

The Supercomputer



By Liam Davie

Tuteur: Tom Goedert

Classe: 5G2

Contents

The Supercomputer.....	0
Introduction	2
1 Computer	2
1.1 The processor	2
1.2 Random access memory.....	4
1.3 Graphics Card/Graphics Processing Unit.....	7
1.4 Storage	8
1.5 The motherboard/mainboard	9
1.6 Bottleneck	10
1.7 Overclocking.....	10
1.8 The lifespan of a computer.....	11
2 Supercomputer.....	12
2.1 How does a supercomputer work?	12
2.2 Uses for supercomputers.	13
2.3 Supercomputer specs	14
2.4 Small supercomputer list.....	15
2.5 Quantum Supercomputers.....	17
Word Explanations	18
Conclusion	18
Sources	19

Introduction

This TraPe is about supercomputers and how they are built and how they work. Here I will also be comparing them to normal desktop computers. I will also explain general desktop computers first. Sometimes words may be used that are not properly explained until a little later, so if you wonder later on what a word means, do not worry it's probably going to be explained eventually. I chose this subject because I am very interested in computers and coding. I also want to be a software engineer when I grow up and finish school.

1 Computer

A computer is made of several components. Each component with its own purpose. It is called a computer since it computes bits of information. You could probably call it an overly complicated calculator. The first computer that is like the machines we know as computers today was the ENIAC (Electronic Numerical Integrator and Computer) built in 1946. It was the size of a room. Ever since then, computer technology has been rapidly advancing.

1.1 The processor

The processor a.k.a. CPU (central processing unit) is the most important part of the computer; it computes all the user's input and is the so-called brain of the computer. The CPU's clock (the number of cycles the CPU can do per second) speed is measured in Hz. Hz is a measurement used to describe the number of times a cycle or process is repeated every second ($100/\text{sec} = 100\text{Hz}$).

Most modern CPUs have speeds that surpass 4 GHz/4 billion Hz. But CPU speed does not mean everything when it comes to processing power. What plays a big role in this is CPU efficiency. You can have a fast CPU but that does not matter if it needs more cycles to solve a problem compared to one with lower speed but higher efficiency.

This means it is more about getting more done per clock cycle than having more cycles. The way a CPU is built and put together effects the efficiency and speed.

Another important part is the size of a processor's transistors. Most modern CPUs have transistors around the size of 5-14nm. The smaller parts, the more power efficient a CPU is. So, by making the parts



Figure 1: Image of an Intel CPU

smaller, one can not only cram more in but also make it more power efficient.

A CPU does not constantly run with the exact same clock speed. The clock speed will adjust to the amount of work the processor has to. Often when looking at the specs of a CPU, you will see “base clock” and “boost clock”. The exact words may vary but the keywords are “base” and “boost”. The base clock is the minimum speed the processor will be running at, so no matter what it is doing it will always at least be running at base clock frequency. The processor raises or lowers its clock speed based on its workload. This means that boost clock is the clock it will run at when it has to handle a lot or complicated tasks.

1.1.1 CPU cache

Note: Memory is explained in a later chapter

The CPU also has its own memory cache. This memory cache contains instructions, and problems for the CPU to execute and solve. While this memory cache often only has a size of a few Megabytes, it's extremely fast. This memory is divided into multiple levels: Level 1(L1), Level 2(L2) and Level 3(L3).

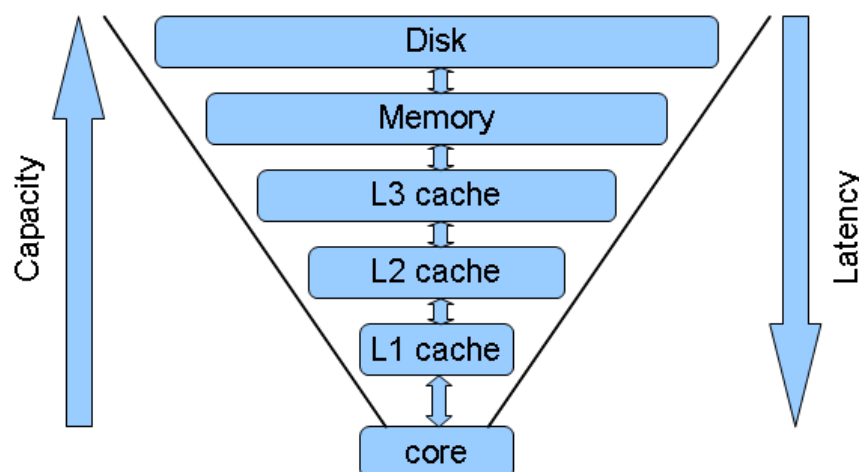


Figure 2: Depiction of memory sequence

The L3 cache is the slowest out of the three levels but it has the largest capacity. Data goes from the L3 cache to the L2 which is faster but has less memory capacity. From the L2 cache it goes to L1 which is the fastest but is often only a few Kilobytes (rarely 1-2 Megabytes) in size. The CPU is divided into processing cores. The L1 cache contains data and instructions that the core is likely to need next.

1.1.1.1 L1

Unlike the L2 and L3 cache, the L1 cache is divided into two sections, L1i and L1d.

L1i contains instructions for a core while L1d contains data that the core needs. The instructions tell the core what it needs to compute/calculate while the data section is the element that is calculated. Each core has its own L1 cache.

[1.1.1.2 Core cache priorities](#)

When a core needs data or instructions, it first looks in the L1 cache, if it does not find anything there it looks in the L2 cache, and then the L3 cache. Then if the CPU still does not get the data it needs, it looks in the RAM (Random Access Memory). The flow of data from the RAM to the CPU caches is controlled by the memory controller on the mainboard/motherboard. Receiving information from the RAM takes far longer than it does for its L caches.

[1.1.2 Cores](#)

A core is the part of the CPU doing the calculating. A single core contains ALU (Arithmetic Logic Unit) and has access to an L1 cache. The ALU does arithmetic calculations, so addition, subtraction, multiplication and division. While a core is far more complicated than I describe it here, I personally do not think that it is necessary to document it at this level of precision. There are a few other controllers in it as well that control the flow of data and instructions. The processes information using logic gates.

[1.1.3 Multithreading](#)

Multithreading is when a core of a CPU divides its attention between two tasks. You could compare it to multitasking. You can also call it parallel processing; the name speaks for itself meaning it runs two processes in parallel on a single core.

[1.1.4 Logic gates](#)

Note: Transistors are explained later.

Transistors are used in computer systems to form logic gates and other mechanisms. There are a few types of logic gates but the idea behind it is that it only lets information flow through it if a certain condition is fulfilled. A logic gate has input lanes, and output lanes. For instance, the AND gate, only outputs 1 if both of its inputs are also 1. Logic gates are used to perform calculations and generally dictate the flow of data.

[1.2 Random access memory](#)

Note: Bits are explained later in the chapter: "Storage"

Random access memory a.k.a. RAM, is memory like the CPU 's cache but in comparison it has a significantly larger storage capacity than the CPU cache. It is also a lot slower than the CPU cache. RAM is often also referred to as just memory.

Ram has several speed indicators. The first is its DDR type. Modern DDR is often referred to as simply DDR4(Double-Data Rate 4). Older types like DDR3 or DDR2 also exist but they are slower compared to DDR4. RAM also has a clock speed or frequency measured in MHz. NEWER RAM types like DDR4 can operate at higher frequencies than the older types of DDR. Its minimum speed is 1600MHz while the

maximum is 3200MHz. The quality of RAM in a computer can affect the CPU's performance. Meaning if you have slow RAM, there is a chance that the CPU is not running as fast it could be due to the RAM not being able to keep up with the CPU. Modern OS' (Operating System) like windows 10 need an absolute minimum of 4GB of RAM to be able to run.

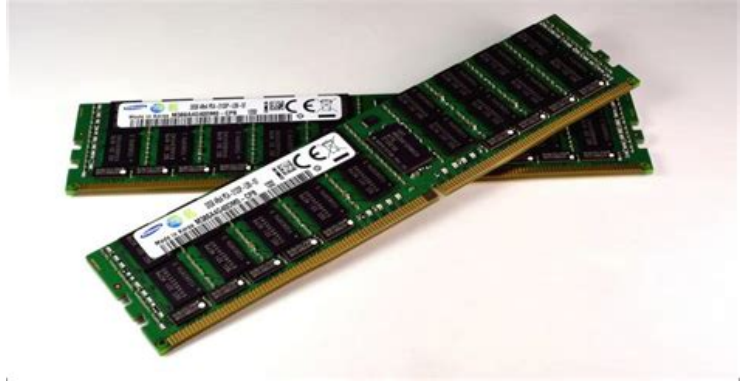


Figure 3: Image of RAM sticks

The two most popular types of RAM are SRAM and DRAM. SRAM meaning Static Random-Access Memory is faster but far more expensive compared to DRAM. DRAM meaning Dynamic Random-Access Memory, is slower than SRAM but it is cheaper and the most common type of RAM. RAM stores data in memory cells that can store one bit of memory, either a 1 or a 0. The way data is stored with SRAM differs from DRAM.

SRAM uses 6 transistors to form a memory cell to store one bit. With this system SRAM uses less power to maintain its stored data for a longer period. CPU caches are often SRAM since they need to be very fast.

DRAM uses a combination of a capacitor and a transistor to form a cell. This type of cell draws more power when it is not being accessed. DRAM has more potential for more storing space compared to SRAM since its cells are much smaller.

Another type of RAM that is not very common at all is non-volatile RAM. Unlike conventional RAM, non-volatile RAM can store information without there being a need for power to be applied to it.

1.2.1 Transistors

A transistor is a small electronic component. Its functions as an on and off switch for electrical currents but it also works as a current amplifier. A typical transistor has three terminal points: The emitter, collector, and base. The way it works is quite simple, the electricity flows from the collector to the emitter. The electricity can only flow if there is a current in the base of 0.7 volts. This is known as an NPN transistor.

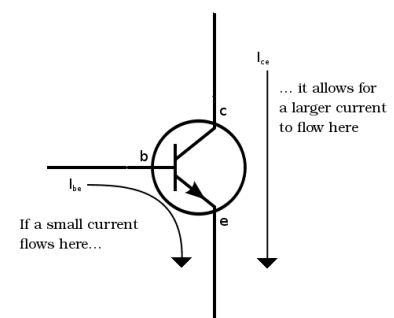


Figure 4: Image that shows how NPN transistors work

Another variant is the PNP transistor. Here the electricity can only flow from the emitter to the collector if the difference in current voltage between the emitter and the base is greater than 0.7 volts.

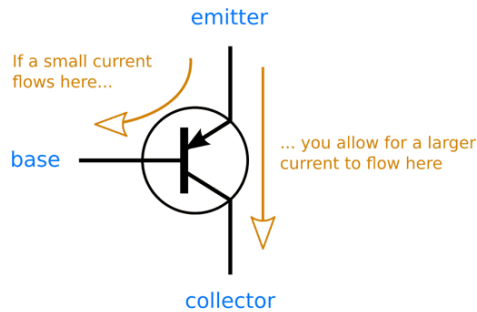


Figure 5: Depiction of the way PNP transistor works.



Figure 6: Image of PNP transistor

1.2.2 Capacitors

A capacitor works like a small rechargeable battery that can quickly discharge all its stored electricity. A capacitor is made up of 2 metallic plates with a non-conducting material in between. By applying a voltage to both plates, a positive charge will gather on one plate while a negative one gathers on the other. Capacitors are divided into two main types, polarised and non-polarised.

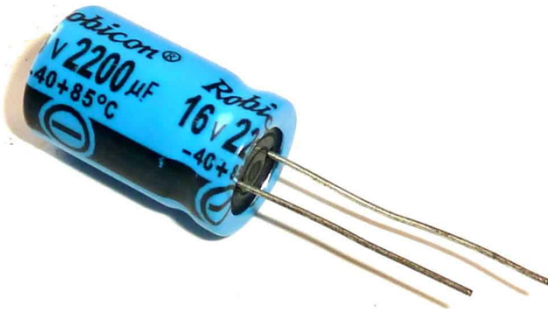


Figure 7: Image of capacitor

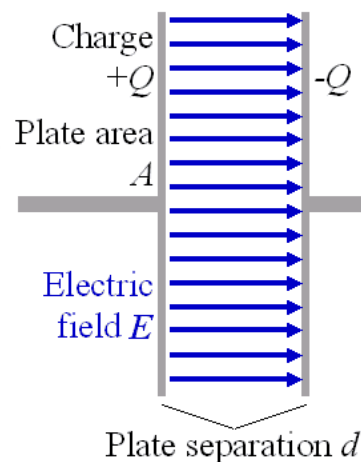


Figure 8: Depiction of how capacitors store power

A polarised capacitor has a positive and negative side which means you must place the positive side where the most positive voltage is. A non-polarised capacitor does not have a positive or negative side, meaning that it does not matter how you place it in your electrical circuit, it will always work.

1.2.3 Memory Cells

A memory cell is a fundamental part of computer storage. It is used to store a bit of information, either a 1 or a 0. There are many types of memory cells that are made up of different components, though the main principal stays the same, it stores a bit.

For example, SRAM uses the MOS memory cell which consists of 6 transistors. These transistors are used to form a Flip-Flop system. This basically is on and off, so 1 and 0.

1.3 Graphics Card/Graphics Processing Unit

The graphics card a.k.a. GPU calculates what is shown on the screen. This is one of the most expensive parts when you buy a computer. The GPU like most other computer parts also has Hz speeds. While they are lower than the GPU's clock speeds, they can still go up to 2 GHz. Graphics cards have the most processing power in the computer. When people look at GPU, they will often look at how many TFLOPS it can do per second. FLOP meaning Floating Point Operation, that means if it can 1TFLOP/sec it is doing 1 000 000 000 000 floating point operations per second. FLOPS cannot be compared to the



Figure 9: Image of a Graphics Card

term Hz. FLOPS describes the amount of float/decimal number calculations it can do every second while Hz does not specify any type of calculation.

The GPU has fast data transfer speeds since it needs to communicate with the CPU.

A floating-point operation is a calculation of decimal numbers. A GPU has thousands of cores doing these calculations. These cores are quite a bit smaller compared to those in the CPU, but they make up for it in quantity. CPUs with integrated graphics do exist but they are not anything special. You would undoubtedly need a GPU in your computer if you were a graphics designer, video editor etc.

The GPU has its own RAM known as VRAM (Video Random Access Memory). VRAM also has Double-Data Rate Types that go all the way to GDDR6(Graphics Double Data Rate 6). The most common type of VRAM is also DRAM. GPUs tend to be made for more specific tasks, like graphic designers or gaming, but just because it was made for work does not mean it cannot game. The opposite is also true.

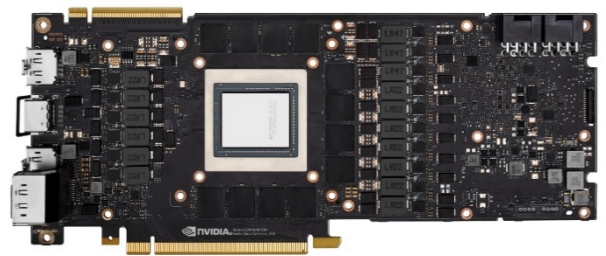


Figure 10: Image of a graphics card without its cooler.

1.4 Storage



Figure 11: Image of SSD

Storage in the computer is important since it contains the OS and all the data stored on the computer. Most computers today should at least 500GB to 1TB worth of storage space. The data is often stored in an HDD (Hard Drive Disk) or SSD (Solid State Drive). SSDs store data similarly to RAM. They also use cells made of capacitors and transistors to store bits of information.

HDD disks store bits using a magnetic film that is attached on each side of the disk. This film has lots of zones that can have different magnetic states. Depending on the state of the magnet, it holds a 1 or a 0. An HDD often has multiple disks in one drive. The data is read by the head of an arm that can read and write the data. Something you must pay attention to when looking at the specs of a storage device is its read and write speed. This is important as it represents the speed at which your device can store and load data.



Figure 12: Image of HDD that is open (They normally don't open this way)

1.4.1 Bits and Bytes

A bit is a piece of information that the computer can comprehend. A single bit can either be a 1 or a 0. Another representation of 1 and 0 is true/false, +/- or on/off. A bit is just a pulse of electricity, depending on its voltage it is a 0 or 1. While a single bit does not mean much, a sequence of bits can mean something. Other units of information like bytes are also commonly used. A byte is equal to 8 bits, meaning each bit can either be a 1 or a 0. While a bit would simply be a 1, a byte can be 10100110.

1000	10^3	kbit	kilobit	1024	2^{10}	Kibit	kibibit
1000 ²	10^6	Mbit	megabit	1024 ²	2^{20}	Mibit	mebibit
1000 ³	10^9	Gbit	gigabit	1024 ³	2^{30}	Gibit	gibibit
1000 ⁴	10^{12}	Tbit	terabit	1024 ⁴	2^{40}	Tibit	tebibit
1000 ⁵	10^{15}	Pbit	petabit	1024 ⁵	2^{50}	Pibit	pebibit
1000 ⁶	10^{18}	Ebit	exabit	1024 ⁶	2^{60}	Eibit	exbibit
1000 ⁷	10^{21}	Zbit	zettabit	1024 ⁷	2^{70}	Zibit	zebibit
1000 ⁸	10^{24}	Ybit	yottabit	1024 ⁸	2^{80}	Yibit	yobibit

Figure 13: Words that indicate the number of bits

Larger decimal units of information like kilobyte or kilobit also exist. One kilobit is exactly 1000 bits while one kilobyte is 1000 bytes. Other binary type units of information also exist like kibibit, which is worth 1024 bits or 2^{10} bits.

1.5 The motherboard/mainboard

The motherboard connects all the computer components. The RAM, CPU and GPU are all connected to the motherboard. This board also controls the flow of data from one place to another.

The CPU is embedded in the CPU socket on the motherboard. It can be removed and replaced by another compatible CPU. Motherboards are manufactured to be compatible with certain CPUs since different companies produce CPUs with differing socket types.



Figure 14: Image of a motherboard

It also has RAM slots; you can also remove the RAM and replace it with other compatible RAM sticks. Low quality motherboards often only have around 2 RAM slots while the medium to high quality motherboards have 4.

Attached at the bottom are one or more PCI or PCIe slots. These slots are where you would insert the GPU or any other extension cards like audio cards or network cards (Wi-Fi card).

The motherboard also has SATA ports that can be used to connect to an SSD or an HDD via a SATA cable. The

mainboard also comes along with NVMe slots, which has data transfer speed that is quite a bit faster than that of a SATA connection. NVMe is only used to connect to an SSD. NVMe can reach very high read and write speeds as it is placed directly placed on the motherboard while conventional storage must be connected via a cable. While I will not explain the exact components of a motherboard, everything I have mentioned will be on any modern motherboard. The exact specifications depend on the model and can greatly vary.

1.6 Bottleneck

Bottleneck is exactly what it sounds like. It is when the power of your computer is limited by one part that is far superior or inferior compared to the other parts.

Having a balanced build is essential to having stable performance. For example, if the GPU in your computer is the latest and the greatest but the rest of your parts are older and a lower grade. Then parts like the CPU or even motherboard would not be able to keep up with your

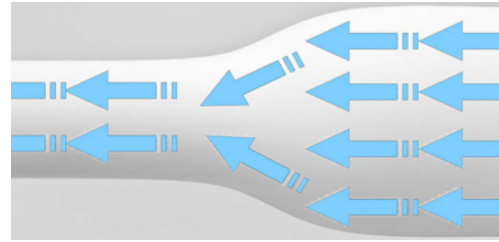


Figure 15: Depiction of a bottleneck

super-fast GPU. As the saying goes “A chain is only as strong as its weakest link”. When building a computer, making sure your parts are compatible and do not have a bottleneck, can be painful. An easy solution for this problem is already there. Parts like GPUs and CPUs often have gradings like low, middle, and high. For instance, one can figure out what grade it is by simply looking at the numbers. An example is NVIDIA where a part of the GPU’s name indicates the grade. The RTX 3070 is a mid-level GPU made by NVIDIA. The 30 indicates which generation it is and 70 refers to the model of the generation. Often the weakest will be 60 then 70 and then 80. Sometimes they hang something else onto the, but these are inherent indications. So, by checking its level, one can draw a conclusion on which parts to buy.

Note: Each company does this differently so checking how it does its naming would not hurt if you were currently looking to build your own computer.

1.7 Overclocking

When it comes to normal the use normal use of desktop computers, overclocking is not necessary, but I decided to throw it in anyways because it is interesting.

Overclocking is when you raise the base core clock to make it go faster. Often you will get promising results, for instance your computer might do complicated calculations in a shorter amount of time or you get more frames per seconds in your video games. But what you must keep in mind when overclocking is that you are at the same time taking a risk. By overclocking certain parts, a big side effect would almost certainly be heat. By raising the clock of your component too high, there also is a possibility of a system crash though this can be one of the smaller risks. By overclocking anything, the



Figure 16: Image of non-overclocked settings in MSI Afterburner

chances of it breaking sooner or even straight away get higher. This is almost certainly proportional to how much you overclock the part. Some people even see this as a contest, where they see how much they can overclock their computer. Then they would test it using a program that would strain the computer. This program gives you results which you can compare to others results via the internet. In some cases, people have even overclocked their computers to the point where they used liquid nitrogen (very cold liquid) to cool them.

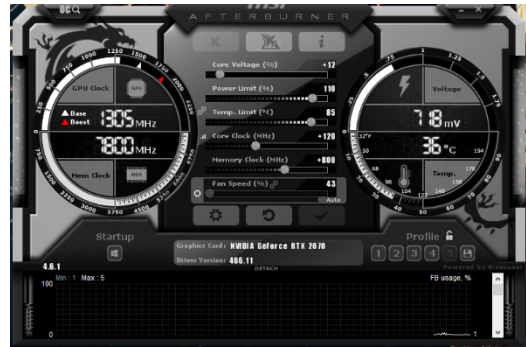


Figure 17: Image of overclocked settings in MSI Afterburner

1.8 The lifespan of a computer.

The lifespan of a computer is greatly influenced by many things like, how often you use it, how you treat it and the strain put on the parts. One thing that can also drive down the numbers of times your computer will live to see another day is heat. Computer part can deteriorate much faster if they are at a higher temperature. This partly why overclocking is not exactly healthy for it. So, keeping your computer as cold as possible should be a priority, so what I would not recommend is buying a cheap low-quality cooler.

2 Supercomputer

A supercomputer is a computer that has a massive amount of processing power. It takes up a huge amount of room and electricity.

These computers are usually used for scientific purposes such as looking for vaccines for HIV, COVID-19 and other hard to cure diseases.

Another very common use for supercomputers is running an AI (Artificial Intelligence) program.

A supercomputer is usually composed of thousands of computers, each with their own components. A computer in a cluster supercomputer is often referred to as a node. Only rich people or companies with a lot of money can



Figure 18: Image of Fugaku supercomputer

afford these massive computers. Most supercomputers are designed so that they perform best at what they are supposed to do. This often makes them have a unique layout and design that is not identical to others, though at a brief glance, they probably look similar.

2.1 How does a supercomputer work?

A supercomputer handles tasks like how modern desktop computers do multi-threading, just on a much larger scale. There two types of supercomputers, a cluster and a grid. A cluster supercomputer is often a facility that holds thousands of computers in a large room.

A grid supercomputer is more complicated; it is a supercomputer that is composed of lots of computers that are connected to the internet. Each of these computers has a software installed on it that connects it to the grid. Anyone can connect to a grid if they download and launch the published software.

Technically the definition of a supercomputer is just “A computer that has far more power than an average person’s computer”. So, this means a supercomputer technically does not have to be a cluster or grid of lots of nodes. Though practically achieving this, is lightly put extremely difficult. Though that does not mean it is impossible. One good example is quantum computers; these machines can calculate things that would take even supercomputers years. I’ll explain these in a bit more detail in a later chapter.

2.1.1 Parallel processing

A supercomputer works by splitting its work among the computers that are connected to it. This is called parallel processing; multi-threading is also a type of parallel processing. The tasks are not divided into cores but into whole computers.

Conventional programs do not use parallel processing, they use serial processing. Serial processing is when the computer does its tasks one after another while in parallel processing, the computer computes each task at the same time. Parallel processing is very effective when it comes to using it in the supercomputers, the time it takes to complete a task is significantly shorter. This is because parallel processing can take advantage of the computer's many nodes.

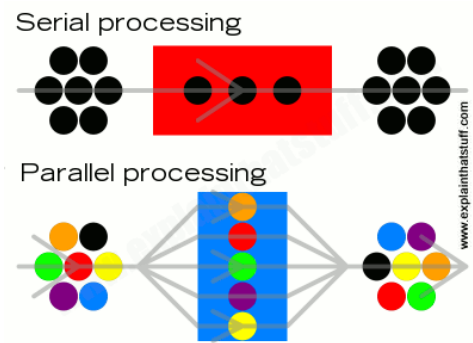


Figure 19: Depiction of the difference between processing methods

Parallel processing can sometimes be slower than serial processing when it comes to some tasks. One major factor that slows it down is that combining the results from each process can take so much time to the point where serial processing is faster. This is however often not the case.

When building a supercomputer, one would have to write a program that distributes its tasks among its nodes. The program must also be able to put the data back together when it receives its results from the individual nodes.

2.2 Uses for supercomputers.

Supercomputers have various uses like looking for cures for diseases or predicting the weather. But some might have a more unexpected usefulness. The ones described in this next part are some very common uses for supercomputers. While there are many more uses, I am trying to keep things simple.

2.2.1 Crypto mining

I am sure you have heard of crypto currencies and their practical use on the internet. Supercomputers can be used to "Mine" these crypto currencies. Since crypto currencies do not have central banks, to do things like confirming the validity of the money that is being transferred, crypto mining was invented. To keep it simple, to check the validity of money, a supercomputer is used. This supercomputer is a grid type which anyone can connect to if they download and run the correct software. Anyone who becomes a part of the grid, will be rewarded for making sure the money is not

somehow being messed with in an illegal way. Why would you need a supercomputer? The way this money is verified, simply put, is through very long and hard guesswork. To achieve the goal of guessing large numbers quickly, supercomputers are the best solution. Most of the time, the main component used in these crypto mining grids is a graphics card. When it comes to raw power the graphics card far outperforms the processor and is therefore better suited for this job. Every year when a new



Figure 20: Image of a crypto mining rig

generation of graphics cards is released into the market, a lot of them go to crypto miners. The year 2020 is a very extreme example where people were writing programs that would buy graphics card as soon as they were in stock. The result was that there were no cards left and that people were re-selling them for more than double the original price.

2.2.2 Simulations

A supercomputer has many uses when it comes to scientific tasks. One of them is simulating things. They have the capability to simulate things like new materials, the human brains and many more things. One example is the testing of nuclear weapons test. Since 1992, the USA has banned the testing of nuclear weapons. The best alternative ever since has been simulating them. By simulating nuclear explosions, they can confirm the weapons “safety” and functionality.

By simulating brains, researchers can figure out how the brain works. This can help them conduct virtual psychiatric treatments. They may even be able to make an AI using this simulated brain. A Swiss supercomputer managed to simulate a 10000-neuron chunk of a rat brain. They hope to someday make a complete model of the human brain.

2.3 Supercomputer specs

Just like normal computers, supercomputers also have specs or specifications. For some this is obvious, but I thought I should point this just to be sure. A supercomputer is often rated based on its FLOP performance. This FLOP is the same kind I explained earlier for the graphics card. Meaning that the more FLOPs a supercomputer can do per second, the better it is. While the performance of a graphics card is measured in Teraflops/s (Tera Floating Point Operations per second), a supercomputer is usually is measured with Petaflops/s (Peta Floating Point Operations per second). To explain the units of measurement used here, Tera means trillion (10^{12}) and Peta means Quadrillion (10^{15}). The sheer power these computers have is simply put insane. This also stands to reason as the price for these machines is also ludicrous. A perfect example is the Fukagu supercomputer. Its performance is around 400 to 500 petaflops. You do not have to be especially bright to know that this is a lot.

This might just be me pointing out the obvious, but components used in the computers are proportional to its awesome performance and price. Using the supercomputer Fukagu again, a single node of this supercomputer already has mind-blowing specs when compared to any home computer. It boasts a large computational capacity with its 32GB memory. While the memory size might not be mind-blowing (it is still much larger than your average computer), its speed most definitely is. It has a speed of 1024GB/s read and write speed. It has 48 processing cores and a large and rapid cache. Its L1 has a load speed of 256GB/s and write speed of 128GB/s. The L1 has even faster speeds with its 4TB/s load and 2TB/s write speed. So, if you think about it, thousands of computers that have the performance and price of a high-end gaming computer; it makes complete sense that supercomputers can reach these ludicrous speeds and prices.

The list called TOP500 lists the world's 500 fastest supercomputers. A supercomputer's place in this list is based on their number crunching power a.k.a. FLOP/s. A benchmark program is used to test the computer's capabilities.

2.4 Small supercomputer list

I decided to compile a list of some popular or relevant supercomputers that exist now.

2.4.1 Fugaku

The Japanese supercomputer Fugaku is the world's fastest supercomputer. It was finished in May 2020 after being conceived in 2014. It was made as a collaboration between Fujitsu and RIKEN. It has slightly



Figure 21: Image of the Fugaku supercomputer

more than double the performance than the second place in the list which is IBM's Summit supercomputer.

It has a total of 158,976 nodes. This amounts to around 488 petaflops (That is a lot). A single node has 48 processor cores. Its memory has the large capacity of 32GB. Its read and write speed is 1024GB/s. It runs a version of the Linux

operating system like most supercomputers. The CPU used these nodes is the A64FX made specifically for this cluster. A single CPU contains 48 cores, which have a base clock of 1.8GHz. Each core has an L1i and L1d cache, which each have 64KB of capacity. It has four sets of L2 caches each assigned to 13 cores. Each L2 cache has 8MB of capacity.

Up until now, it has mainly been used for scientific and medical purposes. It has worked on solutions like corona mask efficiency and tsunami wave simulations to forecast floods in Japan. It was also able to fully analyse cancer genes in less than one day which used to take a significantly longer time with any other supercomputer.

2.4.2 Meluxina

This supercomputer is not actually completed yet and it is not going to be the new best one. The reason I am putting this in the list is because it is something quite relevant to those who live in Luxembourg as it will be built there. It will be the country's most powerful supercomputer. With a speed of 10 petaflops per second, this supercomputer cannot be called slow by any means. The company building this computer is called LuxConnect. The government also has a hand in the supercomputer's construction. They plan to make the cooling mechanism as environmentally friendly as possible. Its main use is going to be to research new technologies for the



Figure 22: Image of the Meluxina supercomputer

industry. A company named IEE plans to use it to simulate things like a material's strength. One example is simulating materials for the automotive industry. Meluxina is a part of Europe's EuroHPC initiative to build 8 supercomputers across Europe. These supercomputers will make Europe more competitive in research fields like science, medicine, climate change, etc.

2.4.3 Summit

Summit is a supercomputer made by IBM and NVIDIA. It is the world's second most powerful supercomputer with 4608 nodes. The supercomputer cost around 400-600 million dollars.

A single node has 2 IBM POWER9 CPUs and 6 NVIDIA Tesla V100 GPUs. A single node has a total of 600GB of DRAM. It also has an additional 800GB of non-volatile RAM. The IBM POWER9 can have up to 24 cores, and it can have up to 96 threads depending on the specific model. It has a L1i and L1d cache of 32KB. Its L2 cache has a size of 512MB. The L3 has a capacity of 120MB. The CPU can have a clock speed of 4GHz. The NVIDIA Tesla V100 can do around 7.8 Teraflops. It has 32 or 16 GB of memory depending on the model. A single card consumes around 300 watts of electricity.

The supercomputer has a total of 5.5296 Petabytes of memory and 27648 GPUs (NVIDIA Tesla V100), and 9216 CPUs (IBM POWER9).

The supercomputer can do around 148 petaflops. It is going to research things like artificial intelligence, fusion energy, cancer, and many more things that can benefit our society.



Figure 23: Image of the Summit supercomputer

2.4.4 HPC5

HPC5 is a supercomputer located in Italy. It currently holds the 8th place in the TOP500 list. Its performance is around 70 petaflops. It was built by the computer company DELL.

The computer has a total of 3640 Intel Xeon Gold 6252 CPUs, 7280 NVIDIA Tesla V100 GPU. The Xeon has a total of 24 cores, its base clock is 2.1GHz while its boost clock is 2.8GHz. Its architecture size is 14nm. Each of core has 32KB of L1 cache and 1MB of L2 cache. The entire CPU has 35MB of L3 cache.



Figure 24: Image of the HPC5 supercomputer

2.5 Quantum Supercomputers

Quantum computing is still a technology that is in its early stages and cannot be applied properly yet. But because it is interesting, I thought I would include it anyways. The way quantum computers work is fundamentally different to conventional computers. Unlike normal computers that use bits to represent information, quantum computers use something known as a qubit. The quality that makes a qubit differ from a normal bit is the way it can represent information. Both types can be either 1 or 0, but a qubit can be something in-between, meaning it can be both a 1 and a 0 at the same time. This type of computing is very

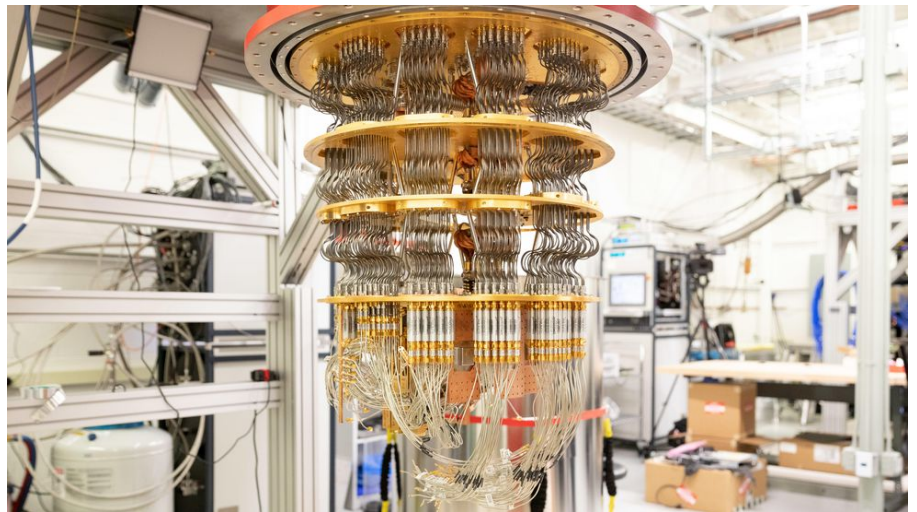


Figure 25: Image of a quantum computer

complicated and hard to understand, but essentially a quantum computer can theoretically calculate things very fast because of the special quality a bit has.

There are cases where researchers successfully completed quantum computations that would have taken any normal computer hundreds if not thousands of years.

Word Explanations

a.k.a. = also known as

CPU = Central Processing Unit

GPU = Graphics Processing Unit

OS = Operating System

FLOP = Floating Point Operation

RAM = Random Access Memory

NVIDIA = A company that makes computer hardware

GB/s = Gigabytes per second; this describes the flow of data, either write or read

Hz = the number of times something repeats every second

Conclusion

To sum everything up, supercomputers are mainly just ultra powerful computers used for science. They are big and expensive, and the average person could not afford one, even if they saved all their money for their entire life. But in return for the size and price, they offer a huge amount of performance and power. They make normal computer pale in comparison. While I did write a lot about normal computers, it was so you could understand all those words I wrote in the second part, like Petaflops. I personally think that all of this is extremely interesting, and if this also caught your attention, then I recommend doing some research on the internet as this does not cover everything by far. Something that really surprised me when I was researching for this, is the sheer computational power these computers have. While I knew these computers were powerful, I didn't think they would be to this extent. This subject is especially interesting to me, as I would like to work with artificial intelligence in my future job. Thank you for reading my travail personnel all the way to the end. This took a lot of time and effort and I hope you learned something new.

Sources

<https://www.explainthatstuff.com/how-supercomputers-work.html> 14/10/2020

<https://www.intel.com/content/www/us/en/gaming/resources/cpu-clock-speed.html> 14/10/2020

<https://www.makeuseof.com/tag/what-is-cpu-cache/> 15/10/2020

<https://en.wikipedia.org/wiki/Computer> 15/10/2020

<https://computer.howstuffworks.com/motherboard1.htm> 21/10/2020

<https://www.redhat.com/sysadmin/cpu-components-functionality> 21/10/2020

[https://en.wikipedia.org/wiki/Multithreading_\(computer_architecture\)](https://en.wikipedia.org/wiki/Multithreading_(computer_architecture)) 21/10/2020

https://en.wikipedia.org/wiki/DDR4_SDRAM 18/11/2020

<https://en.wikipedia.org/wiki/FLOPS> 07/01/2021

https://en.wikipedia.org/wiki/GDDR6_SDRAM 07/01/2021

<https://en.wikipedia.org/wiki/FLOPS> 22/01/2021

https://en.wikipedia.org/wiki/Random-access_memory 21/01/2021

https://en.wikipedia.org/wiki/Solid-state_drive 21/01/2021

https://en.wikipedia.org/wiki/Hard_disk_drive 21/01/2021

<https://en.wikipedia.org/wiki/Motherboard> 21/01/2021

<https://en.wikipedia.org/wiki/Bit> 22/02/2021

https://en.wikipedia.org/wiki/Units_of_information 22/02/2021

<https://www.explainthatstuff.com/howtransistorswork.html> 05/03/2021

<https://www.build-electronic-circuits.com/how-transistors-work/> 05/03/2021

<https://www.build-electronic-circuits.com/pnp-transistor/> 05/03/2021

<https://www.build-electronic-circuits.com/how-does-a-capacitor-work/?sfw=pass1615190124>
08/03/2021

<https://www.build-electronic-circuits.com/capacitor-types/> 08/03/2021

[https://en.wikipedia.org/wiki/Memory_cell_\(computing\)#MOS_memory_cells](https://en.wikipedia.org/wiki/Memory_cell_(computing)#MOS_memory_cells) 08/03/2021

<https://www.investopedia.com/tech/how-does-bitcoin-mining-work/> 21/03/2021

https://en.wikipedia.org/wiki/Video_card 14/04/2021

https://www.reddit.com/r/explainlikeimfive/comments/50b59v/eli5_what_means_node_in_supercomputer/ 14/04/2021

<https://www.fujitsu.com/global/about/innovation/fugaku/specifications/> 14/04/2021

<https://en.wikipedia.org/wiki/Supercomputer> 15/04/2021

<https://en.wikipedia.org/wiki/TOP500> 15/04/2021

<https://www.techrepublic.com/article/fugaku-the-fastest-supercomputer-in-the-world-is-finally-at-full-operational-status/> 15/04/2021

<https://www.iee-sensing.com/en/blog/details/2019/06/iee-will-use-the-supercomputer-meluxina-in-2020-for-innovative-simulation> 15/04/2021

<https://www.youtube.com/watch?v=TdGKkvLmmb0> 15/04/2021

<https://en.wikichip.org/wiki/ibm/microarchitectures/power9> 16/04/2021

<https://www.nvidia.com/en-gb/data-center/tesla-v100/> 16/04/2021

[https://en.wikipedia.org/wiki/Summit_\(supercomputer\)](https://en.wikipedia.org/wiki/Summit_(supercomputer)) 16/04/2021

<https://www.pcgamer.com/ornl-summit-is-the-most-powerful-supercomputer-in-the-world/>
16/04/2021

<https://wccfttech.com/nvidia-ibm-summit-supercomputer-official-volta-gpus/> 17/04/2021

<https://www.technipages.com/base-clock-and-boost-clock-explained> 17/04/2021

<https://www.fujitsu.com/global/about/resources/publications/technicalreview/2020-03/article03.html> 19/04/2021

<https://www.cpu-world.com/CPUs/Xeon/Intel-Xeon%206252.html> 21/04/2021

https://www.youtube.com/watch?v=g_laVepNDT4 21/04/2021

<https://www.quora.com/Is-it-correct-to-say-a-data-center-is-a-big-supercomputer?share=1>
21/04/2021