

Course name : Advanced Microprocessor - Elective

Topic - 1

- * Computer system performance and its evaluation : A brief account
- * CPU performance equation and differentiation of CISC & RISC architectures

Performance Summarization:

Example 1

Suppose we have two programs or process - P1 and P2 and three computers : A, B, C

Suppose

- A is 10 times faster than B for P1
- B is 10 times faster than A for P2
- A is 20 times faster than C for P1
- C is 50 times faster than A for P2
- C is 5 times faster than B for P2 and so on

From here : Performance of a program on a m/c depends on type of program (Application)

: Taking individually, one statement may be of some use; but relative performance is unclear w.r.to m/c's.

Example 2

| | Comp A | Comp B | Comp C |
|-------------------|--------|--------|--------|
| P1 (secs) | 1 | 10 | 20 |
| P2 (secs) | 1000 | 100 | 20 |
| Total time (secs) | 1001 | 110 | 40 |

(Execution time of 2 programs on 3 m/c's)

A summary of the table is as!

B is $\frac{1001}{110}$ times faster than A for P1 and P2.

C is $\frac{110}{40}$ times faster than B for P1 and P2.

C is $\frac{1001}{40}$ times faster than A for P1 and P2.

The above table can give relative performance of machines if P1 and P2 run equal no. of times.

In that case C is faster for P1 and P2.

Suppose: Time taken for P1 on m/c A = T_1

" " " P2 " " B = T_2

" " " P1 " " C = T_3

Av. Execution time = $\frac{T_1 + T_2 + T_3}{3}$;

Similarly Av. execution time of P2 can be found

Thus the av. execution time that tracks total execution time is the arithmetic mean:

$$\frac{1}{n} \sum_{i=1}^n \text{Time}_i \quad (\text{Time}_i \text{ is execution time of } i\text{th program of total of } n \text{ work load})$$

Weighted execution time: As shown above, for equal mix i.e. if programs run equal no. of times execution time is governed by arithmetic mean.

When there is unequal mix of programs in the work load, then we assign weightage (w_i) to each program to indicate relative frequency of program in work load

eg. If 20% of tasks in work load are P1 \Rightarrow weightage = 0.2
80% of tasks in work load are P2 \Rightarrow weightage = 0.8

Weightage Arithmetic mean $\sum \text{weightage}_i \times \text{Time}_i$ (Weightage is freq. of i th program in work load.)

Conclusion of above examples:

"Performance of a computer is application specific"

∴ For a part this leads us to an important design principle for performance enhancement

"MAKE COMMON CASE FAST"

In general purpose computers used for all types of applications; having invariably conflicting requirements attainment of this design principle is not possible.

Therefore this has led us to design ~~computers~~ different computers for different applications:

eg. For I/O applications

For computational problems

and ~~of~~ Embedded processors for Systems/~~for~~

v. specific applications.

Amdahl's law

The law states, "The performance improvement to be gained by using some faster mode of execution is limited by fraction of t_c the faster mode can be used."

The law defines speed up that can be gained by using a particular feature.

Suppose enhancement has been made

$$\text{Speed up} \triangleq \frac{\text{Performance of entire task using enhancement when possible}}{\text{Performance of entire task without using enhancement}}$$

$$= \frac{\text{Execution time for entire task without enhancement}}{\text{Execution } t_c \text{ of entire task using enhancement when possible}}$$

Speed up tells us how much faster ~~info~~ ^{task} will run using the machine as opposed to original m/c.

Execution enhancement depends on two factors:

1. The fraction of computation time in original machine that can be converted to take advantage of enhancement

eg. If unutilized time of m/c is 60 secs & suppose '10' secs of m/c are utilized out of 60 secs. ~~out of~~ by virtue of enhancement
 \therefore Fraction enhanced = $10/60$

2. Speed up Enhanced:

This is improvement gained by enhanced execution mode, i.e., how much faster the task would run if the enhanced mode was used for a program. This may be given as:

$$\frac{\text{Time of original mode (Ex. time)}}{\text{Time of Enhanced mode (Ex. t.e)}}$$

eg. Suppose prog in its original mode 5 secs are required to run a program. If the m/c is enhanced and in enhanced mode the program runs only in 2 secs

$$\text{Then improvement (Enhancement)} = 5/2 > 1$$

Exple: Suppose using enhancement processor is 10 times faster than original processor. Assume original processor is busy with computation 40% i.e. and waiting for I/O 60% i.e.

What is overall Speed up gained by incorporating enhancement.

Some important definitions.

1. Execution time: This can be defined in different ways depending on what we count. Most straight forward definition is: Wall clk time, Response time or Elapsed time. This time is latency to complete a task.

Elapsed time includes:

- Disk access
- Memory access
- I/O activities
- Operating System overheads etc.

Thus elapsed time is time as seen by the user.

2. CPU time: This is time during which CPU is executing. This time does not include waiting time etc. (as may be case in elapsed time).

CPU Time $\left\{ \begin{array}{l} \text{CPU time spent on program (user CPU time)} \\ \text{CPU time spent on O.S.} \end{array} \right.$

Example: If we give a command in Unix: Time

Time: 90.7 U 12.9 S 2:39 65%.

\downarrow \downarrow \downarrow
 User CPU time System CPU time Elapsed time \nearrow % of CPU time actually used (excluding waiting time)

So more than $\frac{1}{3}$ rd of time was spent on waiting for I/O or running other programs (in multi program env) or both. Many times we ignore CPU time spent on O.S. because of inaccuracies in O.S.

- \therefore System Performance = Elapsed time on unloaded system
 CPU Performance = User CPU time on an unloaded system.

CPU Performance Equation:

Computers are constructed using a CLK running at a constant rate. These discrete time events are called clock periods, CLK Cycles etc.

CLK Period is designated by

- a: Duration (n.s.) or
- b: Rate (GHz)

$$\begin{aligned} \text{CPU Time} &= \text{CPU CLK cycles per program} \times \text{CLK cycle time} \\ &= \frac{\text{CPU CLK cycles for program}}{\text{CLK rate}} \end{aligned}$$

In addition to ~~CLK cycles~~ number of CLK cycles needed to execute a program, we can also count the no. of instructions executed; This is also called Path length or Instruction Count (IC)

If we know no. of CLK cycles & IC we can calculate No. of CLK cycles

per Instruction (CPI). Designers also use Instructions per CLK cycle (IPC) which is inverse of CPI

$$\text{CPI} = \frac{\text{CPU CLK cycles for a program}}{\text{IC}}$$

↳ Figure of merit, provides insight into different styles of Instruction sets & implementations.



Total clk cycles = IC x Av. no. of clk cycles per instruction

CPU time = Instruction Count x CLK cycle time x cycles per inst.

$$\text{CPU time} = \frac{\text{Instruction Count} \times \text{CLK cycle time}}{\text{CLK rate (IPC)}} \quad \text{cycles/inst.}$$

Expanding the formula in the units of measurement

$$\left(\frac{\text{Inst}}{\text{Program}} \right) \times \left(\frac{\text{CLK cycles}}{\text{Inst}} \right) \times \left(\frac{\text{Seconds}}{\text{CLK cycle}} \right) = \frac{\text{Seconds}}{\text{Program}} = \text{CPU time}$$

or $\text{CPU time} = \frac{\text{Time}}{\text{Program}} = \frac{\text{Time}}{\text{Inst}} \times \frac{\text{Inst}}{\text{Program}}$

$$\text{or CPU Time} = \frac{\text{Time}}{\text{Program}} = (\text{IC}) (\text{CPI}) \left(\frac{1}{\text{CLK rate}} \right) \quad \text{cycles/sec.}$$

↓ ↓
Reduce Reduce
for CISC for RISC.

$$\frac{\text{Inst}}{\text{Program}} \times \frac{\text{Sec}}{\text{cycles}}$$

CPU time or CPU performance is dependent on

- a) CLK cycles or CLK rate (Processor specific)
- b) CLK cycles per inst
- c) Instruction Count

At top A x% improvement in any one of them leads to x% improvement in CPU time.

It is difficult to change one of the parameters in total because basic technology is involved in changing each characteristic are independent

- CLK cycle time : HW technology & organization
- CPI : organization - inst. set architecture
- IC : Inst set architecture & compiler tech.

Examples to solve and submit

- Suppose we have two implementations of some instructions set architecture. Computer A has clk cycle t_c of 250 ps - and CPI of 2.0 for some program. Computer B has CLK cycle t_c of 500 ps. and CPI of 1.2 for same program. Which computer is faster for this program and by how much?
- A program runs in 10 sec on computer A which has 4 GHz CLK. We are trying to help computer designer to build a computer B that will run this prog in 6 Sec. The designer has ~~desired~~ determined that a substantial increase in CLK rate is possible; but this increase will effect rest of CPU design, causing computer B to require 1.2 times as many CLK cycle of computer A for this program. What CLK rate should designer target?
- Suppose we are considering enhancement of server system. The new CPU is 10 times faster on computation ~~is~~ than original. Assuming that the original CPU is busy with computation 40% of t_c and is waiting for I/O 60% of the t_c . What is overall speedup gained by incorporating enhancement.
- What is significance of MIPS ^{rate}? A 40 MHz processor was used to execute a bench mark program with following instr. mix

| Inst. type | Inst Count | CLK cycle count |
|--------------------|------------|-----------------|
| Integer Arithmetic | 45000 | 1 |
| Data Tx | 32000 | 2 |
| FP op | 15000 | 2 |
| Control Tx | 8000 | 2 |

Determine:
CPI, MIPS rate
and Execution t_c
for program.

- What are bench marks? Discuss SPEC bench mark
- Differentiate: Price Performance, Reliability Performance, Power performance, ~~in Green - efficient~~