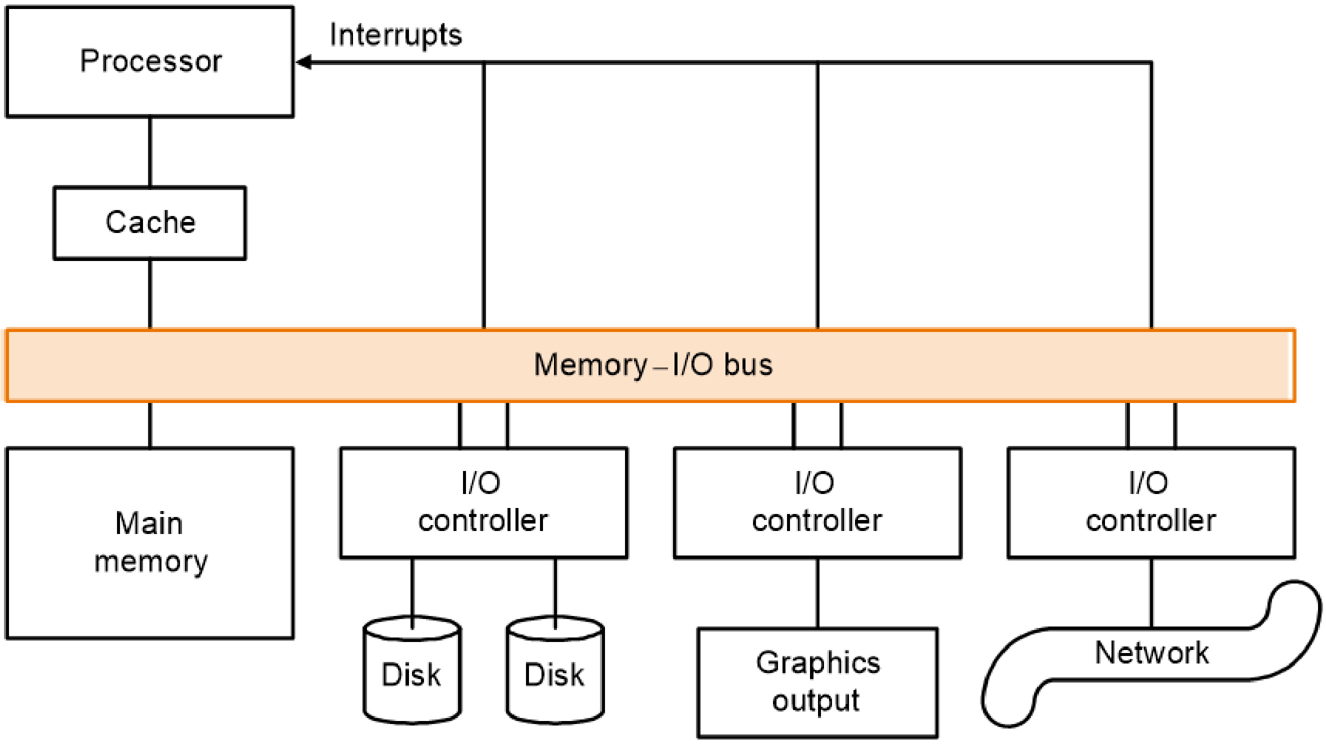
# 存储器和IO设备

## I/O系统

I/O系统的表现难以估量，与设备与系统间的关系、内存的等级制度、操作系统均有关。




三个特点：

- Behavior

  Input (read once), output (write only, cannot read) ,or storage (can be reread and usually rewritten)

- Partner

  Either a human or a machine is at the other end of the I/O device, either feeding data on input or reading data on output.

- Data rate

  The peak rate at which data can be transferred between the I/O device and the main memory or processor.

I/O性能与应用有关：

- throughout 吞吐量——与bandwidth 带宽有关。
  - 特定时间内我们能够移动多少数据
  - 一个单位时间内我们能做多少I/O操作。
- 响应时间

## 硬盘存储和可靠性

floppy disk

**hard disk**

- 磁盘platters: disk consists of a collection of *platters*, each of which has two recordable disk surfaces
- 磁道tracks: each disk surface is divided into concentric circles同心圆
- 扇区 sectors: each track is in turn divided into *sectors*, which is the smallest unit that can be read or written

**时间**

磁道->扇区->磁盘

- Seek: position read/write head over the proper track
    - minimum seek time
    - maximum seek time
    - average seek time (3 to 14 ms) 固定
- Rotational latency: wait for desired sector（需计算）
    - 转半圈是要找单个扇区用的平均时间，如果你要读一整圈上的所有扇区后续应该还要加一个读一圈的时间。
    - RPM:Rotation per minute
    - Average——转半圈，故时间=半圈/一分钟转多少圈
    - $$\textit{Average rotational latency} = \frac{0.5\ \text{rotation}}{5400RPM} = \frac{0.5\ \text{rotation}}{5400RPM \ \Big/ \left(60\ \frac{\text{seconds}}{\text{minute}}\right)} = 0.0056\ \text{seconds} = 5.6\ ms$$
- Transfer: time to transfer a sector (1 KB/sector): function of rotation speed, Transfer rate of today's drives - 30 to 80 MBytes/second （需计算）
- 传输时间是一个扇区的大小/速度=0.5KB（一个扇区512K）/50MB/sec=0.01ms
- Disk controller, which controls the transfer between the disk and the memory固定。假设0.2

{ Access Time = Seek time + Rotational Latency + Transfer time + Controller Time }

MTTF mean tine to failure

MTTR mean time to repair

MTBF (*Mean Time Between Failures*)* = *MTTF+ MTTR*

{Availability=MTTF/(MTTF+MTTR)}

(Availability:service still provided to user, even if some components failed)即使部件失败仍能听歌服务

RAID: Redundant Arrays of Inexpensive Disks以代替更大的硬盘

Array\ Reliability:

Reliability of N disks = **Reliability of 1 Disk ÷ N**

Too unreliable->Redundant Arrays of (Inexpensive) Disks-->容量和带宽上的罚时

## Buses and Other Connections between Processors Memory, and I/O Devices

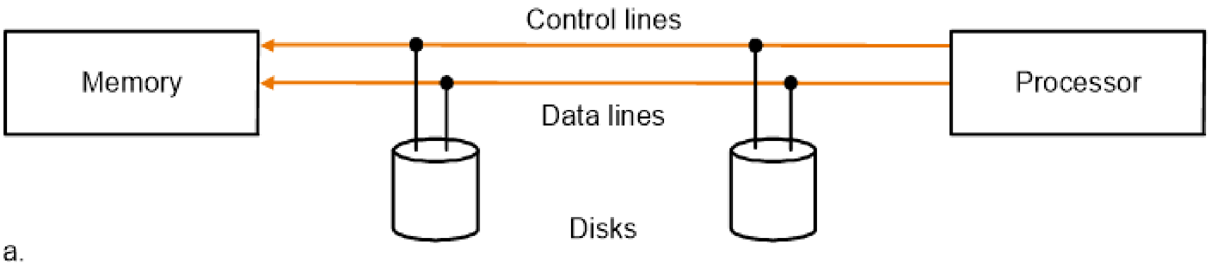
Bus->总线 Shared communication link， 总线有两条线，一条控制线传送信号和请求，一条数据线传送信息（数据、地址、命令）。

- Control lines：which are used to **signal requests** and

  **acknowledgments**, and to indicate what types of

  information is on the data lines.
- Data lines：which **carry information** (e.g., data, addresses, and complex commands) between the source and the destination
- Transaction:事务——input&output，传送地址，收取或发送数据
  - input: inputting data from the device to memory

    控制线发送写请求，数据线发送地址；内存准备好后发送信号；处理器数据线发送数据；内存接受数据并存储。
  - output: outputting data to a device from memory

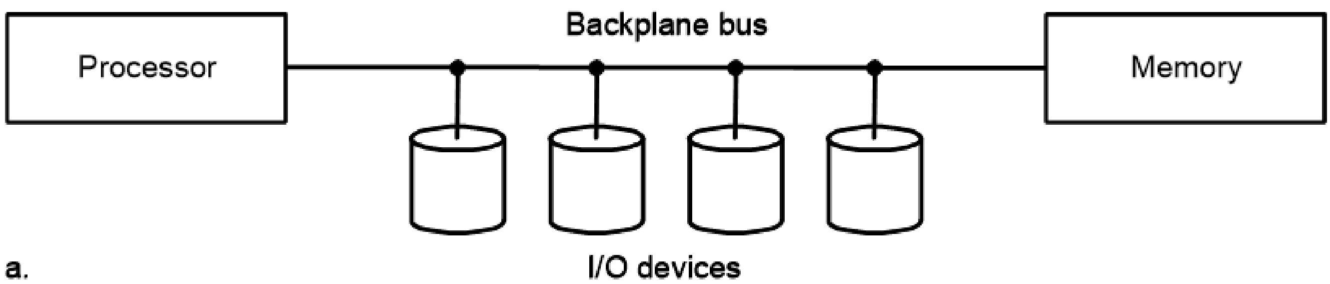
    
    
    a.

    控制线向内存发送读请求，数据线发送读取地址；内存获取数据，内存通过数据线返回数据。

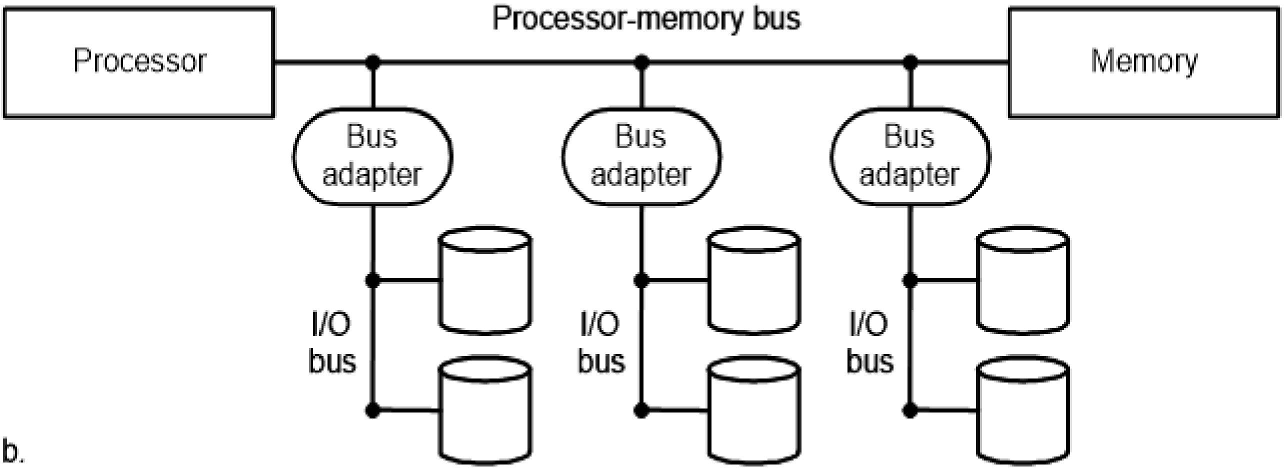
## 总线类型

- processor-memory (short high speed, custom design)
- backplane (high speed, often standardized, e.g., PCI)底板
- I/O (lengthy, different devices, standardized, e.g., SCSI)

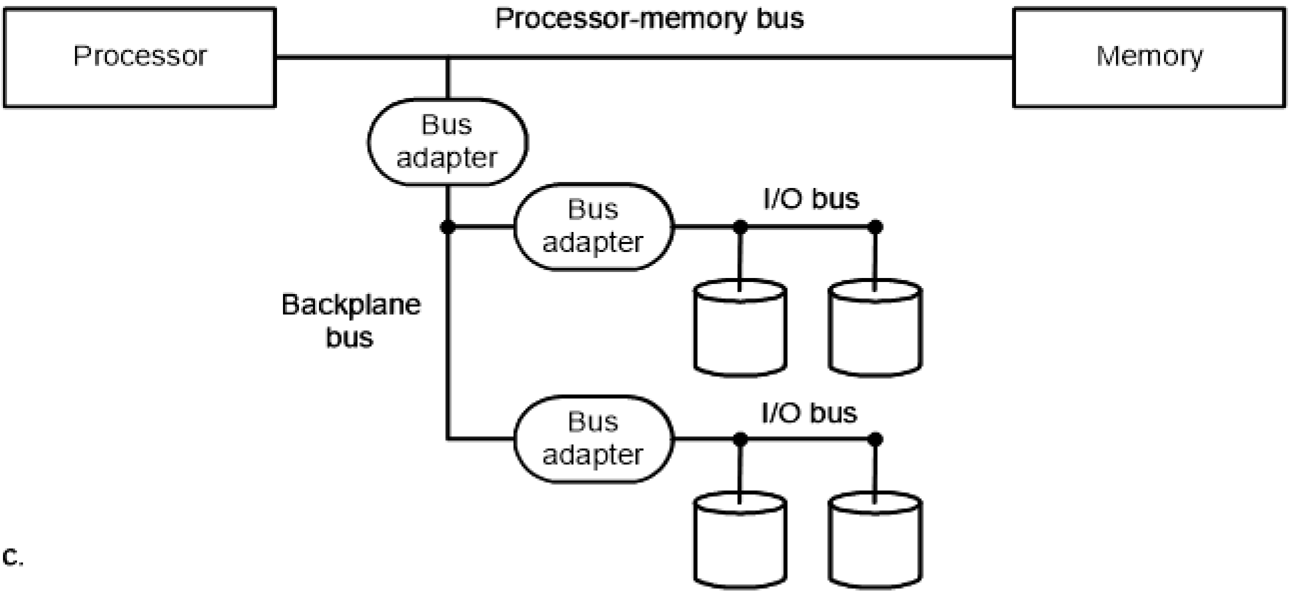
Old——single bus 同时进行内存和处理器、内存和IO设备的交流




separate bus。 IO bus 使用 bus adapter介入Processor-memory bus




b.

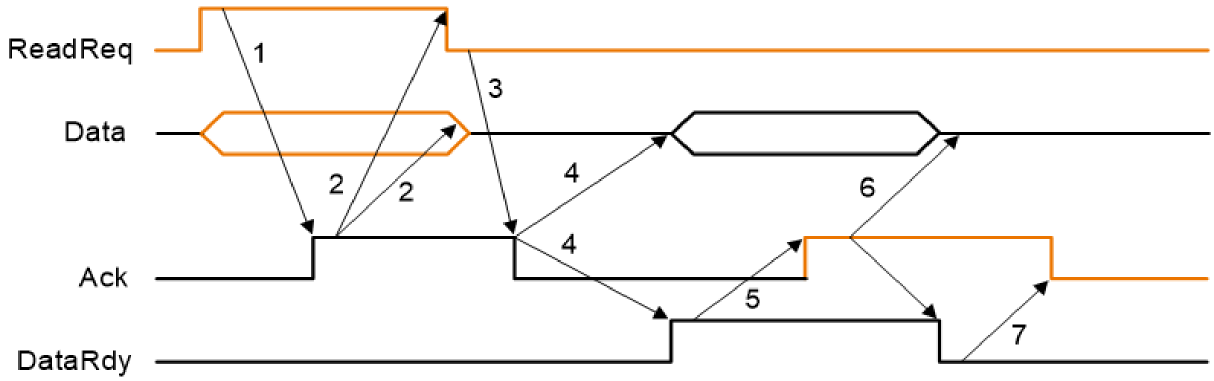



c.

**Synchronous vs. Asynchronous**

- 同步：时钟+同步协议，每个设备必须以相同的速度运行，时钟要求总线短
- 异步：握手
  - 三种控制流程：ReadReq（***address is***

    put on the data lines),DataRdy（***data word is now ready on the data lines***），Ack(acknowlege the signal)
  - 七步：



Data中第一块是IO要求地址，第二块是Memory找出的数据

高低：I/O device saw the Ack line high and releases the ReadReq data lines.Memory sees that ReadReq is low and drops the Ack line.

4. When the memory has the data ready, it places the data on the data lines and raises ***DataRdy***.
5. The I/O device sees DataRdy, reads the data from the bus , and signals that it has the data by raising ***ACK***.
6. The memory sees Ack signals, drops ***DataRdy***, and releases the data lines.
7. Finally, the I/O device, seeing DataRdy go low, drops the ***ACK*** line, which indicates that the transmission is completed.

**获取总线控制 ,需要一个管理者控制所有请求**

bus master:processor is always a bus master. e.g. IO设备向Processor发送请求，Processor响应并产生对应的控制信号，告诉IO设备你的请求正在处理，IO设备随后将数据放在数据line上

## Bus Arbitration总线仲裁

多个设备共享总线，同时发送数据时将产生总线竞争最终通信失败，需要引入仲裁机制决定什么时间谁来占用总线。决定哪个bus master来使用bus。

四种模式：

- daisy chain 阻塞；不公平
- centralized parallel (如PCI) 中心化，并行，需要一个arbiter（仲裁者）
- self selection NuBus
- collision detection Ethernet

决定因素：bus priority，fairness

## 接口 Interfacing I/O Devices to the Memory, Processor, and Operating System

Three **characteristics** of I/O systems

- shared by multiple programs using the processor.多个程序共享
- often use **interrupts** to communicate information about I/O operations.进程终端来交流信息
- The low-level control of an I/O devices is complex底层控制复杂

Three types of **communication** are required:

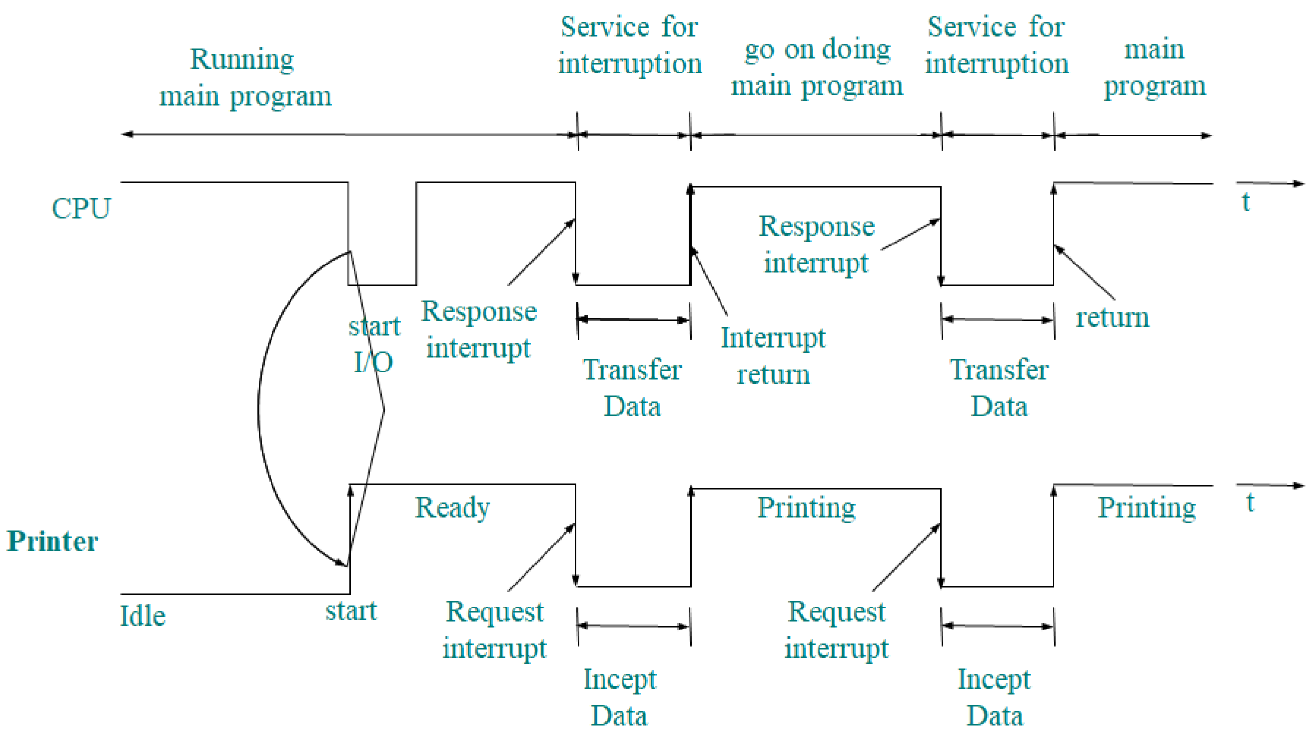
- The OS must be able to give commands to the I/O devices.操作系统需要传递命令给IO设备
  - 两种方法：
  - memory-mapped I/O:portions of the memory address space are assigned to I/O devices,and lw and sw instructions can be used to access the I/O port.内存映射
  - special I/O instructions：Give a command to an I/O device
  - command/data port:**The Status register** (a done bit, an error bit……),The Data register, The command register
- The device must be able to notify the OS, when I/O device completed an operation or has encountered an error. 设备需要有能力告知操作系统
- Data must be transferred between memory and an I/O device数据需要在内存和IO设备间传输

## 重点：IO 设备和处理器间的交流

I/O SYTEM DATA TRANSFER CONTROL MODE

- Polling轮询: The processor periodically checks status bit to see if it is time for the next I/O operation. 浪费时间。周期性检查状态位看是否需要进行下一个IO操作
- Interrupt: When an I/O device wants to notify processor that it has completed some operation or needs attentions, it causes processor to be interrupted.中断：IO完成操作后发送信号使Processor中断。在读写数据时可以做其他工作
    - 优点：并行操作
- DMA *(direct memory access):* the device controller transfer data directly to or from memory without involving processor.有一个设备管理器，直接传输数据给IO和Processor，不需要处理器的控制，节省处理器的时间

## 中断驱动的IO模式（？）




## DMA传输模式

CPU初始化DMA后（需要设备身份、操作、内存地址、数据大小），由Memory和IO Device直接传输，IO和Memory亦可直接与CPU交流，即不存在Memory和IO DEVICE之间直接交互。当DMA传输完成后，控制器发送中断信号给处理器检查是否出现错误。

## Overhead（经费，代价？） of polling in an I/O System

假设：

- number of clock cycles for a polling operation is 400 一次轮询所需要的clock cycle为400
- processor executes with a 500-Mhz（兆赫兹） clock.， 500Mhz M是10^6，即一秒钟有500*10^6的clock cycle

求：Determine the **fraction of CPU time** consumed for the mouse, floppy disk, and hard disk.

假设：

1. The mouse must be polled 30 times per second to ensure that we do not miss any movement made by the user.

   鼠标一秒轮询30次。

2. The floppy disk transfers data to the processor in 16-bit units and has a data rate of 50 KB/sec. No data transfer can be missed.

   软盘16位数据，50KB/s

3. The hard disk transfers data in four-word chunks and can transfer at 4 MB/sec. Again, no transfer can be missed.

   四个word=八个字节=32位数据，4MB/s

计算：

> 鼠标：
>
> 30*400=12000 cycles/sec
>
> Fraction =12000/(500*10^6)=0.002%

> 软盘
>
> 16 bit =2B
>
> 50KB/s 每次轮询传输2B数据，故每秒需要25K次轮询
>
> 一次轮询需要400 cycles
>
> 故一秒需要400*25K=10*10^6
>
> Fraction：10*10^6/(500*10^6)=2%

> 硬盘
>
> （全错）4MB/s 4B 每次 故4*2^18 次轮询/s
>
> 每次轮询400 cycle
>
> 故100*2^22个cycle每秒。
>
> 故（50（全错）
>
> 正确：
>
> 4MB/s 4-word （一个word四个字节！！ 傻逼！！！）16B 故250K
>
> 250K*400=100000K=100*10^6个cycle
>
> 20%

所以鼠标可以使用轮询，但硬盘不可以。

标准答题格式：

- The number of polling access per second
- Clock cycles per second for polling
- Fraction of the processor clock cycles consumed

## Overhead of Interruprt-Driven I/O

假设：

- 一次中断需要500个clock cycle，如果硬盘只在5%的时间里传输数据，求Fraction。
- processor executes with a 500-Mhz（兆赫兹） clock.， 500Mhz M是10^6，即一秒钟有500*10^6的clock cycle
- The hard disk transfers data in four-word chunks and can transfer at 4 MB/sec. Again, no transfer can be missed.

  四个word=八个字节=32位数据，4MB/s

**注意：这边的K和M我们在最终计算时用的是10^3/10^6来做**

解题：

> 如果100%传输数据。
>
> 250K*500=125*10^6 cycles per sec
>
> 25%
>
> 5%的时间内传，则25%*5%=1.25%
>
> 不传输数据时不消耗CPU时间。

## Overhead of I/O Using DMA

初始化需要1000 cycle， DMA完成发送中断需要500 cycle。平均一次需要传送的数据大小为8KB。4MB/s。100%时间传送。

Fraction

解题：

> 每次 8KB 一秒4MB，那么传输一次需要2*10^-3S，则一秒内传输1/2*10^3次
>
> 一次传输需要500（中断）+1000（初始化）
>
> 1500*1/2*10^3=750*10^3 cycle
>
> 总共500*10^6
>
> 故0.2%。

相邻磁道寻道时间=(最大时间-最小时间)/磁道数=(8ms-4ms)/2K=0.002ms

读取连连看的时间=控制时间+(0磁道寻道时间+半圈时间+读取1圈的所有扇区)+(0磁道换到1磁道时间+半圈时间+读取1圈的所有扇区）+…+(5磁道换到6磁道时间+半圈时间+读取61/64圈的扇区)+传输时间

这个半圈是什么

17.937ms读取了1.78M数据，带宽为一秒能传输的数据量=1.78MB/17.937ms=99.2MBps=793.9Mbps

17.937ms读取了一个文件，即进行了一次IO操作，吞吐率为1秒能进行几次IO操作，因此吞吐率=1/17.937=55.75(tps：每秒事务数)