Computer Organization

Douglas Comer

Computer Science Department
Purdue University
250 N. University Street
West Lafayette, IN 47907-2066

http://www.cs.purdue.edu/people/comer

XIX

Assessing Performance

Measuring Computational Power

Measuring Computational Power

- Difficult to assess computer performance

Measuring Computational Power

- Difficult to assess computer performance
- Chief problems

Measuring Computational Power

- Difficult to assess computer performance
- Chief problems
  - Flexibility: computer can be used for wide variety of computational tasks

Measuring Computational Power

- Difficult to assess computer performance
- Chief problems
  - Flexibility: computer can be used for wide variety of computational tasks
  - Architecture that is optimal for some tasks is suboptimal for others

Measuring Computational Power

- Difficult to assess computer performance
- Chief problems
  - Flexibility: computer can be used for wide variety of computational tasks
  - Architecture that is optimal for some tasks is suboptimal for others
  - Memory and I/O costs can dominate

The Point About Performance

Because a computer is designed to perform a wide variety of tasks and no architecture is optimal for all tasks, the performance of a system depends on the task being performed.

Consequences

Consequences

- Many groups try to assess computer performance

Consequences

- Many groups try to assess computer performance
- A variety of performance measures exist

Consequences

- Many groups try to assess computer performance
- A variety of performance measures exist
- No single measure suffices for all situations

Measures Of Computational Power

- Two primary measures
- Integer computation speed
  - Pertinent to most applications
  - Example measure is millions of instructions per second (*MIPS*)
- Floating point computation speed
  - Used for scientific calculations
  - Typically involve matrices
  - Example measure is floating point operations per second (*FLOPS*)

Average Floating Point Performance

- Assume
  - Addition or subtraction takes Q nanoseconds
  - Multiplication or division takes $2Q$ nanoseconds
- Average cost of floating point operation is:

$$T_{avg} = \frac{Q + Q + 2Q + 2Q}{4} = 1.5\,Q \quad ns\ per\ instr.$$

- Note: addition or subtraction costs 33% less than average, and multiplication or division costs 33% more

A Note About Average Execution Times

Because some instructions take substantially longer to execute than others, the average time required to execute an instruction only provides a crude approximation of performance. The actual time required depends on which instructions are executed.

Application Specific Instruction Counting

- More accurate assessment of performance for specific application
- Examine application to determine how many times each instruction occurs
- Example: multiplication of two $N \times N$ matrices
  - N^3 floating point multiplications
  - $N^3 - N^2$ floating point additions
  - Time required is:

$$T_{total} = 2 \times Q \times N^3 + Q \times (N^3 - N^2)$$

Weighted Average

- Alternative to precise count of operations
- Typically obtained by instrumentation
- Program run on many input data sets and instruction counts averaged over all runs

Example Of Instruction Counts

Instruction Type	Count	Percentage
Add	8513508	72
Subtract	1537162	13
Multiply	1064188	9
Divide	709458	6

Computation Of Weighted Average

- Uses instruction counts and cost of each instruction
- Example

$$T_{avg'} = .72 \times Q + .13 \times Q + .09 \times 2\ Q + .06 \times 2\ Q$$

- Or

$$T_{avg'} \;=\; 1.16\ Q \;\; ns\ \ per\ instruction$$

- Note: weighted average is %23 less than uniform average obtained above

Instruction Mix

- Measure a large set of programs
- Obtain relative weights for each type of instruction
- Relative weights used to assess the performance of a given architecture
- Try to choose set of programs that represent a typical workload

Use Of Instruction Mix

An instruction mix consists of a set of instructions along with relative weights that have been obtained by counting instruction execution in example programs. An architect can use an instruction mix to assess how a proposed architecture will perform.

Standardized Benchmarks

- Provides workload used to measure computer performance
- Represent ‘‘typical’’ applications
- Independent corporation formed in 1980s to create benchmarks
  - Named *Standard Performance Evaluation Corporation* (*SPEC*)
  - Not-for-profit
  - Avoids having each vendor choose benchmark that is tailored to their architecture

Examples Of Benchmarks Developed By SPEC

- SPEC cint2000
  - Used to measure integer performance
- SPEC cfp2000
  - Used to measure floating point performance
- Result of measuring performance on a specific architecture is known as the computer's *SPECmark*

I/O And Memory Bottlenecks

- CPU performance is only one aspect of system performance
- Bottleneck can be
  - Memory
  - I/O
- Some benchmarks focus on memory operations or I/O performance rather than computational speed

Increasing Overall Performance

To optimize performance, move operations that account for the most CPU time from software into hardware.

Which Items Should Be Optimized?

- Adding additional hardware increases cost
- Architect cannot use high-speed hardware for all operations
- Computer architect Gene Amdahl observed that it is a waste of resources to optimize functions that are seldom used

Amdahl's Law

The performance improvement that can be realized from faster hardware technology is limited to the fraction of time the faster technology can be used.

Amdahl's Law And Parallel Systems

- Amdahl's law
  - Applies directly to parallel systems
  - Explains why adding processors does not always increase performance

Summary

- A variety of performance measures exist
- Simplistic measures include MIPS and FLOPS
- More sophisticated measures use a weighted average
- Weights can be derived by counting the instructions in a program or set of programs
- Set of weights corresponds to an instruction mix

Summary
(continued)

- Benchmark refers to a standardized program or set of programs used to measure performance
- Best-known benchmarks, known as SPECmarks, are produced by the SPEC Corporation
- Amdahl’s Law helps architects select functions to be optimized (moved from software to hardware)

Questions?