第四章 性能评价

作业参考答案

第四章 性能评价

4.1 [5] <§4.1> We wish to compare the performance of two different computers: M1 and M2. The following measurements have been made on these computers:

Program	Time on M1	Time on M2
1	2.0 seconds	1.5 seconds
2	5.0 seconds	10.0 seconds

Which computer is faster for each program, and how many times as fast is it?

答：对于程序1：M2比M1快2.0/1.5=1.3倍
对于程序2：M1是M2快10.0/5.0=2倍

4.2 [5] <§4.1> Consider the two computers and programs in Exercise 4.1. The following additional measurements were made:

Program	Instructions executed on M1	Instructions executed on M2
1	5×10^9	6×10^9

Find the instruction execution rate (instructions per second) for each computer when running program 1.

答：在M1上运行程序1时，指令吞吐率为：5000/2=2500MIPS
在M2上运行程序1时，指令吞吐率为：6000/1.5=4000MIPS

第四章 性能评价

4.3 [5] <§4.1> Suppose that M1 in Exercise 4.1 costs $500 and M2 costs $800. If you needed to run program 1 a large number of times, which computer would you buy in large quantities? Why?

答： 大多数情况下运行程序1，所以主要考虑程序1：
对于程序1：M2比M1快2.0/1.5=1.3倍，而M2的价格是M1的800/500=1.6倍
偶尔运行程序2的话，M1比M2快2倍。
所以，应该选择购买M1。

4.6 [5] <§4.1> Another user has the following requirements for the computers discussed in Exercise 4.1: P1 must be executed 1600 times each hour. Any remaining time is used to run P2. If the computer has enough performance to execute program 1 the required number of times per hour, then performance is measured by the throughput for program 2. Which computer is faster for this workload? Which computer is more cost-effective?

答： 1小时为3600秒， P1执行1600次后，剩下时间执行P2，只要比较P2的效率
对于M1：1600x2.0=3200秒，还剩400秒， P2共执行400/5=80次
对于M2：1600x1.5=2400秒，还剩1200秒， P2共执行1200/10=120次
所以，这种负载下，M2更快。但M1的性价比更好，因为：
M2的价格是M1的1.6倍，而M2只比M1快120/80=1.5倍。

第四章 性能评价

4.7 [10] <§4.2> Suppose you wish to run a program P with 7.5×10^9 instructions on a 5 GHz machine with a CPI of 0.8.

a. What is the expected CPU time?

b. When you run P, it takes 3 seconds of wall clock time to complete. What is the percentage of the CPU time P received?

答： 程序P的CPU执行时间为：

Instruction count x CPI x Clock time = $7.5x10^9x0.8x1/(5x10^9)$ = 1.2秒

执行程序P时用户感觉到的执行时间为3秒，

其中真正花在程序P的用户CPU时间只有1.2秒，仅占1.2/3=40%.

其他60%的时间执行其他用户程序、操作系统程序或等待I/O完成。

第四章 性能评价

4.8 [10] <§4.2> Consider two different implementations, P1 and P2, of the same instruction set. There are five classes of instructions (A, B, C, D, and E) in the instruction set.

P1 has a clock rate of 4 GHz. P2 has a clock rate of 6 GHz. The average number of cycles for each instruction class for P1 and P2 is as follows:

Class	CPI on P1	CPI on P2
A	1	2
B	2	2
C	3	2
D	4	4
E	3	4

Assume that peak performance is defined as the fastest rate that a computer can execute any instruction sequence. What are the peak performances of P1 and P2 expressed in instructions per second?

答： （同一指令集的不同实现，可能具有不同的**CPI**和不同的时钟频率。）

P1中**A**类指令的**CPI**为**1**，所以，如果一段时间内执行的都是**A**类指令，则可达到峰值性能。其值为：$4x10^9$=**4000MIPS**

P2中**A**、**B**和**C**类指令的**CPI**为**2**，所以，若一段时间内执行的都是**A/B/C**类指令，则可达到峰值性能。其值为：$6x10^9/2$=**3000MIPS**

第四章 性能评价

4.11 [5] <§4.2> Consider program P, which runs on a 1 GHz machine M in 10 seconds. An optimization is made to P, replacing all instances of multiplying a value by 4 (mult X, X,4) with two instructions that set x to $x + x$ twice (add X,X; add X,X). Call this new optimized program P′. The CPI of a multiply instruction is 4, and the CPI of an add is 1. After recompiling, the program now runs in 9 seconds on machine M. How many multiplies were replaced by the new compiler?

答： 乘法指令的CPI为4，加法指令的CPI为1， 1条乘4指令变成2条加法指令，

故每条乘4指令的执行时间缩短了4x1-1x2=2个时钟

程序P的执行时间为10秒，程序P'的执行时间为9秒，缩短了1秒=10^9个时钟

相当于$10^9/2=5x10^8$条乘4指令。

所以，在新编译的程序中有5亿条乘4指令被替换为加法指令。

第四章 性能评价

4.12 [5] <§4.2> Your company could speed up a Java program on their new computer by adding hardware support for garbage collection. Garbage collection currently comprises 20% of the cycles of the program. You have two possible changes to the machine. The first one would be to automatically handle garbage collection in hardware. This causes an increase in cycle time by a factor of 1.2. The second would be to provide for new hardware instructions to be added to the ISA that could be used during garbage collection. This would halve the number of instruction needed for garbage collections but increase the cycle time by 1.1. Which of these two options, if either, should you choose?

答： 方式**1**下，取消了原来用于垃圾收集的指令，而使得时钟周期变为**1.2**倍，所以，程序执行时间是原来的 **(1-0.2)x1.2=0.96**倍

方式**2**下，原来用于垃圾收集的指令条数较少了一半，但时钟周期为原来的**1.1**倍，所以，程序的执行时间是原来的**(1-0.2/2)x1.1=0.99**倍

由此可见，方式**1**比**2**好。

第四章 性能评价

4.14 [5] <§4.2> The table below shows the number of floating-point operations executed in three different programs and the runtime for those programs on three different computers:

Program	Floating-point operations	Execution time in seconds		
		Computer A	Computer B	Computer C
Program 1	$5x10^9$	2	5	10
Program 2	$20x10^9$	20	20	20
Program 3	$40x10^9$	200	50	15

Which computer is fastest according to total execution time? How many times as fast is it compared to the other two computers?

答：从程序总执行时间来看，计算机A为222秒，B为75秒，C为45秒。

所以，C的执行速度最快, 比A快约222/45=4.9倍，比B快约75/45=1.7倍。

第四章 性能评价

4.15 [15] <§§4.2, 4.3> One user has told you that the three programs in Exercise 4.14 constitute the bulk of his workload, but he does not run them equally. The user wants to determine how the three computers compare when the workload consists of different mixes of these three programs. (You know you can use the arithmetic mean to find the relative performance.)

Suppose the total number of floating-point operations (FLOPs) executed in the workload is equally divided among the three programs. That is, program 1 runs 8 times for every time program 3 runs, and program 2 runs twice for every time program 3 runs. Find which computer is fastest for this workload and by what factor. How does this compare with the total execution time with equal numbers of program executions?

答：假设A、B和C中浮点运算平均分布，则各程序所占比例为：

A: 8/11=72.7%、B: 2/11=18.2%、C: 1/11=9.1%

可用算术平均值来比较其性能, 各程序的算术平均值为：

计算机A为: **2x72.7%+20x18.2%+200x9.1%=23.3**秒

计算机B为: **5x72.7%+20x18.2%+50x9.1%=11.8**秒

计算机C为: **10x72.7%+20x18.2%+15x9.1%=12.3**秒

所以，B的速度最快, 比A快约**23.3/11.8=1.97**倍，比C快约**12.3/11.8=1.04**倍。

假设每个程序执行次数相同，则各计算机的平均值为：

A: 222/3=74秒； **B: 75/3=25**秒；**C: 45/3=15**秒

所以，C的速度最快，比A快**4.9**倍，比B快**1.7**倍。

从上述分析来看，不同场合，性能比较的结果不同。