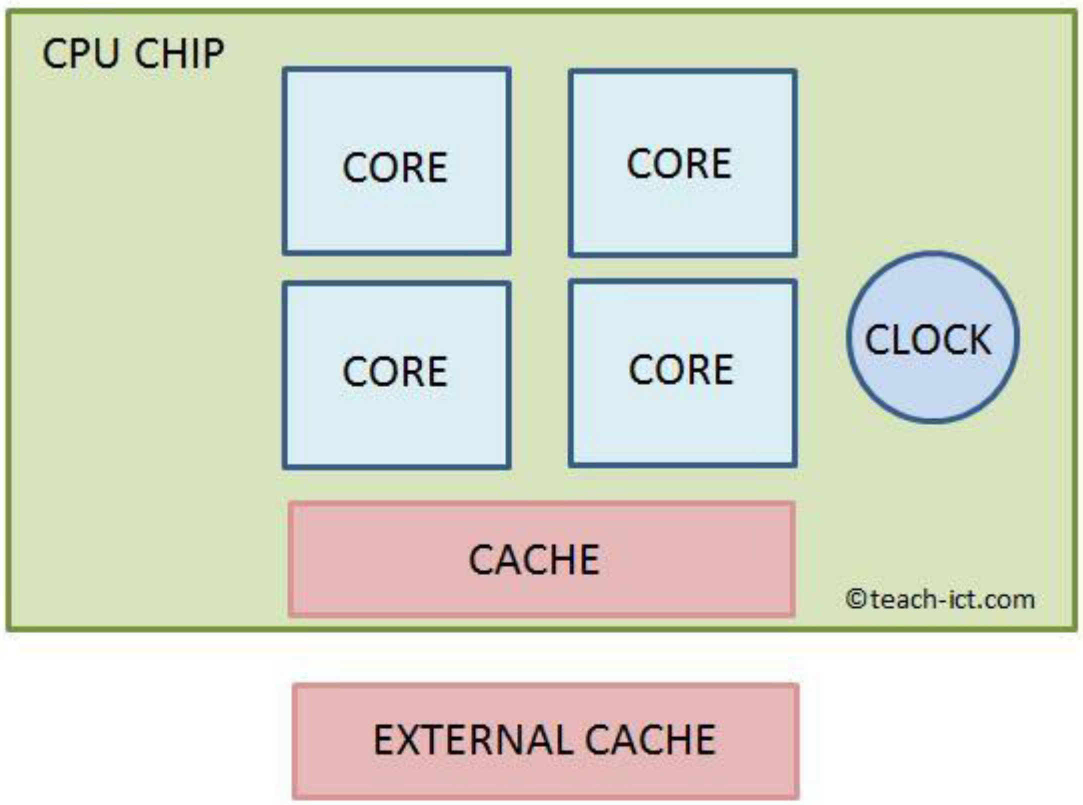
CPU performance factors

1. CPU Performance

There are a number of things which determine the performance level of a CPU.

The diagram below shows the main ones.



These are:

- Clock
- Cores
- Cache (internal and external)

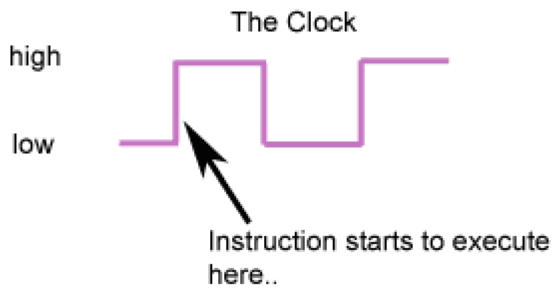
CPU performance factors

2. Clock

Every operation inside the CPU needs to be synchronised precisely.

To do that, an internal clock is used that 'ticks' at a constant rate.

Every part of the CPU uses the ticks to time when they need to do something. This tick rate is called the **clock frequency** or more commonly referred to as 'clock speed'.



source: www.teach-ict.com

With each tick of the clock, the CPU can process one instruction. This might not sound a lot but in a CPU with a typical clock speed of 3.5 GHz, this means that 3.5 billion instructions can be processed every second. That is a lot of instructions!

However, if you could upgrade the clock from 3.5 GHz to 4 GHz then the CPU would be able to process an extra 1/2 billion instructions every second.
So the higher the clock frequency, the faster the CPU can run and the more instructions that can be processed every second.

Every time the clock ticks, a bit of power is used and that power causes heat to be generated. If you go higher than a clock frequency of around 3.5-4GHz, the heat will damage the chip itself. This is why clock speeds have remained relatively stable at around a maximum of 3.5GHz - 4GHz.
In order to keep the temperature to a minimum and to stop damage to the CPU a large **heat sink** is required. A heat sink is a device designed to take heat away from the CPU, it includes metal fins to dissipate heat and a fan to drive air across the fins. The heatsink is placed right on top of the cpu chip, with heat-conducting paste joining the two.

See the large heat sink in this desktop computer.

(c) teach-ict.com

Another reason clock speed has not increased is that the building block of the CPU - the transistor - has not been getting any faster.

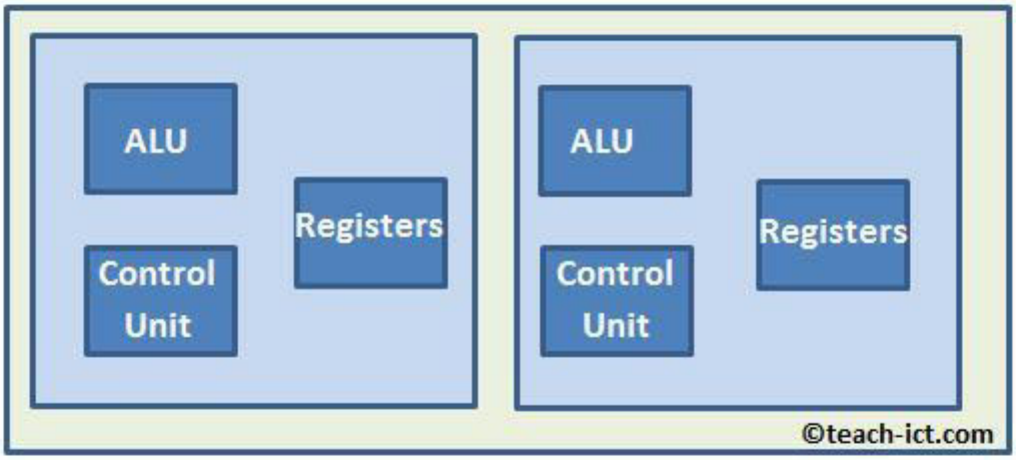
It is possible to run a CPU faster than its normal design speed - this is called '**overclocking**'. But to do so requires some serious cooling systems. For example keen gamers might invest in a water-cooled computer. There are some even more exotic scientific computers cooled by liquid nitrogen.

CPU performance factors

3. Core

A 'core' is a complete processing unit within the CPU - it has an ALU, Control Unit and Registers

Originally there was only one processing core in a CPU. But as time moved on the processing core could be made even smaller, and around 2006 it became possible to put two cores inside a single CPU chip.



Now it is common for a CPU chip to have two, four, eight cores or more. A cpu with two cores is called a 'dual core' and one with four cores is called a 'quad-core'.

Increasing the number of cores improves performance in two ways.

Multi-tasking

Multi-tasking is the ability to carry out more than one task at the same time.

Clearly, with two cores, a CPU can run two tasks simultaneously.

A quad core can handle four simulateous tasks. So for example, one core could be running a photo editing application whilst another is handling a word processing application.

Parallel Processing

Parallel processing is when a single task (program) is split into two or more parts and each part is processed at the same time.

In theory this would double performance on a dual core cpu as each part is processed independently.

However, you do not always get twice the performance because many programs cannot be split neatly into two independent parts. This is because they are mostly sequential i.e. each task depends on the outcome of a previous task and so it has to wait for the first task to be completed.

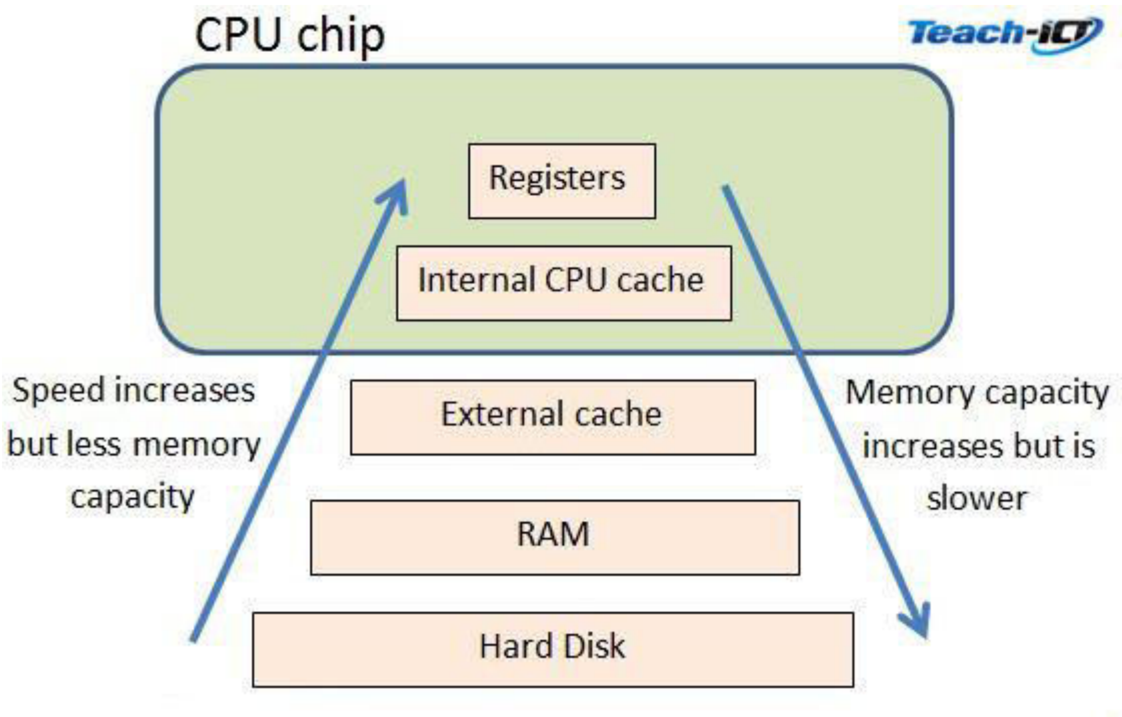
An example where parallel processing is possible is a task to update all the pixels on the screen. The colour of one pixel does not tend to depend on another pixel.

CPU performance factors

4. Cache performance

In an ideal world, when the CPU needs data, it should be available instantly so as not to slow down the processing operation.

A cache is a small, but extremely fast type of RAM, often located inside the CPU chip itself. The diagram below explains why it is needed.



Having a larger cache will speed up processing because the CPU can access data and instructions faster than they can from RAM.

However, cache memory is more expensive than standard RAM so it is a balance between performance and cost.

Data and programs currently in use are stored in RAM. But RAM is comparatively slow to access when compared to the speed at which the registers work.

So to help speed up the processing time, cache memory is used to store instructions or data that are either frequently used, have recently been used or are about to be used. This means that they don't have to be fetched directly from RAM.

There are different levels of cache. Level 1 (L1) cache is the fastest but the smallest size and resides inside the CPU chip. Level 2 (L2) cache is a bit slower but is larger, and so can hold more data, then Level 3 (L3) cache is slower still but can hold the most data. A cache though is always faster than the main RAM.