Factors Effecting CPU Performance

CPU Performance

Performance is about more than just speed. It's like an F1 car – there's no point having excellent top speed if it can't brake or corner well.

The **performance** of the CPU is how "well" it works to complete tasks. This includes how quickly it works, but also how effectively it goes about completing the tasks it has.

You need to know about 3 factors which effect CPU performance:

- clock speed
- number of processor cores
- cache size

1. Clock speed

A clock pulse is called a **clock cycle**

The **clock** is a component within the CPU. It sends out an electronic pulse on a regular basis – billions of times per second. Each electronic pulse allows the CPU to carry out one process within the **Fetch–Decode–Execute cycle**.

The clock speed of my home computer is 2.6GHz. My school machine has a 3.4GHz clock. In theory this makes the school machine quicker, but this is running all sorts of network software which slows it down - so, in practice my home machine is quicker for me.

The faster the clock speed, the more electronic pulses per second, so the more instructions can be executed each second. This, in theory, means that the computer will complete a task quicker.

It is possible for a processor to temporarily increase its clock speed. Intel processors do this using a technology called **Intel Turbo Boost**. This allows the control unit to increase clock speeds when it is doing something particularly complex. This happens at the cost of increased heat. Once heat builds up the processor will need to slow down again.

2. Number of Processor Cores

A **core** is a processing component within the CPU. Modern machines might be described as "dual core" or "quad core".
Each core can process a set of instructions on each clock pulse. So, a dual core machine can process two, independent sets of instructions on each clock cycle – one in each core. A quad core machine can process 4 instructions on each clock cycle etc…

If a machine needs to run multiple programs at once then a multi-core processor will help to do this. For example, programs such as anti-virus or firewalls may be running "in the background". This still uses processor resources, and a multi-core machine will make this more efficient.

So, the more cores you have, the more instructions can be processed per clock cycle, so the better the performance of the machine. In theory.

The problem is that a program is a set of instructions designed to be worked on in order. If programs have been specifically written to use a number of cores then they can be run quicker, but this is not always the case. It's often the case that one program needs to use just one of the cores available. As a result it will not run any quicker even if you have a "6-core" machine.

What is Cache?

One of the things that slows down a computer is accessing main memory – i.e. bringing data from RAM into and out of the processor.

One solution to this is to make all of the RAM use faster memory chips and buses, but these are expensive.

The computer I'm writing this on has 8GB of main memory (of RAM chips), but only 3MB of cache memory (note the difference in units here)

An alternative solution is to use a **memory cache**.

A cache is an area of faster memory where data and instructions which need to be used frequently can be stored temporarily. They are close to the CPU and use a faster Bus. This means data transfer is quicker.

Essentially caches are an area of memory that sit between the CPU and RAM. The data and instructions stored in caches are accessible much more quickly by the CPU, which speeds up processing speeds.

Cache memory is quicker so it's more expensive.

3. Cache size

The greater the size of the cache memory used, the more data and instructions that can be stored. If a program requires lots of repetition then this will certainly make it run quicker.

For example, a video game such as Call of FIFA will often have lots of repeated processes.

The bigger the cache, the more data can be stored in it – so a bigger cache should make the CPU perform better.

Caches just store the data that's being used regularly. The process works something like this:

1. data or instruction used by CPU
2. dump it on the top of the cache
3. if the cache is full the data at the bottom of the cache falls out

There are actually different types of cache as well. The machine I'm writing this on has a 256KB level 2 cache for each processor core and a 3MB level 3 cache which is shared by the cores.

Then the next time the CPU needs to retrieve data or instructions, it looks in the cache for it first. If it's in the cache, it retrieves it and after it's finished with it dumps it back on the top of the cache. This means that stuff that the CPU is using lots stays at the top of the cache and never falls out - so it's always quicker to retrieve it.

Activities:

a) List the three ways to improve the performance of a CPU
b) What does clock speed mean?
c) How many instructions per second are possible at each clock speed:
   (i) 2.6Ghz (ii) 3.4Ghz (iii) 1.6Ghz (iv) 340MHz
d) What will be the impact of changing a processor to a "quad-core" type?
e) What is cache memory? Why is it used?
f) There are all sorts of ways to modify the performance of a CPU. This can involve "over-clocking". Find out what this is, how it works and what the impacts might be.