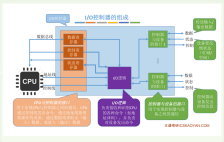
IO管理

I/O管理概述

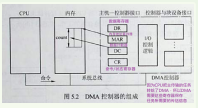
- 定义 - I/O设备是将数据输入输出计算机的外部设备
- 分类 -
  - 按使用特性分类 -
    - 人机交互外部设备 - 用于与用户交互，如打印机、显示器、鼠标键盘
    - 存储设备 - 用于存储程序和数据，如磁盘、光盘、磁带
    - 网络通信设备 - 用于与远程设备通信，如各种网络接口、调制解调器
  - 按传输速率分类 -
    - 低速设备 - 传输速率每秒几字节到几百字节，如：鼠标键盘
    - 中速设备 - 每秒数千字节到数万字节，如打印机
    - 高速设备 - 数百千字节到千兆字节，如磁带机、磁盘、光盘
  - 按信息交换的单位分类 -
    - 块设备 - 信息存取以数据块为单位，每块均设备，速率高，可寻址，如磁盘(可随机读写)
    - 字符设备 - 用于数据输入输出的设备，传输基本单位是字符，无结构类型，传输速率低，不可寻址，ASII码等中断驱动方式及DMA模式，打印机
- I/O控制器 -
  - 设备控制器的功能 -
    1. 接收和识别CPU或通道发来的命令，如磁盘控制器能接收读、写、查找命令
    2. 实现数据交换，通过数据寄存器实现CPU发来的数据，通过数据总线或通道完成与主存数据传输
    5. 向CPU报告设备的状态，通过状态寄存器用于记录I/O设备的当前状态
    4. 设置地址识别，为区分设备控制器的各个寄存器，为各寄存器设置地址，IO控制器通过CPU提供的参数地址来判断CPU要读写哪个寄存器
  - 编址方式 -
    - 内存映像I/O - 控制器中的寄存器与内存地址统一编址
      - 优点 - 简化了指令，可以采用访存指令实现对设备的控制
    - 寄存器独立编址 - 控制器中的寄存器使用各自单独的地址
      - 缺点 - 需要设置专门的指令来操作控制器，需指明寄存器地址，控制器编号
  - 组成 -
    1. 设备控制器与CPU的接口 - 如：数据线、地址线、控制线
    2. 设备控制器与设备的接口 - 一个接口连接一台设备，用于传输数据
    3. I/O控制逻辑 - 用于实现对设备的控制：1.CPU想启动设备时，将启动命令通过控制线发送给控制器。2.控制器对地址线识别地址发送给控制器。3.控制器对从CPU收到的IO命令译码，对所选设备进行控制



- I/O控制方式 -
  1. 程序直接控制方式 -
     - 定义 - 计算机从外部设备读取数据到存储器，每个字CPU必需对外设状态进行循环检测，直到该字传输到CPU控制器的数据寄存器中
     - 缺点/特征 - CPU和IO设备只能串行工作，因为CPU同每个字需要一段时间检查，导致CPU长期忙等，利用率低；CPU干预频繁，每次读写一个字

  2. 中断驱动方式 -
     - 定义 - 引入中断机制，IO速度慢，当CPU发出读写命令后，将等待IO的进程阻塞，切换到别的进程执行。中断机制允许I/O设备控制器发出中断信号至处理器CPU的设计开源思路，从而解放CPU做别的工作；CPU检测到中断信号后保存当前进程的运行环境，专门执行中断处理程序，IO完成后，CPU恢复CU继续执行被中断的进程
     - 细节补充 -
       1. CPU会在每个指令周期的末尾检查中断(内中断)
       2. 中断过程要保存恢复进程的运行环境，开销大，所以中断频率不能太高
     - 优缺点/特征 - 比程序直接控制方式有改，但数据的每个字在内存缓冲区与I/O控制器之间的传输都要经过CPU，消耗大多CPU时间

  3. DMA方式 -
     - 定义 -
       DMA：直接存储器存取
       - 设IO模块发送读块命令 - CPU→DMA；…→ 继续处理 - DMA模块可以通过发送请求信号，请求CPU暂停一个或多个周期，使用系统总线，CPU暂停期间，DMA模块直接向内存写入一个字，CPU暂停期间，CPU可以继续处理其他任务，不会因为等待IO而空闲
       - 设DMA模块接收到状态 - …← 中断；DMA→CPU - 当设置中断DMA模块向CPU发送中断信号，CPU暂停当前程序，转入中断处理程序，DMA模块通知CPU数据已经写入内存，CPU可以继续使用该数据
       - 下一条指令
     - 与中断驱动方式的区别及改进点 -
       1. 数据的传送单位是"块"，不再是一个字传送
       2. 数据的流向是从设备直接送入内存，或者从内存直接到设备，不需要CPU做"中转"的最小单位
       5. 仅在传输一个或多个数据块的开始时，才需要CPU干预，数据传送由DMA控制器控制下完成，不再是CPU控制
     - 优缺点 -
       - 优点 - 数据传输以块为单位，CPU介入频率降低
       - 缺点 - 数据传输不再经过CPU再写入内存，传输效率高；CPU和DMA设备的并行性得提升
       - 缺点 - CPU每发出一条IO指令，只能读写一个或多个连续的数据块。如果要读写离散的数据块，或写到不同内存区域，CPU就要分别发出多个IO指令，经过多次中断处理才能完成
       - I/O的本质仍是"内存"与"设备"
     - DMA控制器的组成 -



  4. 通道控制方式 -
     - 定义 - 通道是一种具有一定处理能力的特殊的IO设备，它能执行一系列通道命令，1.当CPU要完成一组或组成的数据的有关控制时，只需向IO通道发送一条IO指令，并给出通道程序的首地址、要访问的设备等信息，2.通道接收指令后，执行任务，5.数据传输结束后向CPU发出中断请求

     - IO通道与一般处理机区别 -
       - 通道指令类型单一
       - 通道没有自己的内存，通道执行的通道程序是放在主机的内存中，即通道共享CPU的内存
     - IO通道与DMA方式的区别 -
       - 数据传送的单位大小不同，每次读写一组数据块
       - DMA是要CPU来控制传输的数据块大小、传输的内存位置，而通道方式中传通道控制应该自主做决定
       - 每个DMA控制器只对应一个设备与内存传送数据，而一个通道可以控制多台设备与内存的数据交换
     - 优点 - CPU、通道、IO设备可并行工作，资源利用率高
     - 总结与对比 -

	完成一次数/字所需时间	CPU干预频率	每次IO传输单位	数据流向	优缺点
程序直接控制方式	CPU花费IO命令后需要不断轮询	很高	字	设备→CPU→内存 内存→CPU→设备	每一个字节的读入或写出都需要一定的CPU时间
中断驱动方式	CPU花费IO命令后可以继续做事，当IO完成后通知设备控制器当中断信号	高	字	设备→CPU→内存 内存→CPU→设备	数据传输以字为单位，每个字节的读写都需要中断处理
DMA方式	CPU花费IO命令后可以继续做事，当IO完成后通知DMA控制器当中断信号	中	块	设备→内存 内存→设备	DMA控制器可以直接访问内存，在DMA控制器控制下，设备与内存的数据交换不再经过CPU，减轻了CPU的负担，适用于大批量数据的传输
通道控制方式	CPU花费IO命令后可以继续做事，通道会执行通道程序可以实现IO，完成后通道向CPU发出中断信号	低	一组块	设备→内存 内存→设备	通道是一种具有一定处理能力的特殊的IO设备，适用于多道及实现数据密集型任务。

- 最大限度的减少CPU的干预
- 减少单位数据的处理时间

I/O子系统的基本结构

- 为什么分层 -
  - IO软件涉及面广，往下与硬件紧密联系，往上与用户直接交互，它完成设备管理、存储器管理、文件管理等的功能系统
  - 为了使复杂的IO软件具有清晰的结构、良好的可移植性、扩展性、适应性，在IO软件中采用层次结构
  - 每层利用其下层提供的服务，完成IO控制中的某些子功能，并屏蔽该层功能的细节，向高层提供自己本层应子功能服务
  - 在层次结构中，只要层次间的接口不变，对某一层次中的软件的修改都不会引起其下层或高层代码的变更，仅要能够不设计硬件的具体特性
  - 易于实现 - 易于调试
  - 易于实现功能"接口"设定
- 层次结构 -
  1. 用户层I/O软件 -
     - 系统调用
     - 高层IO模块
     - 1. 实现与用户交互的接口，如IO等函数printf
     - 2. 用户层软件将用户请求翻译成IO请求，通过"系统调用"来请求操作系统内核下层软件服务
     - 1. 向用户提供同一接口，如read/write系统命令
  2. 设备独立性软件 -
     - 3. 执行所有设备的共有操作：如对设备的分配与回收，将逻辑设备名映射为物理设备名，对设备进行用户权限保护，缓冲管理，差错管理，提供统一大小的逻辑块以屏蔽各设备之间信息交换的单位与传输速率的差异
  3. 设备驱动程序 -
     - 1. 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
     - 2. 不同IO设备有不同的硬件特性，所以需要不同的设备驱动程序来以正确的IO命令控制主机
  4. 中断处理程序 - 用于保存被中断进程的CPU环境，转入相应的中断处理程序进行处理，处理完后恢复中断进程的现场，唤醒阻塞中断进程
  5. 硬件设备 - 机械部件设备外部和电子部件设备控制器组成

知识点回顾与重要考点
I/O请求
用户层软件
系统调用
设备独立性软件
设备驱动程序
中断处理程序
硬件
I/O设备
I/O软件的层次
内核态
用户态
IO设备
IO设备
IO设备