Computer Organization and Architecture - Performance Analysis.

Shriram K Vasudevan

Session – 16

Introduction

- How can someone say that “the CAR A has better economy than CAR B?”. They will have to measure the performance. One should measure, summarize the factors which actually governs the performance. Here, if you take a CAR, its mileage, efficiency, engine capacity, maintenance etc. all help in determining the performance.
- A simple example can be Car A gives 30 KMPL whereas CAR B gives 25 KMPL. So, it is obvious with respect to mileage Car A is better.
- Now comes the crux. How to measure performance of a computer? One should measure and summarize the key factors which govern the performance of a computer.
- The next question to be answered is “Why do we measure performance?” Answer is simple and straight forward. To understand how much effective the system is, one should measure the performance.
- More about measuring performance is to be seen in this session.

It is challenging – Let us understand why!

- We know very well that computer is not just hardware or software. It is a mix of both.
- The moment you have both in picture, complexity comes in as an non-invited guest.
- That too, with the amount of features and sophistication the recent systems have, measuring performance is not an easy task. How to measure the performance itself is a research area and the designers really have a tough time in this sector.
- One can think about analysing the performance of the system by reading through the existing instruction set details and other technical information pamphlets. It is next to impossible to read through all of these and to assess the performance.
- Based on the applications installed, based on the type of application, performance measurement metrics may be different and certainly, it is not an easy task.
- Again, measuring the performance while also carefully analysing the limitations would help in selecting the most appropriate computer.

Contd.,

- One should not just think that performance is connected to external parameters of a computer. It is much more than that.
- Before going any deeper to analyse the performance it is important to understand that what governs the performance of a typical computer.
- This discussion is aimed at it and will help the readers in understanding the importance of performance monitoring while also highlighting the measurement criteria.

Let us define performance

- When someone looks abstract, the performance could appear simple. But, once digged deeper, it will really be challenging.
- Well, the simple question one should try answering now is "How can someone say Computer A is performing better than Computer B?".
- We will see an example to understand how and why it challenging to measure the performance.
- Let us take two different car manufacturers - A and B.
- Metrics for Maker A's and Maker B's Car are summarized below as a table 1 and 2. (All assumptions)

1. CC - 3000	3000
2. Highest Speed	280 KMPH
3. Economy	10 KMPL
4. Gears	Manual – 6 Gears
5. Safety Features	2 Airbags + ABS
6. Maintenance cost per year	500 USD

1. CC - 3000	2000
2. Highest Speed	180 KMPH
3. Economy	18 KMPL
4. Gears	Manual – 4 Gears
5. Safety Features	No Airbags
6. Maintenance cost per year	100 USD

Contd.,

- So, let us see which of these two cars is a better performing car.
- Car A is better than CAR B when it comes to
  - Speed
  - Pickup
  - Safety Features
- Car B is better than CAR A when it comes it
  - Fuel economy.
  - Maintenance cost

1. CC - 3000	3000
2. Highest Speed	280KMPH
3. Economy	10 KMPL
4. Gears	Manual – 6 Gears
5. Safety Features	2 Airbags + ABS
6. Maintenance cost per year	500 USD

1. CC - 3000	2000
2. Highest Speed	180 KMPH
3. Economy	18 KMPL
4. Gears	Manual – 4 Gears
5. Safety Features	No Airbags
6. Maintenance cost per year	100 USD

Now, it could be understood! Performance can't be blindly measured! Based on what you measure, performance varies!

If you want Virat Kohli to perform for bowling you can't get the best performance and vice versa with bat for Bumrah!

Contd.,

- Assume you have 2 computers (PCs). PC named Sachin and PC named Sehwag.
- How do we compare the performance now?
  - We can say that, whichever PC completes the task first can be regarded as the fastest one. Correct? Here, if Sachin completes the task faster, then, he is the fastest.
- Assume there is a group of 11 PCs (11 Players) in a lab. How could performance be measured in this scenario?
  - Simple, how much work has been done as a team (all PCs together) is the performance measurement criteria here.
- Given this example, one should understand that "Performance metrics are different for different conditions".

One Question For You !!!

Throughput and Response Time

Do the following changes to a computer system increase throughput, decrease response time, or both?

1. Replacing the processor in a computer with a faster version
2. Adding additional processors to a system that uses multiple processors for separate tasks—for example, searching the World Wide Web.

Decreasing response time almost always improves throughput. Hence, in case 1, both response time and throughput are improved.

In case 2, no one task
gets work done faster, so only throughput increases. If, however, the demand for processing in the second case was almost as large as the throughput, the system might force requests to queue up. In this case, increasing the throughput could also improve response time, since it would reduce the waiting time in the queue. Thus, in many real computer systems, changing either execution time or throughput often affects the other.

Let us do some math!

- An Athlete is running 100 meters in 15 Seconds. With glucose being consumed before race, he covers it by 12 Seconds.
- With decrease or reduction in the time consumed to cover the race distance, one can see improved performance.
- Relating this to computers, one can say, with the execution time getting reduced, one can maximize the performance.

Contd.,

- Performance is generally defined by the formula presented below:

  Performance = 1 / Execution Time.

- As discussed earlier, when someone has to compare the performance of two diffferent PCs say A and B, the following would be the method.
- **Assume performance of A is better than B. (Disclaimer : Assume)**

  A's Performance > B's Performance

  1/ A's Execution Time > 1 / B's Execution Time

  Execution Time of B > Execution Time of A

- Or in other words, **Execution Time of A < Execution Time of B.**

Contd.,

- Instead of saying just Performance of A is greater than performance of B, it would always be better if someone can say it as "A is n times faster than B". So, 'n' is to be found.
- To find 'n', one can use the following expression:

n = **A's Performance / B's Performance.**

$$\frac{\text{A's Performance}}{\text{B's Perfomance}} = \frac{\text{Execution time of B}}{\text{Execution time of A}} = \text{'n'}$$

Contd.,

A Simple problem for you to solve:

Assume that the computer 'Sachin' executes a program in 5 seconds. Whereas, computer 'Dravid' runs the same program in 15 seconds. Find out how much Sachin is faster than Dravid?

We have already read the formula for the problem to be solved.

$$\frac{\text{A's Performance}}{\text{B's Perfomance}} = \frac{\text{Execution time of B}}{\text{Execution time of A}} = \text{'n'}$$

15/5 = 3 = 'n'

Hence, Sachin is 3 times faster than Dravid.

People may also want to say this way " Dravid is 3 times slower than Sachin".

How do we analyse the performance of Embedded Systems?

- First, let us answer this question.
- What is an “embedded system”?
  - An embedded system is combination of software + hardware meant to do a dedicated task.
  - Simple example: Projector, Washing Machine etc.
- Now, the target is to measure the performance of the Embedded Systems and it is not as other general purpose systems.
- Recollect the term “Real Time”.
  - Embedded Systems are all expected to perform the task real time. Real time is “Logical correctness of the operation within deterministic deadlines”.
  - Real time can be hard real time or soft real time.
  - So, to conclude, for embedded systems, “Meeting the deadlines” is the very important performance measure.
    - Relate this. If car braking system does not work in time, what will happen?
  - Rest all performance measures are looked into after looking at the real time behaviour.

Performance Measurement – Let us get better.

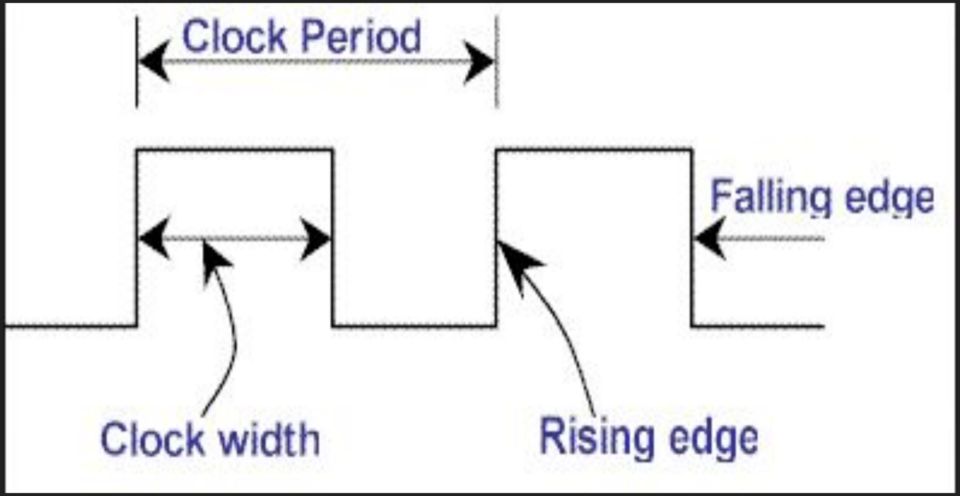
Session - 17

Let us understand the definitions, first!

- First, let us understand what is **CPU execution time. (AKA CPU TIME)**
- **Amount of time, the CPU actually spends on executing a task is referred as "CPU Execution Time".**
- **CPU could spend time on "the tasks" and "OS" tasks. The tasks given by the user, when CPU spends time for it, it is called, user CPU time. Meanwhile, if spent time for OS tasks, it could be termed "System CPU time".**

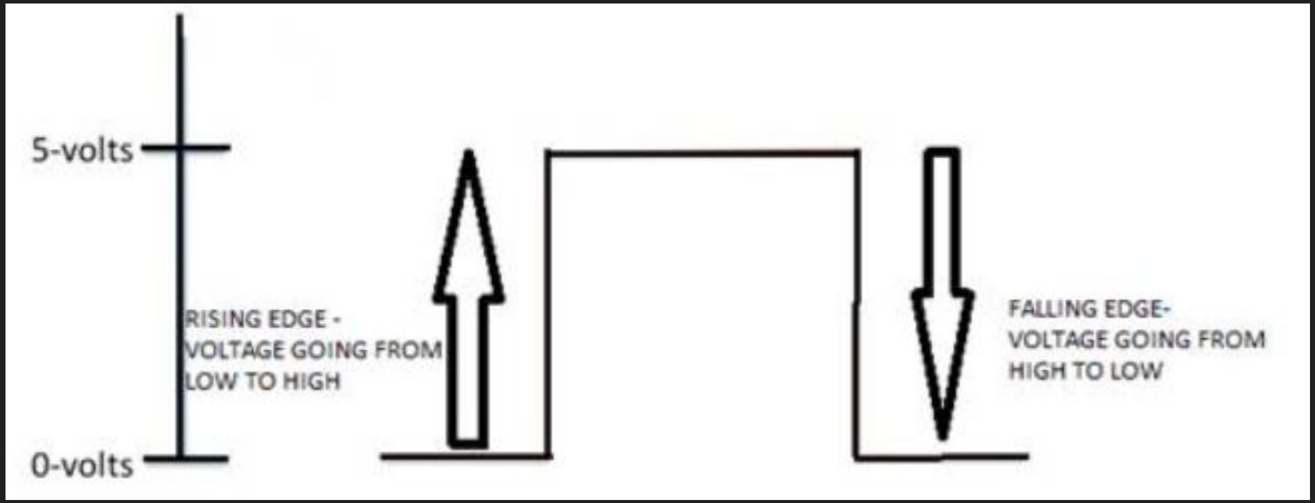
Contd.,

- How do you defined the speed of a computer (i.e. CPU)? You could have read it as 4 GHZ, 2 GHZ etc. There is a term for the same. It is called **clock cycle**. It is the amount of time spent between the two pulses of an oscillator.
- Unit of the Clock Cycle is Hz. We are never used the slower processors in the recent days and the unit has not grown to GHz. Typically a 2GHz processor go for 2,00,00,00,000 clock cycles per second. (People call this clock rate)



Contd.,

- Let us understand **Falling edge and Rising edge.**
- If **a circuit gets rising edge triggered**, the circuit will become active when its **clock signal turns from LOW TO HIGH. This is also referred as "Rising Edge or positive edge".**
- Similarly for a falling is referred as high to low transition!



Factors related to CPU performance – Some formulas.

- There are so many metrics followed by designers. Let us follow one metric – CPU execution time.
- Let us learn some formula:

$$\boxed{\text{CPU execution time} = \text{No. of CPU Clock Cycles (CC) for a program} \times \text{Clock Cycle Time}}$$

Also, one could use the below formula:

$$\text{CPU execution time} = \frac{\text{No. of CPU Clock Cycles (CC) for a program}}{\text{Clock Rate}}$$

I.e. Clock Rate = Inverse of Clock Cycle Time.

A Problem …

Finding @ calculation for A

CPU Clock Cycles $_A = 40 * 10^9$ Cycles

Finding @ calculation for B

B Requires 2 times as clock rate as A. (Means, 8 Ghz)

Lets get deeper!

- Let me ask a question to you guys. I can read a book with 250 pages in 3 hours. Should it be same as your speed?
  - I could be faster or slower. Am I right? – Think about it.
  - This is the point.
  - Remember this formula, This could be useful.

Number of CLOCK CYCLES reqd. for a program :

CPU CLOCK CYCLES = Number of instructions for a program * Average clock cycles per instruction.

(makes sense, right?)

NEXT STEP:

LET'S COMPUTE CPU TIME FOR 'A' & 'B'

CPU TIME FOR A = CPU CLOCK CYCLES-A * CLOCK CYCLE TIME-A

= Z * 2.0 * 250 PS

= 500 * Z PS.

CPU TIME FOR B = CPU CLOCK CYCLES-B * CLOCK CYCLE TIME-B

= Z * 1.2 * 500 PS

= 600 * Z PS.

TO GET FINAL ANSWER:

$$\frac{\text{CPU PERFORMANCE A}}{\text{CPU PERFORMANCE B}} = \frac{\text{EXEC. TIME B}}{\text{EXEC. TIME A}} = \frac{600 * Z\text{ PS}}{500 * Z\text{ PS}} = 1.2$$

'A' is 1.2 times faster than 'B'.

Let us get a step better.

- CPI = Clock Cycles Per Instruction (Remember this, without fail).
- How do we get CPI?. It is possible through simulation of the implementation. Or, one could choose hardware over the simulation to get the details of CPI.

CPI = clock cycles per instruction. — REMEMBER THIS

A designer has a challenge:

* he has to choose between one of the two code sequences.

Sequence : A X
Sequence : B Y

'Don't worry now'

DESIGNER IS PROVIDED WITH:

	INSTRUCTION CLASSES		
	A	B	C
CPI	1	2	3

CODE SEQUENCES:

	INSTRUCTION COUNTS		
SEQUENCE	A	B	C
X	2	1	2
Y	4	1	1

QUESTION :-

→ WHICH CODE SEQUENCE EXECUTES MOST INSTRUCTIONS??

→ ~~Which~~ WHO WILL BE FASTER??

→ FIND 'CPI' FOR X, Y?

SEQUENCE 1 = 2 + 1 + 2 = 5 instructions. ← MINIMUM INSTRUCTIONS

SEQUENCE 2 = 4 + 1 + 1 = 6 Instructions.

Hence, it is obvious 'SEQ-1' Executes Minimum inst.

LEARN A NEW FORMULA :

$$\text{CPU CLOCK CYCLES} = \sum_{i=1}^{n} (CPI_i * C_i)$$

REMEMBER

C_i = No/.. (Count of) of instructions of class i.

CPI_i = Avg. no/.. of cycles per instruction

n = no/.. of instruction classes

So,

CPU clock cycles$_1$ = [2x1] + [1x2] + [2x3]
= 2 + 2 + 6
= 10 CYCLES

FOR SEQUENCE-1

CPU CLOCK CYCLES$_2$ = [4x1] + [1x2] + [1x3]
= 4 + 2 + 3
= 9 CYCLES

FOR SEQUENCE-2

OBSERVATIONS:

SEQ.2 = 9 CYCLES [6 instructions]
SEQ.1 = 10 CYCLES [5 instructions]

SEQ2 IS FASTER THAN SEQ1. [THOUGH INSTRUCTIONS ARE MORE]

$$CPI = \frac{\text{CPU CLOCK CYCLES}}{\text{INSTRUCTION COUNT}}$$

$$CPI_1 = \frac{10}{5} = 2 \qquad CPI_2 = \frac{9}{6} = 1.5$$

Home work – Solve More Problems with the inputs given

Shriram K Vasudevan