4.7 Historical Perspective and Further Reading

This section, which reviews the history of performance measurement and benchmarking, appears as Section 4.7on the CD.

4.8 Exercises

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?

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.

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?

4.4 [10] <§4.1> For More Practice: Cost-Effective Computing

4.5 [5] <§4.1> For More Practice: Cost-Effective Computing

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?

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?

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?

4.9 [5] <§§4.1–4.2> If the number of instructions executed in a certain program is divided equally among the classes of instructions in Exercise 4.8 except for class A, which occurs twice as often as each of the others, how much faster is P2 than P1?

4.10 [12] <§4.2> Consider two different implementations, I1 and I2, of the same instruction set. There are three classes of instructions (A, B, and C) in the instruction set. I1 has a clock rate of 6 GHz, and I2 has a clock rate of 3 GHz. The average number of cycles for each instruction class on I1 and I2 is given in the following table:

Class	CPI on M1	CPI on M2	C1 Usage	C2 Usage	C3 Usage
A	2	1	40%	40%	50%
B	3	2	40%	20%	25%
C	5	2	20%	40%	25%

The table also contains a summary of average proportion of instruction classes generated by three different compilers. C1 is a compiler produced by the makers of I1, C2 is produced by the makers of I2, and the other compiler is a third-party product. Assume that each compiler uses the same number of instructions for a given program but that the instruction mix is as described in the table. Using C1 on both I1 and I2, how much faster can the makers of I1 claim I1 is compared to I2? Using C2, how much faster can the makers of I2 claim that I2 is compared to I1? If you purchase I1, which compiler would you use? If you purchased I2, which compiler would you use? Which computer and compiler would you purchase if all other criteria are identical, including cost?

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?

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?

4.13 [5] <§4.2> For the following set of variables, identify all of the subsets that can be used to calculate execution time. Each subset should be minimal; that is, it should not contain any variable that is not needed.

{CPI, clock rate, cycle time, MIPS, number of instructions in program, number of cycles in program}

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	5×10^9	2	5	10
Program 2	20×10^9	20	20	20
Program 3	40×10^9	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?

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?

4.16 [15] <§§4.2, 4.3> An alternative weighting to that of Exercise 4.15 is to assume that equal amounts of time will be spent running each program on one of the computers. Which computer is fastest using the data given for Exercise 4.14 and assuming a weighting that generates equal execution time for each of the benchmark programs on computer A? Which computer is fastest if we assume a weighting that generates equal execution time on computer B? Computer C? Explain the results.

4.17 [5] <§§4.2–4.3> If the clock rates of computers M1 and M2 in Exercise 4.1 are 4 GHz and 6 GHz, respectively, find the clock cycles per instruction (CPI) for program 1 on both computers using the data in Exercises 4.1 and 4.2.

4.18 [5] <§§4.2–4.3> Assuming the CPI for program 2 on each computer in Exercise 4.1 is the same as the CPI for program 1 found in Exercise 4.17, find the instruction count for program 2 running on each computer using the execution times from Exercise 4.1.

4.19 [5] <§4.3> **In More Depth:** Amdahl's Law

4.20 [10] <§4.3> **In More Depth:** Amdahl's Law

4.21 [10] <§4.3> **In More Depth:** Amdahl's Law

4.22 [5] <§4.3> **In More Depth:** Amdahl's Law

4.23 [5] <§4.3> **In More Depth:** Amdahl's Law

4.24 [20] <§4.3> **In More Depth:** Amdahl's Law

4.25 [5] <§4.3> **In More Depth:** Choosing the Right Mean

4.26 [15] <§4.3> **In More Depth:** Choosing the Right Mean

4.27 [3 hours] <§4.3> **In More Depth:** Synthetic Benchmarks: Attempts to Replicate "Typical" Behavior

4.28 [3 hours] <§4.3> **In More Depth:** Synthetic Benchmarks: Attempts to Replicate "Typical" Behavior

4.29 [4 hours] <§4.3> **In More Depth:** MIPS, MOPS, and Other FLOPS

4.30 [5] <§4.3> **In More Depth:** MFLOPS as a Performance Metric

4.31 [15] <§4.3> **In More Depth:** MFLOPS as a Performance metric

4.32 [4 hours] <§4.3> **In More Depth:** MFLOPS as a Performance Metric

4.33 [5] <§4.3> **In More Depth:** Embedded Benchmarks

4.34 [20] <§§4.2, 4.3> **In More Depth:** Using Hardware-Independent Metrics to Try to Predict Performance

4.35 [10] <§§4.2, 4.3> **For More Practice:** Analyzing a Processor with Floating-Point Implemented in Hardware or Software

4.36 [10] <§§4.2, 4.3> **For More Practice:** Analyzing a Processor with Floating point implemented in Hardware or Software

4.37 [5] <§§4.2, 4.3> **For More Practice:** Analyzing a Processor with Floating Point Implemented in Hardware or Software

4.38 [10] <§§4.2, 4.3> **For More Practice:** Analyzing Enhancements to a Processor

4.39 [5] <§§4.2, 4.3> **For More Practice:** Analyzing Enhancements to a Processor

4.40 [10] <§§4.2, 4.3> **For More Practice:** Analyzing Enhancements to a Processor

4.41 [5] <§§4.2, 4.3> **For More Practice:** Analyzing Enhancements to a Processor

4.42 [5] <§§4.2, 4.3> **For More Practice:** Analyzing Enhancements to a Processor

4.43 [10] <§§4.2, 4.3> **For More Practice:** Analyzing Enhancements to a Processor

4.44 [10] <§§4.2, 4.3> **For More Practice:** Analyzing Enhancements to a Processor

4.45 [5] <§4.3> Assume that multiply instructions take 12 cycles and account for 15% of the instructions in a typical program, and the other 85% of the instructions require an average of 4 cycles for each instruction. What percentage of time does the CPU spend doing multiplication?

4.46 [5] <§4.3> Your hardware engineering team has indicated that it would be possible to reduce the number of cycles required for multiplication to 8 in Exercise

4.45, but this will require a 20% increase in the cycle time. Nothing else will be affected by the change. Should they proceed with the modification?

4.47 [10] <§4.4> Look at the current list of SPEC programs in Figure 4.5 on page 260. Does it include applications that match the ways you typically use your computer? What classes of programs are irrelevant or missing? Why do you think they were or were not included in SPEC? What would have to be done to include/exclude such programs in the next SPEC release?

4.48 [5] <§4.4> If benchmark suites are designed to provide a real-world metric for a specific computing task, explain why benchmarks suites need to be updated.

4.49 [5] <§§4.2, 4.3, 4.4> **In More Depth:** The Difficulty with Kernel Benchmarks

4.50 [15] <§§4.2, 4.3, 4.4> **In More Depth:** The Difficulty with Kernel Benchmarks

4.51 [10] <§§4.1–4.5> Consider the following hypothetical news release:

> *"The company will unveil the industry's first 5 GHz version of the chip, which offers a 25% performance boost over the company's former speed champ, which runs at 4 GHz. The new chip can be plugged into system boards for the older original chip (which ran at 1 GHz) to provide a 70% performance boost."*

Comment on the definition (or definitions) of performance that you believe the company used. Do you think the news release is misleading?

4.52 [indefinite] <§§4.1–4.5> Collect a set of articles that you believe contain incorrect analyses of performance or use misleading performance metrics to try to persuade readers. For example, an article in the *New York Times* (April 20, 1994, p. D1) described a video game player "that will surpass the computing power of even the most powerful personal computers" and presented the following chart to support the argument that "video game computers may be the supercomputers of tomorrow":

Computer	Approximate number of instructions per second	Price
1975 IBM mainframe	10,000,000	$10,000,000
1976 Cray-1	160,000,000	$20,000,000
1979 Digital VAX	1,000,000	$200,000
1981 IBM PC	250,000	$3,000
1984 Sun 2	1,000,000	$10,000
1994 Pentium-chip PC	66,000,000	$3,000
1995 Sony PCX video game	500,000,000	$500
1995 Microunity set-top	1,000,000,000	$500

The article never discussed how the nature of the instructions should impact the definition of "powerful." For each article you collect, describe why you think it is misleading or incorrect. Good places to look for material include the business or technology sections of newspapers, magazines (both articles and ads), and the Internet (newsgroups and the Web).

Answers to Check Yourself

§4.1, page 246: 1. a: both, b: latency, c: neither. 2. 7 sec.
§4.2, page 253: 6.
§4.3, page 258: 1. F, F, F, T. 2. T, F, F, T
§4.4, page 265: b.
§4.5, page 270: a. Computer A. b. Computer B.