**第一章**

**1.1［2］<1.1>列举和描述除智能手机之外的4种类型的计算机。**

解：

①个人移动设备(PMD，包括平板电脑):PMD是电池驱动的，通过无线连接到互联网，通常要花费数百美元。而且，就像个人电脑一样，用户可以下载软件(“应用程序”)在上面运行。与个人电脑不同的是，它们不再有键盘和鼠标，而且更有可能依靠触摸屏甚至语音输入。

②服务器:用于运行大型问题的计算机，通常通过网络访问。

③仓库规模计算机:由多个处理器组成一个大集群。

④超级计算机:由数百至数千个处理器和tb级内存组成的计算机。

⑤嵌入式计算机:设计用来运行一个或一组相关应用程序并集成到一个系统中的计算机。

**1.2 [5]<1.2>计算机系统结构中的8个伟大思想与其他领域的思想相同。将计算机系统结构中的8个伟大思想“面向摩尔定律的设计”、“使用抽象简化设计”、“加速大概率事件”、“采用并行提高性能”、“采用流水线提高性能”、“采用预测提高性能”、“存储器层次”、“通过冗余提高可靠性”与其他领域的下列思想进行匹配：**

**a.汽车制造中的组装生产线**

**b.吊桥缆索**

**c.采用风向信息的飞机和船舶导航系统**

**d.高楼中的高速电梯**

**e.图书馆的预定台**

**f.通过增大CMOS晶体管的栅极面积来减小翻转时间**

**g.增加电磁飞机弹射器（不同于流体驱动模型，它采用电驱动），允许有新型反应堆技术才生成更多的能量**

**h.制造自动驾驶汽车，其控制系统是安装在汽车上的传感器系统，例如车道偏离检测系统和智能导航控制系统**

解：

a.通过流水线提高性能 b.通过冗余提高可靠性

c.通过预测提高性能 d.加速大概率事件

e.存储器层次 f.层次通过并行提高性能

g.面向摩尔定律的设计 h.使用抽象简化设计

**1.4[2]<1.4>一个彩色显示器中的每个像素由三种基色（红，绿，蓝）构成，每种基色用8位表示，分辨率为1280×1024像素。**

**a.为了保存一帧图像最少需要多大的缓存（以字节计算）？**

**b.在100Mbit/s的网络上传输一帧图像最少需要多长时间？**

解：

1. 8\*3\*1280\*1024b=31457280 bit/帧 （/ 8）= 3932160 B/帧
2. 31457280/(100\*106) = 0.31 S

**1.5[4]<1.6>有3种不同的处理器P1、P2和P3执行同样的指令集，P1的时钟频率为3GHz，CPI为1.5；P2的时钟频率为2.5GHz，CPI为1.0；P3的时钟频率为4GHz，CPI为2.2。**

**a.以每秒钟执行的指令数目为标准，哪个处理器性能最高？**

**b.如果每个处理器执行一个程序都花费10秒钟时间，求它们的时钟周期数和指令数。**

**c.我们试图把执行时间减少30%，但这会引起CPI增加20%。问：时钟频率应该是多少才能达到时间减少30％的目的？**

解：

1. （MIPS=时钟频率/CPI）每秒百万指令

MIPS1= 3\*109/1.5 = 2\*109

MIPS2= 2.5\*109/1 = 2.5\*109 性能最高✔

MIPS3= 4\*109/2.2 = 1.8\*109

1. （CPU时间= 时钟周期时间\*CPI\*指令数 时钟周期=1/时钟频率

10 = (1/(3\*109))\*CPI\*指令数 CPI\*指令数 即时钟周期数）

P1时钟周期数 = 10\*3\*109=3\*1010

P1的指令数 = 10\*3\*109/1.5= 2\*1010

P2时钟周期数= 10\*2.5\*109=25\*109

P2指令数= 10\*2.5\*109/1=25\*109

P3时钟周期数= 10\*4\*109

P3指令数= 10\*4\*109/2.2= 18.18\*109

1. CPI new = CPI old \* 1.2

CPI(P1)= 1.5\*1.2 =1.8, CPI(P2)= 1\*1.2 =1.2, CPI(P3)= 2.2\*1.2 =2.6

（f = 指令数\*CPI/CPU时间）

f(P1) = 20\*109\*1.8/(10\*(1-30%)) = 5.14GHz

f(P2) = 25\*109\*1.2/(10\*(1-30%)) = 4.28GHz

f(P3) = 18.18\*109\*2.6/(10\*(1-30%)) = 6.75GHz

( f1new/f1old=I1\*1.5\*1.2/(10\*(1-30%))/ I1\*1.5/(10) =1.2/0.7 =1.71，同理P2P3新旧频率比例也是1.71 ，改进前后CPU时间是1/0.7，CPI是1/1.2，指令数是1/1，能算时钟频率之比)

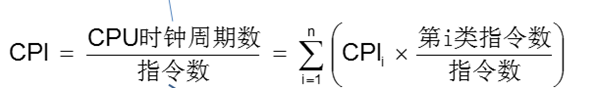
**1.6[20]<1.6>同一个指令集体系结构有两种不同的实现方式。根据CPI的不同将指令分成4类（A、B、C和D），P1的时钟频率为2.5GHz，CPI分别为1、2、3和3；P2时钟频率为3GHz，CPI分别为2、2、2和2。**

**给定一个程序，有1.0×106条动态指令，按如下比例分为4类：A，10%；B，20%;C，50%;D，20%。**

**a.每种实现方式总的CPI是多少？**

**b.计算两种情况下的时钟周期数。**

解：

1. 

CPI1 = (106\*0.1\*1+106\*0.2\*2+106\*0.5\*3+106\*0.2\*3)/106=2.6

CPI2 = (106\*0.1\*2+106\*0.2\*2+106\*0.5\*2+106\*0.2\*2)/106=2

1. 时钟周期数= CPI\*指令数

P1时钟周期数 = 2.6\*106

P2时钟周期数 = 2\*106

**1.8 2004年发布的Pentium 4 Prescott处理器时钟频率为3.6GHz，工作电压为1.25V。假定平均情况下静态功耗为10W，动态功耗为90W。**

**2012年发布的Core i5 Ivy Bridge时钟频率为3.4GHz，工作电压为0.9V。假定平均情况下静态功耗为30W，动态功耗为40W。**

**1.8.1 [5]<1.7>分别求出每个处理器的平均电容负载。**

**1.8.2 [5]<1.7>对于每种工艺，求出静态功耗占总功耗的比例和静态功耗相对于动态功耗的比率。**

**1.8.3 [15]<1.7>如果要将整体功耗降低10%，求出在保持漏电流不变的情况下电压要降低多少？注意：静态功耗定义为电压与电流的乘积。**

解：动态功耗 = 0.5\*负载电容\*电压2\*开关频率

1. C = 2 \* 动态功耗DP / (电压V2 \* 开关频率F)

Pentium 4 Prescott：2\*90/（1.252\*3.6\*109）=3.2\*10-8

Core i5 Ivy Bridge：2\*40/（0.92\*3.4\*109）=2.9\*10-8

1. Pentium 4 Prescott：静/总= 10/（10+90）=90%

静/动= 10/90 =11.1%

Core i5 Ivy Bridge：静/总= 30/（30+40）=42.9%

静/动= 30/40 =75%

1. (S new + D new)/(S old +D old) = 0.9 D new =0.9\*(S old +D old)-S new

D new = 0.5\*C\*Vnew2 \*f V new = (2\*D new/(C\*f))1/2

S = V \*I 漏电电流不变S new=V new\*(S old/V old)

Pentium的 S new = V new \* (10/1.25) = V new \* 8

D new = 0.90 \* 100 - V new \* 8 = 90 - V new \* 8

V new = [2\*(90 - V new \* 8)/( 3.2\*10-8 \* 3.6\*109)] 1/2

V new = 1.182V

corei5 的 S new = V new \* (30/0.9) = V new \* 33.3

D new = 0.90 \* 70 - V new \* 33.3 = 63 - V new \* 33.3

V new = [2\*(63 - V new \* 33.3)/( 2.9\*10-8 \* 3.4\*109)] 1/2

V new = 0.84 V

**1.12 1.10节引证了一个用性能公式的一个子集去计算性能的陷阱。为了说明它，考虑下面两种处理器。P1的时钟频率为4GHz，平均CPI为0.9，需要执行5.0×109条指令；P2的时钟频率为3GHz，平均CPI为0.75，需要执行1.0×109条指令。**

**1.12.1［5］<1.6,1.10>一个常见的错误是，认为时钟频率最高的计算机具有最高的性能。这种说法正确吗？请用P1和P2来验证这一说法。**

**1.12.2［10］<1.6,1.10>另一个错误是，认为执行指令最多的处理器需要更多的CPU时间。考虑P1执行1.0×109条指令序列所需的时间，P1和P2的CPI不变，计算一下P2用同样的时间可以执行多少条指令？**

**1.12.3［10］<1.6,1.10>一个常见的错误是用MIPS(每秒百万条指令)来比较两台不同的处理器的性能，并认为MIPS最大的处理器具有最高的性能。这种说法正确吗？请用P1和P2验证这一说法。**

**1.12.4［10］<1.10>另一个常见的性能标志是MFLOPS(每秒百万条浮点指令)，其定义为**

**MFLOPS=浮点操作的数目/（执行时间×（1×106））。**

**它与MIPS有同样的问题。假定P1和P2上执行的指令有40%的浮点指令，求出程序的MFLOPS。**

解：cpu时间=IC\*cpi\*ccT

1. P1 CPU时间 = 5\*109\*0.9\*(1/(4\*109)) = 1.125 S

P2 CPU时间 = 109\*0.75\*(1/(3\*109)) = 0.25 S

P1 时钟频率＞P2 时钟频率 但是P2性能优于P1

1. P1 CPU时间 = 109\*0.9\*(1/(4\*109)) = X\*0.75\*(1/(3\*109))=0.225

解得P2用同样的时间可以执行X = 0.9\*109条 指令

P1同样的时间执行1.0\*109多于P2，但P2性能高

1. MIPS = f \*10-6/CPI

MIPS(P1) = 4 \* 109 \* 10-6 /0.9 = 4.44 \* 103

MIPS(P2) = 3 \* 109 \* 10-6 /0.75 =4.0 \* 103

MIPS(P1)＞MIPS(P2) 但是P2性能优于P1（第一问结论）

1. MFLOPS=浮点操作的数目/（执行时间×（1×106））

MFLOPS（P1）= 5\*109\*40%/ (1.125\*106)= 1777.78

MFLOPS（P2）= 109\*40%/ (0.25\*106)=1600

MFLOPS(P1)＞MFLOPS(P2) 但是P2性能优于P1（第一问结论）

**1.13 1.10节引证另一个易犯的错误是通过只改进计算机的一个方面来改进计算机的总体性能。假如一台计算机上运行一个程序需要250s，其中70s用于执行浮点指令，85s用于执行L/S指令，40s用于执行分支指令。**

**1.13.1[5]<1.10>如果浮点操作的时间减少20%，总时间将减少多少？**

**1.13.2[5]<1.10>如果总时间减少20%，整数操作时间将减少多少？**

**1.13.3[5]<1.10>如果只减少分支指令时间，总时间能否减少20%？**

解:

1. T fp = 70\*0.8 =56s

T new = 56+85+40+55 = 236s 250-236= 14s

1. T new = 250\*0.8 = 200s T fp+ T L/S+ T 分支=70+85+40=195s

即改进后加法只能用5秒，原来加法用250-70-85-50=55秒，减少50秒

1. 250\*0.8 = X +70+85+55

200 = X +210 不可能

**1.14假定一个程序需要执行50×106条浮点指令、110×106条整数指令、80×106条L/S指令和16×106条分支指令。每种类型指令的CPI分别是1、1、4和2。假定处理器的时钟频率为2GHz。**

**1.14.1［10］<1.10>如果我们要将程序运行速度提高至2倍，浮点指令的CPI需如何改进？**

**1.14.2［10］<1.10>如果我们要将程序运行速度提高至2倍，L/S指令的CPI需如何改进？**

**1.14.3［5］<1.10>如果整数和浮点指令的CPI减少40%，L/S和分支指令的CPI减少30%，程序的执行时间能改进多少？**

解：

1. 

CPU时钟周期数 = 106\*(50\*1+110\*1+80\*4+16\*2) = 512\*106

CPU时间 = 时钟周期数/f = 512\*106/(2\*109) = 0.256s

通过改进FP指令的CPI，将程序运行速度提高至2倍（即CPU时钟周期数减半）

可列式：106\*(50\*x+110\*1+80\*4+16\*2)= 512\*106/2=256\*106

解得 x = -4.12 < 0 不可能

1. 通过改进L/S指令的CPI，将程序运行速度提高至2倍（即CPU时钟周期数减半）

可列式：106\*(50\*1+110\*1+80\*x+16\*2) =512\*106/2=256\*106

解得 x = 0.8

1. CPU时钟周期数（新）= 106\*(50\*1\*0.6+110\*1\*0.6+80\*4\*0.7+16\*2\*0.7) =342.4\*106

CPU时间（新）= 时钟周期数/f =342.4\*106/(2\*109) =0.171s

程序的执行时间改进了 0.256s – 0.171s = 0.089s