第八章 外存、网络和其他外设

作业参考答案

8.1 [10] <§§8.1–8.2> Here are two different I/O systems intended for use in transaction processing:

- System A can support 1500 I/O operations per second.
- System B can support 1000 I/O operations per second.

The systems use the same processor that executes 500 million instructions per second. Assume that each transaction requires 5 I/O operations and that each I/O operation requires 10,000 instructions. Ignoring response time and assuming that transactions may be arbitrarily overlapped, what is the maximum transaction-per-second rate that each machine can sustain?

参考答案：

处理器如果**100%**进行处理处理，则每秒钟最多处理的事务个数为：

500M/（5x10000）= 10000

系统**A**每秒钟最多处理的事务个数为：

1500/5= 300 (< 10000)

系统**B**每秒钟最多处理的事务个数为：

1000/5= 200 (< 10000)

8.8 [10] <§8.3> Consider two RAID disk systems that are meant to store 10 terabytes of data (not counting any redundancy). System A uses RAID 1 technology, and System B uses RAID 5 technology with four disks in a "protection group."

a. How many more terabytes of storage are needed in System A than in System B?

b. Suppose an application writes one block of data to the disk. If reading or writing a block takes 30 ms, how much time will the write take on System A in the worst case? How about on System B in the worst case?

c. Is System A more reliable that System B? Why or why not?

参考答案:

a. 系统**A**为**RAID1,** 采用磁盘镜像，所以，所用磁盘容量为**10+10=20TB**
系统**B**为**RAID5,** 采用一个奇偶校验盘，所用磁盘容量为**10+10x1/4=12.5TB**
所以，**A**比**B**多**20-12.5=7.5TB**

b. 最坏的情况下，系统**A**写一块数据的时间为**2x30=60ms;** 系统**B**写一块数据的时间为**2**次读和**2**次写，所以为**4x30=120ms**

c. 系统**A**比系统**B**更可靠，直观上，系统**A(RAID1)**的冗余信息比系统**B (RAID5)**多，可靠性应更高。例如，当同一组数据发生两位以上错误时，系统**B**无法恢复，但系统**A**能恢复。

8.9 [15] <§8.3> What can happen to a RAID 5 system if the power fails between the write update to the data block and the write update to the check block so that only one of the two is successfully written? What could be done to prevent this from happening?

参考答案：

　　对于**RAID5**来说，如果在写完数据块但未完成校验块时发生断电，则写入的数据和对应的校验信息不匹配，无法正确地恢复数据。

　　这种情况可以避免，只要同时写数据和校验两个块就行。

8.12 [5] <§8.3> A secret agency simultaneously monitors 100 cellular phone conversations and multiplexes the data onto a network with a bandwidth of 5 MB/sec and an overhead latency of 150 μs per 1 KB message. Calculate the transmission time per message and determine whether there is sufficient bandwidth to support this application. Assume that the phone conversation data consists of 2 bytes sampled at a rate of 4 KHz.

参考答案：

每秒100个会话对应的数据量为：**4K x 2B x100 =800KB/s**

每个**message**为**1KB**，所花的传输时间为：**1KB/5MB+150/10**^{6} **=0.35x10**$^{-3}$**s**

（如果传输**800KB**的时间比**1**秒钟多，则说明带宽不够。）

传输**800KB**的时间为：

800KB / 1KB x 0.35x10^{-3} **= 0.28 s < 1s**

说明带宽足够！

8.14 [5] <§8.3> There are two types of parity: even and odd. A binary word with even parity and no errors will have an even number of 1s in it, while a word with odd parity and no errors will have an odd number of 1's in it. Compute the parity bit for each of the following 8-bit words if even parity is used:

a. 01100111   1

b. 01010101   0

8.15 [10] <§8.3>

a. If a system uses even parity, and the word 0111 is read from the disk, can we tell if there is a single-bit error?

b. If a system uses odd parity, and the word 0101 appears on the processor-memory bus, we suspect that a single-bit error has occurred. Can we tell which bit the error occurs in? Why or why not?

c. If a system uses even parity and the word 0101 appears on the processor-memory bus, can we tell if there is a double-bit error?

参考答案：**a.** 不能说明，也可能是三位出错（可说明有奇数位出错）

b. 不能确定哪位出错，而只能判断有奇数位出错

c. 可能无错或偶数位出错

8.16 [10] <§8.3> A program repeatedly performs a three-step process: It reads in a 4 KB block of data from disk, does some processing on that data, and then writes out the result as another 4 KB block elsewhere on the disk. Each block is contiguous and randomly located on a single track on the disk. The disk drive rotates at 10,000 RPM, has an average seek time of 8 ms, and has a transfer rate of 50 MB/sec. The controller overhead is 2 ms. No other program is using the disk or processor, and there is no overlapping of disk operation with processing. The processing step takes 20 million clock cycles, and the clock rate is 5 GHz. What is the overall speed of the system in blocks processed per second?

参考答案：

操作步骤：读 - 处理 – 写， 每块4KB。

平均旋转等待时间为：**60000 / (2x10000) = 3ms**

每块数据的平均访问时间为：**3+8+2+1000x4KB/50MB=13.08ms**

处理所花时间为：**1000 x 20M x 1/5G = 4ms**

整个操作所花时间为：**13.08x2+4 ≈30ms**

所以，整体上每秒钟处理的块数为：**1/30=33.3**

8.18 [5] <§§8.3, 8.5> Suppose we have a system with the following characteristics:

1. A memory and bus system supporting block access of 4 to 16 32-bit words.
2. A 64-bit synchronous bus clocked at 200 MHz, with each 64-bit transfer taking 1 clock cycle, and 1 clock cycle required to send an address to memory.
3. Two clock cycles needed between each bus operation. (Assume the bus is idle before an access.)
4. A memory access time for the first four words of 200 ns; each additional set of four words can be read in 20 ns.

Assume that the bus and memory systems described above are used to handle disk accesses from disks like the one described in the example on page 570. If the I/O is allowed to consume 100% of the bus and memory bandwidth, what is the maximum number of simultaneous disk transfers that can be sustained for the two block sizes?

参考答案：对于具有**1～4**特点的系统，我们得到结果如下：

4-字块传送的总线带宽为： **71.11MB/s**

16-字块传送的总线带宽为： **224.56MB/s**

给定磁盘的传输速度为**50MB/s**，当整个总线带宽都为**I/O**服务时，情况如下：

对于**4-**字块传送的情况，最多只能接到**[71.11/50]=1**个磁盘进行传送

对于**16-**字块传送的情况，最多能同时接**[224.56/50]=4**个磁盘进行传送。

复习：数据块大小对带宽的影响

假定有一个系统具有下列特性：

(1)系统支持**4**～**16**个**32**位字的块访问。

(2)64位同步总线，时钟频率为**200MHz**，每个**64**位数据传输需一个时钟周期，地址发送到存储器需**1**个时钟周期。

(3)在每次总线操作**(**事务**)**间有两个空闲时钟周期。

(4)存储器访问对于开始的**4**个字是**200ns**，随后每**4**个字是**20ns**。假定最近读出的数据的总线传送和随后**4**个字的存储器读操作可以重叠进行。

请求出分别用**4-**字块和**16-**字块方式读取**256**个字时的持续带宽和等待时间。并且求出两种情况下每秒钟内的有效总线事务数。一个总线事务是由一个地址传送后跟一个数据块传送组成的。

复习：数据块大小对带宽的影响

分析 **4-**字块传送情况：

对于**4-**字块传送方式，一次总线事务由一个地址传送后跟一个**4-**字块的数据传送组成。也即每个总线事务传送一个**4**个字的数据块。

每个数据块所花时间为：

(1) 发送一个地址到主存花一个时钟周期

(2) 从主存读**4**个字花：**200ns/(5ns/Cycle)=40**个时钟周期

(一个周期是10^9**ns/200MHz=1000/200=5ns)**

(3) 4个字（**128**位）的传输需**2**个时钟周期

(一个**64**位数据传输需**1**个时钟周期**)**

(4) 在这次传送和下次之间有**2**个空闲时钟周期

所以一次总线事务总共需**45**个周期，**256**个字需**256/4=64**个事务，所以整个传送需**45x64=2880**个时钟周期，因而总等待时间为：**2880**周期**x 5ns/**周期**=14400ns**。每秒钟的总线事务数为：**64 x (1s/14400ns) = 4.44M**个。总线带宽为：**(256 x 4B)/14400ns =71.11MB/s**。

复习：数据块大小对带宽的影响

分析 **16-**字块传送情况：

对于**16-**字块传送，一次总线事务由一个地址传送后跟一个**16-**字块的数据传送组成。也即每个总线事务传送一个**16**个字的数据块。

第一个**4-**字所花时间为：

(1) 发送一个地址到主存花一个时钟周期

(2) 从主存读开始的**4**字花：**200ns/(5ns/Cycle)=40**个时钟周期

(3) 4个字传输需**2**个时钟周期，在传输期间存储器开始读取下一个**4**字

(4) 在本次和下次之间有**2**个空闲时钟，此期间下一个**4**字已读完

所以，**16**字中其余三个**4**字只要重复上述最后两步。因此对于**16-**字块传送，一次总线事务共需花费的周期数为：**1+40+4 x (2 +2) = 57**个周期，**256**个字需**256 / 16=16**个事务，因此整个传送需**57 x 16 = 912**个时钟周期。故总等待时间为：**912**周期**x 5ns /** 周期**=4560ns**。几乎仅是前者的**1/3**。每秒钟的总线事务个数为：**16 x (1s / 4560ns) = 3.51M**个。总线带宽为：**(256 x 4B)x (1s/4560ns) =224.56MB/s**，比前者高**3.6**倍。

由此可见，大数据块传输的优势非常明显。

8.19 [5] <§8.5> In the system described in Exercise 8.18, the memory system took 200 ns to read the first four words, and each additional four words required 20 ns. Assuming that the memory system takes 150 ns to read the first four words and 30 ns to read each additional four words, find the sustained bandwidth and the latency for a read of 256 words for transfers that use 4-word blocks and for transfers that use 16-word blocks. Also compute the effective number of bus transactions per second for each case.

参考答案：假定存储器访问对于开始的4个字是150ns，随后每4个字是30ns。

分析 4-字块传送情况：

对于4-字块传送，一次总线事务由一个地址传送后跟一个4-字块的数据传送组成。也即每个总线事务传送一个4个字的数据块。

每个数据块所花时间为：

(1) 发送一个地址到主存花1个时钟周期

(2) 从主存读4个字花：150ns/(5ns/Cycle)=30个时钟周期

(一个周期是10^9ns/200MHz=1000/200=5ns)

(3) 4个字（128位）的传输需2个时钟周期

(一个64位数据传输需一个时钟周期)

(4) 在这次传送和下次之间有2个空闲时钟周期

所以一次总线事务总共需35个周期，256个字需256/4=64个事务，所以整个传送需35x64=2240个时钟周期，因而总等待时间为：2240周期x 5ns/周期=11200ns。每秒钟的总线事务数为：64 x (1s/11200ns) = 5.71M个。总线带宽为：(256 x 4B)/11200ns =91.43MB/s。

分析 16-字块传送情况：

对于16-字块传送，一次总线事务由一个地址传送后跟一个16-字块的数据传送组成。也即每个总线事务传送一个16个字的数据块。

第一个4-字所花时间为：

(1) 发送一个地址到主存花1个时钟周期

(2) 从主存读开始的4字花：150ns/(5ns/Cycle)=30个时钟周期

(3) 4个字传输需2个时钟，在传输期间存储器开始读取下一个4字

(4) 在本次和下次之间有2个空闲时钟，但下一个4字需6个时钟（30ns）才能完成存取，所以还要6-2-2=2个时钟

综上所述，16字中其余三个4字要重复上述最后两步。因此对于16-字块传送，一次总线事务共需花费的周期数为：1+30+3x(2+4)+2+2=53个周期，256个字需256/16=16个事务，因此整个传送需53x16=848个时钟。故总等待时间为：848周期x 5ns/周期=4240ns。每秒钟的总线事务个数为：16 x (1s / 4240ns) = 3.77M个。总线带宽为：(256 x 4B)x (1s/4240ns) =241.5MB/s。

比前者高2.64倍。

由此可见，大数据块传输的优势非常明显。但因为随后的4字需要30ns（6个时钟），比7.18题多了两个时钟，因此，大数据块的优势没有7.18题明显。

8.20 [5] <§8.5> Exercise 8.19 demonstrates that using larger block sizes results in an increase in the maximum sustained bandwidth that can be achieved. Under what conditions might a designer tend to favor smaller block sizes? Specifically, why would a designer choose a block size of 4 instead of 16 (assuming all of the characteristics are as identified in Exercise 8.19)?

参考答案：

从8.18和8.19来看，小数据块和大数据块的区别在于：

小数据块的每个事务所花的时钟数比大数据块的少，但传送一块特定大小的数据所需事务数比大数据块的多。所以当需要传输的数据块较小时，采用小数据块更好。

例如，若每次传送只有4个字时，则4-字块比16-字块好。

8.27 [15] <§8.5> We want to compare the maximum bandwidth for a synchronous and an asynchronous bus. The synchronous bus has a clock cycle time of 50 ns, and each bus transmission takes 1 clock cycle. The asynchronous bus requires 40 ns per handshake. The data portion of both buses is 32 bits wide. Find the bandwidth for each bus when performing one-word reads from a 200-ns memory.

参考答案见下页。

复习：同步和异步总线的最大带宽比较

举例：假定同步总线的时钟周期为**50ns**，每次总线传输花**1**个时钟周期，异步总线每次握手需要**40ns**，两种总线的数据都是**32**位宽 ，存储器的取数时间为**200ns**。要求求出从该存储器中读出一个字时两种总线的带宽。

分析如下：

同步总线的步骤和时间为：

（**1**）发送地址和读命令到存储器：**50ns**

（**2**）存储器读数据：**200ns**

（**3**）传送数据到设备：**50ns**

所以总时间为**300ns**，故最大总线带宽为**4B/300ns**，即：**13.3MB/s**。

异步总线的步骤和时间为：

第**1**步为：**40ns**；

第**2**、**3**、**4**步为：**Max(3x40ns，200ns)=200ns**；

（第**2**、**3**、**4**步都和存储器访问时间重叠）

第**5**、**6**、**7**步为：**3x40ns=120ns**。

总时间为**360ns**，故最大带宽为**4B/360ns=11.1MB/s**

由此可知：同步总线仅比异步快大约**20%**。要获得这样的速度，异步总线上的设备和存储器必须足够快，以使每次在**40 ns**内能完成一个子过程

8.29 [10] <§8.5> Let's determine the impact of polling overhead for three different devices. Assume that the number of clock cycles for a polling operation—including transferring to the polling routine, accessing the device, and restarting the user program—is 400 and that the processor executes with a 500-MHz clock.

Determine the fraction of CPU time consumed for the following three cases, assuming that you poll often enough so that no data is ever lost and assuming that the devices are potentially always busy:

1. The mouse must be polled 30 times per second to ensure that we do not miss any movement made by the user.
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.
3. The hard disk transfers data in four-word chunks and can transfer at 4 MB/sec. Again, no transfer can be missed.

参考答案：**CPU**的速度为：每秒**500×10^6**个时钟

1. 鼠标的速度：每秒**30x400=12000**个时钟

占**CPU**时间的 **$12000/500\times10^6$=0.002%**

2. 软盘的速度为：每秒**50KB/2B=25000**次查询（**polling**）

每秒的时钟数为：**$25000\times400=10^7$**

占**CPU**时间的 **$10^7/500\times10^6$=2%**

3. 硬盘的速度为：每秒**4MB/16B=0.25×10^6**次查询（**polling**）

每秒的时钟数为：**$250000\times400=10^8$**

占**CPU**时间的 **$10^8/500\times10^6$=20%**

由此可见，查询方式不适合快速设备的输入输出！

8.34 <§8.6> An important advantage of interrupts over polling is the ability of the processor to perform other tasks while waiting for communication from an I/O device. Suppose that a 1 GHz processor needs to read 1000 bytes of data from a particular I/O device. The I/O device supplies 1 byte of data every 0.02 ms. The code to process the data and store it in a buffer takes 1000 cycles.

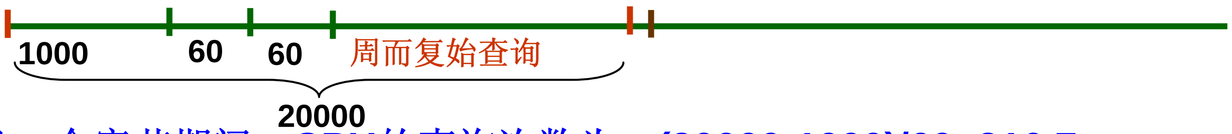
a. If the processor detects that a byte of data is ready through polling, and a polling iteration takes 60 cycles, how many cycles does the entire operation take?

b. If instead, the processor is interrupted when a byte is ready, and the processor spends the time between interrupts on another task, how many cycles of this other task can the processor complete while the I/O communication is taking place? The overhead for handling an interrupt is 200 cycles.

参考答案：**CPU**的速度为：每秒**10^9**个时钟，即：每个时钟为**10^{-9}秒=1ns**

I/O设备每隔**$0.02x10^{-3}/10^{-9}$=20000**个时钟提供一个字节

a. 查询方式：



在**I/O**设备准备一个字节期间，**CPU**的查询次数为：**(20000-1000)/60=316.7**

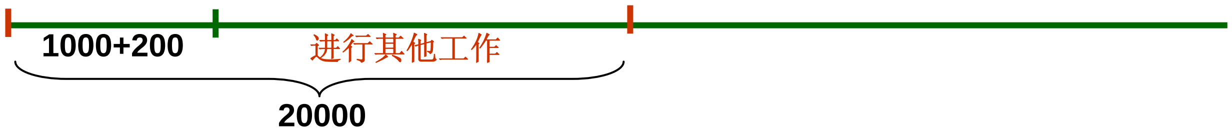
因为，**I/O**设备必须准备好才能查询到，所以，应该在第**317**次查询时才能查到

所以每个字节所花的时钟数为：**1000+317x60= 20020**

整个**1000**字节所花时钟数为：**20020x1000=20,020,000**个

（每**3**个字节可多**60**个时钟，正好进行一轮查询，所以共少**333x60=19980**个）

中断方式：



中断处理的时钟数为：额外开销**200**+处理**1000=1200, 20000-1200=18800**用来完成其他工作

用于其它工作的时钟总共有：**18800x1000=18,800,000**

8.39 [8] <§8.6> Suppose we have the same hard disk and processor we used in Exercise 8.18, but we use interrupt-driven I/O. The overhead for each transfer, including the interrupt, is 500 clock cycles. Find the fraction of the processor consumed if the hard disk is only transferring data 5% of the time.

参考答案：

8.38题以查询方式进行硬盘传送：

速度为：每秒4MB/16B=0.25×10^6次查询（polling）

每秒内查询所用的时钟数为：250000x400=10^8

占CPU时间的 $10^8/500 \times 10^6$=20%

8.38题以中断方式进行磁盘传送：

磁盘每准备好 4 字（=16B）就发出中断请求一次，为了保证没有任何数据传输被错过，传送的速率应达到每秒4MB/16B=250K次中断的速度；

每秒钟用于中断的周期数为250Kx500=125×10^6；

在一次传输中所消耗的处理器时间的百分比为：$125 \times 10^5/(500 \times 10^6)$=25%；

硬盘仅用其中5%的时间来传送数据，所以：

处理器消耗的百分比为25%x5%=1.25%

8.40 [8] <§8.6> Suppose we have the same processor and hard disk as in Exercise 8.18. Assume that the initial setup of a DMA transfer takes 1000 clock cycles for the processor, and assume the handling of the interrupt at DMA completion requires 500 clock cycles for the processor. The hard disk has a transfer rate of 4 MB/sec and uses DMA. If the average transfer from the disk is 8 KB, what fraction of the 500-MHz processor is consumed if the disk is actively transferring 100% of the time? Ignore any impact from bus contention between the processor and DMA controller.

参考答案：

8.38题以**DMA**方式进行磁盘传送：

每个**DMA**传送将花**8KB/(4MB/Sec)=2x10^{-3}**秒；

一秒钟有**1/(2x10^{-3})=500**次**DMA**传送；

如果硬盘一直在传送数据的话，处理器必须每秒钟花**(1000+500)x500=750x10^{3}** 个时钟周期来为硬盘**I/O**操作服务；

在硬盘**I/O**操作上处理器花费的时间占：

750x10^{3}/500x10^{6}=1.5x10^{-3}=0.15%

比较：**polling: 20%, interrupt: 1.25%, DMA: 0.15%**，能得出什么结论？