考点一: IO软件 ★ ★ ★ ★

考点二: IO缓冲 ★ ★ ★ ★

考点三: 假脱机技术 ★★★

考点四: IO方式 ★



"I/O" 就是"输入/输出" (Input/Output)

- ▶ I/O设备就是可以将数据输入到计算机,或者可以接收计算机输出数据的外部设备,属于计算机中的硬件部件。
- > 他是人机交互的主要的界面和部件





# **U** IO设备的分类

- 1) 按传输速率分类
- 第一类是低速设备,这是指其传输速率仅为每秒钟几个字节至数百个字节的一 类设备。属于低速设备的典型设备有键盘、 鼠标器、语音的输入和输出等设备。
- 第二类是中速设备,这是指其传输速率在每秒钟数千个字节至数万个字节的一类设备。典型的中速设备有行式打印机、激光打印机等。
- 第三类是高速设备,这是指其传输速率在数百千个字节至数十兆字节的一类设备。备。 典型的高速设备有磁带机、 磁盘机、 光盘机等。

# U IO设备的分类

- 2) 按信息交换的单位分类
- 》 第一类是块设备(Block Device),这类设备用于存储信息。由于信息的存取总是以数据块为单位,故而得名。它属于有结构设备。典型的块设备是磁盘,每个盘块的大小为512 B~4 KB。
- ➤ 第二类是字符设备(Character Device),用于数据的输入和输出。 其基本单位是字符, 故称为字符设备。

# **U** IO设备的分类

- 3) 按照使用方式
- 独占设备:一段时间内只允许一个用户/进程访问的设备,例如:临界资源必须 互斥访问,且有可能引起进程死锁
- 共享设备:一段时间内允许多个进程同时访问的设备,每一时刻仅允许一个 进程访问,如磁盘为可寻址、可随机访问的设备
- 虚拟设备:通过虚拟技术将一台独占设备变换为若干台逻辑设备;供若干个用户/进程同时使用



# 设备分配与回收





# 常识补充 设备分配与回收

	设备控制表DCT
	控制器控制表COCT
设备分配中的数据结构	通道控制表CHCT
	系统设备表SDT
	逻辑设备表LUT

### 设备分配中的数据结构

- ▶ 设备控制表 DCT:系统为每个设备配一个DCT,记录其状态
- ➤ 控制器控制表 COCT: 系统为每个控制器配一个COCT
- ➤ 通道控制表 CHCT: 系统为每个通道配一个CHCT
- > 系统设备表 SDT: 记录系统全部设备情况的一个表格, 每个设备占一个表目
- 逻辑设备表 LUT: 存放逻辑设备名、物理设备名、设备驱动程序入口地址的表格,用于完成逻辑设备到物理设备的转换

# 设备分配与回收 设备分配中的数据结构

设备类型 type DCT 1 设 设备标识符: deviceid 备 控 设备状态: 等待/不等待 忙/闲 制 DCT 2 指向控制器表的指针 表 集 重复执行次数或时间 合 DCT n 设备队列的队首指针 进程等待

# 设备分配中的数据结构

控制器标识符: controllerid

控制器状态: 忙/闲

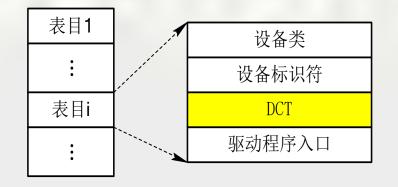
与控制器连接的通道表指针

控制器队列的队首指针

控制器队列的队尾指针

通道标识符: channelid 通道状态: 忙/闲 与通道连接的控制器表首址 通道队列的队首指针 通道队列的队尾指针

(b) 通道表CHCT



(c) 系统设备表SDT

(a) 控制器表COCT



设备分配与回收

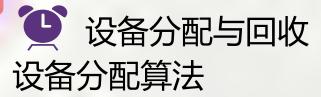
### 设备分配中的需要考虑的因素

□ 设备固有属性: 不同属性采取不同的分配策略

独占设备:得到设备的进程独占,直至完成或其他原因释放设备

> 共享设备:可分配给多个进程,并合理分配各进程的访问顺序

▶ 虚拟设备:共享设备,可将其同时分配给多个进程



#### □ 设备分配算法

- > 先来先服务: 进程按照申请设备的先后顺序, 插入设备请求队列
- > 高优先级优先: 进程按照优先级, 插入设备请求队列



设备分配与回收

设备分配中的需要考虑的因素

□ 设备回收: 释放设备、设备控制器、通道; 修改相应数据结构

□ 安全性

> 安全分配: 进程发出 I/O 请求后进入阻塞状态, 串行工作

> 不安全分配: 进程发出 I/O 请求后仍可运行, 并行工作

#### 逻辑设备名到物理设备名的映射:逻辑设备表 LUT

- □ LUT个数
  - 》 整个系统一个: 不允许具有重名的逻辑设备

逻辑设备名。	物理设备名。	设备驱动程序入口地址。
/dev/console	2.	2000₽
/dev/printer₀	5₽	5000₽
/dev/sd[a-z]	6₽	6000₽
••••	••••	••••

➤ 每个用户一个:适合多用户系统,放入进程 PCB

逻辑设备名。	系统设备表指针
/dev/console	2₽
/dev/printer	5₽
/dev/sd[a-z]	6₽
٠٠٠٠	••••



# **设备分配与回收**

分配的流程,从资源多的到资源紧张的:LUT->SDT->DCT-> COCT->CHCT

设备分配中过程

在申请设备的过程中,根据用户请求的I/O设备的逻辑名,查 找逻辑设备和物理设备的映射表; 以物理设备为索引, 查找 SDT, 找到该设备所连接的DCT; 继续查找与该设备连接的 COCT和CHCT,就找到了一条通路。

显然, 在有通道的系统中, 一个进程只有获得了通道, 控制 器和所需设备三者之后,才具备了进行I/O操作的物理条件

考点一:

IO软件

# 考点框架

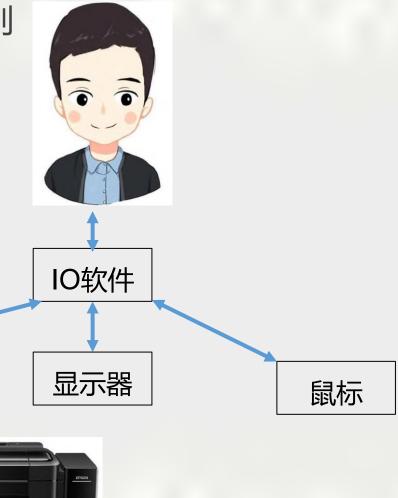


IO软件的设计目标和原则



IO系统的层次结构

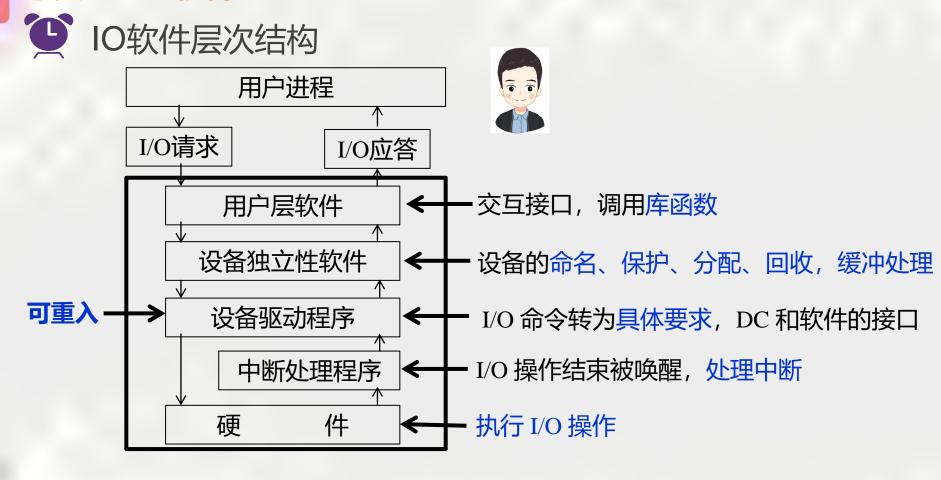
- U/O软件的设计目标和原则
  - > 与具体的设备无关
  - > 统一命名
  - > 对错误的处理
  - > 缓冲技术等



打印机





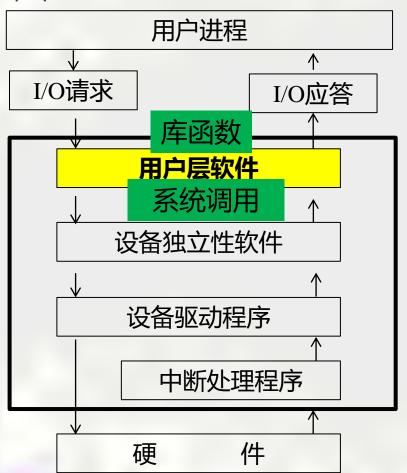




- 用户层软件:实现与用户交互的接口,用户可直接调用在用户层提供的、与I/O操作有关的库函数,对设备进行操作。
- 设备独立性软件:负责实现和设备驱动器的统一接口、设备命名、设备的保护以及设备的分配与释放,同时为设备管理和数据传送提供必要的存储空间。
- 设备驱动程序:与硬件直接相关,负责具体实现系统对设备发出的操作 指令,驱动I/O设备工作的驱动程序。
- 中断处理程序: 用户保存被中断进程的CPU环境, 转入相应的中断处理程序进行处理, 处理完后再恢复被中断进程的现场后返回到被中断进程。



# IO软件层次结构

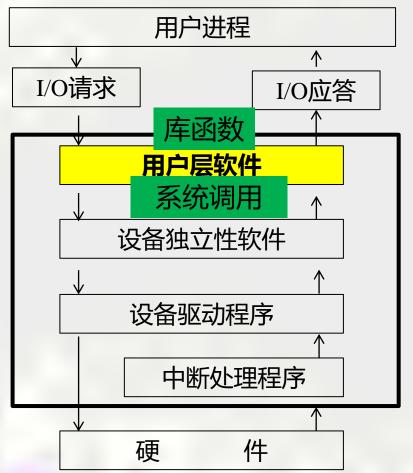


#### 用户层软件

用户层软件:实现与用户交互的接口,用户可直接调用在用户层提供的、与I/O操作有关的库函数,对设备进行操作。



# IO软件层次结构

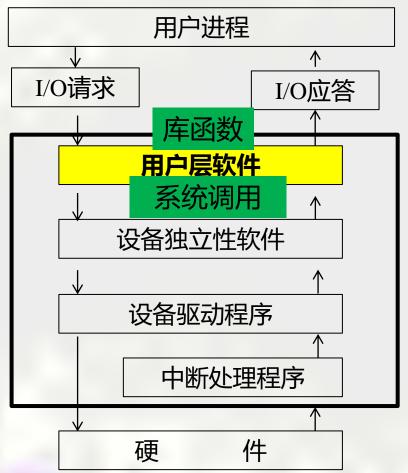


#### 用户层软件

一般而言,大部分的I/O软件都在操作系统内部,但是仍有一小部分在用户层,包括与用户程序链接在一起的库函数,以及完全运行在内核之外的一些程序。用户层软件必须通过一组系统调用来获取操作系统服务。



# IO软件层次结构

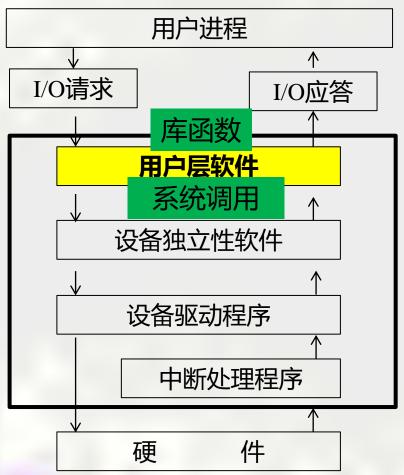


#### 用户层软件

实现与用户交互的接口,用户可直接调用在用户层提供的、与I/O操作有关的库函数,对设备进行操作。用户层软件必须通过各系统调用——对应的库函数。



# IO软件层次结构

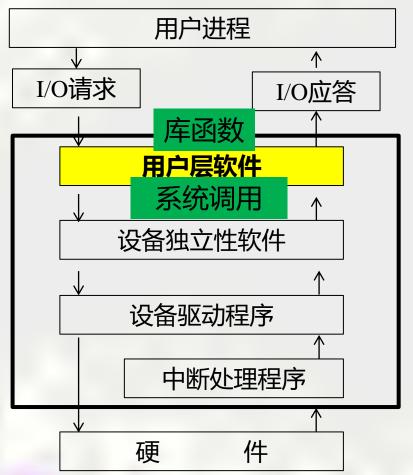


#### 设备无关性软件

为了提高操作系统的可适应性和可扩展性,在现代操作中都毫无例外地实现了设备独立性,也称为设备无关性。



# **U** IO软件层次结构



#### 设备无关性软件



基本含义是: 应用 程序独立于具体使 用的物理设备。



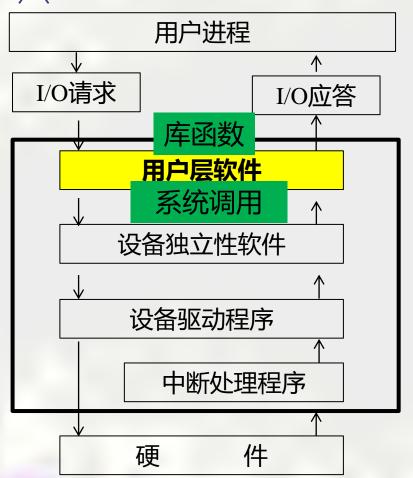
在应用程序中, 使 用逻辑设备名称来 请求使用某类设备;



系统在实际执行时, 还必须使用物理设 备名称。



# IO软件层次结构



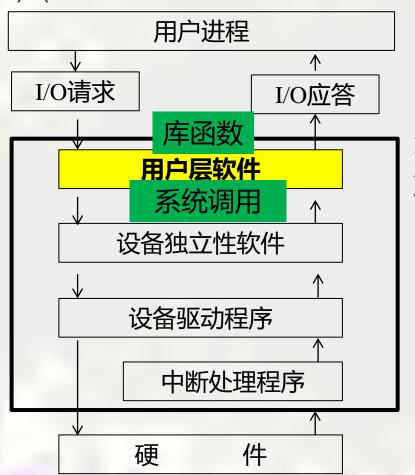
#### 设备无关性软件

系统须具有将逻辑设备名称转换为某物理 设备名称的功能。在实现了设备独立性的 功能后,可带来以下两方面的好处:

- 1) 设备分配时的灵活性
- 2) 易于实现I/O重定向



# IO软件层次结构

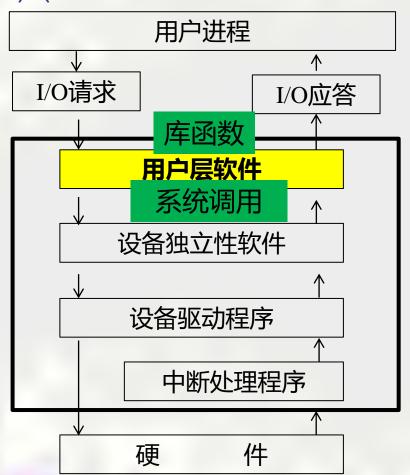


#### 设备无关性软件

为了实现设备独立性,必须再在驱动程序至上设置一层软件,称为设备独立性软件。



# IO软件层次结构



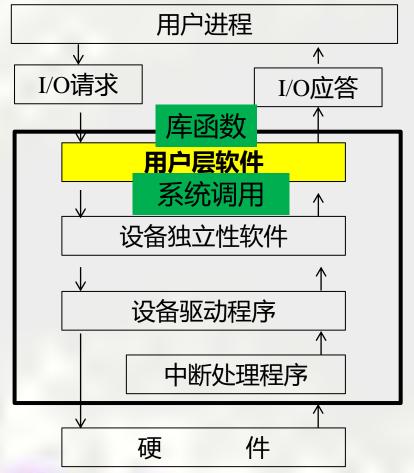
#### 设备无关性软件

设备独立性软件主要功能:

- (1) 执行公共设备的共有操作:
- 1) 对独立设备的分配与回收
- 2) 将逻辑设备名映射为物理设备名,进一步可以找到对应的物理设备的驱动程序
- 3) 对设备进行保护,禁止用户直接访问设备
- 4) 缓冲管理
  - (2) 向用户层软件提供统一接口。



# IO软件层次结构



#### 设备驱动程序

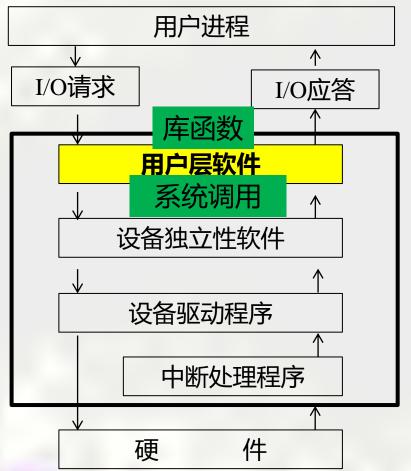
设备驱动程序通常又称为设备处理程序, 它是I/O进程与设备控制器之间的通信程序, 又由于常以进程的形式存在, 故以后就称之为设备驱动程序。







# IO软件层次结构

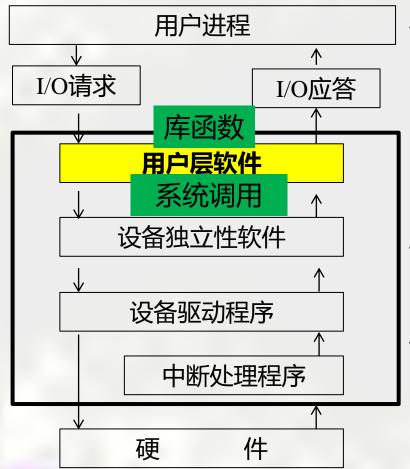


#### 设备驱动程序

其主要的任务是接收上层软件发来的抽象的I/O要求,在把它转换为具体要求后,发送给设备控制器,启动设备去执行;此外它也将设备控制器发来的信号发送给上层软件。



# IO软件层次结构



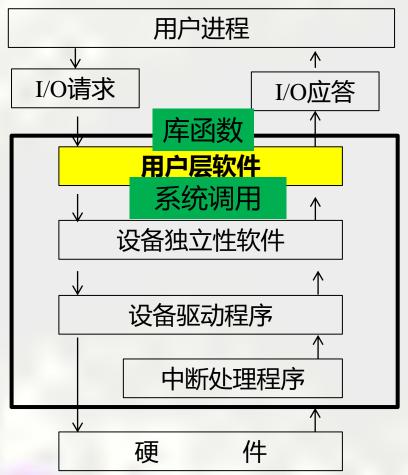
#### 设备驱动程序

设备驱动程序的功能

- (1) 接收由设备独立性软件发来的命令和参数
- (2) 检查用户I/O请求的合法性
- (3) 发出I/O命令
- (4) 及时响应由控制器或者通道发来的中断请求, 并根据其中的中断类型调用相应的中断处理程序
- (5) 对于设置有通道的计算机系统,驱动程序还应能根据用户的I/O的请求,自动构建通道程序
  - (6) 启动IO设备



# IO软件层次结构



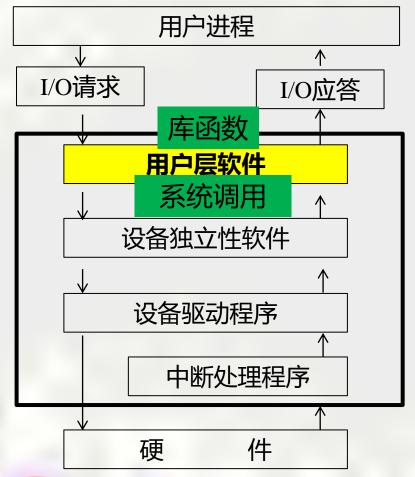
#### 设备驱动程序

#### 设备处理方式

- (1) 为每一类设备设置一个进程,专门用于执行这类设备的I/O操作
- (2) 在整个系统中设置一个I/O进程,专门用于执行系统中所有各类设备的I/O操作
- (3) 不设置专门的处理程序,而只为各类设备设置相应的设备处理程序(模块),供用户进程或系统进程调用。



# IO软件层次结构



#### 设备驱动程序

#### 设备处理方式

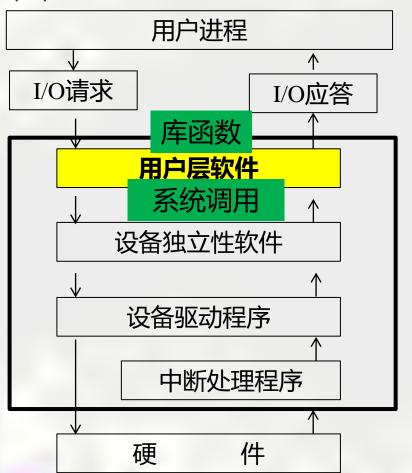
(1) 为每一类设备设置一个进程,专门用于执行这类设备的I/O操作







# IO软件层次结构

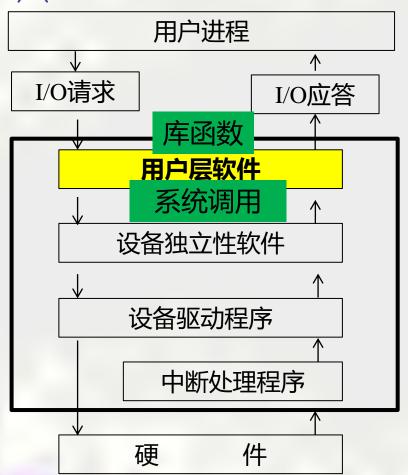


#### 中断处理程序

中断处理层的主要工作有:进行进程上下文的切换,对处理中断信号源进行测试,读取设备状态和修改进程状态。



# IO软件层次结构



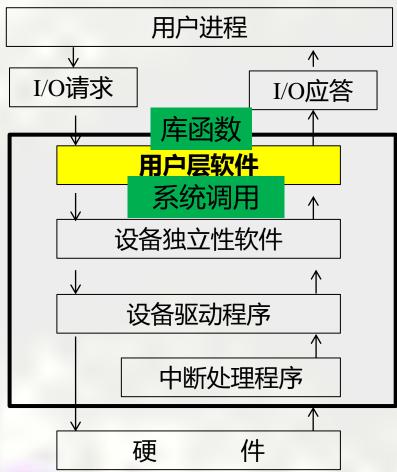
#### 中断处理程序

#### 中断处理程序的处理过程:

- 1、唤醒被阻塞的驱动进程
- 2、保护被中断进程的CPU环境
- 3、载入相应的设备处理程序
- 4、中断处理
- 5、恢复被中断进程的现场

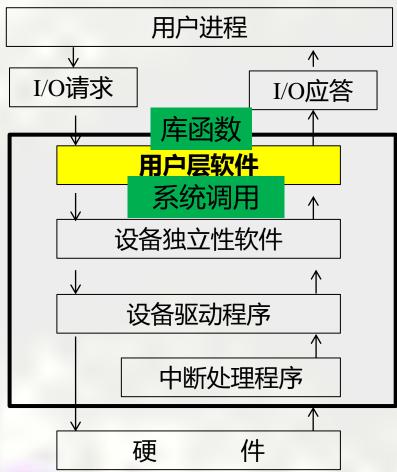


# IO软件层次结构





# IO软件层次结构



【政哥点拨】

1. 一个计算机系统配置了2台绘图机和3台打印机,为了正确驱动这些设备,系统应该提供()个设备驱动程序。

A. 5 B. 3 C. 2 D. 1

C【解析】因为绘图机和打印机属于两种不同类型的设备,系统只要按设备类型 配置设备驱动程序即可,即每类设备只需一个设备驱动程序。

- 2. 将系统调用参数翻译成设备操作命令的工作由()完成。
  - A. 用户层I/O软件 B. 设备无关的操作系统软件
  - C. 中断处理程序 D. 设备驱动程序
- B【解析】系统调用是操作系统提供给用户程序的通用接口,不会因为具体设备的不同而改变。而设备驱动程序负责执行操作系统发出的I/O命令,它因设备不同而不同。

【牛刀小试】

- 1. 【广东工业大学 2019】在单用户系统中可为( )设置一张逻辑设备表,在多 用户系统中应为()设置一张逻辑设备表。

  - A. 整个系统 B. 每个用户(进程)
  - C. 每种逻辑设备 D. 每种物理设备
- A, B【解析】在单用户系统中可为整个系统设置一张逻辑设备表, 在多用户系统 中应为每个用户(进程)设置一张逻辑设备表。因此选择A, B。

- 2. 【南京理工大学 2016】向设备寄存器写命令是在I/O软件的哪一层完成的? ( )
  - A. 用户层软件 B. 设备独立性软件
  - C. 设备驱动程序 D. 中断处理程序
- C【解析】I/O软件一般分为用户层软件、设备独立性软件、设备驱动程序和中断处理程序。向设备寄存器写命令是在设备驱动程序中完成的。因此选择C。

- 3. 【南京理工大学 2017】寻找设备驱动程序接口函数的任务由()完成。
  - A. 用户层I/O软件 B. 设备无关的操作系统软件
  - C. 中断处理程序 D. 设备处理程序
- B【解析】寻找设备驱动程序接口函数的任务由设备无关的操作系统软件完成。 因此选择B。

- 4. 【华东师范大学 2013】以下哪个功能不是由设备驱动程序提供的?()
  - A. 提供中断处理程序
  - B. 提供DMA控制功能
  - C. 提供标准的设备访问系统调用(如open(), read()等)
  - D. 提供内核直接访问设备的接口
- B【解析】设备驱动程序的功能包括:将接收到的抽象要求转换为具体要求;检查用户I/O的合法性,了解设备的状态,传递有关参数,设置设备的工作方式;发出I/O操作命令,启动I/O设备,完成I/O操作;响应通道发来的中断请求,根据中断类型调用相应的中断处理程序;构造通道程序。可见B不是设备驱动程序的功能。因此选择B。

- 5. 【北京交通大学 2017】设备独立性软件层应提供()。
  - A. 缓冲机制
- B. 设备启动功能
- C. 设备驱动程序统一接口 D. 被中断进程的现场保护
- C【解析】设备独立性软件用于实现用户程序与设备驱动器的统一接口、设备命名、设备的保护以及设备的分配与释放等,同时为设备管理和数据传送提供必要的存储空间。因此选择C。

- 6. 【北京交通大学 2017】 (多选) 一般而言,设备驱动程序的功能应包括()。
- A. 检查用户输入输出请求合法性,了解设备工作状态并传递操作控制参数和设置工作方式
- B. 接收来自设备无关性软件层的命令和参数,并转化为与设备相关的低级操作序列
- C. 根据设备状态阻塞请求者进程或发出输入输出命令启动设备
- D. 及时响应和处理源自设备控制器的中断请求

ABCD【解析】为了实现I/O系统的高层与设备控制器之间的通信,设备驱动程序 应具有以下功能:接收来自设备无关的软件发来的命令和参数,并将命令中的抽象要求转换为与设备相关的低层操作序列;检查用户I/O请求的合法性;发出I/O命令;及时响应由设备控制器发来的中断请求。

- 【安徽大学 2014】在计算机系统中,通常把财务管理程序看作是()。

- A. 系统软件 B. 支援软件 C. 接口软件 D. 应用软件
- D【解析】在计算机系统中,通常把财务管理程序看作是应用软件。因此选择 D.

- 8. I/O中断是CPU与通道协调工作的一种手段, 所以在( )时, 便要产生中断。
  - A. CPU执行"启动I/O"指令而被通道拒绝接收
  - B. 通道接收了CPU的启动请求
  - C. 通道完成了通道程序的执行
  - D. 通道在执行通道程序的过程中
- C【解析】CPU启动通道时不管启动成功与否,通道都要回答CPU,通道在执行通道程序的过程中,CPU与通道并行执行,当通道完成了通道程序的执行,便产生I/O中断向CPU报告。因此选择C。

- 9. 设备的打开/关闭、读/写等操作是由()完成的。
  - A. 用户程序
- B. 编译程序
  - C. 设备分配程序 D. 设备驱动程序
- D【解析】设备驱动程序是I/O进程与设备控制器之间的通信程序,主要任务是接 收上层软件发送来的抽象要求,转化为具体要求,发送给设备控制器,启动设备 去执行。设备的打开/关闭、读/写等操作都是由设备驱动程序完成的。

- 10. 对设备映射表来说,不能实现具体的功能。下列关于设备独立性的叙述中,正确的是()。
  - A. 设备独立性是指I/O设备具有独立执行I/O功能的一种特性
  - B. 设备独立性是指用户程序独立于具体使用的物理设备的一种特性
  - C. 设备独立性是指独立实现设备共享的一种特性
  - D. 设备独立性是指设备驱动独立于具体使用的物理设备的一种特性
- B【解析】设备独立性是指用户不指定特定的设备,而指定逻辑设备,使得用户作业和物理设备独立开来,再通过其他途径建立逻辑设备和物理设备之间的对应关系的特性,即用户程序独立于具体使用的物理设备的一种特性。

- 11. I/O软件一般分为用户层软件、设备独立性软件、设备驱动程序和中断处理程序4个层次。下面的工作在设备驱动程序层完成的是()。
  - A. 向设备寄存器写命令
  - B. 检查用户是否有权使用设备
  - C. 将二进制整数转换成ASCII的格式打印
  - D. 缓冲管理

A【解析】(1)向设备寄存器写命令是在设备驱动程序中完成的;(2)检查用户是否有权使用设备属于设备保护,是在设备独立性软件中完成的;(3)将二进制整数转换成ASCII码的格式打印是通过I/O库函数完成的,因此属于用户层软件;(4)缓冲管理属于I/O的公有操作,是在设备独立性软件中完成的。因此选择A。

谢谢大家

考点二:

IO缓冲

# 考点框架



IO缓冲的引入



单缓冲



双缓冲



IO控制方式指的是CPU实现与IO设备进行数据交换的方式

CPU和外设之间交换数据,实质上是通过IO端口进行的

CPU ← IO接口/端口 ← IO设备

常用的IO控制方式 DMA产

程序查询方式程序中断方式DMA方式I/O通道控制方式IO处理机

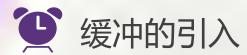


# 缓冲的引入

- ▶ 减少 CPU 被中断的频率
- ▶ 放宽CPU的中断响应时间限制
- ▶ 提高 CPU 和 I/O设备 之间的并行性
- ➤ 缓解 CPU 与 I/O 设备之间的速度矛盾



如果是采用中断IO方式,则每输出完一 个字符就要向CPU发送一次中断信号



- ▶ 减少 CPU 被中断的频率
- ➤ 放宽CPU的中断响应时间限制
- ➤ 提高 CPU 和 I/O设备 之间的并行性
- ➤ 缓解 CPU 与 I/O 设备之间的速度矛盾



CPU可以把要输出的数据快速地读出缓冲区,之后就可以做后续的处理

慢速的I/O设备可以慢慢从缓冲区取 走数据



- 》 缓冲区是一个存储区域,可以由专门的硬件寄存器组成,也可利用内存作为缓冲区。
- 一般情况下,更多的是利用内存作为缓冲区,"设备独立性软件"的缓冲区管理就是要组织管理好这些缓冲区



CPU可以把要输出的数据快速地读出缓冲区,之后就可以做后续的处理

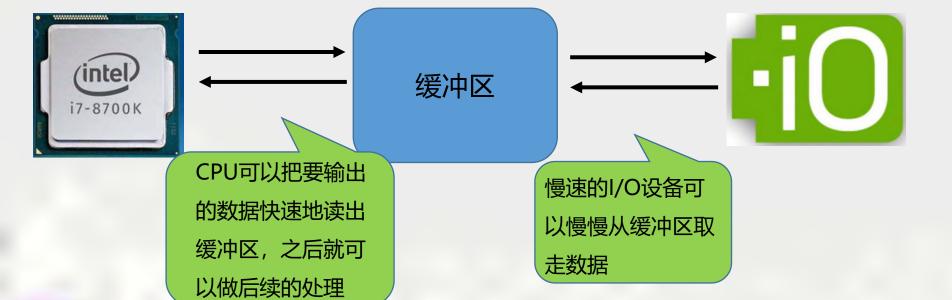
慢速的I/O设备可以慢慢从缓冲区取 走数据

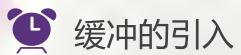


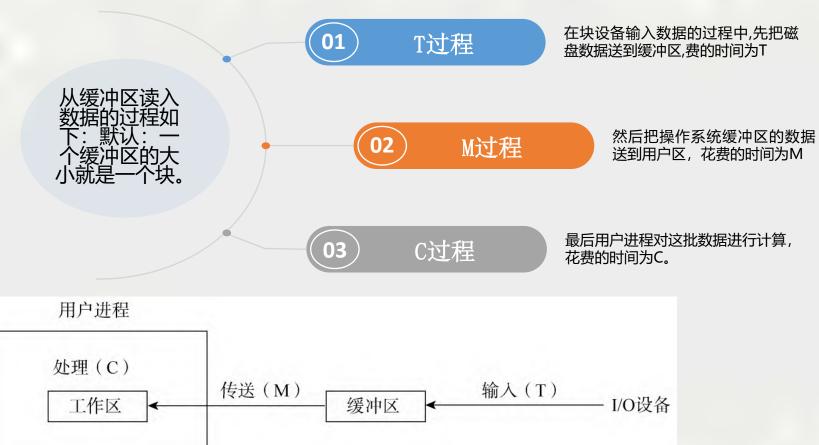
# 货 缓冲的引入

当缓冲区为空时,可以往缓冲区冲入 数据,但必须把缓冲区充满以后,才 能从缓冲区把数据传出。

当缓冲区数据非空时,不能往缓冲区 冲入数据,只能从缓冲区把数据传出;



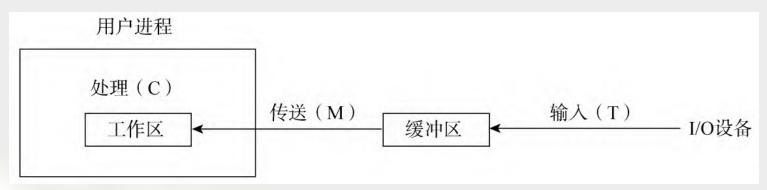






# 单缓冲

单缓冲情况下,每当用户进程发出I/O请求时,操作系统便在主存中为之分配一 个缓冲区。在块设备输入时,假定从磁盘把一块数据输入到缓冲区的时间为T, 操 作系统将该缓冲区中的数据传送到用户区的时间为M,而CPU对这一块数据的处 理时间为C, T和C是可以并行的,





# 单缓冲

01

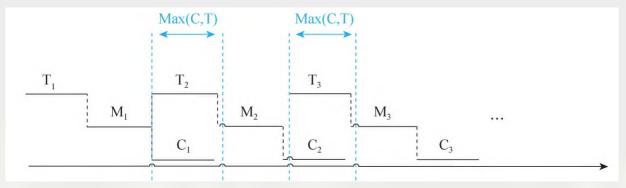
假如有多个数据需要传送,T 过程和M过程不能并行,C过 程和M过程不能并行。大致过 程: 如图所示, 先读入磁盘块 T1,然后送至用户区M1。M1 执行完成以后, 处理磁盘块 C1和读入下一个磁盘块T2可 以同时执行。

02

如果T>C(即T2>C1),第一个磁 盘块处理完成后,还需要等待 第二个磁盘块的读入, 所以这 段时间用时为T2。

03

如果T<C(即T2<C1),第二个磁 盘读入后,还需要第一个磁盘 处理完成以后再传送到用户工 作区,所以这段时间用时为C1。



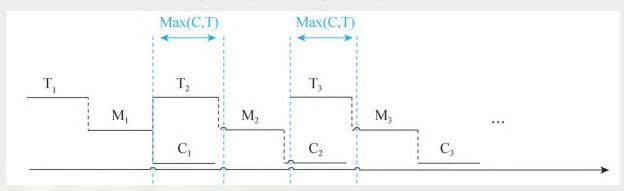


根据上述过程,如图

□ 当T > C 时, T过程和M过程不能并行, 如果有n个数据块要传送, 因此

总时间Time = n \* (T + M) + C

平均时间 = (n \* (T + M) + C) / n ≈ T + M

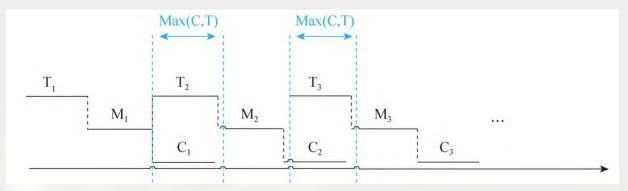




根据上述过程,如图

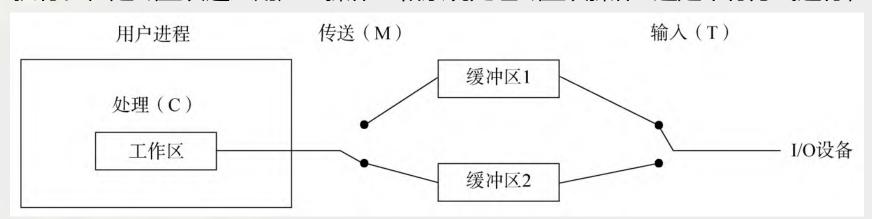
□ 当T < C 时, T过程和M过程不能并行, 如果有n个数据块要传送, 因此

□ 综上所述, 单缓冲区的平均时间是max(T,C) + M。





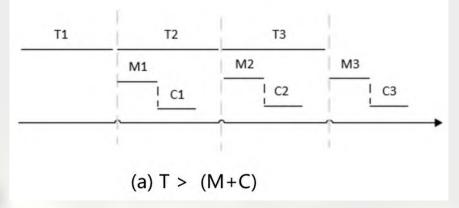
双缓冲区方式,是先将第一个缓冲区填满,紧接着填满第二个缓冲区。不同于单缓冲区,双缓冲区将磁盘块读入缓冲区操作T和磁盘块送入用户区操作M可以并行执行。但是磁盘块送入用户区操作M和系统处理磁盘块操作C还是串行方式进行,





双缓冲区方式,是先将第一个缓冲区填满,紧接着填满第二个缓冲区,根据上 述过程,如果有n个数据块要传送。

- □ T > (M+C),可以采用二级流水线的工作模式,
- □ 总时间time = n \* T + M + C
- 平均时间是 = (n\*T+M+C) / n ≈ T





# 双缓冲

双缓冲区方式,是先将第一个缓冲区填满,紧接着填满第二个缓冲区,根据上 述过程,如果有n个数据块要传送。

- □ T < (M+C),可以采用二级流水线的工作模式,
- □ 总时间time = T + n \* (M + C)
- □ 平均时间是 = (T + n \* (M + C)) / n ≈ M + C

综上所述,双缓冲区的平均时间是max(T,C+M)。



# 雙 缓冲区和缓存的对比

	高速缓存Cache	缓冲区
位置	主存和CPU之间	CPU和IO设备之间 (高速设备和低速之间)
原理	缓解主存和CPU之间速度的不	缓解高速设备和低速设备之间速度不匹配,
	匹配,存放经常访问的副本	用于暂存高速设备和低速设备之间的数据
数据	高速设备Cache中存放的只是	缓冲中存放的数据孤本,缓冲中存在,但
特点	副本	是在低速设备中可能不存在
特性	用于提高CPU访问内存的效率,	较少IO对CPU的干预,提高系统并行性;
	Cache的核心作用是加快取用	Buffer的核心作用是用来缓冲,缓和冲击。
	的速度	

【政哥点拨】

1. 设从磁盘将一块数据传送到缓冲区所用的时间为80µs,将缓冲区中的数据传送到用户区所用的时间为40µs, CPU处理一块数据所用的时间为30µs。若有多块数据需要处理,并采用单缓冲区传送某磁盘数据,则处理一块数据所用的总时间为()。

A. 120μs B. 110μs C. 150μs D. 70μs

A【解析】采用单缓冲区传送数据时,设备与处理机对缓冲区的操作是串行的。当进行第i次将磁盘数据送至缓冲区时,系统再同时读出用户区中第(i-1)次数据进行计算,此两项操作可以并行,并与数据从缓冲区传送到用户区的操作串行进行。所以系统处理一块数据所用的总时间为Max(80μs, 30μs)+40μs=120μs。

2. 某操作系统采用双缓冲区传送磁盘上的数据。设从磁盘将数据传送到缓冲区所 用的时间为T<sub>1</sub>,将缓冲区中的数据传送到用户区所用的时间为T<sub>2</sub>(假设T<sub>2</sub>远小于  $T_1$ ),CPU处理数据所用的时间为 $T_3$ ,则处理该数据,系统所用的总时间为()。

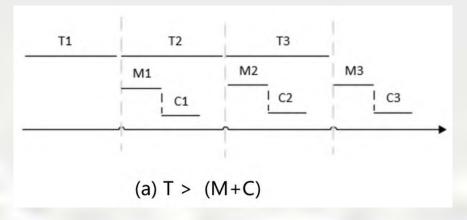
A. 
$$T_1 + T_2 + T_3$$

A. 
$$T_1+T_2+T_3$$
 B.  $Max(T_2, T_3)+T_1$ 

C. 
$$Max(T_1,T_3)+T_2$$
 D.  $Max(T_1,T_2)$ 

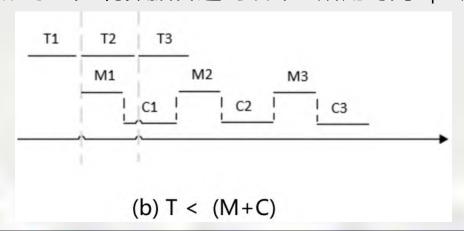
D. Max(
$$T_1,T_2$$
)

- D【解析】处理该数据所用的总时间,即可以默认初始状态缓冲区1已将数据传送 到用户区。然后分情况讨论:
- □ 若T<sub>3</sub>>T<sub>1</sub>, 即CPU处理数据块比数据传送慢, 意味着I/O设备可连续输入, 磁盘 将数据传送到缓冲区,再传送到用户区,与CPU处理数据可视为并行处理,时 间的花费取决于CPU最大花费时间,则系统所用总时间为T3;



2. 某操作系统采用双缓冲区传送磁盘上的数据。设从磁盘将数据传送到缓冲区所 用的时间为T<sub>1</sub>,将缓冲区中的数据传送到用户区所用的时间为T<sub>2</sub>(假设T<sub>2</sub>远小于  $T_1$ ),CPU处理数据所用的时间为 $T_3$ ,则处理该数据,系统所用的总时间为()。

- A.  $T_1+T_2+T_3$  B. Max $(T_2, T_3)+T_1$
- C.  $Max(T_1,T_3)+T_2$  D.  $Max(T_1,T_2)$
- D【解析】处理该数据所用的总时间,即可以默认初始状态缓冲区1已将数据传送 到用户区。然后分情况讨论:
- □ 若T<sub>3</sub><T<sub>1</sub>,即CPU处理数据块比数据传送快,此时CPU不必等待I/O设备,磁盘 将数据传送到缓冲区,与缓冲区中数据传送到用户区及CPU处理数据可视为并 行执行,则花费时间取决于磁盘将数据传送到缓冲区所用时间T<sub>1</sub>。所以选D。



3. 假设磁盘块与缓冲区大小相同,每个盘块读入缓冲区的时间为10µs,由缓冲区送至用户区的时间是5µs,系统对每个磁盘块数据的处理时间为2µs。若用户需要将大小为10个磁盘块的Doc1文件逐块从磁盘读入缓冲区,并送至用户区进行处理,那么采用单缓冲区需要花费的时间为多少µs?

采用双缓冲区需要花费的时间为多少µs?

【解析】单缓冲区: (10+5)×10+2=152;

 10us
 5us

 10us
 5us

3. 假设磁盘块与缓冲区大小相同,每个盘块读入缓冲区的时间为10µs,由缓冲区送至用户区的时间是5µs,系统对每个磁盘块数据的处理时间为2µs。若用户需要将大小为10个磁盘块的Doc1文件逐块从磁盘读入缓冲区,并送至用户区进行处理,那么采用单缓冲区需要花费的时间为多少µs?

采用双缓冲区需要花费的时间为多少µs?

【解析】双缓冲区: 10×10+5+2=107。

10us 5us 2us

【牛刀小试】

- 1. 【广东工业大学 2014】操作系统中采用缓冲技术的目的是增强系统()的能力。
  - A. 串行操作 B. 控制操作 C. 重复操作 D. 并行操作
- D【解析】引入缓冲技术就是为了匹配内存和处理器的运行速度,提高并行操作能力和系统运行效率。因此选择D。

- 2. 【南京工业大学 2013】引入高速缓冲的主要目的是( )。
  - A. 提高CPU的利用率
  - B. 提高I/O设备的利用率
  - C. 改善CPU和I/O设备之间速度不匹配的情况
  - D. 节省内存
- C【解析】CPU与I/O设备执行速度通常是不对等的,前者快,后者慢,通过高速缓冲技术来改善两者速度不匹配的问题。因此选择C。

- 3. 【广东工业大学 2019】操作系统使用缓冲的目的是增强系统(1)的能力,为了使多个进程能有效地同时处理输入和输出,最好使用(2)。
  - (1)A. 串行操作 B. 并行操作 C. 控制操作 D. 中断操作
  - (2)A. 缓冲池 B. 单缓冲 C. 双缓冲 D. 循环缓冲
- B, A【解析】操作系统使用缓冲的目的是增强系统并行操作的能力, 为了使多个进程能有效地同时处理输入和输出, 最好使用缓冲池。因此选择B, A。

- 4. 【燕山大学 2015】双向设备应该使用( )。
  - A. 单缓冲区 B. 双缓冲区
  - C. 多缓冲区 D. 缓冲池
- B【解析】双缓冲区是一种能够实现设备和设备、CPU和设备并行操作的简单模型,当其中一个缓冲区数据装满后,可向另一个缓冲区写入数据,同时前一个缓冲区可被访问。因此双向设备应该使用双缓冲区。因此选择B。

- 5. 【南京工业大学 2017】在现代操作系统中采用缓冲技术的主要目的是( )。
  - A. 改善用户编程环境
  - B. 提高CPU的处理速度
  - C. 提高CPU和设备之间的并行程度
  - D. 实现与设备无关性
- C【解析】在现代操作系统中采用缓冲技术的主要目的是提高CPU和设备之间的并行程度。因此选择C。

- 6. 【北京交通大学 2018】在设备管理中,引入缓冲的主要原因不包括()。
  - A. 缓和处理器和外设之间访问速度不匹配的矛盾
  - B. 缓和处理器和内存之间访问速度不匹配的矛盾
  - C. 减少对处理器的中断频率,放宽对处理器中断响应时间的限制
  - D. 提高处理器和外围设备之间的并行性
- B【解析】引入缓冲的主要原因:
- ①缓和CPU与I/O设备间速度不匹配的矛盾;
- ②减少对CPU的中断频率,放宽对处理器中断响应时间的限制;
- ③提高CPU与I/O设备之间的并行性。

因此选择B。

- 7. 考虑单用户计算机上的下列I/O操作,需要使用缓冲技术的是()。
  - I. 图形用户界面下使用鼠标
  - Ⅱ. 多任务操作系统下的磁带驱动器(假设没有设备预分配)
  - Ⅲ. 包含用户文件的磁盘驱动器
  - IV. 使用存储器映射I/O, 直接和总线相连的图形卡

  - C. 仅<sub>I</sub>、<sub>I</sub>、<sub>I</sub>V D. 全选
- D【解析】I项正确,在鼠标移动时,如果有高优先级的操作产生,为了记录鼠标 活动的情况,必须使用缓冲技术。

Ⅱ项正确,由于磁带驱动器和目标或源I/O设备间的吞吐量不同,必须采用缓冲技 术。

- 7. 考虑单用户计算机上的下列I/O操作,需要使用缓冲技术的是()。
  - I. 图形用户界面下使用鼠标
  - Ⅱ. 多任务操作系统下的磁带驱动器(假设没有设备预分配)
  - Ⅲ. 包含用户文件的磁盘驱动器
  - IV. 使用存储器映射I/O, 直接和总线相连的图形卡

  - C. 仅<sub>I</sub>、<sub>I</sub>、<sub>I</sub>V D. 全选
- D【解析】Ⅲ项正确,为了能使数据从用户作业空间传送到磁盘或从磁盘传送到 用户作业空间,必须采用缓冲技术。

IV项正确,为了便于多幅图形的存取及提高性能,缓冲技术是可以采用的,特别 是在显示当前一幅图形又要得到下一幅图形时,应采用双缓冲技术。

8. 在某系统中, 若采用双缓冲区(每个缓冲区可存放一个数据块), 将一个数据块从磁盘传送到缓冲区的时间为80µs, 从缓冲区传送到用户的时间为20µs, CPU 计算一个数据块的时间为50µs。总共处理4个数据块, 每个数据块的平均处理时间是多少?

【解析】系统采用双缓冲区,可采用二级流水线结构进行传送,传送4个数据块的时间是80×4=320μs,最后一个缓冲区的数据传送到用户区加上CPU处理时间为20+50=70μs,总时间为320+70=390μs,平均处理时间为390/4=97.5μs。

80us	20us 50us						
	80us	20us	50us				
		80us		20us	50us		
				80us		20us	50us

谢谢大家

考点三:

假脱机技术

# 考点框架



SPOOLing概念



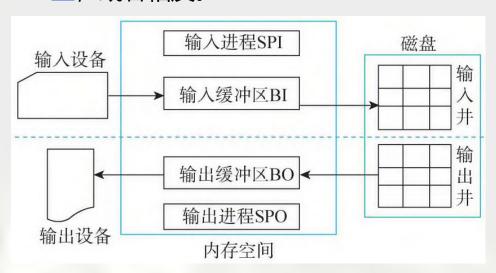
SPOOLing实现和组成



SPOOLing输入和输出过程



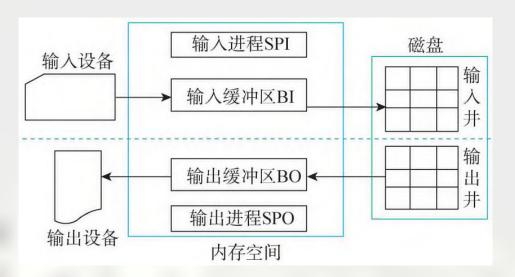
- □ 为了缓和CPU的高速性与I/O设备低速性间的矛盾而引入了脱机输入、 脱机输出技术。
- □ 该技术是利用专门的外围控制机, 将低速I/O设备上的数据传送到高速磁盘上; 或者相反。





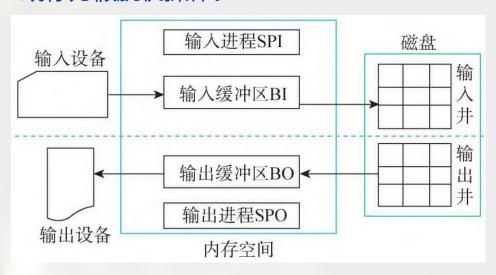
### **SPOOLing概念**

- □ 当系统中引入了多道程序技术后,完全可以利用其中的一道程序,来模拟脱机 输入时的外围控制机功能,把低速I/O设备上的数据传送到高速磁盘上;
- □ 再用另一道程序来模拟脱机输出时外围控制机的功能, 把数据从磁盘传送到低 速输出设备上。这样,便可在主机的直接控制下,实现脱机输入、输出功能。





此时的外围操作与CPU对数据的处理同时进行, 我们把这种在联机情况下实现的同时外围操作称为SPOOLing(Simultaneaus Periphernal Operating On-Line), 或称为假脱机操作。

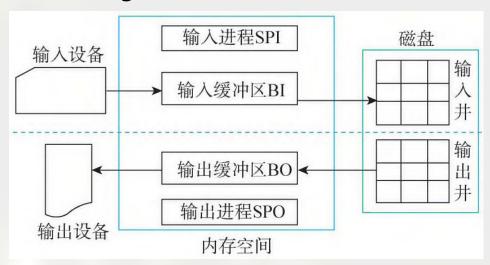




### **SPOOLing实现和组成**

SPOOLing 技术是对脱机输入/输出系统的模拟,SPOOLing 系统建立在通道技 术和多道程序技术的基础上,以高速随机外存(通常为磁盘)为后援存储器。

#### SPOOLing 系统主要由以下四部分构成:



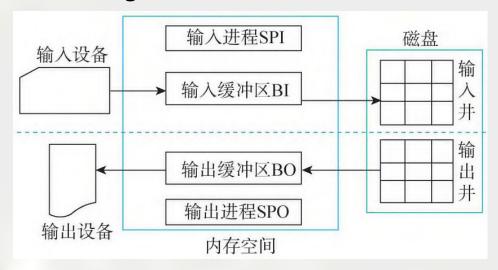
- (1)输入井和输出井。这是在磁盘上开辟 出来的两个存储区域。
- □ 输入井模拟脱机输入时的磁盘,用于 收容I/O设备输入的数据。
- □ 输出井模拟脱机输出时的磁盘,用于 收容用户程序的输出数据。



### **SPOOLing实现和组成**

SPOOLing 技术是对脱机输入/输出系统的模拟,SPOOLing 系统建立在通道技 术和多道程序技术的基础上,以高速随机外存(通常为磁盘)为后援存储器。

#### SPOOLing 系统主要由以下四部分构成:

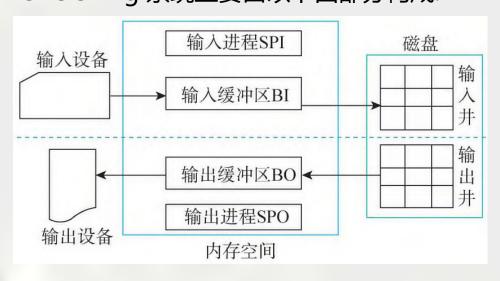


- (1)输入井和输出井。这是在磁盘上开辟 出来的两个存储区域。
- □ 输入/输出井中的数据一般以文件的 形式组织管理, 我们把这些文件成为 井文件。
- □ 一个文件仅存放某一个进程的输入或 者输出数据,所有进程的输入或者输 出文件链接成为一个输入(或输出) 队列。



### **SPOOLing实现和组成**

SPOOLing 技术是对脱机输入/输出系统的模拟,SPOOLing 系统建立在通道技 术和多道程序技术的基础上,以高速随机外存(通常为磁盘)为后援存储器。 SPOOLing 系统主要由以下四部分构成:



(2)输入缓冲区和输出缓冲区。这是在 内存中开辟的两个缓冲区,用于缓和 CPU和磁盘之间速度不匹配的矛盾。

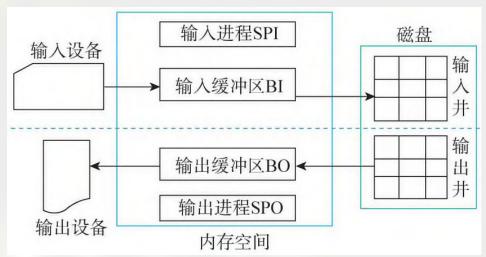
- □ 输入缓冲区用于暂存由输入设备传 送的数据,之后再传送到输入井。
- □ 输出缓冲区用于暂存从输出井传送 的数据,之后再传送到输出设备。



### **SPOOLing**实现和组成

SPOOLing 技术是对脱机输入/输出系统的模拟,SPOOLing 系统建立在通道技 术和多道程序技术的基础上,以高速随机外存(通常为磁盘)为后援存储器。

#### SPOOLing 系统主要由以下四部分构成:



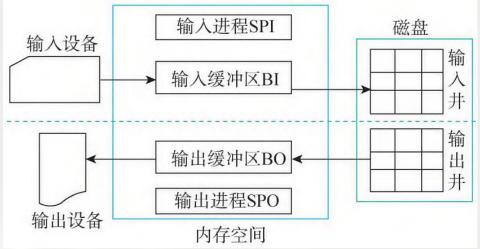
(3)输入进程和输出进程:输入进程也 称预输入进程,用于模拟脱机输入时的 外围控制机,将用户要求的数据从输入 设备传送到输入缓冲区, 再存放到输入 井。



### **SPOOLing实现和组成**

SPOOLing 技术是对脱机输入/输出系统的模拟,SPOOLing 系统建立在通道技 术和多道程序技术的基础上,以高速随机外存(通常为磁盘)为后援存储器。

## SPOOLing 系统主要由以下四部分构成:



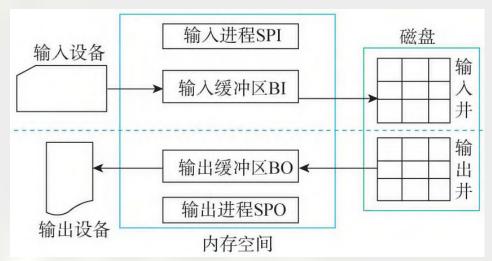
- (3) 当CPU需要输入设备时,直接从输 入井读入内存。
- □ 输出进程也称缓输出进程,用于模 拟脱机输出时的外围控制机,把用 户要求输入的数据从内存传送并存 放到输出井, 待输出设备空闲时, 再将输出井中的数据经过输出缓冲 区输出至输出设备上。



### **SPOOLing实现和组成**

SPOOLing 技术是对脱机输入/输出系统的模拟,SPOOLing 系统建立在通道技 术和多道程序技术的基础上,以高速随机外存(通常为磁盘)为后援存储器。

#### SPOOLing 系统主要由以下四部分构成:

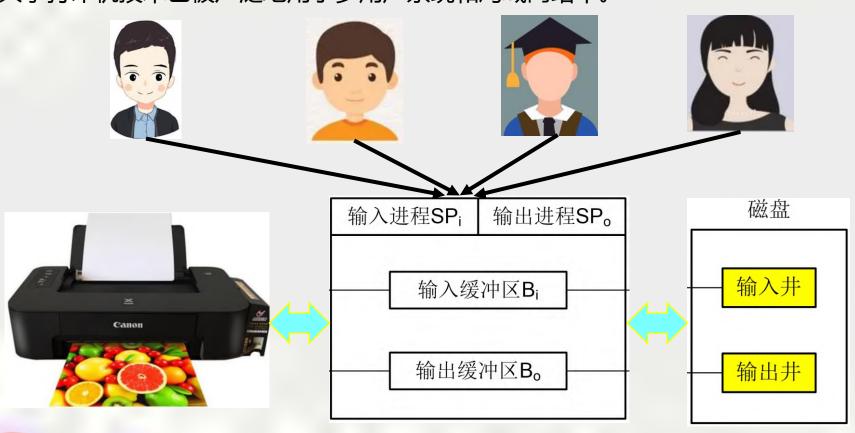


- (4)井管理程序:用于控制作业与磁盘 之间信息的交换。
- □ 当作业执行过程中向某台设备发出 启动输入或输出操作请求时,由操 作系统调用井管理程序, 由其控制 从输入井读取信息或将信息输出至 输出井。



# **SPOOLing输入和输出过程**

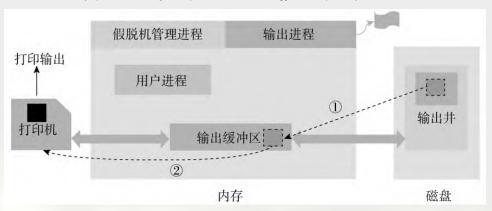
共享打印机技术已被广泛地用于多用户系统和局域网络中。





### **SPOOLing输入和输出过程**

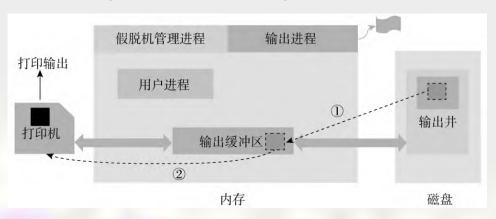
- □ 在联机的条件下, 进行两个方向的操作, 如图所示, 在数据输入时, 将数据从 输入设备传送到磁盘或磁带(块设备),然后把这些块设备与主机相连;
- □ 输入进程SPI是模拟脱机输入时的外围控制机,它将用户要求处理的数据从输 入设备通过输入缓冲区再送到输入井(磁盘上开辟的一块区域),当CPU处理 这些数据时,就直接从输入井读入内存。





### SPOOLing输入和输出过程

- □ 在数据输出时,将输出数据传送到磁盘或磁带上,再从磁盘或磁带传送到输出设备。这样,可以将一台独占的物理设备虚拟为并行使用的多态逻辑设备,从而使该物理设备被多个进程共享。
- □ 输出进程SPOOLing是模拟脱机输出时的外围控制机,把用户要求输出的数据, 先从内存送到输出井,待输出设备空闲时,再将输出井中的数据通过输出缓 冲区(内存中一块区域)传送到输出设备上。





### 实例——利用打印机实现打印机共享

- □ 当用户进程请求打印输出时, SPOOLING系统立即同意为该进程执行打印输 出,但并不是真正地把打印机分配给该用户进程,而只是为该进程做两项工作:
- □ 一项是由输出进程SPO在输出井中为之申请一个空闲的存储空间,并将要打印 的数据传送其中存放;
- □ 另一项工作就是由输出进程SPO再为用户进程申请一张空白的用户请求打印表, 并将用户的打印请求填入其中,然后将该表挂到打印机的请求队列上。
- □ 这时,如果还有另一个进程请求打印机时,则系统仍同意为该进程执行打印输 出, 当然, 系统所做的工作仍是以上两项内容。



### 实例——利用打印机实现打印机共享

- □ 在打印机执行实际打印时,如果打印机空闲,输出进程SPO将从请求打印队列 的队首取出一张打印表, 根据打印表中的要求将要打印的数据从输出井传送到 内存输出缓冲区,再传送到打印机打印。
- □ 打印完后,输出进程SPO将再检查请求打印队列中是否还有待打印的请求表, 若有则再取出一张请求打印表,将新的但因要求继续打印。
- □ 如此反复, 直到请求打印队列空为止, 输出进程才将自己阻塞起来, 并在下次 再有打印请求时被唤醒。



### **SPOOLing的特点**

- □ 提高了I/O的速度,这里, 对数据所执行的I/O操作, 已从对低速I/O设备执行 的I/O操作演变为对磁盘 缓冲区中数据的存取,如 同脱机输入输出一样。
- □ 将独占设备改造为共享设 备,在假脱机打印系统中 实际上并没有为任何进程 分配设备,而只是在磁盘 缓冲区中为进程分配一个 空闲盘块和建立一张I/O 请求表,这样,便把独占 设备改造为共享设备。
- □ 实现了虚拟设备功能, 宏观上,虽然是多个进 程在同时使用一台独占 设备,而对于每一个进 程而言,它们都会认为 自己是独占了一个设备。

# 【牛刀小试】

- 1. 【广东工业大学 2014】假脱机技术是将输入输出控制工作大部分交由相应的通道来承担,利用磁盘作为后援存储器,实现了外设同时联机操作。采用假脱机技术的目的是()。
  - A. 提高外设和主机的利用率 B. 提高内存和主机效率
  - C. 减轻用户编程负担 D. 提高程序的运行速度

A【解析】采用假脱机技术的目的是提高外设和主机的利用率。因此选择A。

- 2. 【中国计量大学 2019】SPOOLing技术中, 其中输入井和输出井是()。
  - A. 硬盘的一部分 B. 内存的一部分
  - C. 缓冲区 D. 以上都是

A【解析】SPOOLing技术中的输入井和输出井实质上是在磁盘中开辟的两大存储区域,输入缓冲区和输出缓冲区实质上是内存中开辟的两大存储区域。因此选择A。

- 3. 【南京理工大学 2016】下列关于SPOOLing系统的说法,正确的是()。
  - A. 只要输入设备空闲,SPOOLing可预先将输入数据从设备传送到输入井中
  - B. SPOOLing是脱机I/O系统
  - C. 当输出设备忙时, SPOOLing系统中的用户程序暂停执行
  - D. SPOOLing系统是虚拟存储技术的体现

A【解析】SPOOLing系统实现了对I/O设备的虚拟,只要输入设备空闲,SPOOLing可预先将输入数据从设备传送到输入井中供用户程序随时读取,选项A正确;SPOOLing是假脱机I/O系统,选项B错误;利用SPOOLing技术,提高了系统和I/O设备的利用率,进程不必等待I/O操作完成,选项C错误;SPOOLing系统采用以空间换取时间的技术,选项D错误。因此选择A。

- 4. 【南京理工大学 2017】在采用SPOOLing技术的系统中,用户作业的打印输出结果首先被送到()。
  - A. 磁盘固定区域 B. 内存固定区域
  - C. 终端 D. 打印机

A【解析】SPOOLing技术是利用一道程序来模拟脱机输入时外围控制机的功能,把低速I/O设备上的数据传送到高速磁盘上,再利用另一道程序来模拟脱机输出时外围控制机的功能,把数据从磁盘传送到低速设备上,这样外围操作与CPU对数据的处理同时进行。采用SPOOLing技术的系统中,用户作业的打印输出结果首先被送到磁盘固定区域。因此选择A。

- 5. 【南京理工大学 2019】SPOOLing技术的输入井和缓冲区分别位于()。
  - A. 内存和磁盘 B. 磁盘和内存
  - C. 内存和内存 D. 磁盘和寄存器
- B【解析】SPOOLing技术中的输入井和输出井实质上是磁盘中开辟的两大存储区域,输入缓冲区和输出缓冲区实质上是内存中开辟的两大存储区域。因此选择B。

- 6. 【北京交通大学 2017】 (多选) 采用假脱机技术来实现打印机共享, 需要()。
  - A. 设置内存缓冲区暂存从输出井传送来的欲打印输出的数据
  - B. 在磁盘上设立输出井用来收容欲打印输出的数据
  - C. 接受用户打印输出请求并将之插入到打印请求队列
  - D. 处理打印请求队列的各项打印请求并实施打印操作

ABCD【解析】输入井和输出井是在磁盘上开辟的两大存储区域。

- □ 输入并是模拟脱机输入时的磁盘,用于收容输入设备输入的数据;
- □ 输出并是模拟脱机输出时的磁盘,用于收容用户程序的输出数据。输入缓冲区和输出缓冲 区是在主存中开辟的两个缓冲区。
- □ 输入缓冲区用于暂存由输入设备送来的数据,以后再传送到输入井;
- □ 输出缓冲区用于暂存从输出井送来的数据,以后再传送给输出设备。

- 6. 【北京交通大学 2017】 (多选) 采用假脱机技术来实现打印机共享, 需要()。
  - A. 设置内存缓冲区暂存从输出井传送来的欲打印输出的数据
  - B. 在磁盘上设立输出井用来收容欲打印输出的数据
  - C. 接受用户打印输出请求并将之插入到打印请求队列
  - D. 处理打印请求队列的各项打印请求并实施打印操作

ABCD【解析】输入井和输出井是在磁盘上开辟的两大存储区域。

- □ 输入进程模拟脱机输入时的外围控制机,将用户要求的数据从输入设备通过输入缓冲区送 到输入井。当CPU需要数据时,直接从输入井读入主存;
- □ 输出进程模拟脱机输出时的外围控制机,把用户要求输出的数据先从主存送到输出井,待输出设备空闲时,再将输出井中的数据经过输出缓冲区送到输出设备上。

系统为每个请求打印的进程建立一张请求打印表,由若干张请求打印表形成打印请求队列。

- 7. ( )是操作系统中采用的以空间换取时间的技术。
  - A. SPOOLing技术 B. 虚拟存储技术
  - C. 覆盖与交换技术 D. 通道技术

A【解析】SPOOLing技术需有高速大容量且可随机存取的外存支持,通过预输入及缓输出来减少CPU等待慢速设备的时间,这是典型的以空间换时间策略的实例。因此选择A。

【经典总结】时间换空间的技术:虚拟存储技术、覆盖与交换技术等。

空间换时间的技术: SPOOLing技术、缓冲技术等。

各种虚拟存储技术都是时间换空间的技术,包括请求分页、请求分段、请求段页式,这些都使访问时间增加了,但是扩充了主存的逻辑容量,使得大于主存容量的程序也可以得到执行。

- 8. 采用假脱机技术,将磁盘的一部分作为公共缓冲区以代替打印机,用户对打印机的操作实际上是对磁盘的存储操作,用以代替打印机的部分由()完成。
  - A. 独占设备 B. 共享设备
  - C. 虚拟设备 D. 一般物理设备
- C【解析】采用假脱机技术,将磁盘的一部分作为公共缓冲区以代替打印机,用户对打印机的操作实际上是对磁盘的存储操作,用以代替打印机的部分由虚拟设备完成。因此选择C。

- 9. 下面关于独占设备和共享设备的说法中,不正确的是()。
- A. 打印机、扫描仪等属于独占设备
- B. 对独占设备往往采用静态分配方式
- C. 共享设备是指一个作业尚未撤离,另一个作业即可使用,但每个时刻只有一个作业使用
- D. 对共享设备往往采用静态分配方式
- D【解析】独占设备采用静态分配方式,而共享设备采用动态分配方式。

【经典总结】独占设备:即不能共享的设备,一段时间只能由一个作业独占,如打印机、读卡机、磁带机等。所有字符型输入输出设备原则上都应是独占设备。共享设备:可由若干作业同时共享的设备,如磁盘机等。共享分配技术保证多个进程可以同时方便地直接存取一台共享设备。共享提高了设备的利用率。块设备都是共享设备。

谢谢大家

谢谢大家

考点四:

IO方式

# 考点框架



基本原理



程序查询方式



程序中断方式



直接存储器存取方式



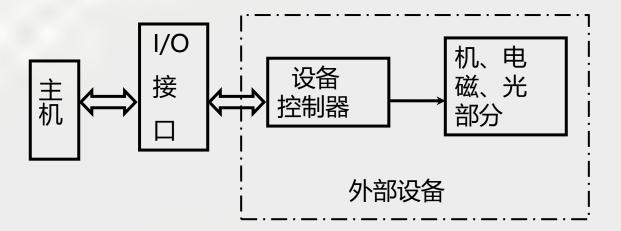
IO通道控制方式

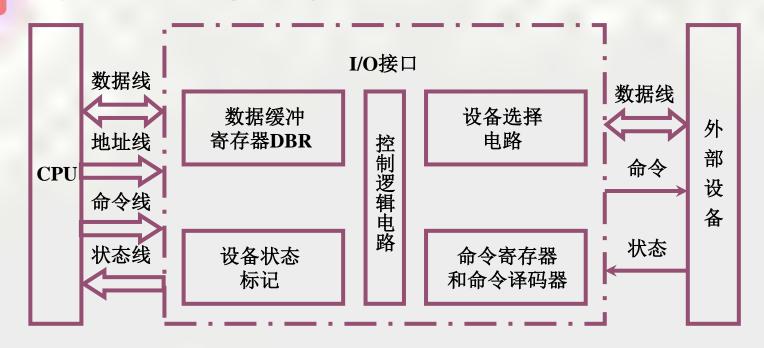


IO处理机



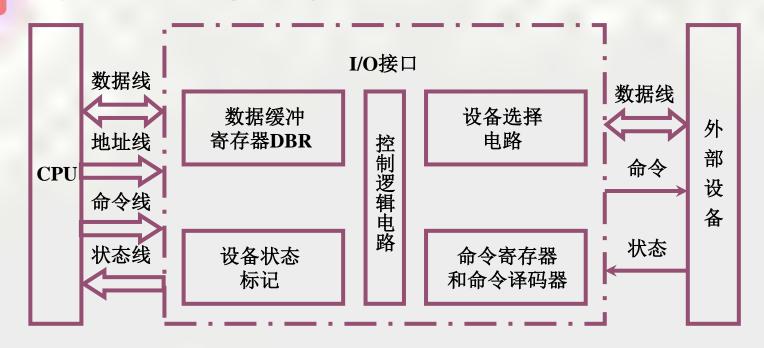
□ I/O系统包括外围设备、输入/输出接口和输入/输出软件,其主要作用是用来 完成数据的输入、输出、成批存储等任务。





CPU和外设之间隔着IO接口

IO接口是CPU和外设的中间件,本质上是IO接口控制IO设备



CPU和外设之间交换数据,实质上是通过IO端口进行的



常用的IO控制方式

速度:

CPU > 内存

内存 > IO设备

程序查询方式

程序中断方式

DMA方式

程序查询方式



□程序查询由CPU执行一段输入、输出程序来实现主存与外设之间的数据传送方式,也叫做程序直接控制方式,它完全通过程序来控制主机和IO设备之间的信息传送。

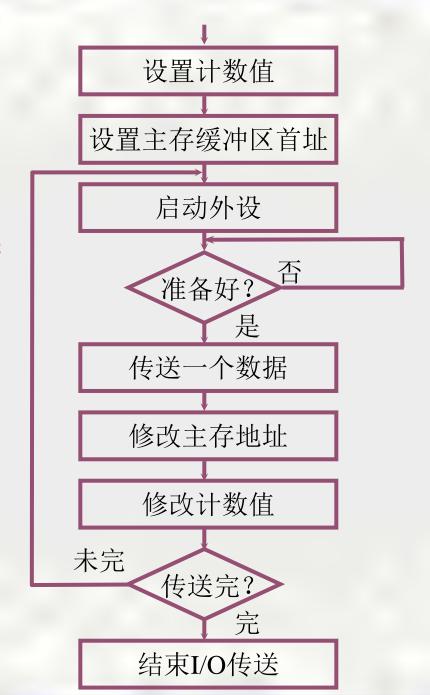


- □ 通常的实现方式是在用户的程序中安排一段由**输入、输出指令和其他 指令**所组成的程序段直接控制外围设备的工作。
- □ 传送时,首先**启动设备**,发出启动命令,接着**CPU**不断地用一条测试指令检测外围设备的工作状态标志触发器。
- □ 一旦测试到标志触发器已置成 "完成"状态,即可进行数据传送。
- □ 为了保证数据传送的正确进行,就要求CPU在程序中查询外设的工作状态,如果外设尚未准备就绪,CPU就等待,只有外设已作好准备,CPU才能执行



## 程序查询方式的工作过程

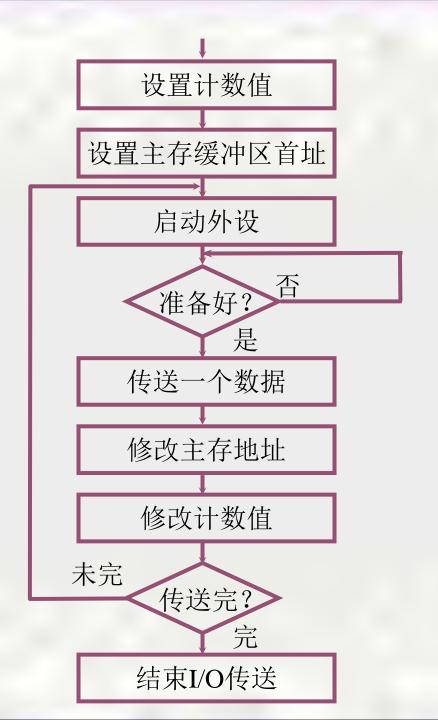
- (1) 由于这种方式传送数据时要占用CPU 中的寄存器,故首先需将寄存器原内容保护起 来(若该寄存器中存有有用信息)。
- (2) 由于传送可能是一批数据,因此需先设 置IO设备与主机交换数据的计数值。
- (3) 设置欲传送数据在主存缓冲区的首地 址。
  - (4) CPU启动I/O设备。





## 程序查询方式的工作过程

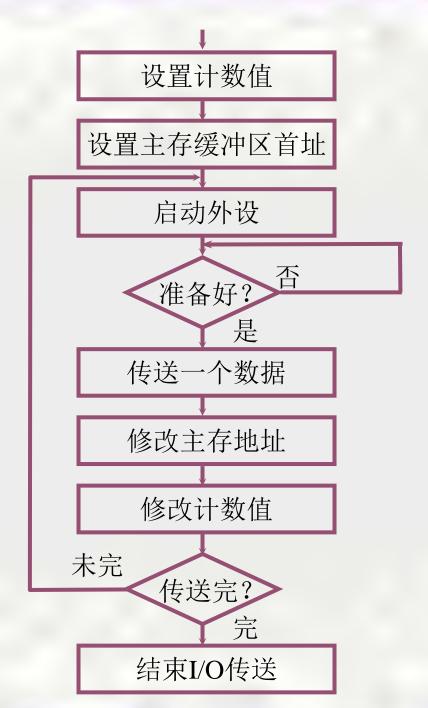
(5) 将IO接口中的设备状态标志取至CPU 并测试IO设备是否准备就绪。如果未准备就 绪,则等待,直到准备就绪为止。当准备就绪 时,接着可实现传送。





## 程序查询方式的工作过程

(5) 对输入而言,准备就绪意味着接口电路 中的数据缓冲寄存器已装满欲传送的数据,称 为输入缓冲满,CPU即可取走数据;





## 程序查询方式的工作过程

(5) 对输出而言,准备就意味着接口电路中 的数据已被设备取走,故称为输出缓冲空,这 样CPU可再次将数据送到接口,设备可再次从 接口接收数据。

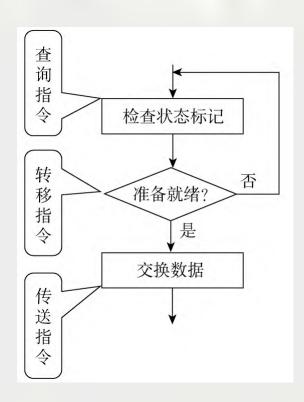




## 程序查询方式的工作过程

这一步一个CPU不断测试的过程,为了实现这种传 送,如图所以,通常要使用3条指令:

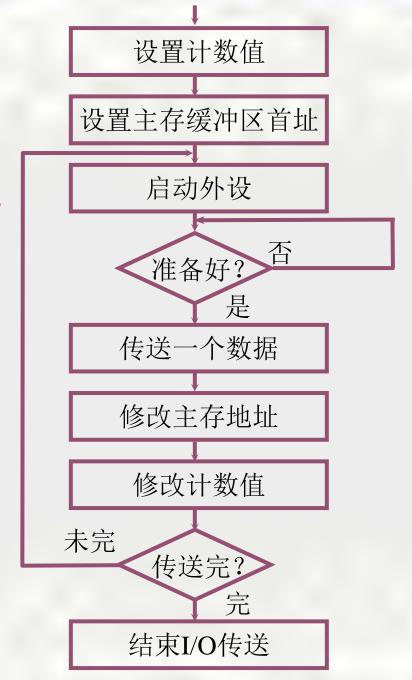
- (1) 查询指令,用来查询设备的状态。
- (2) 传送指令, 当设备就绪时, 执行数据交换。
- (3) 转移指令, 当设备未就绪时, 执行转移指令转 向查询指令继续查询。





## 程序查询方式的工作过程

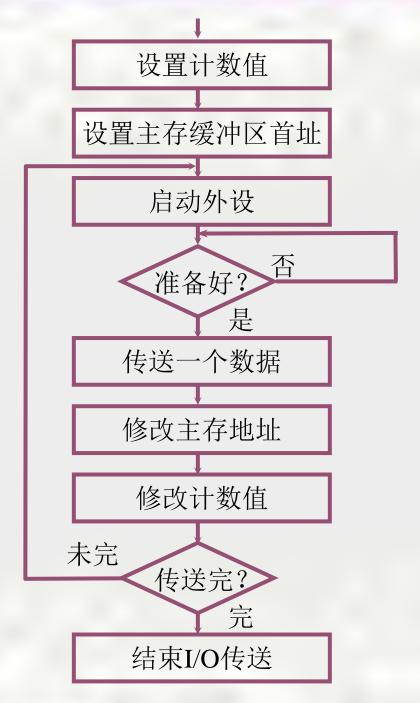
在IO接口中至少有两个寄存器,一个是数据缓冲寄 存器,即数据端口,用来存放与CPU进行传送的数 据信息,另一个是供CPU查询的设备状态寄存器, 即状态端口。





## 程序查询方式的工作过程

- (6) CPU执行IO指令,或从IO接口的数据缓冲 寄存器中读出一个数据,或把一个数据写入IO接 口中的数据缓冲寄存器内,同时将接口中的状态 标志复位。
  - (7) 修改主存地址。
  - (8) 修改计数值。





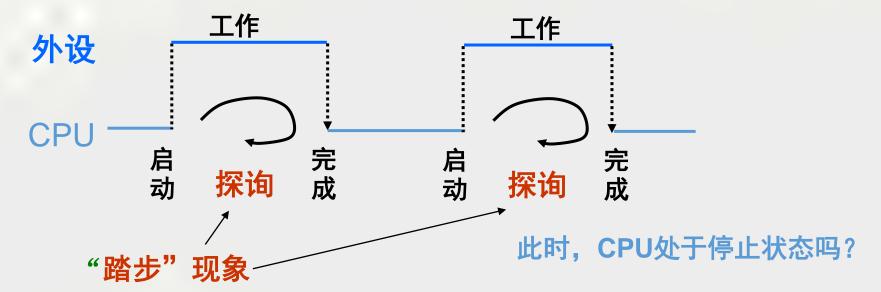
## 程序查询方式的工作过程

- (9) 判断计数值。若计数值不为0,表示一批数 据尚未传送完,重新启动外设继续传送;若计数值 为0,则表示一批数据已传送完毕。
  - (10) 结束I/O传送,继续执行现行程序。





程序查询方式的特点



工作方式: 完全串行工作方式或部分串行, CPU用100%的时间为I/O服务!



## 程序查询方式的特点

程序查询方式的优点是硬件简单,容易实现。但是缺点也比较突出,主要有以下 三点:

- (1) 在查询过程中,CPU需要不断的查询,也就是说CPU完全和IO设备串行工 作,即CPU需要全程参与IO过程的两个阶段,白白浪费了CPU资源,CPU利用率 低, 系统效率大大降低。
  - (2) CPU在一段时间内只能和一台外设交换信息,其他设备不能同时工作。
  - (3) 不能发现和处理预先无法估计的错误和异常情况。

程序中断方式



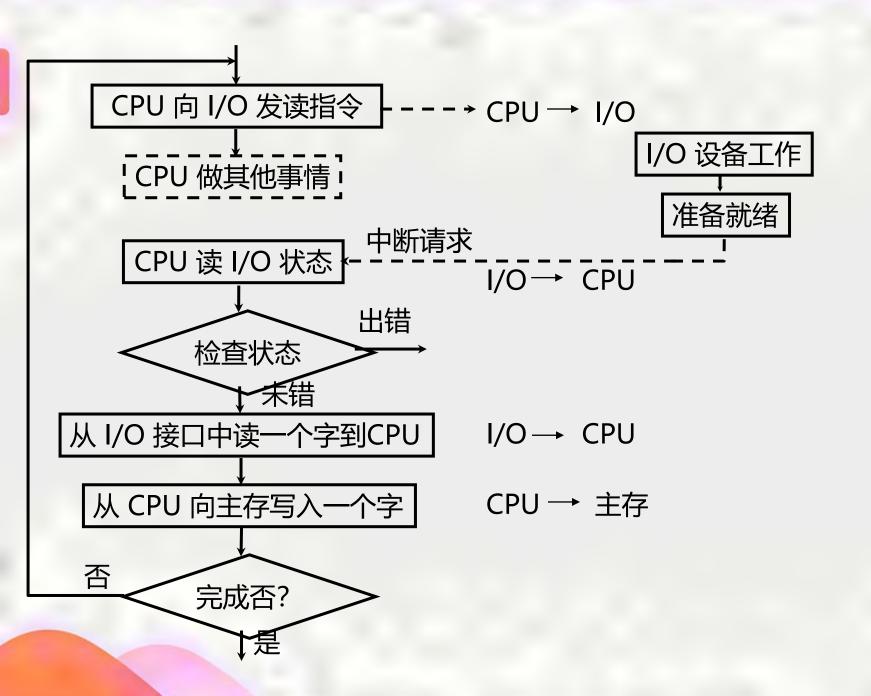
## 程序中断方式基本原理

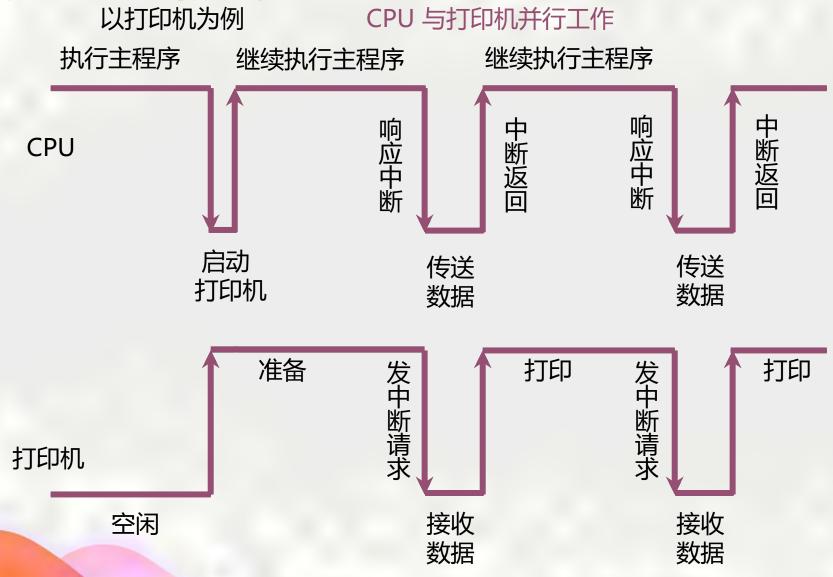
- □ 程序中断IO方式是在程序中断传送方式中,通常在程序中安排一条指令,发 出START信号启动IO设备,然后机器继续执行程序,此时IO处于准备数据的 过程。
- □ 当IO设备在做好输入/输出准备时,向主机发中断请求,主机接到请求后就暂 时中止原来执行的程序,转去执行中断服务程序,在中断服务程序中完成输 入/输出操作,
- □ 在中断处理完毕后, CPU将自动返回原来的程序继续执行。

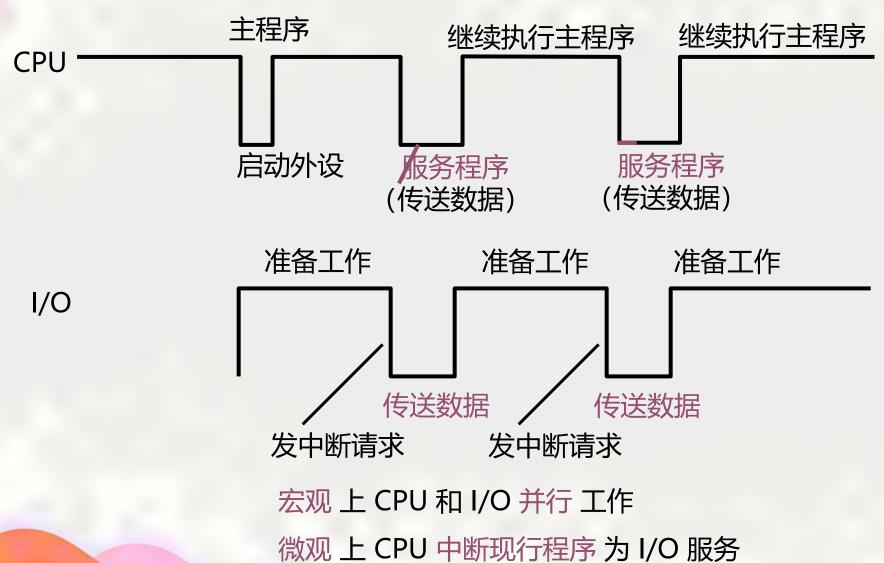


## 程序中断方式的过程

- □ 程序中断IO方式是在程序中断传送方式中,通常在程序中安排一条指令,发 出START信号启动IO设备,然后机器继续执行程序,此时IO处于准备数据的 过程。
- □ 当IO设备在做好输入/输出准备时,向主机发中断请求,主机接到请求后就暂 时中止原来执行的程序,转去执行中断服务程序,在中断服务程序中完成输 入/输出操作,
- □ 在中断处理完毕后, CPU将自动返回原来的程序继续执行。



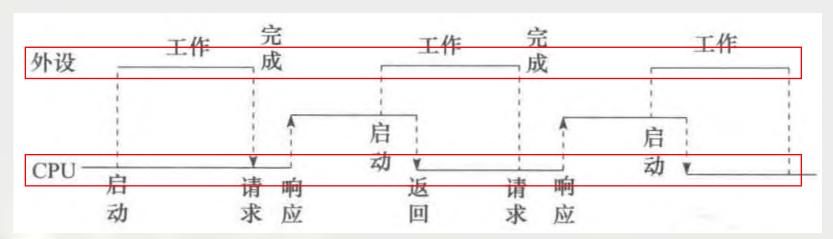






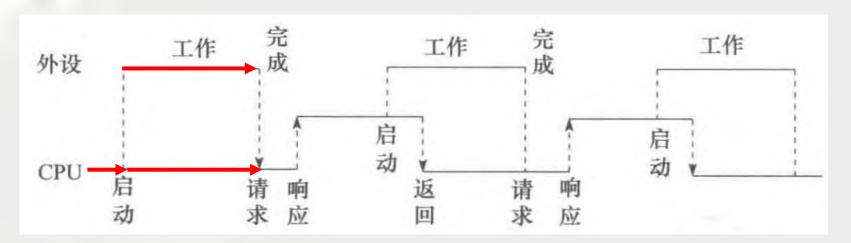
# 程序中断方式的过程

图给出了程序中断IO方式下CPU与外设的工作时空图,在这个时空图中,有两 条轨迹,第一条是IO设备(外部设备,外设)的时间图,第二条是CPU的时间 图。从图中可以看到,如果不进行IO操作,CPU一直执行主程序,

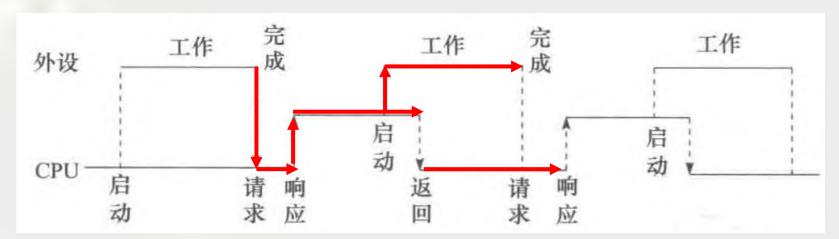


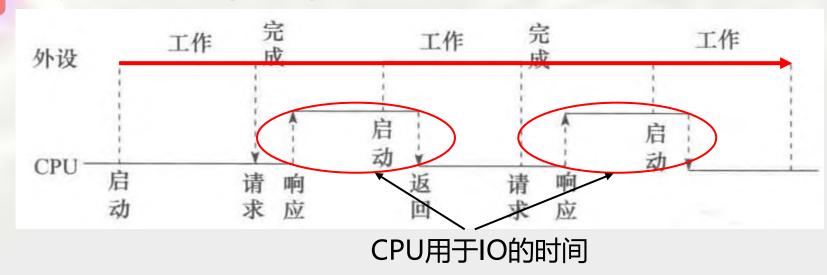
当CPU需要IO输入时, CPU通过IO指令启动IO设备;

接着CPU继续执行主程序,IO设备开始工作,进行数据的准备,也就是完成数据从IO设备到IO接口(端口)的准备工作,此阶段CPU与IO设备并行工作;



当IO设备将数据放入到IO接口后,将会向CPU发出中断请求,CPU响应中断, 执行中断服务程序,完成从IO接口到CPU的通用寄存器的数据传送,且当需要 继续使用IO时,启动IO设备,并重复上述过程。





要想传输一个数据(往往是一个字或者字节,因为程序中断IO方式的传输单位是字或者字节),

- □ IO设备的工作时间是: 准备数据 (工作) +中断响应过程+中断处理过程;
- □ CPU参与IO过程的时间是:中断响应过程+中断处理过程。



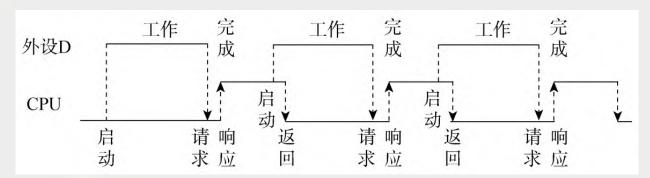
# 程序中断方式的特点

- (1) 数据传输的基本单位是字节或者字, 每传输一个字或者字节都需要一次中断。
- (2) I/O设备的启动需要CPU参与,但是CPU处理其他事情和I/O设备准备数据可 以同时并行,提高CPU利用率和系统效率。



# 程序中断方式的特点

(3) 整体而言, CPU是通过响应中断、执行中断处理程序完成I/O操作。换句 话说,程序中断IO方式请求的是CPU中断响应和中断服务程序的时间。





# 程序中断方式的特点

(4) 程序中断方式克服了程序查询方式中的CPU "踏步" 现象,实现了CPU与IO 设备的并行工作,提高了CPU的资源利用率。但从微观操作分析,发现CPU执行中 断服务程序时仍需暂停原程序的正常运行,尤其是当高速IO设备或辅助存储器频繁 地、成批主存交换信息时,需不断地打断CPU执行主程序而执行中断服务程序。所 以由于数据中的每个字在存储器与I/O控制器之间的传输都必须经过CPU,这就 导致了中断驱动方式仍然会消耗较多的CPU时间。

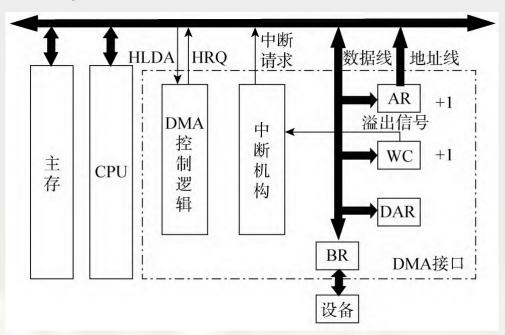
直接存储器存取方式



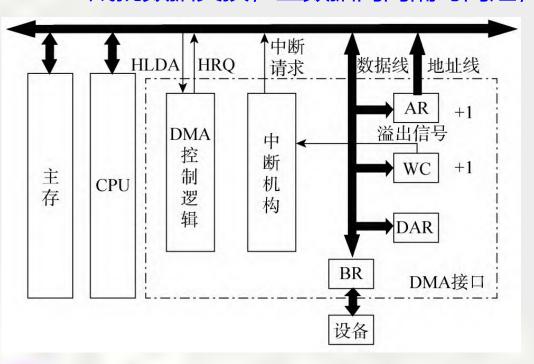
DMA方式的基本思想

在高速外设和主存间直接传送数据

由专门硬件(即: DMA接口)控制总线进行传输



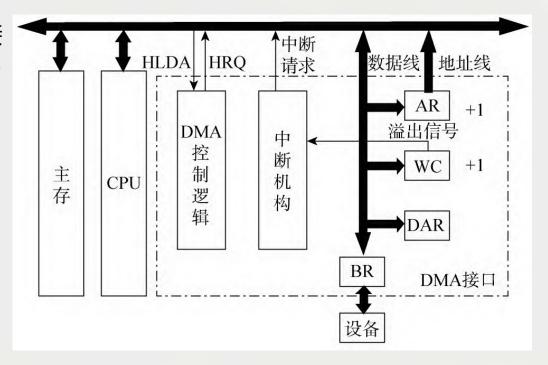
- DMA方式适用场合
  - 高速设备(如:磁盘、光盘等)
  - 成批数据交换,且数据间间隔时间短,一旦启动,数据连续读写





## **DMA与CPU的访存控制方法**

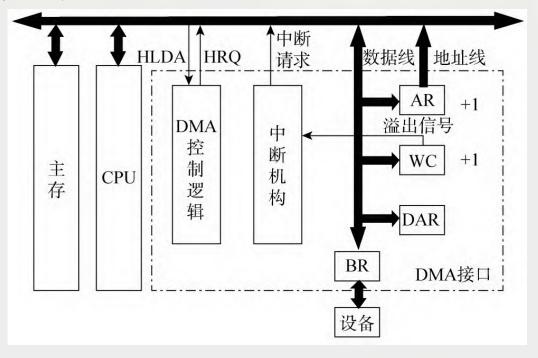
□ 在DMA方式中,在主存和DMA接 口之间有一条数据通路,, DMA可 以直接访问主存,而CPU也可以 直接访问主存,那么就会出现一 个问题, 当CPU和DMA同时访问 主存时,该如何安排访问顺序呢?





## DMA与CPU的访存控制方法

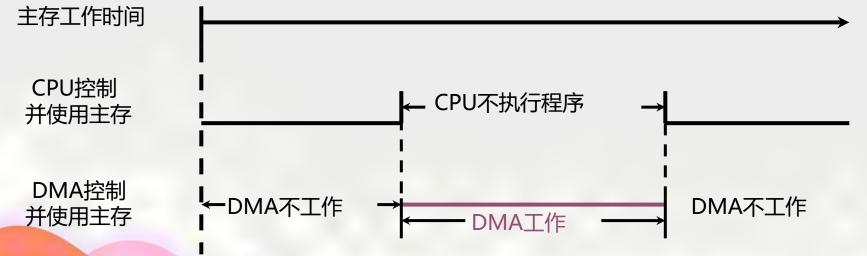
□ DMA的优先级较高,也就是需要 让DMA先访问主存,因为DMA连 接的是高速设备,如果不及时将 数据放入内存,有可能导致数据 丢失。





## **DMA与CPU的访存控制方法**

- □ CPU停止访问主存法
- 每次DMA请求获得批准后,DMA控制器获得总线控制权,连续进行成组数 据传送,直至批量传送结束,DMA控制器才把总线控制权交回CPU。
- 在DMA操作期间,CPU处于保持状态,停止访问主存,仅能进行一些与总 线无关的内部操作。这种方法只适用于高速外设的成组传送。

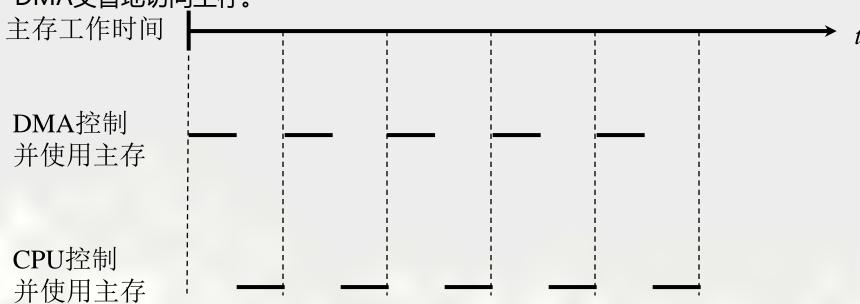




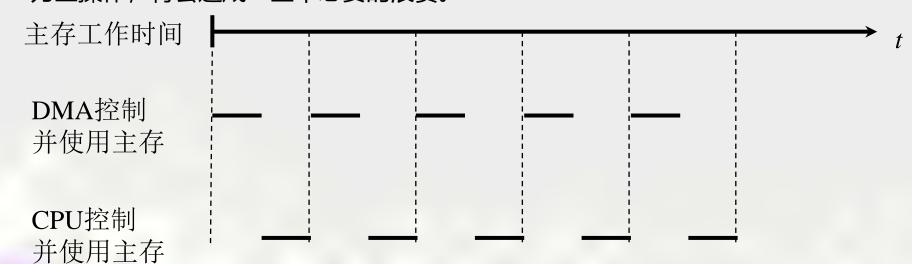
# **DMA与CPU的访存控制方法**

#### □ 存储器分时法

• 把原来的一个存取周期分成两个时间片,一片给CPU,一片给DMA,使CPU和 DMA交替地访问主存。



- □ 存储器分时法
- 这种方法不需要申请和归还总线,总线控制权的转移几乎不需要什么时间,所以对DMA传送来说效率很高,但这种方法如果要维持CPU的访存速度不变,就要求主存的工作速度提高一倍。
- 由于大多数外设的速度都不能与CPU 相匹配,所以供DMA使用的时间片可能成为空操作,将会造成一些不必要的浪费。





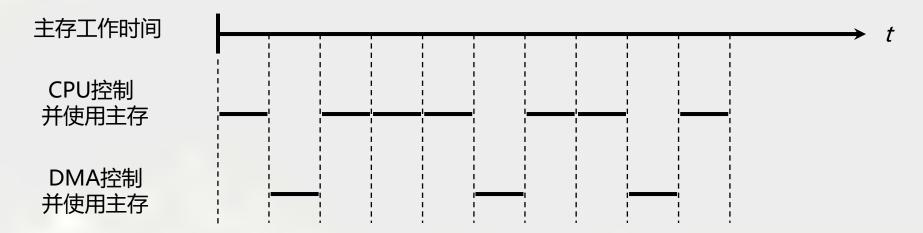
# **DMA与CPU的访存控制方法**

- □ 周期挪用法 (周期窃取法)
- DMA和CPU的访存频率是不一样, 且CPU的访存频率远远高于DMA的访存频 率,如果我们将DMA和CPU的访存频率看作一样,将会导致大量的CPU时间 的浪费。所以就提出了周期窃取法,

- □ 周期挪用法 (周期窃取法)
- 主要的访存对象依然是CPU,而一旦外设有DMA请求并获得CPU批准后, CPU让出一个周期的总线控制权,让DMA控制总线使用权进行一次数据传送, 该传送过程传送一个字节或一个字;
- 当传输结束后,DMA控制器将总线控制权交回CPU,CPU继续进行自己的操作,并等待下一个DMA请求的到来;重复上述过程,直至数据块传送完毕。



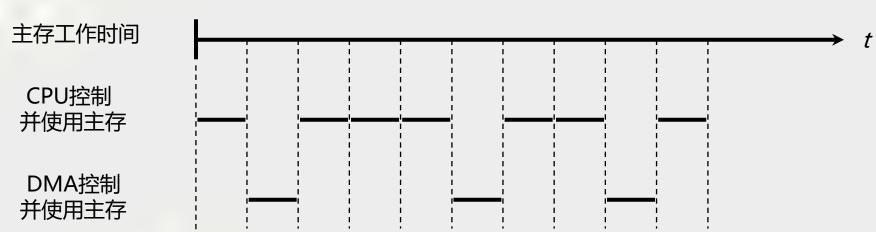
- □ 周期挪用法 (周期窃取法)
- 如果在同一时刻,发生CPU与DMA的访存冲突,那么优先保证DMA工作,使CPU等待一个存取周期。若DMA传送时CPU不需要访存,则外设的周期挪用对CPU的执行程序无任何影响。





# DMA与CPU的访存控制方法

□ 周期挪用法 (周期窃取法)



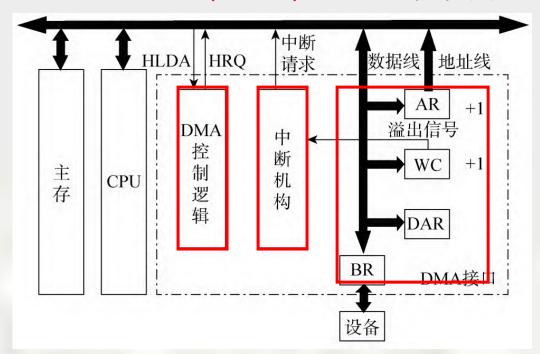
周期窃取: 窃取的是什么周期

- 指令周期
- 机器周期 (CPU周期, 存取周期)
- 时钟周期



## **DMA**的组成和功能

- □ 在主存和DMA接口之间有一条数据通路,因此主存和设备交换信息时,不通过 CPU, 那么就需要相应的部件来控制传送过程, 在DMA传送方式中, 对数据 传送过程进行控制的硬件称为DMA控制器。
- □ DMA控制器由3部分构成: DMA控制逻辑 (DMAC) 、寄存器和中断机构。





## **DMA**的组成和功能

#### DMA控制器逻辑的功能包括:

- □ (1) 接受外设发出的DMA请求,并向CPU发出总线请求。这里请大家注意, IO设备向DMAC发出DAM请求, DMAC向CPU发出总线使用请求。
- □ (2) 当CPU响应此总线请求时,发出总线响应信号后,接管对总线的控制, 进入DMA操作周期。
- □ (3) 确定传送数据的主存单元地址及传送长度,并能自动修改主存地址计数 值和长度计数值。
- (4) 识别传送数据方向,发出读/写或其他控制信号,并执行数据传送的操作。
- (5) 向CPU报告DMA操作的结束。



## **DMA**的组成和功能

- □ 为了实现在主机与控制器之间成块数据的直接交换,在DMA控制器中设置如 下四类寄存器,要想与主存直接数据交换,必须要依赖于寄存器。
- □ (1) 命令/状态寄存器CR, 用于接收从CPU发来的I/O命令或有关控制信息或 设备的状态。
- □ (2) 内存地址寄存器MAR, 在输入时, 存放把数据从设备传送到内存的起始 目标地址; 在输出时, 存放由内存到设备的内存源地址。
- (3) 数据寄存器DR,用于暂存从设备到内存,或从内存到设备的数据。
- (4) 数据计数器DC,存放本次CPU要读或写的字(节)数。用来存放主存 中要交换数据的地址, 当DMA传送时, 每传送一个数据, 将地址计数器加1。



## **DMA**的组成和功能

- □ DMA传送方式是异步的方式, CPU并不直接参与数据的传输过程, 那么 CPU也不知道DMA什么时候传输结束,
- □ 当DMA的数据块传输结束后,需要DMA主动向CPU报告,就需要中断机构, 中断机构的功能是一个数据块传送完毕后,中断机构向CPU提出中断请求,
- □ CPU通过执行中断服务程序,进行DMA传送的后处理。



## DMA 传送过程

(1) 预处理

通过几条输入输出指令预置如下信息

- 通知 DMA 控制逻辑传送方向 (入/出)
- 设备地址 → DMA 的 DAR
- 主存地址 → DMA 的 AR
- 传送字数 → DMA 的 WC

#### CPU

#### 预处理:

主存起始地址 → DMA 设备地址 → DMA 传送数据个数 → DMA 启动设备



## DMA 传送过程

(2) 数据传输阶段,这个阶段不需要CPU的 参与,由DMA控制器控制完成。

CPU

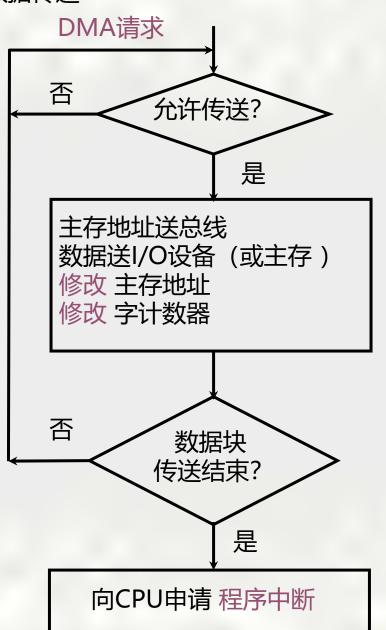
#### 预处理:

主存起始地址 → DMA 设备地址 → DMA 传送数据个数 → DMA 启动设备

#### 数据传送:

继续执行主程序 同时完成一批数据传送

#### 数据传送





(3) 后处理

校验送入主存的数是否正确

是否继续用 DMA

测试传送过程是否正确, 错则转诊断程序

由中断服务程序完成

#### 预处理:

主存起始地址 → DMA 设备地址 → DMA 传送数据个数 → DMA 启动设备

#### 数据传送:

继续执行主程序 同时完成一批数据传送

#### 后处理:

中断服务程序 做 DMA 结束处理

继续执行主程序

#### DMA的过程:

- (1) 当IO设备要传输数据时, IO设备向DMA控制器发送DMA请求;
- (2) DMA控制器接受外设发出的DMA请求,并向CPU发出总线请求;
- (3) 当 CPU 响应此总线请求,发出总线响应信号后,接管对总线的控制,进入DMA操作周期;
- (4) 确定传送数据的主存单元地址及传送长度,并能自动修改主存地址计数值和传送长度计数值
- (5) 规定数据在主存与外设之间的传送方向,发出读/写或其他控制信号,并执 行数据 传送的操作
  - (6) DMA中断机构向CPU报告DMA操作的结束
  - (7) CPU决定是否继续使用DMA



- □ (1) 它使主存与CPU的分开,主存既可被CPU访问,又可被外设访问。
- □ (2) 基本单位是数据块,DMA控制器直接与存储器交互,传送整个数据块,每次传送一个字,所传送的数据,是从设备直接送入内存的,或者相反,这个过程不需要CPU参与。CPU就把控制操作委托给DMA控制器,由该控制器负责处理,在数据块传送时,主存地址的确定、传送数据的计数等都用硬件电路直接实现,整块数据的传送是在 DMA控制器的控制下完成的。
- □ (3) DMA传送速度快,CPU和外设并行工作,提高了系统的效率。
- □ (4) DMA在开始前和结束后要通过程序和中断方式进行预处理和后处理, 也就是仅在传送一个或多个数据块的开始和结束时,才需CPU干预。



# **DMA和中断的区别**

	中断方式	DMA 方式
(1) 数据传送	程序	硬件
(2) 响应时间	指令执行结束	存取周期结束
(3) 处理异常情况	能	不能
(4) 中断请求	传送数据	后处理
(5) 优先级	低	高

# IO通道控制方式

- □ 通道的基本概念
- ①DMA控制器是通过专门设计的硬件控制逻辑来实现对数据传送的控制;而通道则是一个具有特殊功能的处理器,它具有自己的指令和程序,通过执行一个通道程序实现对数据传送的控制,故通道具有更强的独立处理数据输入/输出的功能。
- ②DMA控制器通常只能控制一台或少数几台同类设备;而一个通道则可以同时控制许多台同类或不同类的设备。

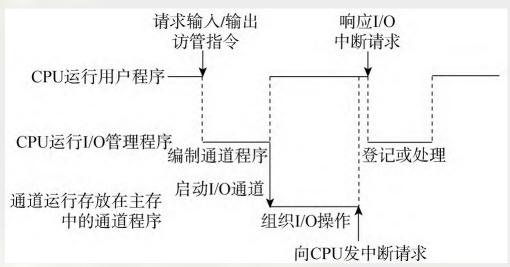
- □ 通道的功能
  - (1) 接受CPU的I/O指令,按指令要求与指定的外设进行联系。
- (2) 从主存取出属于该通道程序的通道指令,经译码后向设备控制器和设备发送各种命令。
  - (3) 实施主存和外设间的数据传送。
- (4) 从外设获得设备的状态信息,形成并保存通道本身的状态信息,根据要求 将这些状态信息送到主存的指定单元,供CPU使用。
  - (5) 外设的中断请求和通道本身的中断请求按次序及时报告CPU。
  - (6) 实施主存和外设间的数据传送。

3、通道的工作过程

通道完成一次数据传输的主要过程分为以下3步。

(1) 在用户程序中使用访管指令进入管理程序,由CPU通过管理程序组织一个

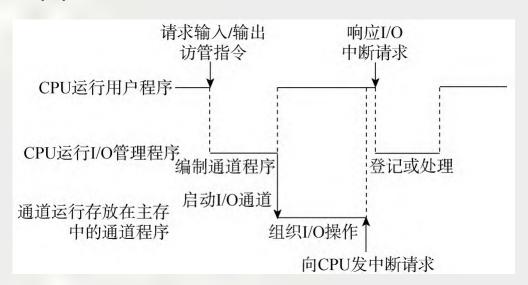
#### 通道程序,并启动通道。



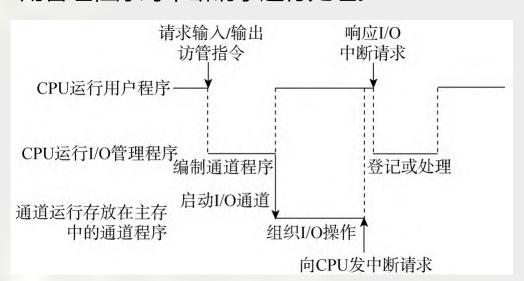
3、通道的工作过程

通道完成一次数据传输的主要过程分为以下3步。

(2) 通道执行CPU为它组织的通道程序,完成指定的数据输入/输出工作。



- 3、通道的工作过程
- 通道完成一次数据传输的主要过程分为以下3步。
- (3) 通道程序结束后向CPU发中断请求,CPU响应这个中断请求后,第二次调用管理程序对中断请求进行处理。



- □ DMA方式与通道方式的主要区别
- (1) 在DMA控制方式中,在DMA控制器控制下设备和主存之间可以成批地进行数据交换而不是CPU干预,这样既减轻了CPU的负担,也大大提高了I/O数据传送的速度。
- (2) 通道控制方式与DMA控制方式类似,也是一种以内存为中心实现设备与内存直接交换数据的控制方式。
- (3) 在通道控制方式中, CPU只需发出启动指令, 指出通道相应的操作和I/O设备, 该指令就可以启动通道并使通道从内存中调出相应的通道程序执行。
- (4)与DMA控制方式相比,通道控制方式所需的CPU干预更少,并且一个通道可以控制多台设备,进一步减轻CPU的负担。对通道来说,可是使用一些指令灵活改变通道程序,这点DMA控制方式无法做到。



- □ I/O处理机通常称作外围处理机。这种外围处理机的结构更接近一般处理机, 甚至就是一般小型通用计算机或微机,它可完成I/O通道所要完成的I/O控制, 还可完成码制变换、格式处理、数据块的检错、纠错等操作。
- □ 它可具有相应的运算处理部件、缓冲部件,还可形成I/O程序所必需的程序转 移手段。

方式	读/写的过程	CPU <b>时间</b>	传输单位	数据流向	
程序查询方式	CPU 发出 I/O 命令后需要不断轮询	极高	字	设备↔CPU↔内存	
程序中断 I/O 方式	CPU 发出 I/O 命令后可以做其他事情,本次 l/O 完成后设备控制器发出中断信号 字 设		设备↔CPU↔内存		
直接存储器存取方式	CPU发出 I/O 命令后可以做其他事情,本次 I/O 完成后 DMA 控制器发出中断信号	中	块	设备↔内存	
I/O 通道控制方式	CPU 发出 I/O 命令后可以做其他事情。通道 会执行通道程序以完成 I/O 操作,完成后通道 向 CPU 发出中断信号	低	多块	设备↔内存	

# 【牛刀小试】

- 1. 【广东工业大学 2014】通道的引入进一步提高整个系统的性能,也有利于系统的安全和保密。关于通道方式,以下说法错误的是()。
  - A. 通道是独立于CPU的专管输入(输出)控制的处理机
  - B. 通道方式在数据传送方式中是对CPU依赖最少的一种方式
  - C. 通道有自己的指令系统和程序
  - D. 通道方式是一种以CPU为中心,实现设备与主存直接交换数据的控制方式
- D【解析】通道是一个独立于CPU的专管输入/输出控制的处理机,它控制设备与主存直接进行数据交换。它有自己的通道指令,这些通道指令受CPU启动,并在操作结束时向CPU发中断信号。通道控制方式是一种以主存为中心,实现设备与主存直接交换数据的控制方式。因此选择D。

- 2. 【广东工业大学 2017】CPU对通道的启动形式是()。

- A. 自陷 B. 中断 C. 通道命令 D. 转移指令
- C【解析】CPU通过通道命令启动通道,指出它所要执行的I/O操作和要访问的设
- 备,通道接到该命令后,便向主存索取相应的通道程序来完成对I/O设备的管理。

- 3. 【广东工业大学 2017】CPU输出数据的速度远远高于打印机的打印速度,为了解决这一矛盾,可采用()。
  - A. 并行技术 B. 通道技术 C. 缓冲技术 D. 虚拟技术
- C【解析】缓冲技术应用在凡是数据到达速度和离去速度不匹配的地方,比如CPU和外设的速度不匹配时。通道技术是实现计算和传输并行的基础,以提高整个系统的效率。虚拟技术主要解决内存不足的问题,当有大文件、程序需要在内存处理时,可以用到虚拟技术。并行技术主要用于多核的CPU或者多I/O总线上并行处理。因此选择C。

- 4. 【重庆理工大学 2015】磁盘属于块设备,磁盘的I/O控制方式主要利用()。
  - A. 程序I/O方式 B. DMA方式
  - C. 程序中断方式 D. SPOOLing方式
- B【解析】磁盘属于块设备,磁盘的I/O控制方式主要利用DMA方式。因此选择B。

- 5. 【重庆理工大学 2016】通道是一种()。
- A. I/O中断口 B. 共享文件 C. I/O专用处理机 D. 数据通道 C【解析】I/O通道的目的是建立独立的I/O通道,使得原来一些由CPU处理的I/O 任务由通道来承担,从而解脱CPU。通道所能执行的命令局限于I/O操作的指令,也就是执行I/O指令集。因此选择C。

- 6. 【中国计量大学 2020】在下面的I/O控制方式中, CPU效率最高的是()。
  - A. 程序I/O方式
- B. 中断驱动I/O方式
- C. 直接存储器访问DMA方式 D. I/O通道控制方式

#### D【解析】

- □ 程序I/O方式是CPU与I/O完全串行工作, CPU效率最低。
- □ 中断驱动I/O方式下,CPU处理程序和外设准备数据并行工作,可以提高CPU 利用率。
- □ DMA方式下,完全由硬件执行I/O交换,其主要优点是数据传送速度很高, CPU根本不参加传送操作,而直接在内存和外设之间进行,DMA以块为单位 进行传送,每块的预处理与后处理需要CPU参与。
- □ I/O通道是DMA的升级版,连续传送若干块才需要CPU的干预。CPU效率最高的是I/O通道控制方式。因此选择D。

- 7. 【燕山大学 2013】()I/O方式中,完全依赖硬件实现了内存和外设间数据的直接传送。
  - A.程序 B.中断驱动 C. DMA控制 D. 通道控制
- C【解析】DMA方式,完全由硬件执行I/O交换,其主要优点是数据传送速度很高, CPU根本不参加传送操作,而直接在内存和外设之间进行。因此选择C。

8. 【南京工业大学 2013】某操作系统中采用中断驱动I/O控制方式,设中断时,CPU用1ms来处理I/O中断请求,其他CPU时间全部用来计算。若系统时钟中断频率为100 Hz,则CPU的利用率为()。

A. 60% B. 70% C. 80% D. 90%

D【解析】首先,根据中断频率确定两次中断的间隔时间: T=1/100=0.01 s=10 ms。其次,在10 ms时间间隔内,CPU需要花费1 ms去处理中断,其他时间(10 ms-1 ms=9 ms)用来进行计算。最后,计算CPU的利用率: p=9 ms/10 ms×100%=90%。因此选择D。

- 9. 【桂林电子科技大学 2018】进程需要读取磁盘上的多个数据块,数据传输方式效率最高的是( )。
  - A. 程序直接控制方式 B. 中断控制方式
  - C. DMA方式 D. 通道方式
- D【解析】程序直接控制方式即循环检测I/O方式,是最古老的方式,数据传输效率最低;中断控制方式每次只能读取一个字,数据传输效率次低;DMA方式每次可以读取一个数据块,数据传输效率较高;通道方式每次可以处理多个数据块,数据传输效率最高。因此选择D。

- 【北京交通大学 2018】 (多选) 就I/O控制方式而言, ()支持内存和外设之 间的直接的数据传输。
  - A. 程序I/O控制方式

- B. 中断驱动I/O控制方式
- C. 直接存储器访问I/O控制方式 D. 通道I/O控制方式

CD【解析】程序控制方式: 特点是依靠程序的控制来实现主机和外设的数据传送, 可分为无条件传送方式和查询方式。中断控制方式:每次输入和输出一个数据, CPU都要检查外设的状态。直接存储器存取控制方式:CPU不参加数据传送,而 是由DMA控制器来实现内存与外设、外设与外设之间的直接传递。通道方式:可 以实现对外围设备的统一管理和外围设备与内存之间的数据传送。因此选择C,D。

- 11. DMA方式是在( )之间建立一条直接数据通路。
  - A. I/O设备和主存 B. 两个I/O设备
  - C. I/O设备和CPU D. CPU和主存

A【解析】直接存储器存取(DMA)是指数据在主存与I/O设备间的直接成块传送,即在主存与I/O设备间传送数据块的过程中,不需要CPU作任何干涉,只需在过程开始启动(即向设备发出"传送一块数据"的命令)与过程结束(CPU通过轮询或中断得知过程是否结束和下次操作是否准备就绪)时由CPU进行处理,实际操作由DMA硬件直接完成,CPU在传送过程中可做其他事情。因此选择A。

- 12. 通道又称I/O处理机,它用于实现()之间的信息传输。
  - A. 内存与外设 B. CPU与外设
  - C. 内存与外存 D. CPU与外存

A【解析】在设置了通道后,CPU只需向通道发送一条I/O指令。通道在收到该指令后,便从内存中取出本次要执行的通道程序,然后执行该通道程序,仅当通道完成了规定的I/O任务后,才向CPU发出中断信号。因此通道用于完成内存与外设的信息交换。因此选择A。

- 13. 在一般大型计算机系统中, 主机对外部设备的控制可以通过通道、控制器和设备3个层次来实现, 以下叙述中正确的是()。
  - A. 设备控制器和通道可以分别控制设备
  - B. 设备控制器、通道和设备可以并行工作
  - C. 通道控制设备控制器、设备控制器控制设备工作
  - D. 以上都不正确
- C【解析】I/O通道设备的引入:独立完成I/O操作的组织、管理及其结束处理,减少CPU对I/O的干预。设备控制器:控制一个或多个I/O设备,以实现I/O设备和计算机之间的数据交换。通道控制设备控制器,设备在控制器控制下工作。因此选择C。

14. 在32位100MHz的单总线计算机系统中(10ns 一个周期),磁盘控制器使用 DMA以40MB/s的速率从存储器中读出数据或向存储器写入数据。假设计算机在 没有被周期挪用的情况下,在每个循环周期中读取并执行一个32位的指令。这样 做,磁盘控制器使指令的执行速度降低了多少?

【解析】首先由题意知DMA的传输速率是40 MB/s, 即4 B/100ns, 也就是平均每100ns传输32bit的数据就能达到DMA的传输要求。

- □ 由于系统总线被CPU和DMA共用,因此要在DMA传输数据时暂停CPU对总线的使用。 为了得到DMA使用总线的频率,需要知道总线的传输速度。
- □ 由题中条件可知, CPU在对总线完全占用的情况下, 每个时钟周期(10ns)可以传输32 bit 的指令, 因此总线的速度是32 bit/10ns。
- □ 而DMA的要求是100ns传输32 bit,也就是说,平均10个时钟周期内,只需挪用1个周期用来传输数据就能达到DMA的传输要求。
- □ 由此可以得到, DMA挪用周期的频率是每10个周期挪用1个, 因此磁盘控制器使指令的执行速度降低了10%。

#### 15. 【西南大学 2014】I/O控制可用哪几种方式,各有什么优缺点?

【解析】I/O控制的实现有4种方式,即程序I/O方式、中断控制方式、DMA控制方式和通道控制方式。

□ 程序I/O方式:其优点是控制简单,不需要很多硬件支持。但CPU和外设之间只能串行工作,并且CPU的大部分时间处于循环测试状态,这使得CPU的利用率大大降低,CPU在一段时间内只能和一台外设交换数据信息,从而不能实现设备之间的并行工作。由于程序I/O方式依靠测试设备状态标志来控制数据传送,因此无法发现和处理因设备或其他硬件所产生的错误。所以,程序I/O方式只适用于执行速度较慢且外设少的系统。

#### 15. 【西南大学 2014】I/O控制可用哪几种方式, 各有什么优缺点?

【解析】I/O控制的实现有4种方式,即程序I/O方式、中断控制方式、DMA控制方式和通道控制方式。

□ 中断控制方式:其优点是能实现CPU和设备、设备与设备间的并行操作,CPU的利用率比程序I/O方式有了很大提高。但I/O控制器的数据缓冲寄存器通常较小,且数据缓冲寄存器装满数据后将会发出中断,因此一次数据传送过程会中断较多次,消耗了大量CPU时间。若系统中配置的外设数目较多,且都以中断方式进行控制,则将消耗大量CPU时间或因为CPU来不及处理而造成数据丢失。

#### 15. 【西南大学 2014】I/O控制可用哪几种方式,各有什么优缺点?

【解析】I/O控制的实现有4种方式,即程序I/O方式、中断控制方式、DMA控制方式和通道控制方式。

□ DMA控制方式:与中断控制方式相比,DMA控制方式的优点是在一批数据传送完成后中断CPU,从而大大减少了CPU进行中断处理的次数,并且DMA控制方式下的数据传送是在DMA控制器控制下完成的,在数据传输过程中无须CPU干预。但DMA方式仍有一定的局限,如对外设的管理和某些操作仍由CPU控制,且多个DMA控制器的使用也不经济。

#### 15. 【西南大学 2014】I/O控制可用哪几种方式,各有什么优缺点?

【解析】I/O控制的实现有4种方式,即程序I/O方式、中断控制方式、DMA控制方式和通道控制方式。

□ 通道控制方式:通道是一个专管输入输出工作的处理器。在通道控制方式下,CPU只需发出I/O指令,通道就能完成相应的I/O操作,并在I/O操作结束时向CPU发出中断信号。由此可见,CPU仅在I/O操作开始和结束时花极短的时间处理与I/O操作有关的事宜,其余时间都与通道并行工作。此外,一个通道还能控制多台设备。但是通道价格较高,从经济的角度出发不宜过多使用。

谢谢大家

# 【真题实战】

- 1、程序员利用系统调用打开 I/O 设备时,通常使用的设备标识是 \_\_\_\_。
- A. 逻辑设 B. 物理设备名
- C. 主设备 D. 从设备号

A【解析】I/O物理设备的种类繁多,为了方便系统对设备的管理和程序对设备的使用,操作系统引入逻辑设备的概念。逻辑设备不是指代某一个具体物理设备,而是对一些设备的统称,比如系统程序中只需要将打印任务给逻辑设备而不需要具体分配到具体的某个打印机,由逻辑设备指定物理设备由操作系统来决定,这样就将CPU专注于程序运行。所以系统要调用I/O设备时使用的是逻辑设备名,操作系统要确定具体设备时使用的是物理设备名。主从设备模式是指主设备将任务进行拆解并下发给从设备执行,从设备执行完毕后将结果发送给主设备进行处理整合成最终结果。这种模式主要用在并行计算中,提高CPU执行效率,而不用于设备调用,所以本题选A项。

- 2、用户程序发出磁盘I/O请求后,系统的正确处理流程是\_\_\_\_。
- A. 用户程序→系统调用处理程序→中断处理程序→设备驱动程序
- B. 用户程序→系统调用处理程序→设备驱动程序→中断处理程序
- C. 用户程序→设备驱动程序→系统调用处理程序→中断处理程序
- D. 用户程序→设备驱动程序→中断处理程序→系统调用处理程序
- B【解析】I/O软件层次结构是指操作系统中用于管理输入输出设备的软件层次结构。包括 以下几个层次:
- □ 用户层软件: 用户层软件通过系统调用来请求I/O操作。
- □ 设备独立性软件:用于实现用户程序与设备驱动器的统一接口、设备命令、设备保护、 以及设备分配与释放等,同时为设备管理和数据传送提供必要的存储空间。
- □ 设备驱动程序: 设备驱动程序负责管理硬件设备,包括初始化、控制和传输数据等操作。
- □ 中断处理程序: 中断处理程序负责处理硬件设备发出的中断信号,以便及时响应设备的请求。

- 2、用户程序发出磁盘I/O请求后,系统的正确处理流程是\_\_\_\_。
- A. 用户程序→系统调用处理程序→中断处理程序→设备驱动程序
- B. 用户程序→系统调用处理程序→设备驱动程序→中断处理程序
- C. 用户程序→设备驱动程序→系统调用处理程序→中断处理程序
- D. 用户程序→设备驱动程序→中断处理程序→系统调用处理程序
- B【解析】当用户使用设备时,首先在用户程序中发起一次系统调用,操作系统的内核接到该调用请求 后请求调用处理程序进行处理,再转到相应的设备驱动程序,当设备准备好或所需数据到达后设备硬件发出中断,将数据按上述调用顺序逆向回传到用户程序中。因此选择B。

- 3、操作系统的 I/O 子系统通常由四个层次组成,每一层明确定义了与邻近层次的接口,其合理的层次组织排列顺序是\_\_\_\_。
- A. 用户级 I/O 软件、设备无关软件、设备驱动程序、中断处理程序
- B. 用户级 I/O 软件、设备无关软件、中断处理程序、设备驱动程序
- C. 用户级 I/O 软件、设备驱动程序、设备无关软件、中断处理程序
- D. 用户级 I/O 软件、中断处理程序、设备无关软件、设备驱动程序
- A【解析】I/O层次结构从上至下为:用户层I/O软件,设备独立性软件,设备驱动程序,中断处理程序。记住即可。

4、用户程序发出磁盘 I/O 请求后,系统的处理流程是:用户程序→系统调用处理程序→设备驱动程序→中断处理程序。其中,计算数据所在磁盘的柱面号、磁头号、扇区号的程序是\_\_\_\_。

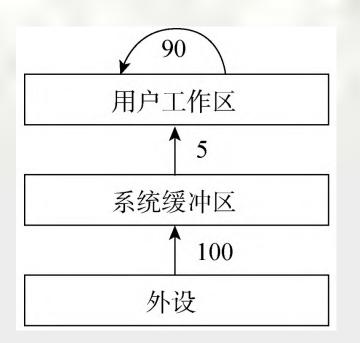
A. 用户程序

B. 系统调用处理程序

C. 设备驱动程序

D. 中断处理程序

C【解析】计算磁盘号、磁头号和扇区号的工作是磁盘设备的地址结构。不同的硬盘其地址 结构因设备硬件的不同而不同,因此应由厂家提供的设备驱动程序实现。所以计算数据所在 磁盘的柱面号、磁头号、扇区号的程序。 5、设系统缓冲区和用户工作区均采用单缓冲,从外设读入1个数据块到系统缓冲区的时间为100,从系统缓冲区读入1个数据块到用户工作区的时间为5,对用户工作区中的1个数据块进行分析的时间为90(如图所示)。进程从外设读入并分析2个数据块的最短时间是



A. 200 B. 295

C. 300 D.390

C【解析】数据块 1 从外设到用户工作区的总时间为 105,在这段时间中,数据块 2 没有进行操作。在数据块 1 进行分析处理时,数据块 2 从外设到用户工作区的总时间为 105,这段时间是并行的。再加上数据块 2 进行处理的时间 90,总共是 300,答案为 C。大家也可以画出草图。

- 6、在系统内存中设置磁盘缓冲区的主要目的是\_\_\_\_\_
- A. 减少磁盘 I/O 次数

B. 减少平均寻道时间

C. 提高磁盘数据可靠性

D. 实现设备无关性

A【解析】磁盘和内存之间存在巨大的速度差异,通常在内存和磁盘之间设置缓冲区,该缓冲区位于内存中。可以将磁盘中的数据先放入到磁盘缓冲区中,之后再从磁盘缓冲区读入到内存,从而减少磁盘 I/O 次数。

请注意,磁盘缓冲区,逻辑上属于磁盘,物理上属于内存。

- 7、下列关于Spooling技术的叙述中, 错误的是\_\_\_\_。
- A. 需要外存的支持
- B. 需要多道程序设计技术的支持
- C. 可以让多个作业共享一台独占设备
- D.由用户作业控制设备与输入/输出井之间的数据传送
- **D**【解析】引入了多道程序技术后,可以利用进程来模拟脱机输入时的外围控制机功能。 这样,外围操作与CPU对数据的处理同时进行,我们把这种在联 机情况下实现的同时外围 操作称Spooling,或称为假脱机操作。因此A正确。

- 7、下列关于Spooling技术的叙述中, 错误的是\_\_\_\_。
- A. 需要外存的支持
- B. 需要多道程序设计技术的支持
- C. 可以让多个作业共享一台独占设备
- D.由用户作业控制设备与输入/输出井之间的数据传送
- D【解析】Spooling系统主要有以下三部分: (1) 输入井和输出井。这是在磁盘上开辟的两个大存储空间。(2) 输入缓冲区和输出缓冲区。为了缓和CPU和磁盘之间速度不匹配的矛盾,在内存中要开辟两个缓冲区: 输入缓冲区和输出缓冲区。(3) 输入进程SPi和输出进程SPo。这里利用两个进程来模拟脱机I/O时的外围控制机,进程SPi模拟脱机输入时的外围控制机,将用户要求的数据从输入机通过输入缓冲区再送到输入井,当CPU需要输入数据时,直接从输入井读入内存; 进程SPo模拟脱机输出时的外围控制机,把用户要求输出的数据先从内存送到输出井,待输出设备空闲时,再将输出井中的数据经过输出缓冲区送到输出设备上。因此B正确,D错误,

- 7、下列关于Spooling技术的叙述中,错误的是\_\_\_\_。
- A. 需要外存的支持
- B. 需要多道程序设计技术的支持
- C. 可以让多个作业共享一台独占设备
- D.由用户作业控制设备与输入/输出井之间的数据传送
- **D**【解析】由输入进程SPi和输出进程SPo控制设备与输入/输出井之间的数据传送。 Spooling系统将独占设备改造为共享设备。因为在Spooling系统中,实际上 并没为任何 进程分配设备,而只是在输入井或输出井中为进程分配一个存储区和建立一张I/O请求表。 这样,便把独占设备改造为共享设备。因此C正确。

- 8、系统将数据从磁盘读到内存的过程包括以下操作:
- ①DMA控制器发出中断请求
- ②初始化DMA控制器并启动磁盘
- ③从磁盘传输一块数据到内存缓冲区
- ④执行 "DMA结束" 中断服务程序

正确的执行顺序是\_\_\_\_。

- A.  $3 \rightarrow 1 \rightarrow 2 \rightarrow 4$  B.  $2 \rightarrow 3 \rightarrow 1 \rightarrow 4$
- C.  $2 \to 1 \to 3 \to 4$  D.  $1 \to 2 \to 4 \to 3$
- B【解析】系统将数据从磁盘读到内存的过程包括以下步骤:
- (1) DMA控制器发出总线请求:在系统将数据从磁盘读到内存的过程中,首先需要通过 DMA控制器向CPU发出总线请求。
- (2) 初始化DMA控制器并启动磁盘:在发出中断请求后,DMA控制器会初始化并启动磁盘。
- (3) 从磁盘传输一块数据到内存缓冲区: 在将数据从磁盘读取到内存缓冲区之前, 需要先将一块数据从磁盘读取到内存中。

- 8、系统将数据从磁盘读到内存的过程包括以下操作:
- ①DMA控制器发出中断请求
- ②初始化DMA控制器并启动磁盘
- ③从磁盘传输一块数据到内存缓冲区
- ④执行 "DMA结束" 中断服务程序

正确的执行顺序是\_\_\_\_。

- A.  $3 \rightarrow 1 \rightarrow 2 \rightarrow 4$  B.  $2 \rightarrow 3 \rightarrow 1 \rightarrow 4$
- C.  $(2) \rightarrow (1) \rightarrow (3) \rightarrow (4)$  D.  $(1) \rightarrow (2) \rightarrow (4) \rightarrow (3)$
- B【解析】系统将数据从磁盘读到内存的过程包括以下步骤:
- (4) DMA控制器发出中断请求:读取完成后,DMA控制器会向系统发出中断请求,通知系统从磁盘读取的数据已经全部读取到内存缓冲区中。
- (5) 执行 "DMA结束" 中断服务程序: 在系统收到中断请求后,会执行 "DMA结束" 中断服务程序,将内存缓冲区中的数据写入磁盘。结合选项,系统将数据从磁盘读到内存的过程为: 初始化DMA控制器并启动磁盘、从磁盘传输一块数据到内存缓冲区、DMA控制器发出中断请求、执行 "DMA结束"中断服务程序。

9、某文件占10个磁盘块,现要把该文件磁盘块逐个读入主存缓冲区,并送用户区进行分析, 假设一个缓冲区与一个磁盘块大小相同,把一个磁盘块读入缓冲区的时间为100us,将缓冲 区的数据传送到用户区的时间是50us, CPU对一块数据进行分析的时间为50us。 在单缓冲 区和双缓冲区结构下,读入并分析完该文件的时间分别是。

A. 1500us, 1000us B. 1550us, 1100us

C. 1550us D. 2000us 2000us

B【解析】单缓冲区下当上一个磁盘块从缓冲区读入用户区完成时下一磁盘块才能开始读入, 也就是当最后一块磁盘块读入用户区完毕时所用时间为150×10 = 1500。加上处理最后一 个磁盘块的时间50,为1550。

100us 50us 50us

100us 50us 50us

100us 50us 50us

100us 50us 50us

9、某文件占10个磁盘块, 现要把该文件磁盘块逐个读入主存缓冲区, 并送用户区进行分析, 假设一个缓冲区与一个磁盘块大小相同,把一个磁盘块读入缓冲区的时间为100us,将缓冲 区的数据传送到用户区的时间是50us, CPU对一块数据进行分析的时间为50us。 在单缓冲 区和双缓冲区结构下,读入并分析完该文件的时间分别是。

A. 1500us, 1000us B. 1550us, 1100us

C. 1550us D. 2000us 2000us

B【解析】双缓冲区下,不存在等待磁盘块从缓冲区读入用户区的问题,采用二级流水线进 行操作,此过程的是100×10+100=1100。

100us	50us 50us						
	100us	50us	50us				
		100us		50us	50us		
				100us		50us	50us

谢谢大家