Outline

Chapter 2 Overview

- How to measure, report and summarize performance?
- What are the major factors that determine the performance of a computer?
- Execution time is the only adequate measure of performance
- Benchmarks, what are they, and how are they used to evaluate performance

Why study Performance?

- Hardware performance is often key to the effectiveness of an entire system of Hardware and Software
- The goal is not just to assess performance but need to understand what affects performance of a machine
- To improve performance of software → understand how hardware affects system performance
  - ✓How well a program uses instructions of the machine?
  - ✓How well underlying HW implements instructions?
  - ✓How well memory and I/O systems perform?

How to define performance?

Airplanes example

- Passenger capacity
- Cruising range (miles)
- Cruising speed (m.p.h)
- Passenger throughput (passengers * m.p.h)

Which airplane has the best performance?

- Highest cruising speed
- Longest range
- Largest capacity
- Speed
  - Highest cruising speed
  - highest throughput

Run a program on two different workstations, which is fastest?

- User: response time (execution time)
- Computer center manager: throughput

(how many tasks were performed during a time interval)

- Relationship between response time and throughput

Performance

Use response time or execution time. To maximize performance minimize execution time for some task

$$Performance = \frac{1}{ExecutionTime}$$

What does it mean that Performance(X) is greater than Performance(Y)?

$$Performance(X) > Performance(Y)$$

$$\frac{1}{ExecutionTime(X)} > \frac{1}{ExecutionTime(Y)}$$

$$ExecutionTime(Y) > ExecutionTime(X)$$

X is *n* times faster than Y $\frac{Performance(X)}{Performance(Y)} = \frac{ExecutionTime(Y)}{ExecutionTime(X)} = n$

Performance Example

Machine A runs a program in 10 seconds and machine B runs the same program in 15 seconds, how much faster is A than B?

$$\frac{Performance(A)}{Performance(B)} = \frac{ExecutionTime(B)}{ExecutionTime(A)} = \frac{15}{10} = 1.5$$

A is 1.5 times faster than B

Measuring Performance 1/4

Time is the measure of computer performance (sec per program)

•Response time or elapsed time

Total time to complete a task including everything (disk access, memory access, operating system overhead, …)

•CPU execution time (CPU time)

Time CPU spends computing for this task and does not include time spent waiting for I/O or running other programs (some computers are timeshared)

•CPU execution time can be divided into

User CPU time: CPU time spent in the program

System CPU time: CPU time spent in operating system performing tasks on behalf of the program

Measuring Performance 2/4

Example of user CPU time and System CPU time

Output of Unix *time* command

90.7u 12.9s 2:39 65%

User CPU time	90.7 sec
System CPU time	12.9 sec
Elapsed time	2:39 = (2 minutes and 39 sec) = 159 sec

% of elapsed time that is CPU time = (90.7 + 12.9)/159 = 65%

Then 100 – 65 = 35% of elapsed time was spent doing something else

(waiting for I/O, running other programs, …)

Measuring Performance 3/4

Express CPU execution time in terms of other metric that relates to how fast the HW can perform basic functions

•Computers governed by a clock that runs at constant rate and determines when events happen in HW

•Length of a clock period is *Clock cycle* (measured in nanoseconds (10^{-9} sec) or picoseconds (10^{-12} sec))

•*Clock rate* is 1/(clock cycle) (measured in Megahertz (MHz = 10^6 Hz), or Gigahertz (GHz = 10^9 Hz))

•1 Hertz is 1 cycle/sec

CPU execution time = CPU clock cycles for a program * clock cycle time

CPU execution time = CPU clock cycles for a program / clock rate

How to improve CPU execution time?

Example on page 60

Measuring Performance 3/4

Question:

A program runs on a 400 MHz computer in 10 seconds. We like the program to run in 6 seconds by designing a faster CPU. Assume that increasing clock rate would mean the program needs 20% more clock cycles. What clock rate should the designer target?

Answer:

The number of clock cycles for the program on the present computer =

$10 * 400 * 10^6 = 4000 * 10^6$ cycles

With 20% increase, the new computer should take $1.2 * 4000 * 10^6 =$

$4800 * 10^6$ cycles

Required execution time = 6 seconds

Then the required clock rate = $4800/6 * 10^6$ cycles/sec = 800 MHz

Measuring Performance 4/4

Relating to Software

- Express CPU clock cycles in terms of program instructions
- CPU clock cycles = Instruction for a program * Average clock cycles per instruction
- Clock cycles per instruction (average number of cycles each instruction takes to execute) is abbreviated as **CPI**
- CPI can be used to compare two implementations of the same instruction set architecture (since instruction count for a program will remain the same)

Measuring Performance 4/4

A program executed in machine A with a 1ns clock gives a CPI of 2.0. The same program with machine B having same ISA and a 2ns clock gives a CPI of 1.2. Which machine is faster and by how much?

Answer: **Let I be the instruction count.**

CPU clock cycles for A = I x 2.0

Execution time on A = 2 I ns

CPU clock cycles for B = I x 1.2

Execution time on B = I x 1.2 x 2 ns = 2.4 I ns

=> CPU A is faster by 1.2 times.