Computer Technical Specifications

Anthony Farrell

15th April 2019
Version: 1.4

Contents

Licence

About this document

This document was made for anybody who wants to either build or upgrade a computer in any way.

In this document I will try to go over all the components of a computer and the technical terms that are used to describe them.

This is not an exhaustive list, however I think anything that is not listed is not going to be relevant or important enough to warrant mentioning, and users are really not going to need to know about it.

Chapter 1

CPUs - the important numbers

1.1 What is a CPU and what does it do?

The CPU (Central Processing Unit) of a PC is often said to be the ‘brain’ of the system, because a CPU controls the basic logic and arithmetic operations, as well as handling interactions between system components such as the system memory.

1.2 The big companies

Whilst this may sound grandiose, and there is lots of discussion to be had about the differences between their products, the realistic short version of this section is as follows:
There are only two companies who currently make CPUs for desktop computers: AMD and Intel, and basically any current product from either will be fine for the average user.

Their products do differ somewhat in what they focus on, and differ hugely in how they work - even though they perform the same function for the end user.
Even though their products work differently underneath, to the average end user these differences are somewhat negligible, since the end result is all that matters.

Whilst this document does [try to] explain the main characteristics of typical PC hardware, the exact specifics of the differences between how AMD and Intel processors work is vastly beyond the scope, intended size, and audience of this document.
In addition, the exact specifics of how they work vary with each new product cycle developed and launched, and trying to keep up with these ever-changing processes is *very much* not the point of this document.

1.3 Socket

A CPU will have a *socket*.
This socket is the physical arrangement of the one side of the CPU that attaches to the motherboard it is to be located on.
They are thus the physical and electrical way that a CPU transfers and receives information to and from other components.
Processor sockets often change as a way of preventing hardware from being installed into systems that are incompatible with them, whilst also allowing newer standards to be used for faster and more advanced features.
As such a processor will need to have an identical socket to the motherboard than that CPU is to be used on, otherwise they can not be used together.

AMD typically does not change sockets particularly often, and thus processors from different product generations will often be compatible because they use the same socket.
This is not a guarantee that they will be compatible however, however firmware updates to motherboards are often released that allow them to use more modern CPUs on older motherboards that have the same socket.
This has the advantage that older systems can be upgraded with newer CPUs to improve performance, but has the disadvantage that older sockets may not be able to make use of newer features that a newer socket could offer.
Examples of AMD sockets include *AM3*, *FM2*, and the current *AM4*.

By comparison Intel uses new sockets regularly, and often uses a different socket with each new product generation or every second product generation.
This means that newer CPUs may have access to more advanced features, however comes with the downside that older systems will be completely and totally incompatible with newer CPUs, meaning a user with an older system will need to buy not only a new CPU, but also a new motherboard as well if they wish to upgrade, as well as any other components of the old PC that the new motherboard does not support.
Examples of Intel sockets include *LGA 1366* a.k.a. *Socket B*, *LGA 1155* a.k.a. *Socket H2*, and *LGA 1150* a.k.a. *Socket H3*.

As you may have noted above, Intel often describes their sockets using the term **LGA**.
Intel uses LGA for their current CPU sockets, and has done so for many years.
LGA is *Land Grid Array*, which means that there are pins protruding from the motherboard socket that make contact with contacts on the underside of the CPU.

By comparison modern AMD AMD desktop processors use *PGA*, which is *Pin Grid Array*, effectively the reverse of LGA, above: The CPU has pins on the bottom, which make contact with contacts on the motherboard in which the CPU is seated in.

Whilst there are advantages to both, both can be used to make great CPUs, and neither will serve as any form of limitation to a CPU.
In addition, companies have switched between them for different product generations.
For example the AMD AM4 socket uses a PGA layout, whereas their TR4 socket uses LGA.

1.4 Instruction sets, x86, and x64

An Instruction set is a specific series of commands that the computers CPU can execute.
This may sound like software, however they are closer to a specific combination of hardware and firmware than conventional software, and as such cannot be changed later on by an update or software patch to the CPU.
In addition, because the CPUs hardware is designed entirely around executing these instruction set commands as efficiently as possible, the CPUs hardware would likely not be able to be patched via software in order to add additional commands anyway.

There are several different instruction sets, the most well known being those of x86 and x64.
x86 was designed by Intel, and is the base set of instructions that desktops processors used for decades.
x86 was inherently limited by its 32 bit design though, that began to show its age, most notably that of RAM - a 32bit architecture can only address (i.e. use) about 3.5GB of RAM, which for a modern high end PC is simply insufficient.
x64 was an update of x86 made by AMD that allowed it to bypass the limitations of the x86's 32bit architecture, and it is these extensions that allow modern computers to run effectively.
x64 should last us well into the future.

All modern CPUs run AMDs x64 instruction set natively, but can also run Intel's x86 if needed.
Note: AMD and Intel cross-licence x86 and x64 to each other, which is what allows modern CPUs to be made to use both. If another company wanted to made modern desktop CPUs they would have to licence these, which would incur licensing costs, and is one of the reasons there are only two players in the CPU manufacturer space.

Some CPUs also run what is known as a *RISC* instruction set, short for *Reduced Instruction Set Computer*.
RISC instruction sets are - as their name implies, reduced, meaning they can not perform all the instructions of a full desktop CPU, such as x64.
This does not mean they can handle less instructions, and in fact many RISC instruction sets actually have more instructions than x86 or x64, they are just much more simple ones.
Almost all ARM CPUs run a RISC instruction set, such as *ARM v7*, and as such pretty much all mobile phones run a RISC instruction set of some form.
Note that most GPUs run RISC instruction sets, instead of x86, hence why the two have no

real overlap in function.

Other instruction sets exist, such as the open source *RED5* instruction set, which is a new project designed to make a free and open source (FOSS) instruction set that is both capable and scalable, as well as being independent of the x86 and x64 instruction sets that must be licenced from Intel and AMD, respectively.

All modern computer operating systems will typically only run on either x64 or ARM though - Desktop operating systems run on x64, and mobile ones typically on ARM.

1.5 Processing Cores

A processor core is all of the hardware that is required in order to correctly execute every possible instruction that that CPUs instruction set supports.
As such a program can run completely on a single processing core.

Modern CPUs typically include numerous CPU cores.
Having many CPU cores potentially allows them to have numerous sets of programs running at the same time, which can be a multiplier to how much work the CPU can perform at once.

Not every task can be spread across multiple cores to make use of all of them at once though.
For example, I start with the number 2, and I want my CPU to continually double that number.
This is a very simple task.
The CPU would start by doubling it to 4, then 8, etc, however this task would likely be impossible, or at least very difficult, to use multiple cores for, because in order to know what number it has to double, it has to have finished doubling the previous number, so you have to wait for the previous calculation to finish before a new one can be started, thus only one CPU core would likely be used.
To put this example simply: *In order to calculate what double the double of* 2 *is, you need to have already calculated what the double of* 2 *is to start with.*

Most modern computer programs are not infinitely multithreaded - the ability to use multiple CPU cores at the same time.
Most conventional programs are limited in cores by how they are programmed, and changing a program to use more CPU cores is often an extremely difficult task, potentially requiring rewriting multiple parts of the program from scratch.

More cores are still useful for the end user though, even if any given program can't make

use of them all.
This is because in the future more advanced programs the user will run may be able to make use of them, or because the user may run multiple programs at a time, which can use different cores, making better use of the CPUs resources, even if a single program can't make use of all of the cores.

Most games typically do not currently make use of more than 4 cores, and only certain very modern examples being able to use more than 8.
8 cores is therefore about the current maximum number that modern games can use as of time of writing, however this will increase as time goes on.
Games that can make use of lots of cores can bottleneck in places if they do not have sufficient amounts of cores, even if the cores they do have are extremely fast.
As such having the maximum amount of cores that the games you want to play can use is heavily recommended, and CPUs that have a higher amount of slower cores can often deliver a much more stable performance, even if their average performance is slightly worse, because they will not see hard bottlenecks due to insufficient core counts.

1.6 SMT a.k.a. *Hyperthreading*, and 'threads'

A physical distinction exists between a computer *core* and a computer *thread*.
For the average end user though, they are extremely similar.
Both a core and a thread outwardly provide a single processing core visible to the computer that it can use to perform calculations, and thus run programs.
The difference though is that a *thread* is not technically a full computer core.

Processors can take different amounts of time to execute different commands, and a processor must have finished executing the previous command before it can start work on the next command to be processed.
There is thus a small delay between instructions executed by a processor and the start of the next instruction to be executed, as the processor must wait for the next instruction and any relevant information for the calculation to be fetched for it.
SMT makes use of this small time delay, adding additional hardware to effectively cache an instruction, ready for the processor to execute the instant it finishes the previous calculation.
As such when the processor finishes an instruction, it has another instruction waiting for it to execute, thus allowing it to drastically reduce the downtime between instruction executions.
The additional instruction that is processed during this time-window between instructions on the real CPU core is known as the CPU *Thread*.
As such a *thread* is a different form a processing core, however to the end user the two are functionally extremely similar.

SMT (*Simultaneous Multi-Threading*) is the generic name for this technique of processing

instructions in the downtime between the normal cycle of instructions the CPU runs.
Intel instead refers to this technique as *Hyperthreading*.

Because the processing thread causes the processor to execute more instructions, it therefore causes the processor to use more power (since executing instructions will use power), which causes more heat to be generated.

A thread is never going to quite match the performance of a full processing core, and as such cores would be chosen over threads if there were ever a choice and all else were (somehow) equal.
However adding a thread onto an existing processing core requires less additional hardware than adding a full CPU core, which means less physical space on the CPU is used, potentially allowing more room for other things, such as more cores, or a lower price of the CPU for the same thread count, and as such “Cores vs Threads" is often not a clear-cut argument.

More threads can dramatically improve the performance of a system if the programs it is running can make use of the additional threads.

1.7 Frequency a.k.a. *Clock Speed*

Frequency is defined to be the number of times a certain thing occurs.
The unit of measurement for frequency is Hertz (Hz) - the number of times per second something occurs.
For example 20Hz means that something happens 20 times per second, or once every 0.05 seconds.

Modern processors have their speeds defined in some multiplier of Hz, these days typically Gigahertz (GHz), with 1GHz representing 1 billion Hz.
The Hz of a CPU indicates the speed at which its cores can execute instructions, with a higher frequency representing more instructions being able to be executed in any given time.
As such *if [and only if] all else were equal*, a processor with a higher frequency would be able to perform more operations than one with a lower frequency, and thus operate faster.

CPUs can also be *Overclocked*, which increases this clock frequency, thus allowing the CPU to execute instructions faster.
Overclocking can be a good way of getting more performance out of a CPU, however a CPU can only be overclocked so far, either because of limitations in what the CPU can handle without instability, or just a limit to how quickly the additional heat an overclocked CPU generates can be dissipated before the CPU overheats.

1.8 Turbo Frequency

Modern CPUs also frequently list a *Turbo Frequency* (or something similar in name to this, such as “Boost Frequency") in addition to a clock speed.

Modern CPUs often do not have applications make effective use of all of their CPU cores, and often have some CPU cores not performing much work. These CPU cores consequently do not generate much heat, whereas other CPU cores are maxed out, due to either bad programming of the programs being run, or technical limitations of the work that must be performed.
This can lead to a CPU imbalance, where some CPU cores are maxed out and being used 100%, whereas others are often at idle, or doing very little work.
Turbo frequencies were thus introduced as a way of increasing performance without bypassing the thermal limitations of the CPU and its cooling setup.

How turbo frequencies work is by dynamically and automatically increasing the clock speed of a CPU core or cores to increase performance of those cores, and is usually able to be maintained indefinitely whilst the CPU has not hit its thermal limits.
Originally Turbo frequencies applied to just one or two cores, but they may now often apply to any cores that require additional performance.
As such some CPUs may list an *all-core* turbo frequency, which is what the CPU will overclock all of its cores to to increase performance if it has the thermal headroom.
Alternatively a *single-core* turbo frequency may also be listed, which is what a single CPU core may be increased to, with other cores at lower frequencies. This is often useful as many applications (for example; most games) use one CPU core to a much larger degree than the others.
Turbo frequencies are thus how far the CPU can push itself if and only if it has thermal headroom to spare.
Turbo frequencies will not activate if the CPU is too hot, or does not have sufficient power delivered to it to sustain the higher amount of power consumed by the CPU performing more work by the increased frequency.

1.9 Overclocking and underclocking

Many consumers enjoy the process of *Overclocking* (a.k.a. *OC'ing*), which is a process whereby users can increase the clock frequency of their processors cores, thus increasing performance.

Overclocking can be a good way of getting a nice performance boost out of a CPU, however has numerous drawbacks:

- **Power consumption** Processors with higher clock speeds can perform more work, and thus can use drastically more power than non-overclocked CPUs.
- **Heat** - Since more power is being drawn by the CPU, more heat will be produced. For most modern CPUs the limit of how far their clock speeds can be overclocked is the rate at which the heat generated can be removed.
  As such overclocked systems often include better cooling systems to help dissipate this extra heat generated.
- **Stability** A CPU core is designed to run a specific set of instructions, and the default clock speed allows a certain time frame for these instructions to be run.
  Computers store information as ones and zeros, and the computer uses voltage to store these discrete bits of information electrically. To swap between these requires an electric current, however overclocking a CPU reduces the time available for this to occur (this is because frequency is a measurement of how often things occur, so increasing frequency means more things occur in any given time, so each thing has less time), which makes it more prone to error as the same voltage must be used to do the same task, but in less time, which can result in errors as if the voltage change is not done completely it may be misinterpreted, as intermediate voltages between 1 and 0 can be misinterpreted.
  More voltage can be used however, which will allow the correct charge levels to be hit in less time (since more is being delivered in that time than it was before), however more voltage will result in more heat. In addition more heat lowers the electrical conductance of silicon which makes hitting the correct voltage harder.
  As such overclocking CPUs can result in them becoming unstable if all the parameters of the CPU are not set up properly, as the CPU may not properly execute instructions.
- **Time** Every CPU is different in the quality of its silicon, and as such will behave slightly differently and have slightly different voltage, current, and thermal properties. As such each CPU will need its properties modified to levels unique to that CPU, and finding the correct parameters for a CPU can take time, since simply loading a set of values that work on another CPU (even of the same type) may not be stable or thermally sustainable in another system.
  Thus, a user must actually spend time on that CPU to find the correct values for it.
  As such, overclocking to high degrees can take a lot of time to do properly.

Do also be aware that the opposite is possible, termed *underclocking*, or *undervolting*, which is the term used for the process of reducing the clock speed of a CPU.
This will of course reduce the performance of the CPU, however the CPU will use less power and thus generate less heat, which for some applications may be desirable.

Overclocking is often used on older hardware to help extend its lifespan, or on new components to get the maximum performance possible from them.
Alternatively many enjoy overclocking just for the fun of it, or for the benefits of learning how to tinker with their systems and figuring how everything works and interacts.

1.10 Locked and Unlocked CPUs

Locked CPUs are designed to prevent modification to them to prevent overclocking, whereas *Unlocked* CPUs are supposed to be open, and thus be modifiable, allowing for overclocking via modification to their operating parameters.

Locked CPUs can often still be overclocked, however this is not guaranteed, is usually more difficult, and the results more limited than what could be achieved with an unlocked CPU.

In the past, unlocked AMD CPUs have often been referred to as “Black Editions", however all of the modern Ryzen series of CPUs are unlocked.
Modern Intel processors that are unlocked will have a “K" suffix.

1.11 Cache

Every CPU features a *Cache*, which it uses to store data.
This cache is orders of magnitude faster than RAM, being located inside the CPU itself as part of the hardware, which it uses to store very frequently accessed information for extremely quick retrieval and use.
Caches are usually split between a *data cache*, and an *instruction cache*, which of course store instructions and data that the CPU frequently uses, respectively.

Modern CPUs typically feature three types of data cache (L1, L2, and L3), however some older or lower-end CPUs may only feature two.
Caches are ordered, with the lower level cache’s being faster, but smaller. Higher level cache’s are much larger, but also slower.
Each core will have its own L1 cache, as sharing these could massively decrease performance of the CPU as a whole. L2 cache is often split between pair of adjacent processors, and L3 cache is usually shared across all cores for the processor.
A processor will check these caches for the data it wants, only going to read the data from the much slower RAM if the data is not found in a cache.

Higher amounts of cache can therefore dramatically speed up system performance, however how much is actually gained in practice by a larger cache will heavily depend on the applications being run and how they are written, and the CPU and its architecture.

The average end-user very unlikely to be able to notice a difference in CPU caches.

1.12 IPC and architecture

IPC, or *Instructions Per Cycle* is a term used to describe the fact that modern processors have different underlying configurations of hardware: They are all built differently, despite the fact that they all perform the same tasks.
This difference in construction means that each processor will process instructions with a different efficiency.
As such the fact that two processors have the same frequency, clock speed, or indeed any other metric, does not mean that each of their cores will be able to execute the same amount of instructions in a given time, as they will execute them in different manners, and thus with different efficiencies.
This efficiency leads to the term *Instructions Per Clock*, which is a somewhat-abstract measure of how quick the CPUs architecture is to execute instructions.
Precisely measuring IPC is *extraordinarily* difficult, however getting a rough idea or ballpark is often not too hard.

This also leads to the term *Architecture*.
A processors architecture is the underlying hardware configuration it was built on.
For example AMDs *Ryzen RX 2XXX* series of processors use the *Zen+* architecture, and Intel's 8th and 9th generation processors both use the *Coffee Lake* architecture.

Typically Intel has had a higher IPC than AMD, and as of their current processor lines: AMDs *Ryzen RX 2XXX* series, and Intel's *Core iX 9XXX* series, Intel has about a 5% better IPC than AMD, meaning that for equal frequency, Intel CPUs will perform about 5% more instructions per given unit of time.
Alternatively: Intel CPUs will process an instruction in about 5% less time than an AMD one.
Note though that this number can vary due to implementation, the application being tested, and the instruction set used - it is by no means a fixed number, an can be extremely hard to accurately test.

For almost all users this difference will not be noticeable, especially as the processors will have different numbers of cores and different frequencies which will usually be the more important factors in overall performance.
For situations though where the processor is executing a program that requires a very fast response time delivered to the user, this difference can be perceived as important, however.
Realistically the only application where this might make a noticeable difference is in gaming, where a small decrease to the time it takes to deliver a frame to the user could be an advantage, as this will decrease frame times, thus deliver information to the user faster.

In truth though, other factors, such as processor frequency, number of cores, and how well the program is written to use both of those will make a *lot* more impact to the users end experience.

It is important to note though that the differences in which processors processors execute instructions has led to some unusual exploits and quirks.
For example, at the beginning of 2018 the *Meltdown* and *Spectre* bugs were discovered in current processors, and have been described as the most serious exploits ever discovered in computing hardware.

Spectre used an exploit of how modern CPUs handled something called *Speculative Execution* - the term used for the fact that modern CPUs could use a technique to process multiple possible outcomes of a calculation at the same time, thus saving time since it would already have the correct result of the calculation ready for immediate use.
However, since AMD CPUs do not use speculative execution, they were immune to this exploit.
The Meltdown bug allowed a process to read all of the computers memory(RAM), even parts that were supposed to be protected, and not able to be read by it. It could do this by bypassing the checks required to occur to test whether it was allowed to read this data.
The meltdown bug affected *almost all* modern CPUs, and required updates to operating systems in order to mitigate. These updates did reduce performance of the CPUs. How much performance was lost depends on the CPU in question, the applications being tested against, and so on, and thus is a hard number to pin down, but the 5-30% range is believed to be approximately accurate.

1.13 iGPU

A GPU is a Graphics Processing Unit - it is the piece of hardware that processes what the user sees, including the computers desktop environment presented to the user, as well as handling calculations related to manipulating 3D objects, such as for gaming.

Every PC will need a GPU of some form in order to display something to the user.

An *iGPU* is an *Integrated GPU*, that is to say; a CPU that contains the hardware for a GPU as well.
As such if a PC contains a CPU that has an iGPU, then the PC will not need a dedicated discrete GPU card plugged in in order to display output to the user, since it already has the GPU hardware in the CPU that can do it.

iGPUs tend to not be very powerful, and even the most powerful iGPUs would struggle to run modern games at high framerates and detail levels.
For the average user though, an iGPU is more than capable of doing basic tasks such as word processing, spreadsheet manipulation, or playing videos. As such having a CPU with an iGPU can be a cheaper option for building a complete system if the performance of a discrete GPU card is not needed. Discrete GPUs are usually only needed for professional workloads, or for gaming.

1.14 TDP

TDP is *Thermal Design Power*, and is a term used to mean the maximum amount of heat that the product is designed to be able to dissipate.
It is usually given in units of *Watts*.
A higher number therefore indicates better cooling capability.

A TDP is *not* the maximum amount of power the CPU can use, and that figure is often higher than the TDP rating of the CPU.

1.15 Cooler

The CPU will require a CPU cooler to cool itself in order to keep cool enough to remain functional, as well as remain cool, which will help the lifespan of the chip, as well as help in functionality such as dynamic overclocking (e.g. turbo boost).

Most CPUs will come with a cooler, and generally only high-end unlocked CPUs will come with no cooler. This is because the manufacturer expects these chips to be used for overclocking, and as such expects the user to add a high-end aftermarket cooler, instead of the stock one they provide.
These stock coolers are typically not great, but are functional, and should keep the CPU cool enough to not cause any thermal issues, however aftermarket coolers can be dramatically better. (Note: they can still suck though, and not all 3rd-part coolers are actually better than the stock ones the CPU comes with.)

1.16 IHSs, TIM, and *Delidding*

The *Integrated Heat Sink* of a CPU is the solid metal covering that goes across the top of the actual CPU die underneath the IHS.
This IHS is always made of Nickel-plated Copper, for very good thermal transfer.
The standard CPU cooler that people mount to their CPUs will attach to the top of this IHS.

The point of the IHS is to allow the heat from all areas of the CPU die underneath the IHS to spread out and be cooled, even if the CPU cooler does not make good contact above it.
IHSs also serve to protect the CPU die underneath from damage.
CPUs IHSs are attached to the CPU die itself with a thermal compound to allow for good thermal transfer, and are usually held in place with glue of some form around their edges.

CPUs can be *Delidded*, which is of course to remove the IHS from on top of the CPU, exposing the CPU die itself underneath.
Many people do this to replace the default thermal compound underneath the IHS, which can dramatically reduce CPU temps, often by 10 or 20 degrees or sometimes even more, however how much of a change will occur will occur on other factors too, such as the efficiency of the attached CPU cooler, the thermal transfer rates of the new and outgoing thermal compound, etc.
Delidded CPUs also sometimes have their default thermal compound (a.k.a "*TIM*" - *Thermal Interface Material*) replaced with *liquid metal*, which is an electrically conductive compound that is often much better than standard thermal compounds.
The disadvantage though is that extreme care must be taken with it, as the electrically conductive nature of the liquid metal can cause shorting of components which can destroy the CPU if it spreads out too far and touches something conductive.

In addition to delidding to replace the thermal paste, it is also possible to buy *direct die coolers*, which are CPU coolers that attach directly to the CPU die itself, rather than the IHS that would usually be above it, and thus require a delidded CPU to be used.
These reduce the volume of thermal medium between the CPU die and its cooling mechanism, and can thus increase thermal transfer rates, which can increase cooling, and thus decreases temperatures of the CPU.

Note that whilst many CPUs can be reliably and safety delidded, and tools are available to buy that assist in the process, all forms of delidding can damage a CPU, and a CPU that has been damaged by delidding will generally be irreparably destroyed.

Also note that certain CPUs will have a soldered TIM underneath. These soldered TIMs will be much stronger and will resist attempts to delid the CPU.
Delidding a CPU with a soldered TIM is extremely ill-advised, as when the soldered TIM breaks it can rip off surface components of the CPU die underneath the IHS, which will

completely destroy the CPU.

1.17 Binning, and the Silicon Lottery

Processors are made of various electrical components connected together. These components may have small impurities or irregularities. The product will still be able to meet the minimum specifications of that product, however the product may be slightly better or worse in certain regards.
Binning refers to the process of categorising chips according to this, thus higher/better binned chips have higher silicon quality.
For most users, this won't make any difference at all, however for overclockers who want to get the absolute maximum out of their components, the binning of their chip can have a relatively major impact on the maximum capabilities of the chip.

The "Silicon Lottery" is a term also used to describe this, whereby certain electronic components (i.e. pieces of silicon) are just inherently better than others.
Thus somebody has "Won" the silicon lottery if they get an extremely good version of that chip, which will perform extremely well.
This term is usually used to refer to CPUs that can sustain very high frequencies via overclocking.

Note that binning is not exclusive to CPUs, and other components such as GPUs may also be binned, however the term is usually used to refer to CPUs.

In addition, whilst the term "binning" refers to chips that are all capable of reaching the minimum requirements set of a specific product, many of the chips produced may fail to meet these minimum requirements, and such chips are often sold as lower-tiered product. This is usually more common at the start of production cycles for chips.
The additional functionality that the higher-tiered chip possessed is usually just disabled via software. It is possible that the user can re-enable this functionality via software as well, however the chip may be unstable, or may be unable to do certain things - there was a reason it wasn't sold as the higher-tiered version to begin with, after all.
Be aware though, that very rarely, a company may just disable features on working chips of higher tiers so they can sell them as a lower tiered product. This has the advantage that the company does not need to design a lower-tiered product - a process that could cost billions, and the savings in R&D could outweigh the additional cost required to produce the higher end chips though. This is very rare though, but has happened.

1.18 Northbridge

The Northbridge is one of two chips that the system uses to talk to external devices.
The Northbridge communicates directly with the CPU via the front-side bus, and allows communications that require extremely high performance, usually between the CPU, RAM, and PCIe lanes.
Modern CPUs usually have the Northbridge built into the CPU itself, whereas on old motherboards it was a separate chip on the motherboard.

1.19 Southbridge

The Southbridge is one of two chips that the system uses to talk to external devices.
The Southbridge communicates with system devices that do not require the high performance of the Northbridge. These include PCI, USB, IDE, or other such interfaces.
The Southbridge communicates with the CPU through the Northbridge, instead of communicating with it directly.

1.20 RAM Channels

The memory controller that handles communication with the systems RAM will be able to operate at a specific number of *memory channels*.
Single channel is the minimum, however for higher channels, such as *dual channel* or *triple channel* there is more channels of communication allowed between the RAM and the memory controller.
This thus effectively multiplies the data rate to the RAM by the channel number.
i.e. Dual channel RAM will have double the data rate to the RAM than single channel.
This is *not* the same as DDR RAM, and the two should not be confused.

From a hardware perspective this is done by increasing the number of RAM sticks that can be accessed at a time, with sticks above the number of channels sharing the bandwidth.
For example if a single channel memory controller has two RAM sticks they will share the data rate, whereas a dual channel memory controller will be able to use each of them at the max rate. If that dual channel controller has four RAM sticks, then it will have the same data rate as two sticks of RAM, with the additional two sticks sharing that data rate.
Thus having more RAM channels is advantageous to data rate, and thus can improve performance if the additional RAM sticks are present to utilize them.

Memory channels are best utilized by RAM that is identical in speed, latency, and capacity.

If memory of different speed are used, then the motherboard will run all the RAM modules at the speed of the slowest (lowest speed) module installed.
If memory modules of different latencies are used, then the motherboard will run all the RAM modules at the latency of the slowest (highest latency) module installed.
If memory of different capacities is used, then the motherboard will often simply run all the RAM in single-channel mode, however some chipsets may be able to use any RAM modules of equal capacity in the appropriate channels for those modules, and then running all other RAM modules in single-channel mode, however this depends heavily on the memory controller for how it is implemented, and many memory controllers will just run everything in single channel mode if the capacities are not equal.
Different RAM models, or RAM that is not from the same manufacturer (or other such parameter) can be used together in multiple-channel memory configurations without limiting each other, however again, they must be the same speed, capacity, and latency to not limit each other.

Multi-channel memory is also the reason why RAM sticks should not be put into just any slot on the motherboard, instead each motherboard will have specific slots that each RAM stick should be put into, based on how many sticks are present, such that the sticks can each use their own data channel if possible, for maximum bandwidth.
On many motherboards these are colour coded, and RAM sticks should be put into the same colour slots if all the RAM slots are not utilized. On some motherboards where the RAM slots are not labelled, the motherboards instruction manual will state which slots RAM is to be plugged into if all the RAM slots are not utilized.

If you are wondering why this section is under "‘CPU" and not "RAM", this is because the memory controller for the system is part of the Northbridge, which is for modern computers part of the CPU itself, and as such the CPU determines the maximum number of RAM channels that can be used, not the motherboard, which merely facilitates them.
The RAM itself doesn't affect the number of memory channels that can be used at all.

Most modern consumer CPUs use Dual channel RAM, with many high end professional platforms able to use up to Quad channel.
In the past, consumer CPUs have varied more in their RAM channels, with some low-end systems using single channel RAM, or some using triple channel.
For example Intel's 1st generation i7 series CPUs used triple channel memory, however all subsequent i7 platforms have used dual channel, whereas all of Intel's HEDT (High-End DeskTop) CPUs use Quad channel memory.

Chapter 2

Motherboard

For the most part, most motherboards available today will usually just list what connectors they have available, such as what CPU socket they have (which determines what CPUs they can accept) RAM slots they have (which determines what type of RAM they can accept), and how many PCI/PCIe or SATA or USB sockets they have available, or how many fan headers they have available.
For the most part you can simply refer to the section of this guide for each specific component that plugs into the motherboard for the specification.
Motherboards do have a few unique elements to them though, as denoted below.

2.1 BIOS and UEFI

BIOS is *Basic Input Output System*.
The computers BIOS is a piece of firmware that is used to initialize the computers hardware, load the bootloader that initializes the computers OS, provide options and modifiable parameters for overclocking various components, communicate between the hardware components, and can also potentially provide hardware abstraction layers, although most modern computers access the hardware directly.

All computers will require a BIOS of some form to function at all, and nothing will run without one.

On many PC, especially older ones, the BIOS was stored on a small removable chip on the motherboard, and the user could remove the BIOS and replace it, perhaps for a newer version, or with a working version in case of a problem with the old version.
For most modern PCs, the BIOS is instead stored on non-removable flash memory.
Flash-based BIOSs have the advantage that they are easily rewritable, and thus can be easily updates or potentially rolled-back, as desired or required, however if the BIOS becomes

corrupted, then the motherboard will usually be bricked, as there is no way to replace the BIOS chip.

For most users, they will never need to interact with their BIOS, and do not really ever need to know it is there.

Most modern BIOS systems are based on very old versions, and the underlying functionality of the BIOS has not really changed in decades.
As such, a newer standard was introduced that aimed to overcome some of the usual shortcomings of most BIOS systems: UEFI (*Unified Extensible Firmware Interface*).
UEFI replaces the BIOS, and provides the interface between the firmware of the systems hardware, and the systems OS.
UEFI is usually backwards-compatible with BIOS services, such that anything that requires a BIOS instead of UEFI (for example for compatibility reasons) will function normally.
UEFI was designed to allow GPT storage disks on large drives (over 2TB), CPU independent architecture and drivers, a modular design, backwards and forwards compatibility, and other advantages.
Since most users do not ever need to interact with their BIOS, UEFI does not provide any noticeable advantages for the average PC user.

UEFI also possesses a “Secure Boot" function that is designed to prevent unauthorized operating systems from being loaded during the systems startup process by preventing loading of any OS or driver that does not have a cryptographic digital signature, and thus can be considered a security measure.
However Secure Boot has come under criticism because it is often enabled by default on some PCs that run Windows, and secure boot can prevent the installation of other operating systems. In addition there are concerns about Secure Boot, such as how the computers hardware must come pre-signed with keys, and this is difficult to do for users who build their own kernels (since they must sign their own keys), as well as the licences that software is released under potentially preventing distribution of that software without its associated key - which renders Secure Boot somewhat redundant.
Concerns have been raised that Microsoft’s secure boot requirements are anti-competitive, obstructive, and unnecessary, and have resulted in legal disputes.
There have also been arguments made that Secure Boot can actually be a large and irreparable security flaw should the private key used to sign and secure the UEFI OSs ever be discovered, as this could allow anybody to create a legitimate and verified binary, regardless of its actual source or contents, and could be used to effectively lock out any changes to it, including legitimate ones, and thus could be exploitable by rootkits or viruses.

2.2 CMOS and the CMOS battery

The *Complementary Metal-Oxide Semiconductor* is a small amount of memory on the computers motherboard that stores all of the settings used by the computers BIOS.
Examples of settings stored include the date and time.
For most users who do not modify their BIOS, these settings are unlikely to be changed, however for more advanced users they may wish to modify their BIOS settings in some way.
If these CMOS settings become incorrect, such that they interfere with the operation of the system, then clearing the CMOS can be used to effectively restore the BIOS to its default settings, which should restore the system to a working configuration.
Many computers will have a dedicated button for clearing the CMOS.

These CMOS settings are kept because the CMOS continually receives power. Whilst the PC is plugged into the mains, it will usually draw a tiny amount of power to power the CMOS (as well as perform other tasks, such as powering status lights), however if the computer stops receiving mains power, then in order to keep the CMOS settings, a battery is used.

CMOS batteries are almost always a single CR2032 coin cell.
Note that these batteries are not rechargeable.
Incorrect system time, or incorrect BIOS settings are signs that the CMOS battery needs to be replaced, however these batteries will last years - often a decade or more, and as such are unlikely to ever need replacing before the motherboard itself is replaced altogether.

In order to delete the CMOS settings, the power to the PC can be unplugged and then the battery removed, which will cause the CMOS settings to be lost.
Alternatively, many modern motherboards also feature a “Clear CMOS" button, that allows the CMOS settings to be cleared without disconnecting the power and/or battery.

2.3 Chipset

A chipset is a specific set of components and protocols that manages interaction between system components, such as CPU, RAM, and PCIe peripherals, such as a GPU.
As time has gone on, modern CPUs have integrated more and more components into themselves that would previously have been components on the motherboard, which means that for most users, their motherboards chipset is somewhat irrelevant.

For the high-end user though, a better chipset can add useful features.
The main benefit usually offered by better chipsets is that of **Overclocking**
The lowest/lower chipsets often can not be overclocked at all, so overclockers will often not want these, and will want more advanced chipsets. Better chipsets also offer other useful

features, such as allowing *Multi-GPU* setups, which can be useful for professionals, as well as offering more USB or SATA or other interface lanes, as well as more RAID options and so on and so forth.

2.4 Form factor

Motherboards come in a range of different *form factors*.
The form factor of a motherboard will determine its physical size, as well as where all the screw-holes are for mounting it, and each form factor will usually have a fixed location for the rear-IO (such as USB sockets) on the motherboard.
The most popular current motherboard form-factors are ATX (9.6x12 inches) and micro-ATX (9.6x9.6 inches). Mini-ITX (6.7x6.7 inches) motherboards are sometimes also seen.
The larger motherboards will have more room on them for components and expansion cards, for example more slots for components such as RAM, or more PCI slots.
For ATX, MicroATX, and MiniITX the same screw-hole layout is used for securing them in place, with larger boards merely having more screw-holes, thus any case or enclosure that can fit a larger motherboard of these will also accept the smaller variations as well.
Computer cases will usually only fit up to a certain size of motherboard in them, and will not accept any larger standards.

2.5 Audio

Whilst dedicated audio cards can be purchased that can dramatically improve the audio quality of a PC, for most users the audio that modern motherboards come with built-in will be more than sufficient.
Basically all modern motherboards will come with audio built-in, even the cheapest motherboards available.

2.6 LAN

Basically all modern motherboards will come with LAN built-in, typically this will be Gigabit (Gb) LAN, almost always via just the one Ethernet slot, however this will be more than sufficient for almost all possible users and use-cases.
Some motherboards can provide more than one Ethernet port, or even 10Gb Ethernet capability, or more, however these are extremely uncommon.

2.7 I/O Shield

The motherboards *I/O Shield* is a metal plate that covers the area behind the motherboard where its ports are exposed on the back of the case. The IO shield is used to cover the gaps between these ports.
IO Shields are always the same size, and differ merely in what cutouts they have.
IO shields snap into a cutout on the back of the PCs case.
The IO shield will therefore have different cutouts depending on the output of the motherboard, and they are therefore specific to each motherboard, since motherboards all have different ports arranged in different ways, which makes their IO shields incompatible.
IO shields are useful to have to help protect the ports on the back of the PC, but are not ever required for the PC to work in any way.

2.8 POST

The *Power On Self Test* is a test that is performed immediately upon powering on a PC, and is usually the first thing that runs on it.
The function of the post is to check CPU registers, verify the size and locations of system memory, check bootable devices, verify critical components such as timers, initialize system buses and devices, and check, verify, then initialize the BIOS.
Post is therefore critical to a computer, however is usually extremely basic in its functionality.

The post is the first thing that loads on a motherboard, and is the first thing that a user will likely see as output of the computer, and as such many users use checking to see if a computer posts (i.e. runs the POST) to check it is powered on correctly, is able to run at a basic level, and display an output, and thus has the very bare minimum requirements of running.
Posting therefore provides useful checks to the system, whilst posting itself gives immediate feedback to the end user that the system is at least somewhat functional.

2.9 POST Code

POSTing is a very low level underlying process, and is usually very simple in both its functionality and execution, despite how important it is.
However, posting is the first thing that a computer runs, therefore if the computer does not post, then that computer will be inoperable, as it will not POST, thus can't load the BIOS, and thus won't load any OS, and thus can't run any programs.
Computers that fail to post must therefore be diagnosed until the problem is fixed.
Because a post requires numerous components working together to work, including the CPU, motherboard, video adapter, and other components, if the computer does not post then there

may be a very large variety of causes, which may make troubleshooting difficult.
Many computers therefore have *post codes*, which are alphanumerical values that are produced when the post fails, and can be used to inform users of the nature of the problem, facilitating an easier fix.

Some good motherboards will have a small screen on them explicitly for diagnosing post codes, and those that do not have this will often produce an audible beep at the user in a specific pattern to give out the post code.
Motherboards that have displays for post codes will almost always implement them as two adjacent seven-segment displays.

Computers that can display post codes will usually cycle through post codes relevant to what they are doing, and will thus display many post codes during posting in series, with the user able to use these to determine which parts of the system have initialized fine, and thus do not need to be checked in case of error, and can usually use the last postcode displayed to find the problem if there is one.

2.10 *Safe Boot*

Most PC users will likely not need to change their BIOS settings at all, and those who do so will likely only change options that do not affect system stability, such as the hard disk boot priority.
System instability usually occurs when attempting to overclock a system, and in extreme cases the severity of the instability may mean that the system is incapable even of posting, in which case the BIOS settings will need to be reverted in order to regain system stability that will allow BIOS or other settings to be changed.
This can provide a permanent fail-state for some systems, where the system is so unstable that the PC won't load sufficiently such that the settings to restore stability could be changed.

Safe Boot is found on many high end motherboards that are designed with overclocking in mind.
Safe Boot is usually a button found on the motherboard, that when pushed will load a default set of values for the BIOS and will effectively cause the system to ignore the saved BIOS settings, however will not delete or modify them like a clear-CMOS button would.
Safe Boot is usually used for overclockers, such that if their overclock causes instability in the system such that it is difficult to boot or post then safe boot can be used to get back into the BIOS to modify settings to revert back to a stable configuration without loosing the modified BIOS settings.
This is useful for overclockers, since they do not have to worry about losing their settings in case the system becomes unstable, and can usually just undo their last change in the BIOS to restore stability, instead of clearing the settings entirely and starting again from scratch.

Chapter 3

RAM

3.1 DDR

RAM has come in many different types over the years.
For the past decade or so though, RAM has used the DDR (*Double Data Rate*) principle.
Computer clock signals use a square wave signal, such that the signal output is at maximum half the time, and at minimum half the time, with a constant frequency.
Instead of activating on either the rising or falling edge of this clock signal, DDR systems instead activate on both the rising and falling edge.
This means that DDR systems activate twice per clock cycle - hence the "Double" in DDR - a system using DDR will activate twice per clock cycle, and thus get twice as many activations as a (very old) *SDR* (*Single Data Rate*) type.

As of writing this document modern RAM uses the *DDR4* standard.

3.2 Socket

RAM has come in many different types over the years.
In order to prevent systems from using hardware that they are not compatible with, each type of RAM uses a different socket, such that it can only fit into a motherboard which can actually accept and use it.
As such, RAM is never backwards or forwards compatible with other RAM standards.

3.3 Speed, JEDEC, and XMP

RAM operates a specific speed, which is the number of cycles per second that the RAM can perform. Since the RAM transfers a fixed amount of data with each cycle, this means the speed is equivalent to the data rate of that RAM.
This speed is usually given in MHz.
Note that unlike other computer components, RAM speed is usually given in MHz, rather than GHz.

Each memory module has a very small section of storage space on it called the SPD (*Serial Presence Detect*) that lists both the JEDEC and XMP profiles the memory module will run at, with each profile listing the speed, all the CAS latencies, and the voltage the memory modules can operate at.
The JEDEC profiles are considered the "safe" standards that the memory should be able to run at without any possible compatibility issues.
These are standard settings that the BIOS can use to automatically set the memory at given speeds, and thus effectively provide the standard operating parameters for the memory module.
When DDR4 launched the highest JEDEC speed possible was 2133MHz, however in late 2018 2666MHz and 2933MHz also became available, and faster speeds still may become available in the future.
If the RAM can operate above the default JEDEC speed in its SPD, then either XMP will need to be activated, or the RAM manually overclocked in order to use the extra potential speed.

Each memory module will usually have numerous JEDEC standards, and the BIOS can select the most appropriate based on various factors, usually the number of RAM slots used and the speeds of the installed memory, however will typically just default to the fastest JEDEC standard available unless XMP is enabled or there is some kind of compatibility error with other memory modules that requires it be run at a slower speed.

XMP is *eXtreme Memory Profile*, which are profiles built into the RAM module which allows the motherboards BIOS to automatically overclock the RAM to preconfigured standards set by the manufacturer, as well as selecting correct voltages, without the user having to manually overclock it.
Thus XMP can be a way of increasing the performance of the RAM without the effort or potential instability of manual overclocking.
XMP will usually work on both AMD and Intel systems.

There are usually two XMP profiles.
XMP profile 1 offers the speeds advertised for the memory module.
XMP profile 2 has more extreme settings for more performance.

XMP profiles are preconfigured standards that are tested by the manufacturers, and are thus usually very stable to run, however this is not actually a guarantee they will be, and other factors such as a CPU overclock might cause instability when using XMP profiles.

How much an increase in memory speed will affect the system can vary a lot depending on the task being performed, the other components, and the difference in speed.
AMD Ryzen CPUs tend to be much more susceptible to changes in memory speed than Intel ones, and benefit more from higher memory speed.

3.4 CAS Latencies

CAS is *Column Access Strobe*, and is sometimes just called “CL" (*CAS Latency*).
When the computer wants to access memory from RAM, the memory controller issues a request to the RAM for the data, and there is a delay whilst the RAM retrieves the data before it can respond to the request with the data. This delay between the memory controller issuing a request for data and the moment the RAM is able to respond with the data is called the CAS Latency.
Thus lower CL’s are faster to respond, which is advantageous to performance of the system.

CAS latencies are usually given either as a single figure (e.g. “CL16"), or as four numbers linked by dashes.
For example a modern DDR4 module might state “16-18-18-35".
If a RAM module only gives a single CL number, e.g. CL16, then it is referring to the first of these four numbers, and does not necessarily mean that the other three numbers are the same as the one listed, however it is possible they are.
For consumer SDRAM this number is a measurement of delay in number of clock cycles, and is so quick it could be measured in nanoseconds. For modern DDR4 RAM this is usually between 7 and 15 nanoseconds - depending on the CL timings and frequency, of course.
Because CAS latencies are measured in clock cycles, running the RAM at a higher clock speed will reduce the CAS latency.

RAM is made in a horizontal matrix package, containing data organised into rows and columns. RAM first activates the row that the data is located in, then the column, after which the data can be read, then the rows and columns are closed, after which other rows and columns can be opened, ready for use.

The four CL numbers represent:

- The first number is the *CAS Latency* itself, and is the time delay between the memory controller asking for data and the moment when the RAM first starts to respond.
- The second number is the tRCD: the *RAS to CAS Delay*, and is the time delay between the activation of the line and the column of where the data is stored in the matrix.
- The third number is the tRP: the *RAS Precharge*. This is the time between disabling access to one line of the memory and beginning to access another.
- The fourth number is the tRAS: The *Active to Precharge Delay*, and is the time delay the memory must wait until the next access to the memory can be initiated.
- Whilst it is usually not listed, RAM also has a fifth number, which is the CMD (*Command Rate*), and is the time it between when the memory chip is activated and when the first command may be sent to the memory. It is usually either T1 or T2, with the 1 and 2 representing number of clock cycles.
- Although manufacturers almost never list it, some RAM reporting programs will also commonly list the tRC, which is the *Row Cycle Time*, and is the minimum delay between successive commands to the same bank.
  The tRC is usually quite a lot higher than the other latencies above.

As such RAM timings can be abbreviated: CL-tRCD-tRP-tRAS-CMD.

Note there are other times that can be reported, such as the tRP, or the tRRD, however the above are the main ones.

Since the CL times are measures of clock speed, RAM with lower CAS latencies can be slower in use than one with higher CL latencies if the latter has a higher clock speed. Thus, CL latencies are not strictly always better, as clock frequency will play a role in end-performance too.
For equal clock speeds, lower CL numbers are always better than higher ones in terms of performance, however.

Overclocking RAM is a complex and hideously time-consuming process that can take days to do correctly, and as such most users would never manually overclock their RAM, instead simply setting it to an XMP profile and leaving it at that.
As such, buying low-CL RAM is beneficial for performance, however overclocking to get the maximum possible clock frequency to minimise RAM access times can be something of a waste of time for all but the most extreme overclockers.

3.5 Channels

Memory channels are fixed data-rate channels of communication between the system and the RAM. Having more memory channels can therefore increase the bandwidth available for communication to the RAM.
Thus the memory channels effectively act as a fixed multiplier to memory bandwidth if the extra memory channels are utilized.
Multiple memory channels will be typically be used to access multiple memory modules, and is not used to increase bandwidth to a single memory stick.
As such, to utilize multi-channel memory, multiple RAM sticks will need to be used, and for many motherboards RAM will need to be installed into the correct sockets to be able to utilize them. motherboard RAM sockets are often marked or colour coded to facilitate each discernment of memory channels
In order to use multiple memory channels the RAM should be equal in capacity, latency, and speed.
Ideally using identical RAM modules (same vendor/part number) would be best, as this should eliminate any compatibility problems.

Please see the section in the “CPU" section for a more in-depth explanation of memory channels, as the memory controller that controls RAM channels is part of the Northbridge, which for modern PCs is part of the CPU, and the RAM itself doesn’t affect the maximum number of memory channels that the systems memory controller can use.

RAM kits (i.e. multiple RAM modules sold together as a single package) are often labelled as “Dual channel", or “Quad channel" (or however many, depending how many modules are part of the kit).
This is because this RAM is intended to be used together in the same system, as all the modules in the kit will equal in speed, latency, and capacity, meaning they can be used together to their maximum.
Different RAM models, or RAM that is not from the same manufacturer (or other such parameter) can be used together in multiple-channel memory configurations without limiting each other, however again, they must be the same speed, capacity, and latency to not limit each other.
For consumers the listed memory channels of a pack of RAM is also an easy way of seeing how many memory modules are available in that RAM package.

3.6 Voltage

The voltage of the RAM is of course how much power the RAM module uses.
Less voltage of course means that the RAM will use less power, and thus create less heat,

which are beneficial.

Low voltage RAM chips are available that use less power than standard ones, and are always backwards compatible, such that the system can use them in standard slots and provide them with the standard voltage and they will still work fine, even if their low voltage ability is not used.
Despite this, mixing ordinary RAM and Low-Voltage RAM in a system is not recommended.
Low voltage RAM can often be worse at overclocking than standard RAM.
For most users this difference in voltage will be almost negligible.

Current DDR4 SDRAM typically operates at 1.2V, and older DDR3 SDRAM typically operated at 1.5V.

3.7 Ranks

RAM sticks are made up of DRAM chips (the circuits that actually hold the data) connected together on the same circuit.
Each *Rank* of the memory is a group of these DRAM chips that are all connected to the same circuit and can be accessed simultaneously.
It is possible for memory modules to have multiple ranks though, resulting in separate groupings of the DRAM modules.
Only one rank can be accessed at a time.

Memory ranks are also sometimes referred to as “Rows" or “Sides", however “Rank" is the more proper and less confusing term, as terms like “side" can lead to confusion about the physical arrangement of the DRAM chips on the die, which is unrelated to their rank.
1 rank is single sided DIMM, 2 ranks would be a double sided DIMM.
Quad and octal rank is also possible.

Increasing the number of ranks per DIMM can increase the memory density, which can allow DIMM modules with higher capacities.

Each memory controller can only handle up to a certain number of ranks, and has a limit to how fast it can handle them. Having too many ranks though can cause excessive load on the channel, which can limit the modules speed, and this can limit the overclocking potential of the memory if too many ranks are used.
Having too many ranks may also cause some memory capacity to not be usable to the computer, as it is unable to handle the number of ranks.

The average user is unlikely to notice a difference between memory of different ranks, however advanced overclockers may find that higher ranked memory can lead to lower overclocked RAM speeds, which can negatively impact performance.
This is more noticeable on systems that are more sensitive to RAM anyway, such as AMDs Ryzen CPUs.

3.8 ECC

When data is written to RAM, the RAM will write that data, and will read it out, back to the system when asked. However, sometimes an error occurs due to electrical or magnetic interference or due to high radiation dosage, in which a part of the data stored is different than the original data that was sent to the RAM to be stored.
ECC is *Error Correcting Code*, and refers to memory that has extra internal components that allow the memory to both detect memory errors and correct for them, such that even if the memory stores the data incorrectly, the original data can still be recovered, as the error can be corrected for.
ECC memory can not necessarily fix all errors, and can usually only fix errors of a single bit, however is still capable of reporting more severe errors.

ECC memory is useful for scenario's where data corruption, even of a single bit, can not be tolerated, such as scientific or financial workloads.
ECC is also used for safety critical workloads, such as on spacecraft, or aeroplanes, where memory corruption could cause system crashes or corruption of data that could prove catastrophic.

ECC memory is usually more expensive than non-ECC memory, as well as slower, and for most consumers, ECC memory is not needed at all.
Many CPUs do not support the use of ECC memory at all.

3.9 Buffered/Registered memory

A *buffer* is a an area of temporary storage that stores data whilst it is being moved between locations.
Memory with buffers is often referred to as *registered* memory, and unbuffered/unregistered for those without.
Unbuffered memory are often called UDIMMs, and registered memory are RDIMMs.
Most registered memory features ECC, however does not strictly have to.

Motherboards will usually have to support registered memory to be able to use it at all. Registered/unbuffered memory modules can not be mixed in a system.

Registered memory is usually slower than unbuffered memory, as well as more expensive.

The buffers have advantages, such as placing less electrical load on the memory controller, which allows for more memory modules than it otherwise would be able to use, they allow a setting time , allowing data to arrive to/from the storage cells at different times, and then be sent off all at once as a single output enable signal.

Consumer RAM is typically unbuffered, and registered memory is not useful for the average users.

Chapter 4

Storage

4.1 HDD/Mechanical Hard Drive

Hard Disk Drives (*HDDs*) are rotating magnetic disks with an arm containing a read/write head that moves over their surface.
This read/write head can modify the surface of the disk directly under it, and the differences in the surface of the disk this creates is a magnetic pattern that can be used to store data.

Hard disks are reliable are storing data for extended periods of time, however their rotating disks and moving read/write arms can be loud, and the fact that the read/write arm can only access the data on the disk that is directly underneath the head of the arm means that only small areas of the disk can be accessed at any given time, with the disk moving the arm or waiting for the disk to spin around to access other areas.
This means that HDDs are often slow to respond, as they must reposition their read/write arm before they can perform the requested operation.
Their moving parts also make the susceptible to shock.
Hard disks have been around for a long time, and are considered safe for long term data storage.

4.2 Platter

Most modern hard disks do not have a single rotating disk inside them, but several.
Each of these rotating disks, called a *platter*, will be mounted to the same central spindle, but will have its own read/write head and arm.
Note that each platter will typically have information stored on both sides of itself, with a read/write arm for both, effectively having two read/write arms per platter.
Having multiple platters can spread out the data to be written/read, and thus can increase

performance, however additional platters are typically used for modern hard disks to increase the total capacity of the drive, not for any performance reasons.

4.3 SSDs

Solid State Drives (*SSDs*) are flash based storage devices, and have large collections of memory cells that each store charge, with the patterns of charge representing the data that is to be stored.

The fact that current consumer SSDs store data in electrical charges means that this charge can potentially leak over time, resulting in data loss as the areas of charge degrade to the point of being unreadable, however SSDS are still relatively new, and this effect has not been well studied over long periods of time.
However, this means that SSDs are likely not suitable for very long term archival storage.
SSDs are also more expensive than HDDs for a given capacity, something that us unlikely to change over the next decade.
Despite this, SSDs use less power than regular hard disks, as they have no moving parts, which also makes them silent to use.
Since SSDs can near-instantly access any area of storage at a time, SSDs have virtually no latency, responding very quickly - orders of magnitude faster than a hard disk.
That they can access more of their storage area at once also means that they have a much higher transfer speed than HDDs, which can only access the area of the disk that is under the read/write head at any given time.
This high transfer rate means that most consumer SSDs can often max out the limitations of the SATA 3 interface that is the standard method of connecting storage drives to computers, and some modern SSDs use other interfaces, such as NVMe, to get around this limitation.

4.4 Caches

Modern hard drives have a *cache* (also known as a *buffer*), which is a very small amount of dedicated but temporary storage space the disk can use for itself, and is, or is very similar to, RAM.
Hard drives use their cache to store data that has recently been read, as well as adjacent data, since that is also likely to be read soon after, in addition to data that the disk needs to write.
By storing data the drive expects the operating system to request, or frequently requested data in cache, the latency for the hard drive sending that data to the operating system is decreased, since the requested data only needs to be read from the very fast cache, instead of the slower disk drive itself.

Having data sent to the cache to be written has the advantage that the operating system can send data to the cache, without having to wait for the data to be written first, potentially allowing it to continue where it otherwise might have to wait for the data to be written, which can prevent lagging.

Caches are usually very small in comparison to the drive as a whole, and the cache does not make the drive faster in any way for data that has not been cached.

The main disadvantage of caching data is that the cached data is left in volatile memory. This means that if the drive loses power before it can *flush* the cache, (i.e. transfer the data from the cache to the drives permanent storage), then the cached data will be lost.
That power loss can cause data loss or corruption of any cached data is considered the main disadvantage of write caching. Despite this, write caching can be useful for improving the performance of a drive, and it is often recommended that SSDs have write caching enabled for better performance.
Write caching is often considered less of a risk with SSDs, as SSDs have much higher transfer rates and lower latencies than mechanical hard disks, allowing incoming data to be written more quickly, whereas the head seek and spin up times of the platters of a mechanical hard drive can result in data being left in the cache for long periods of time, which then becomes vulnerable to loss of the cached data in the event of a power loss to the drive whilst the data remains in the cache.

Write caching therefore provides a trade-off between improved performance and data reliability.

Note that even if a disk has a cache, some applications may wish to *write-through*, which would allow them to bypass the cache and write directly to the disk instead. This may be desired by, for example, anti-virus programs.

4.5 *Hybrid drives* and SSHDs

A hybrid drive is a type of drive that uses both HDD and SSD technologies.
A hybrid drive setup can be composed of an HDD and an SSD as two separate drives, working together via software on the PC.
Alternatively, the hybrid drive may combine the HDD and SSD components into a single storage drive, in which case it is referred to as an *SSHD* (Solid state Hybrid Drive).
The SSD effectively functions as a second form of cache, beyond what the HDD normally has.
The SSD part of the SSHD is used to cache frequently used files or programs, such as the operating system, which can afford SSD-like performance to anything that the SSD portion of the SSHD deals with, with very fast transfer speeds and low latency as a result.

If the SSHDs SSD component is used to cache the OS and/or many commonly used programs, then this can make the computer feel dramatically more responsive due to the reduced latency and increased transfer speeds resulting from the SSD component.
As the SSD functions as a type of cache, the files will be present on both the HDD and the SSD parts of the drive if they are loaded by the SSD part of the SSHD.

SSHDs will be more expensive than an HDD of equivalent size, as they are more complex and have the SSD built-in as well.
The SSHD will not operate faster than the standalone HDD component if the data being accessed is not cached.
SSHDs have not proven particularly popular with consumers, since they are more expensive than traditional HDDs for any given capacity.
SSHDs are a good choice if the user is lacking in drive bays for their PC, however an SSD that is used to store frequently used programs and files and a separate data disk is likely be the better option in many scenarios.

4.6 MOLEX

MOLEX is the term used for a specific pin/socket interface standard, and is named after the Molex Connector Company that pioneered it.
Molex is typically used for older disk drives which predate the SATA power used for modern drives. Molex is also often used for PC expansion cards, such as USB expansion cards.
Molex connectors have 4 pins, and a chamfered top connector edge that prevents it from being used in the wrong orientation.

The Molex pins are +12V, Ground, Ground, and +5V, and are usually colour coded Yellow, Black, Black, and Red, respectively, however for very old cables this may not be correct.

Molex connectors are held in place by friction, and thus can often be hard to attach or detach, and the pins can skew, or the pins on the female connector spread, which can cause problems with the interface, and are some of the reasons why Molex was mostly superseded by SATA, however most modern power supplies still have some Molex connectors available as outputs.

4.7 SATA

Serial AT Attachment is a connection standard that has been used for connecting storage drives to computers for over a decade.
SATA is generally the only connection method that modern internal HDD, SSD, and disk

drives (CD, DVD, and Blu-Ray) use to connect to a PC.

SATA has several versions:

- SATA 1 was the original version, and has a maximum rate of 1.5Gb/s, or 1.2Gb/s (150MB/s) taking transmission overhead into account.
- SATA 2 has a maximum transfer speed of 3.0Gb/s, or 2.4Gb/s (300MB/s) when taking transmission overhead into account.
- SATA 3 has a maximum transfer speed of 6.0Gb/s, or 4.8Gb/s (600MB/s) when taking transmission overhead into account.

All SATA versions use the same connector, and are thus fully backwards compatible, allowing modern cables to be used with older drives without issue.
However since the cables have identical connectors, this means that they are not visually distinguishable, and users must rely on labelling or testing to know the speed of the cable.

Modern hard disks are unlikely to max out even SATA 2, however even old SSDs can easily max out SATA 3, hence the recent introduction of other standards, such as NVMe.

4.8 SATA Power

The SATA specifications also specify the power connector that modern SATA drives use.
The SATA-data connector is a 7-pin wide connector, whereas the SATA-power connection is a visually similar 15-pin wide connector. This difference in number of pins and connector size prevents confusion between the two.

4.9 NVMe

Modern hard disks are unlikely to max out even SATA 2, however even old SSDs can easily max out SATA 3.
This meant that newer connection methods were required to make use of the maximum potential of modern SSD, and hence NVMe was introduced.
NVMe (*Non-Volatile Memory Express*) is an interface specification used for accessing non-volatile memory types via a PCIe bus.
The main draw of NVMe over SATA is that it can reduce latency, transfer overhead, and improve transfer speeds.

Each lane of the PCIe bus can support up to 1GB/s, which allows current M.2 NVMe drives a theoretical maximum speed of 4GB/s, since current M.2 NVMe drives can use up to 4

PCIe lanes.
Currently, NVMe drives are relatively rare, and only newer motherboards will accept them, however their increased performance makes them popular for high-end users, despite their high prices.

4.10 2.5", 3.5", and form factors

Typical consumer drives come in the standard 3.5" form factor, and all computer cases will include slots for 3.5" hard drives.
3.5" drives are 4" x 5.75" x 1" (101.6mm x 146mm x 25.4mm) [WxDxH]

The 2.5" drive standard is popular with modern hard drives used for laptops, and almost all modern SSDs use the 2.5" form factor.
2.5" drives are 2.75" x 3.945" x 0.197-0.75" (69.85mm x 100mm x 5-19mm) [WxDxH]
Note that whilst most 2.5" drives are 9.5mm high, and this is the unofficial standard, not all of them are.

The 2.5" and 3.5" that these drive sizes are referred to as do not refer to the overall size of the drive, but instead the standard size of the rotating disk platters that are used in them. The naming scheme remains for modern drives though - even for SSDs, which have no platters, or even moving components at all.

4.11 RPM, 5400, and 7200

Hard Disk Drives (*HDDs*) are rotating magnetic disks with an arm containing a read/write head that moves over their surface.
The disks rotate at a maximum speed, and hard drives will list this maximum speed as their RPM value.
Hard drives that rotate slower will have less disk moving under their read/write arm at any given time, and as such can have slower transfer speeds, although because other factors will affect the transfer speed of a hard drive, such as the time required to move the read/write arm and so on, drives with lower RPM values are not linearly slower than faster ones, even given the same areal density of the disk.

Modern consumer hard disks will usually rotate at either 5,400RPM or 7,200RPM.
Some hard drives will use different RPMs however, and before the advent of SSDs, faster drives were acquired by using faster speeds, with 10,000RPM or even 15,000RPM drives available, and were for a long time the fastest high-capacity storage available.

These faster RPM drives are less common these days, as SSDs provide even faster transfer speeds still.
Many people often refer to a disk by its rotation speed though, and as such “5200", “7200", and “10k" are often used for this.

4.12 mSATA

mSATA is a drive form factor designed for small and low-powered devices.
mSATA has all the same bandwidth limitations as the standard SATA interface. (i.e. 6.0Gb/s = 600MB/s after transport overhead.)
mSATA uses a different connector than standard SATA, and as such is not compatible with it, requiring a different connector, however they are usually just plugged straight into the motherboard.

mSATA drives are 30mm x 50.95 mm in size [WxL].
Note that the mSATA connector is identical in size to the *PCIe Mini Card* format, however they are not electrically compatible, since the mSATA requires a SATA host controller connection, not the PCIe host controller connection that PCIe Mini cards use.

mSATA is still a current standard, however is less popular than M.2 drives for modern computers, as they occupy similar drive sizes and use-cases, but M.2 drives are able to have higher maximum transfer rates.

4.13 M.2

M.2 is a drive form factor size.

M.2 drives are 12, 16, or 22mm wide, and either 30, 42, 60, 80, or 110mm long.
M.2 drive sizes are often referred to by their width and length, so for example a 1642 is 16mm wide x 42mm long, and a 2260 is 22mm wide x 60mm long.
All motherboard M.2 slots are 22mm wide, however the lengths can vary, and many motherboards will have screw holes that allow M.2 drives of several different lengths to be attached.
M.2 drives will connect to the PC via PCIe 3.0 (up to four lanes), SATA 3 (one port), and/or USB 3.0 (one port).

M.2 drives have a single 75-pin connector, but are differentiated by function by which of these 75 pins are notched (missing).

This prevents incorrect devices from being put into incompatible sockets.
There are currently four connector types in use for M.2 drives.

- Type A has pins 8-15 notched, and can interface via PCIex2, USB 2.0m I^2C and DisplayPort x4.
  Type A is mostly used for Wi-Fi, Bluetooth, or cellular cards.
- Type B has pins 12-19 notched, and can interface via PCIe x2, SATA, USB 2.0, USB 3.0, audio, I^2C, and UIM (Sim Card), as well as PCM, IUM, HSIC and SSIC.
  Type B is mostly used for SATA and PCIe x2 SSDs.
- Type E has pins 24-31 notched, and interface via PCIex2, USB 2.0, I^2C, SD card, as well as UART and PCM.
  Type E is mostly used for Wi-Fi, Bluetooth, or cellular cards.
- Type M has pins 59-66 notched, and interfaces via PCIe x4 and SATA.
  Type M is usually used for fast PCIe x4 SSDs.

B+M is a common interface as well, having both B and M notches present, for greater compatibility at the cost of performance.

The fastest M.2 drives will therefore connect via the Type M connector, however many M.2 SSDs will connect via B+M methods, however, since only M-drives provide all 4 PCIe lanes, B+M drives will be slower, as they only have half the maximum PCIe lanes.

4.14 RAID

Redundant Array of Independent Disks is a method for combining multiple physical storage drives into a single logical drive that is visible as a single drive to the computer.
RAID is used to increase the performance, reliability, availability, or capacity of the data storage system, or some combination thereof.
RAID works by distributing the data that is to be stored on multiple disks, instead of just one like the conventional data storage method.
How this data is distributed depends on the RAID method used, often referred to as RAID level.

Common RAID configurations are as follows:

- **RAID 0** a.k.a. **RAID Stripe**
  RAID 0 works by writing consecutive bits of the files to be stored on sequential drives, thus spreading out the files to be stored on multiple drives in the RAID 0 array.
  As such, when the data is to be read back again, the drive can read from multiple drives simultaneously, which increases performance above that of a single drive linearly with the number of drives in the RAID 0 array.

A RAID 0 array will have a total capacity the same as the sum of all the disks in the array.
RAID 0 stores data on multiple disks, which means that a failure of any one disk will cause data to be lost, as all files, and thus the entire RAID 0 volume, will be lost.

RAID 0 therefore offers increased performance, but at the cost of total data loss of all data on all drives in the case of a drive failure.

- **RAID 1** a.k.a. **RAID Mirror**
  RAID 1 works by writing data identically to two (or more) drives, thus “mirroring" the contents of each drive with each other.
  This means that the RAID 1 array must write data at the speed of the slowest disk in the RAID 1 array, however it can potentially read data at the speed of the fastest drive in the RAID 1 array, and *theoretically* can read data from multiple disks simultaneously, however this is not often done in practice.
  However if any disk in the RAID 1 array fails then the RAID 1 array will still have all files complete and undamaged, since an identical copy of the failed disk will exist.
  RAID 1 arrays will operate normally and with no data loss whatsoever, regardless of how many drives in the array fail, provided a single drive still remains functional.

  RAID 1 therefore offers no performance boost, but complete data reliability in the case of total drive failures, provided at least one drive remains.

- **RAID 5** a.k.a. **RAID Stripe+Parity**
  RAID 5 works by performing block-level striping, but with parity information spread out amongst the drives.
  This means that if a single drive in a RAID 5 array fails, then the data can be reconstructed using the parity information stored on the other drives.
  However a RAID 5 array requires 3 drives at minimum to work, and if another drive fails whilst the array is rebuilding itself after a first drive fails, then all the data on the entire array will be lost, akin to a drive failure on a RAID 0 array.
  Rebuilding a RAID 5 array in the case of a drive failure can be time consuming, as all data from all other disks must be read in order for the last drive to be rebuilt.
  This is unlikely for home users with small arrays of drives, however for large arrays of drives or for critical data this makes RAID 5 highly undesirable, as a drive failure whilst the drive is rebuilding itself from a separate error will cause total data loss.

- **RAID 10** a.k.a **RAID Stripe+Mirror**
  RAID 10 (1+0) is the combination of RAID levels 1 and 0, and is therefore a nested array (multiple RAID levels at the same time on the same system).

  RAID 10 works by mirroring drives, and then striping those mirrors, effectively apply-

ing RAID 0 to RAID 1 (hence 1+0).
This has the advantage of having performance above that of a single drive, due to the striping, as well as data reliability, due to the mirroring.
As long as a single mirror of drives has all of its drives intact, no data loss will occur, regardless of the number of drives lost.
RAID 10 requires a minimum of 4 drives to work.

RAID 10 is popular for disk performance, while still offering high reliability, however does require more disks to use than some other RAID options.

RAID levels can be used in conjunction with each other, for example RAID 5+0 (a.k.a RAID 50) can be used, however RAID 10 is by far the most common and well known implementation of this.

RAID can be achieved either by a hardware card installed into the computer, which handles all the data transfers to/from the storage drives, and distributes data onto them as appropriate.
Alternatively software can be used, with the drives plugged into the computer normally, and the software on the PC distributing data to each drive correctly.
Each has their advantages, however software RAID is generally preferred, as with hardware RAID if the hardware RAID controller fails then the data can still be lost, as the card itself is often required to read the data, depending on RAID levels used.
There are numerous pros and cons to each level, and it is up to the end user to determine which is better for their setup.

Chapter 5

GPU

A graphics card is a dedicated piece of hardware that deals with the output of the computer that is visually displayed to the user.
It controls the operating systems desktop, programs that run, as well as 3D manipulation of objects, as used by for example games.
Graphics cards range from small, low-powered devices to huge, power-hungry ones as used for high-end gaming or demanding professional workloads.

5.1 dGPU

A graphics card is a dedicated piece of hardware.
This typically takes the form of a dedicated, *discrete GPU* card that plugs into the motherboard and has its own video outputs - which the systems monitors should be plugged into.
dGPUs are the main form that GPUs occur in PC, and for gamers or professionals who want GPU performance there is no other option.

High-end gaming GPUs can draw lots of power, and often require their own power cables to work at all, however many low-end ones will work with just the power they are able to draw from their interface slot - usually PCIe, which allows up to 75W max from the PCIe slot.
This high power draw also creates lots of heat, and powerful GPUs will often have very large heatsinks as well as fans to dissipate it.

5.2 iGPU

A graphics card is a dedicated piece of hardware.
This does not mean it is a standalone piece of hardware though.

Many modern CPUs also come with a dedicated GPU built into them, referred to as an *Integrated GPU, or iGPU.* (because it is integrated into the CPU, instead of a standalone component)

iGPUs are mostly used because a CPU with an iGPU is often less expensive than the combined price of a conventional CPU (with no iGPU) and a discrete GPU.
iGPUs are fine for basic computing tasks, such as spreadsheet, text editing, and browsing webpages and the like, and modern iGPUs should easily handle video playback.
iGPUs are not recommended for gaming though, with really only the most powerful iGPUs being able to handle modern games, and even then, gaming at high settings or at resolutions above 1080p are really beyond iGPUs, and gamers should look at actual dedicated gaming GPUs for gaming, rather than iGPUs.
Instead, iGPUs tend to be used to provide a more simple system than would typically be the case with a discrete GPU, as well as potentially one that uses less power, and thus generate less heat and noise.
Despite this, older or less demanding games (such as most eSports titles) will often run fine on the most powerful iGPUs (although certainly not on the least powerful ones) at 1080p and medium or maybe even high graphical settings (depending on the game being tested, as well as other limitations such as RAM, and CPU, and so on...)

Integrated GPUs feature on both CPUs made by AMD and Intel.
AMD iGPUs tend to be drastically more capable than those from Intel, and this has really never not been true.
Not all CPUs have an iGPU, and they are usually only found on lower-end CPUs, as anybody who wants high performance would likely prefer the better performance that a CPU+dGPU combo can offer.
iGPUs remain popular with average PC users though, due to their simplicity and lower cost.

5.3 PCIe and x16

PCIe (*Peripheral Component Interface Express*) is a type of both interface bus and an associated hardware slot.
PCIe slots are the main, and usually only, connection type used to connect modern graphics cards to PC.
PCIe is different from PCI, and the two are incompatible.
PCIe offers a very high-bandwidth, scalable architecture that modern devices have not yet managed to max out, and it is therefore becoming more common for devices other than GPUs to be routed through PCIe.

PCIs is scalable, and uses *lanes* to transfer data.
Each lane can carry a maximum amount of data, and the number of lanes thus determines

how much bandwidth is available to the PCIe device.
Lanes are usually written with an “x" notation.
For graphics cards a PCIe x16 (16-lane PCIe) slot is always used, as this provides the most bandwidth.
Other lanes sizes are sometimes used though, for example many USB expansion cards, or Wi-Fi cards will often only use PCIe x1 (1-lane PCIe).
PCIe devices can use 1, 2, 4, 8, 16, or 32 lanes.

PCIe can act as a bus for other devices to route though, including other devices not using a PCIe slot on the motherboard, for example M.2 drives can connect through the PCIe bus, using either 2 or 4 lanes, depending on their connection type and specific design.
PCIe can also be used to route devices to other controllers - for example M.2 drives can connect to the PCs SATA controller, but physically be routed through the PCIe bus.

PCIe lanes are hardware - there must be hardware present for the lanes to transfer data through, as well as sufficient bus and controller capability for the receiving device to be able to handle the data.
Being hardware, PCIe lanes cannot be increased later by software methods.
Many PCIe controllers can only support so many PCIe lanes at a given time, and if too many PCIe devices are plugged in, some of the devices will have to share lanes. As such, it is possible that a card can be plugged into a high-lane slot, but is not able to use all the lanes available.
The most common example of this is sharing a single PCIe x16 lane into two PCIe x8 lanes.
It is also possible for a larger slot to not operate at the equivalent number of lanes that would be expected given its size - for example a PCIe x16 slot might only ever operate at 8 lanes maximum.
As such, gamers should pay particular attention to which slot they plug their GPU into, as getting insufficient bandwidth can be extremely detrimental for GPU performance.
Despite this however, PCIe x16 is not a limitation for any modern GPU, and many GPUs can run unrestricted in an x8 slot.
Most devices that can run using multiple PCIe lanes can operate normally in lower numbers of lanes, they will simply not have as much bandwidth, and as such can underperform as a result, however will not necessarily experience instability or other problems.

PCIe has come in numerous versions, with the most current version as of the time of writing this document being PCIe 3.0, which has each PCIe lane operate at about 985MB/s (in both directions simultaneously, after taking transmission overhead into account).
PCIe 4.0 has been standardised, having double the bandwidth of PCIe 3.0, but has not yet become massively available to consumers.
PCIe specs start at 250MB/s for PCIe 1.0 (or 200MB/s when taking 8/10b transmission overhead into account), and doubles for each subsequent PCIe version.
PCIe 3.0 changed the transmission protocol from 8/10b to 128/130b, dramatically reducing transmission overhead, and thus increasing available lane bandwidth.

PCIe is both backwards and forwards compatible between versions, and as such and PCIe device can operate in any PCIe slot with any PCIe controller or lane configuration, provided the PCIe devices physically fits into the slot that would accept it, however older PCIe versions will limit the bandwidth available to the device, which can heavily degrade performance if the device depends on the performance of newer PCIe versions.

5.4 PCIe power, and power consumption

Many low-powered graphics cards can be powered simply by the power made available via the PCIe slot itself - a max of 75W.
For other, more-power demanding cards however, dedicated power connectors for the GPU must be used.
These PCIe power connectors come in two forms: a 6-pin, and an 8-pin connector.
The 8-pin connector are rated to provide up to 150W, and the 6-pin up to 75W of additional power.
(However in practice can often deliver more than this just fine, but don't rely on this!)
The most powerful GPUs can require 3 connectors for maximum power delivery, however most GPUs only require 1 or 2 connectors, with 2 being popular for modern cards.
Whilst most GPUs are not likely to max out the specs for the PCIe power connectors maximum draw, when overclocked it is possible.
Many power supplies, especially older or lower-end ones, might not have many PCIe power connectors, however Molex to PCIe power connectors are easily able to supply the power require if needed because the PCIe power connector only delivers 12V, which Molex also supplies directly at up to 132W per Molex cable - this is why Molex to PCIe power connectors often requires 2 Molex cables, because the 132W rated max of Molex is less than the 150W max rating of PCIe 8-pin power connectors, however in practice one molex cable will often be fine as many GPUs will not max out their power draw.

As for the actual amount of power that modern GPUs use, the answer is *"it's variable"*, of course.
The lowest power consumption will come from iGPUs, with dGPUs typically using more, however since dGPUs cost more and are usually more capable, they will often be used for more demanding tasks, such as games, so will probably have a much higher average power draw anyway.
Note that TDP is not power draw, and the two should not be confused, and power draw of a maxed out GPU can often be higher than the TDP rating of that GPU.
The average GPU itself though is unlikely to draw more than 150-200W under max load, with only much more demanding ones drawing more.
During the average gaming scenario, I would expect the average gaming PC to draw between 200 and 250W, however up to 350W might be seen, depending on the components, and the

highest-end systems can draw more.
When selecting a PSU for a PC, remember that most PSUs tend to be the most efficient when delivering around half their rated output.
Overclocking a GPU (or...any other component) can *dramatically* increase the power consumption in a non-linear matter (i.e. performance per power consumption will very likely decrease with overclocking), and overclockers will likely want the highest Wattage PSU they can find, however depending on the components and the degree of overclocking done lower output ones might also be fine, it all depends on the users system and desired results.

5.5 VRAM; GDDR5, GDDR5X, GDDR6, HBM, HBM2

GPUs have VRAM (*Video RAM*) built into them.
This VRAM allows the GPU to store information on board itself, and not have to use system memory (RAM) to store information.
VRAM is used for normal operation of a GPU, however some computers GPUs, especially on laptops, low-end, or older devices may use the systems RAM instead of having their own dedicated memory.

The main use of lots of VRAM for modern GPUs is to store the textures used for games.
How much VRAM will vary depending on the game, as each game will use its own textures.
In addition many GPU reporting programs will list the amount of VRAM the game has requested to use, and the GPU has reserved for that game, instead of how much of that it is actually using.

A GPU with more VRAM will generally not perform better than one with less unless the textures for the game in question are sufficient to fill the VRAM of the GPU with less VRAM, in which case the juggling of textures can cause an extreme performance hit whilst the GPU has insufficient VRAM.
More VRAM will make the GPU more expensive, and cause it draw slightly more power, and thus create more heat, however more VRAM is generally always better to have, if possible, although cards with more VRAM are not always the best option financially, as the extra VRAM may not be useful for the useful performance range of the card.

GPUs have used a wide variety of different types of VRAM over the years (or, well ... decades).
Modern GPUs will typically use GDDR5, GDDR5X, GDDR6, HBM, or HBM2.
HBM *(High Bandwidth Memory)* tends to have higher bandwidth than GDDR, but costs more.
In general most consumers do not need to know about the differences between VRAM types, and the architecture or the GPU will generally have a bigger impact to the performance of the card than its VRAM type.

Note that this is not always the case, and certain video card manufacturers have released video cards that have different memory configurations that can massively degrade performance of the card, simply by using different VRAM.
The Nvidia GT1030 GGR4 is an example of a modern (2018) GPU that had severely degraded performance simply from using a worse type of VRAM, with the change from GDDR5 to GDDR4 resulting is an average performance loss of around 50% in most benchmarks.

Note that unlike system ram, VRAM is built-in as part of the GPU, and is not modifiable at all.

See the *RAM* chapter for more information about GPU VRAM, since much of the technical details are the same.

5.6 Frequency

Frequency is defined to be the number of times a certain thing occurs.
The unit of measurement for frequency is Hertz (Hz) - the number of times per second something occurs.

Modern GPUs have their speeds defined in some multiplier of Hz, these days typically Gigahertz (GHz), with 1GHz representing 1 billion Hz.
The Hz of a GPU indicates the speed at which its cores can execute instructions, with a higher frequency representing more instructions being able to be executed in any given time.
As such *if [and only if] all else were equal*, a GPU with a higher frequency would be able to perform more operations than one with a lower frequency, and thus operate faster.

GPUs can also be *Overclocked*, which increases this clock frequency, thus allowing the CPU to execute instructions faster.
Overclocking can be a good way of getting more performance out of a GPU, however a GPU can only be overclocked so far, either because of limitations in what the CPU can handle without instability, a limit to how much power the GPU is able to draw, or just a limit to how quickly the additional heat an overclocked GPU generates can be dissipated before the CPU overheats.

Unlike CPUs, many modern GPUs come with software that allows them to be easily overclocked from within the OS, thus making GPU overclocking relatively easy compared to CPU overclocking.
GPU overclocking is often much simpler than CPU overclocking.

5.7 Cores

GPUs have cores in the same way that CPUs do.
The difference though is that GPUS are designed to be RISC, rather than the full x86 of CPU Cores.
This means that GPUs are designed more for performing large quantities of simple work in parallel, rather than CPUs, which focus more of smaller amounts of more complex work.

Whilst they are of course useful, GPUs cores are less important of a metric than CPU cores, as things like the GPU architecture, clock speed, and memory will also determine GPU performance.
The main reason that GPU cores cannot be compared to CPU cores though is that GPUs change architecture frequently, and even within the same brand of GPU, GPU core counts can vary quite a lot, let alone the extreme differences in architecture that exist between AMD and Nvidia, which means comparing core counts of GPUs is often not particularly useful, except when within the same product family.

5.8 APIs, DirectX, and Vulkan

APIs are *Application Programming Interfaces*, and are large collections of tools, protocols, and definitions that can be used to build software.
APIs can abstract the underlying hardware information, which makes it much easier for software developers to utilize.

The most well known API for game development is DirectX, however other APIs exist, including OpenGL, and its intended replacement; Vulkan.
Vulkan offers better performance than other APIs, with lower overhead, better CPU and GPU balancing, as well as better distribution of the workload over multiple CPU cores, however is a much lower-level API than many competitors, and thus can be more complex and harder to work with.

DirectX has come in numerous version, with newer versions offering more advanced feature.
As of time of writing, the most modern version of DirecX is version 12.
DirectX is not backwards compatible beyond DirectX 10, due to changes in how DirectX works. Despite this, DirectX 9 was very popular for games due to its long lifespan, ease of use, and good performance and featureset (for the time), and many games still run DirectX 9 and have not been updated for later versions.
Switching to a more modern version of DirectX requires modifying the games underlying engine, so it can introduce problems, as well as potentially take a lot of work regardless, which means that most games are not rewritten for newer DirectX versions, regardless of

the performance or feature benefits of doing so.
Since most games tend to have a relatively short development window beyond their release data, this is most commonly seen in MMOs, which tend to have longer lifespans than most genres of game, and many popular MMOs run old DirectX versions, with DirectX 9 still being popular for many current MMOs.

GPUs themselves can usually only support up to a fixed DirectX version, and will be unable to use newer versions that that.

5.9 CrossFire, SLI, and Multi-GPU

Multi-GPU is the term to describe having multiple GPUs in a system with the intention that programs (typically professional or gaming) will use several GPUs at the same time to deliver a better output.
Multi-GPU is usually called *CrossFire* when referring to AMD GPUs, and *SLI* (*Scalable Link Interface*) when referring to Nvidia GPUs.
Multi-GPU setups are usually not possible using a combination of AMD and Nvidia GPUs - all the cards must usually be from the same brand, and often from the same product series.

Multi-GPU is typically used to provide more GPU-performance per price, and because at the low-medium end of GPU hardware a small increase in price can lead to a good step up in GPU performance, multi-GPU setups are usually only used with high-end GPUs, although in practice many people do run a wide variety of cards in multi-GPU setups. However, many gamers will use multi-GPU setups to add performance to an existing system, perhaps due to limit budget when buying the system, or other limitations that are no longer in place.

Whilst Multi-GPU almost exclusively refers to multiple dGPUs, it can also work using iGPUs - with both an iGPU and a dGPU working together.
In addition, multi-GPU setups do not always have to use the same GPU, and multiple different GPUs in a multi-GPU setup is possible, although is often not advised, as the lower-performing GPU may bottleneck the higher-performing one, leading to little performance gain for the price of GPU, as well as leading to compatibility issues.

Multi-GPU usually uses one of two techniques; AFR, or SFR:

- **AFR**
  Alternate Frame Rendering is the term used to describe when a multi-GPU setup will run a games frames in sequence, and will run alternate frames on different GPUs.
  So for example in a system with two GPUs, the first frame will be computed on the first GPU, the second frame on the second GPU, then the third frame on the first

GPU, the fourth frame on the second GPU, etc.
AFR can increase frame rates quite a lot, however can result in *Micro-stuttering* - which is when the output time of the two GPUs is not quite the same, resulting in there being different times between frames delivered to the user.

- **SFR**
  Split Frame Rendering is the term used to describe when a multi-GPU setup will run a games frames on both GPUs at once. Each frame is run on both GPUs, with the GPUs splitting the workload between them.
  How the frame is split will depend on the scene and technique used, however geometry will need to be fully calculated on both the GPUs regardless.
  SFR tends to not work as well as AFR, however does not exhibit the same micro-stuttering problems.
- **AA**
  Some multi-GPU setups can be used to perform *Anti-Aliasing* on the scene computed. This will not increase fps of the output, but will increase effective resolution of the image, resulting in a higher quality output.
  This can be potentially more pleasing to the user than simply more fps.

The advantage of a multi-GPU setup is more potential GPU performance, which can increase performance in professional or gaming workloads, and whilst for professional users multi-GPU setups are often used, multi-GPU is generally never recommended for gamers to actually use, and there are several reasons for this;

The **advantages** of a multi-GPU setup would be as follows:

- **Startup cost**
  A multi-GPU setup can be used as a way of upgrading an existing setup to increase performance, thus lowering the barrier to entry of a given performance level.
- **Performance**
  Multi-GPU setups can deliver higher framerates, higher output quality, or at higher resolution, depending on technique used.

The **disadvantages** of a multi-GPU setup would be as follows:

- **Micro-stuttering**
  When using AFR, micro-stuttering is a somewhat common problem, and can dramatically degrade the users experience.
- **Scaling** Multi-GPU setups will usually not scale perfectly, and the multi-GPUs fps will be lower than the combined fps that each individual card could get, sometimes even as low as a 10% increase to fps from an additional GPU.
  In some cases, a multi-GPU setup may actually provide less fps than one of the GPUs would on its own.
  For professional workloads, however, multi-GPU often scales well, sometimes even perfectly, depending on the application, and is more recommended.

- **Price to performance ratio**
  Multi-GPU setups do not usually scale linearly with number of GPUs [See above.], and this reduces the price to performance ratio of the multi-GPU system, such that it may be more cost effective to instead buy a single, more powerful GPU, rather than several less powerful ones.

5.10 Backplate

Many modern GPUs come with a *backplate*.
The backplate covers the rear face of the GPU PCB, and is usually screwed onto it in some fashion.
Backplates are often used for aesthetics - existing solely to make the back of the GPU look aesthetically pleasing, instead of the PCB, which many people would think doesn't look as good. It can also help provide a more complete visual look, matching the aesthetics of the cooler on the front of the GPU.
For some GPUs however, the backplate is used thermally. For these GPUs the backplate will be made of metal, and will have thermal pads connecting it to components on the back of the PCB, allowing it to function as a heatsink to aid in cooling.
Some backplates that are used for cooling can include heatpipes in them to transfer heat around the backplate more efficiently, allowing for better thermal dissipation.
GPU backplates are almost always made of metal, and if they are then they will include a cover between the metal backplate and the PCB, which will prevent the GPU PCB/components from shorting with the backplate, which will damage them.

Note that the back of the GPU will not produce much heat at all, and there will typically be no components on the back of the PCB that will require active cooling, so backplates will not make any real different to the actual thermals of the important GPU components (i.e. GPU core, memory modules, VRM).
All backplates however will block the back of the PCB, so it will prevent airflow to the back of the PCB, so if the backplate is not being used to sink heat from the GPU in some fashion then it will negatively impact thermals of the GPU, but these will likely only be very minor, often lower than a single degree change at absolute maximum, since the back of the GPU PCB doesn't really do much.

Chapter 6

Monitor

6.1 Resolution

One of the main specifications of a monitor will be its *resolution*.
The resolution will be specified as two numbers, the first being horizontal pixels, and the second being vertical pixels, with one pixel being the minimum size displayable, representing a single point of information (or single point of colour, if you prefer).
Monitor resolutions have often just been referred to by their vertical resolution though.
The typical size of modern monitors is 1920x1080, also referred to as 1080p, or “Full HD".
Monitors with a resolution of 2560x1440, a.k.a. 1440p, are also becoming more common, and are often referred to as QHD, as they are 4x (Quad) the size of 1280x720 panels, which themselves are “HD" - hence, Quad HD (QHD).

“4k" monitors - those with a resolution of 3840x2160, are also starting to gain some traction, however have not really become particularly popular yet.

Note: the “p" in 1080p, 1440p, etc, refers to a *progressive scan*, as opposed to say 1080i, which is sometimes used on cheap TVs.
A progressive scan is how you would expect displayed images to work - where each image displayed is received from the source, and displayed, as a whole.
This is opposed to 1080i - which refers to an *interlaced scan*, where the odd rows are displayed in one frame, and the even rows displayed in the next, such that only half of the actual image frames are used to produce the video.
This was common on very old CRT TVs, and some early or very cheap TVs, where the nature of TVs, which tend to show moving images, means that the problems this creates is less noticeable than it is with computer work, which often display lots of stationary elements, such as text, where interlacing can cause annoying issues.
Interlaced scans can use less bandwidth to transmit, hence why they were used for TV in the first place, however other than that progressive scans are flatly better than interlaced.

All modern computer monitors should use a progressive scan, and you should not use an interlaced monitor for such a task - usually you will only find interlacing on TVs though, dedicated computer monitors should never use it, and you should not buy or use an interlaced screen as a computer monitor unless you absolutely have to.

6.2 Inputs

Computer monitors often have several inputs, often of varying types.
Whilst each monitor can only display one image at a time, having several input options is useful for consumers.
The main input types are as follows.

- **VGA**
  VGA, or *Video Graphics Array* is the oldest input type still used in modern monitors, as well as being the simplest, and is often perceived to be the most reliable.
  VGA is an analogue display type, using a connector consisting of 3 offset rows of 15 pins.
  Because VGA is an analogue technology, it can be prone to issues such as crosstalk, or ghosting, in very long cables.
  Many modern monitors still include a VGA adapter as it is simple, reliable, and regarded to be universal.
  VGA cables and connectors always have two screws adjacent to them to hold them in place.
- **DVI**
  DVI is *Digital Visual Interface*, a digital standard that followed VGA, and became extremely popular.
  DVI is the oldest digital display technology still in use, and is often used on high end monitors because its age and simplicity make it robust and reliable, as well as relatively universal.
  DVI cables and connectors always have two screws adjacent to them to hold them in place.
- **HDMI**
  High Definition Multimedia Interface is a more modern digital standard than DVI, and allows for more bandwidth, thus allowing higher resolutions and/or higher monitor refresh rates.
  HDMI cables and interfaces can also transfer audio through them as well, unlike VGA or DVI, which can simplify connection systems.

  There are several HDMI connector types; the Type A (standard connector), the Type C (the “Mini" connector,), and the type D (“Micro" connector).
  The smaller connectors were sometimes used on devices such as smartphones, or games

consoles.

Unlike the VGA or DVI standards, HDMI cables do not have screws to hold them in place as standard, and does not have latches or other mechanisms to hold them in place.

HDMI is becoming less common with time as DisplayPort becomes more common, and may be due in part to the fact that HDM is not a free technology, and manufacturers of tech that uses it must pay to licence it.

- **DisplayPort**
  DisplayPort (a.k.a. *DP*) is a digital interface that is usually seen on modern, high-end graphics cards.
  DP cables and interfaces can carry audio through them as well, as well as USB or other forms of data.

  Unlike the VGA or DVI standards, DP cables do not have screws to hold them in place as standard, but does sometimes have latches to hold them in place.

  DisplayPort is becoming more common with time as HDMI becomes less common, and may be due in part to the fact that HDM is not a free technology, and manufacturers of tech that uses it must pay to licence it.

- **USB Type-C**
  USB Type-C specifically can output a display stream or streams.
  Note that not all USB-C connectors can do this - it is up to the manufacturer to determine if they support that functionality.
  USB-C is not actually a display technology by itself, and instead transmits one of the above standards (i.e. HDMI or DisplayPort), as the USB-C output will require a cable that terminates in one of the connector types listed above, however USB-C is itself a different plug type.

  Like DisplayPort, USB-C cables and connectors never feature a screw-locking mechanism to secure them in place.

The cables have maximum data throughputs as follows:
First Note: The maximum frame rates will vary according to resolution and colour depth, so actual results might vary from those given depending on the setup.
Also note:All the frame rate examples below are using 8bit colour and $Y'C_BC_R$ $4:4:4$ with 8/10b encoding as the bandwidth overhead and are meant as examples of commonly-used capacities the cables are capable of only, not what will max out the given data rate - some capabilities may have some overhead remaining beyond what is given.

Connector	Max Bandwidth	Max Data Rate	Examples of Capability
VGA	N/A	N/A	1440 75Hz
DVI [Single Link]	4.95Gb/s	3.96Gb/s	1080p 60Hz
DVI [Dual Link]	9.90Gb/s	7.92Gb/s	1440p 75Hz or 1080p 120Hz
DisplayPort 1 & 1.1	10.8Gb/s	8.64Gb/s	1080p 144Hz or 1440p 75Hz
DisplayPort 1.2	21.6Gb/s	17.28Gb/s	4K 75Hz or 1440p 144Hz
DisplayPort 1.3 & 1.4	32.4Gb/s	25.92Gb/s	4K 120Hz or 1440p 240Hz
HDMI 1.1 - 1.2	4.95Gb/s	3.96Gb/s	1080p 60Hz
HDMI 1.3 - 1.4b	10.2Gb/s	8.16Gb/s	1440p 75HZ or 1080p 144Hz
HDMI 2.0	18.0Gb/s	14.4Gb/s	4K 60Hz or 1440p 144Hz
HDMI 2.1	48.0Gb/s	42.6Gb/s	8K 30Hz or 4K 144Hz

VGA connectors do not have a maximum bandwidth or data rate because they are analogue, and their data rate will be dependant on their clock speed. Higher clock speed cables will have higher bandwidths. In addition, shorter cables and those with better internal wiring can have higher bandwidths as there is less signal degradation that would prevent higher bandwidths.
This also means that the maximum capability given, whilst certainly possible, may not be possible by certain cable configurations, or may be exceeded by others, and is meant as a rough comparative figure only.

6.3 Refresh Rate

Monitors will have a *Refresh Rate*, which specifies the maximum number of frames that the monitor can display.
This number usually refers to both the maximum number of frames that the monitor displays, as well as the framerate that the monitor always operates at.
Refresh rates are always listed in Hz (Hertz) - which is the SI metric for how many times something occurs in a second.
E.g. a 60Hz monitor will display 60 frames per second, and a 144Hz one will display 144 frames per second.
Most modern monitors are 60Hz, with 144Hz (or even 240Hz) being popular for high-end gaming systems.

6.4 Response Time

Monitors will have a *response time*, which is the amount of time it takes a monitor to change the image that is displayed.
Longer refresh times will therefore have more delay between the action on the computer,

and the output being displayed to the user.
Longer refresh times can result in a blurring effect on the image, called *ghosting*, where the monitor cannot change colours quickly enough, resulting in trails on moving objects.
Monitors that can change faster (i.e. have a lower response time) will have less visible ghosting, and thus should provide a sharper image to the user.

Lower refresh times are very useful for gamers, and some professional workloads, but are not as useful for the average user.
5ms is considered a good refresh time, and 1ms is extremely good. Some bad monitors can be 25ms or more.

Some [typically gaming-oriented] monitors claim “Anti-Ghosting" technology, that should help reduce the effect of ghosting on images displayed, and thus should improve the image quality.

6.5 Adaptive sync: FreeSync and G-sync

Typically how monitors work is by displaying a fixed amount of images (frames) every second. For example a 60Hz monitor will show 60 frames a second, a 144Hz one will display 144 frames per second (Hz being the unit for how many times things occur per second)
This number is fixed, and thus the graphics card that displays these frames must either always output this number of frames, or if it outputs less, then the monitor will typically just continually display the last frame delivered until a new one is delivered, all without changing the frequency at which they are displayed.
Adaptive-sync is the term used to describe a specific type of technology that does not work this way.

Adaptive-sync is sort of a reverse form of hardware based Vsync - which is a form of software where the GPU is forced to wait until the monitors next refresh-interval before displaying a new frame, thus preventing the scenario where the monitor receives a new frame as it is displaying the old one, which can result in screen-tearing as the new frame and the old frame differ, but parts of each are displayed in the same frame displayed to the user. However since the GPU is forced to wait, this delays the image the user sees, thus increasing input lag, and can create stuttering in images as previous frames are repeated.
Thus by doing this on a hardware-level, adaptive-sync eliminates this stuttering and tearing issue, which can result in a smoother, more fluid experience.

Adaptive-sync monitors will allow the graphics card to output as many frames as it wants, and the monitor will simply output the frames it receives, as it receives them.
Adaptive-sync monitors are designed to work at a variety of frame-rates, and will usually

only work within a specific framerate range (typically 40Hz-75Hz).
This allows the monitor to adapt to increases and decreases in frame rates as delivered by the GPU, but usually there is a minimum framerate allowable (see the 40Hz minimum threshold, above) before the monitor reverts to a fixed framerate to maintain consistency, however there does not necessarily have to be a minimum framerate.
However many adaptive-sync monitors can double the framerate if it dips below this threshold (e.g. if you have 25fps then the monitor can double that to 50, which gets it back into the range of the adaptive-sync), which can increase input lag a little, but reduces screen tearing and stuttering.

Whilst adaptive-sync monitors often have a wide range of refresh rates they can display (e.g. 9-240Hz), adaptive-sync only works in a specific range of this band (e.g. 40-75Hz).
This is because the monitor will need to be able to display frames faster than the average framerate delivered from the GPU, as the time required to deliver each frame is not the same.
Note: Problems such as screen-tearing and inconsistent frame-delivery are a problem that are not usually seen when comparing strict fps numbers, as these gloss over delivery.
That these can still be an issue that is problematic to gameplay are one reason why reviews that only give fps numbers are often not that useful, as it is possible to have high fps numbers but a bad user experience if there is a lot of screen tearing and stuttering. Instead comparing frame delivery times is a much more useful metric, and having a smoother frame delivery is typically a better experience than simply having more fps.

Adaptive-sync usually does not have a performance penalty to use, and merely requires that the GPU and the monitor it is outputting to support the same tech.

There are current two different implementations of adaptive sync: FreeSync, and G-Sync.

FreeSync is AMDs free and open-platform implementation of the adaptive-sync specification of DisplayPort.
FreeSync, being free, and thus costing less to implement (often considerably so) means that in the adaptive-sync space it is the more popular option, and AMD has a large market share in this area as a result.

G-Sync is a limited, closed-source version of adaptive-sync made by Nvidia that requires a specific hardware module to be bought from Nividia. The cost of this module was often why G-Sync monitors cost more than FreeSync ones, and is why Nvidia has less market share in the adaptive-sync market segment, with many gamers choosing AMD setups due to the lower cost, and many because they felt Nvidia was trying to force customers to pay more for technology when there was no reason for it, thus trying to abuse its position in the industry.
Nvidia did very recently state that it would allow G-Sync to work with non G-Sync monitors (a.k.a. FreeSync monitors). This is very useful for the industry, however the long-term effect

on AMDs monopoly in this space remains to be seen.

Despite the technical and cost differences between Freestyle and G-sync, the two are mostly functionally-identical.
The main differences between them are that FreeSync can be used over HDMI as well as DP, whereas G-sync can only be used on DP, and G-Sync monitors often have a wider refresh interval where G-sync works, which can potentially reduce input lag. AMD has also introduced *FreeSync 2*, which further reduces input lag.

At the beginning of 2019 Nvidia also announced that they would allow G-Sync to work with non G-Sync displays - Since the only adaptive sync technologies are G-Sync and Freesync, this means that G-Sync can now work with FreeSync displays via a software update provided by driver version 417.71 or later.
G-Sync was not originally designed to be used with FreeSync though, and as such some inter-compatibility problems may arise as a result.
Nvidia thus therefore tests adaptive sync displays according to how well they work with G-Sync, and certifies them accordingly.

- **“G-Sync Compatible"** means that the FreeSync monitor being tested produces no abnormal results when used with G-Sync, such as flickering, pulsing, blanking, or other strange effects that would be detrimental to the users experience.
- **“G-Sync"** (with no suffix) means that the monitor was designed for G-Sync instead of FreeSync, and these monitors that have the G-Sync certification have a refresh rate from 1Hz to the panels max display rating, 1000nit HDR, DCI-P3 colour support (which provides a wider range of colour than the usual sRBG), as well as other features. Nvidia state displays have to pass over 300 compatibility and quality tests to be certified for G-Sync.
  G-Sync monitors also have *variable overdrive*, which is where the monitor tries to anticipate the next refresh, preventing a frame being drawn as a new one becomes available, which can help reduce lag and/or stutter by waiting for the next frame to be completed.
- **“G-Sync Ultimate"** (previously known as *G-Sync HDR*) are 4K panels with 1000 nit peak HDR brightness, 95% of DCI-P3 colour, 384-zone contrast brightness, and a variable refresh rate from 0Hz to the monitors maximum refresh rate, which is a minimum of 144Hz, as well as all the benefits of regular G-Sync [above].

6.6 Panel technology

Monitors can have a variety of different technologies that display things in different ways, as below:

- **CRT**

Cathode Ray Tube monitors used an electron gun and a phosphorescent screen to display images.
CRT monitors were always large, as the electron gun would need to be mounted behind the front screen, resulting in a large monitor. This also made them very heavy.
CRT have almost entirely vanished from the PC ecosystem, having been replaced with other panel technologies such as LCD or OLED.
Despite this, CRTs are still sometimes used for hardcore gamers as they have almost no input lag at all, making them superior to all other panel technologies in that regard.

- **LCD**
  Liquid Crystal Displays use liquid crystals to modulate light, having a backlight to provide light that illuminates them.
  LCDs have two sheets of polarising material with a liquid crystal between them. An electric current causes the crystals to align, which will either block light or allow it to pass through, depending on the current applied.
  LCDs thus work by blocking out the backlight and modulating the light from it to produce images.

  LCDs can be produced that have no backlight, instead requiring an external light source to be visible. A well-known example of this is the original Gameboy Advance, which had a TFT LCD with no backlight, requiring external light to make the display visible to the user.
  LCDs can be much thinner than CRTs, and use dramatically less power, as well as being lighter and easier/safer to dispose of meant their introduction did somewhat kill off CRTs.

- **TFT**
  A *Thin Film Transistor* is a manufacturing technique that doesn't use a typical substrate, such as in a silicon wafer, but instead has a thin film of semiconductor deposited on the substrate directly, instead of the semiconductor material actually being the substrate itself.
  A TFT is thus a type of LCD if the substrate is the LCD glass, forming a TFT-LCD, though they are usually just referred to as LCDs instead of TFT-LCDs.

- **LED**
  True *Light Emitting Diode* displays have a matrix of LEDs that can display images, however computer monitors that use LEDs exclusively as pixels are not a real thing that exists [yet], and instead an “LED monitor" will usually refer to a TFT panel with an LED backlight, which can reduce power consumption and increase brightness compared to say a cold-cathode, however this terminology never really caught on as a selling point, and all modern monitors will use LEDs as their backlights.
  Despite this many monitors will list themselves as LEDs, and this always refers to the backlight used for the LCD. They can thus be referred to as an *LED backlit LCD-TFT monitor.*
  An LCD-LED monitor is very much not the same as an OLED monitor [see below for those], which function very differently and have vastly different properties.

- **TN**
  Twisted Nematic is a type of TFT panel that have faster response times that can help eliminate ghosting, making them popular for gamers, however can suffer from very extremely poor viewing angles, as well as often having a poor colour range.

- **IPS**
  In Plane Switching is a type of TFT panel that have the crystal molecules move parallel to the panels plane, rather than perpendicular to it as in TN panels. This dramatically reduces the scattering of the light in the matrix, increasing the viewing angles of the monitor considerably.

- **OLED**
  Organic LED displays have the light-emitting compound made of a film of organic compound.
  As such, OLEDs do not need a backlight, and thus can have deeper blacks than LCDs, as well as potentially be lighter and thinner, as well as potentially higher contrast ratios.
  However because they use organic components, OLEDs can suffer burn in as these organic components degrade in uneven manners, resulting in permanent areas of degradation, in which the image that was displayed on the screen becomes permanently displayed as a ghost-type superposition on the screen.
  Concerns about burn-in have hurt OLED adoption for conventional PC monitors, where common elements of the screen, such as the taskbar and program menus are commonly displayed and thus are potentially causes of burn-in. However OLEDs do have numerous advantages that make them desirable for many uses.
  Their organic nature also makes the quality of the display more variable, and colour accuracy can vary more between OLEDs than LCDs.
  OLEDs can also be relatively transparent, however transparent screens have not seen mainstream production or adoption yet.

- **AMOLED and PMOLED**
  AMOLED is *Active Matrix OLEDs*.
  Conventional displays have a complex system to control individual pixels, where charges are sent down the correct row and column to each pixel.
  Active matrix systems attach a TFT and capacitor to each LED, allowing the capacitor at each pixel to retain its charge in-between refresh cycles, allowing for faster and more precise control.
  Active matrix technology is often used for ordinary LCD TFTs instead of AMOLEDs, however is not usually marketed as such.

  Active matrix technology opposes *Passive Matrix* (*PMOLED*) technology, which controls each pixel in a row and column sequentially, instead of individually, which can reduce complexity, but is more limited in frame rate for larger displays than active matrix technologies, and as such AMOLED has become the prevalent type rather than PMOLED.

6.7 Stand and VESA Mounts

All computer monitors sold will come with some form of stand built in that allows the monitor to stand freely on its own and thus be used.
Many monitors also come with a *VESA Mount*, which is a specific screw-hole section on the back of the monitor that allows either objects to be mounted to it, or more usually; allows the monitor to be attached to a separate stand, instead of the default one.
VESA mounts for monitors allow them to be mounted to a wider variety of mounting mechanisms, such as swivel arms, or stands designed to hold multiple monitors as once.
A standard VESA mount is four screw-threaded holes on the back of the monitor.
There are multiple standards of VESA mounting available, which dictate the locations of the screw holes relative to each other, and VESA stands will often be able to accommodate several of these standards.
75mm spacing of square-placed holes is the common VESA standard for modern monitors.

Chapter 7

Power Supply a.k.a. *PSU*

A PSU (*Power Supply Unit*) is what is used to convert the AC power input from the mains electricity into multiple lower voltage DC lines that the PC can use.
Modern power supplies also often include protection systems that will trip if needed to prevent damage to the PC, usually by shutting down or blowing a fuse.
High quality units will often also include low-ripple outputs and other things that are useful or beneficial to a PC

7.1 Output

Power supplies are primarily rated by their output, usually a figure given in Watts.
For modern power supplies this is usually in the 300-700W range, although it is possible to get very high or very low output power supplies, for example a 160W PicoPSU, or a 2,000W PSU.
This figure denotes the *maximum* amount of power that the power can supply.

This output figure is not a single universal figure though, and is made up of the sum of several smaller output figures, most notably the 3.3V, 5V and 12V rails.
The PSU delivers certain amounts of power to each of these rails, and this total is then added up to get the total power rating of the power supply.
However, because of this, each power supply may be able to supply different amounts to each rail, and as such some power supplies may be more suited to certain types of PC than others.

For the average user getting any low-output (300W or so) power supply from a reputable manufacturer should be sufficient.

7.2 Efficiency and *80 plus*

Power supplies take their input from mains electricity, and convert that AC power to DC - since computer components use DC. (Most electronics use DC power)
This process is not 100% efficient - that is to say, you will get out less DC power than the AC power used to make it, with the wasted electricity being produced as heat in the PSU system.
More efficient units will therefore use less power, and generate less heat, which will be beneficial to the end-user.

Modern power supplies are all constructed differently and as such have a variety of outputs and efficiencies.
Power supplies are often tested and given a rating of their efficiency, usually according the the *80 Plus* standard, which can be determined as follows:

- **No rating**
  Obviously the worst efficiency rating is *not having a rating at all.*
  In truth any power supply that falls into this category because it does not meet even the 80 plus white spec below is probably best avoided.
- **80 Plus [White]**
  80 Plus (also known as *80 Plus White*, due to the white logo of this certification level) PSUs are rated to be a minimum of 82% efficient at 20% and 100% loads, and 85% efficient at 50% loads, and with a power factor of 0.9 or higher at 100% load. Higher power factors reduce peak current draw, which reduces load on the circuit.
- **80 Plus Bronze**
  80 Plus Bronze power supplies are 85% efficient at 20% and 100% loads, and 88% efficient at 50% loads, as well as having a 0.9 or better power factor at these load levels.
- **80 Plus Silver**
  80 Plus Silver power supplies are 87% efficient at 20% and 100% loads, and 90% efficient at 50% loads, as well as having a 0.9 or better power factor at these load levels.
- **80 Plus Gold**
  80 Plus Gold power supplies are 90% efficient at 20% load, 89% efficient at 100% load, and 92% efficient at 50% loads, as well as having a 0.9 or better power factor at these load levels.
- **80 Plus Platinum**
  80 Plus Platinum power supplies are 92% efficient at 20% load, 90% efficient at 100% load, and 94% efficient at 50% loads, as well as having a 0.95 or better power factor at these load levels.

- **80 Plus Titanium**
  80 Plus Bronze power supplies are 94% efficient at 20% and 100% loads, and 96% efficient at 50% loads, as well as having a 0.95 or better power factor at these load levels.

Note that the efficiency levels are different for 115V than the 230V levels listed above, and 115V has lower efficiency at all levels than 230V.
The “80 Plus" name is derived from this: 80 Plus White is only required to be 80% efficient at 20, 50, and 100% load levels for 115V.

A list of manufacturers that make power supplies that meet the above standards can be seen at:
https://www.plugloadsolutions.com/80PlusPowerSupplies.aspx

Note that there are no other 80 Plus standards, so if you see anything else listed than the above six (80 Plus, 80 Plus Bronze, 80 Plus Silver, 80 Plus Gold, 80 Plus Platinum, 80 Plus Titanium) then you should probably avoid it like the plague.
Also note that many dodgy power supplies might be named something similar to the above, such as an “80P" or something - naming something similar to a certification does not make it certified though, of course.
Always check the efficiency level stated if buying a PSU, and remember that having a bad efficiency rating is an awful lot better than not having an efficiency rating at all.

As an example calculation:
If a PSU is rated for 1000W at 80 Plus White, and the PC draws 200W, then the PC will draw 200/0.82=243.9W of AC power from the wall, thus wasting 43.9W of power.
If you were to use the PC for 2 hours a day then you will waste 2h*43.9W*365=32,047Wh=32kWh of power each year.
By comparison, if using a 400W 80 Plus Gold PSU, then the PC would draw 200/0.92=217.4W of AC power from the wall, thus wasting 17.4W of power.
If the PC is used for 2 hours a day then it will waste 2h*17.4W*365=12,702Wh=12.7kWh of power each year.
This means it would waste less than 40% of the power if it had a more efficient and more suitably rated power supply than a less-efficient and over-rated one.
Alternatively, the first example wastes over twice as much power as in the second one.
This is quite a difference in wasted power, even between 80 Plus units, let alone compared to a non-80 Plus unit that may not be 75% or even 70% efficient.

7.3 Modular

Power supplies deliver power to the components of a PC. They do this by having wires go from the PSU to the components.
Many users however, do not have that many components in their PC, and as such will not make use of most of these cables.
For this reason *modular power supplies* exist, which can allow the users to only have plugged in the cables that they actually use.
There are effectively three types of power supply modularity:

- **Fully Wired**
  Fully Wired power supplies have all the cables for the power supply permanently wired up to it, and connected at all times.
  If a user does not use a specific cable then it must still remain attached, and merely moved or secured out of the way.
- **Semi-Modular**
  Semi-modular power supplies come with usually only the essential 24-pin motherboard cable attached, with the rest being detachable, however some might come with a fixed 4/8-pin CPU connector, and/or an 8-pin GPU connector.
  This allows the user to only select the cables that they will actually use, and not have others cluttering up the inside of their PC.
  Semi-modular power supplies are often more expensive than fully wired ones, however.
  Because the connectors that are fixed are not always identical between semi-modular power supplies, it is a good idea to check before you buy one if which ones are permanently attached is important.
- **Fully-Modular** *Fully-modular* power supplies have every cable as detachable.
  This means the user can plug in any cable or combination of cables that they want., allowing for the minimum number of cables to be connected at a time, which reduces cable clutter to a minimum.
  Fully-modular power supplies are typically more expensive than semi-modular though.

Do note however, that even high-end fully-modular power supplies do often have all connectors of the same type coming off of the same cable, which means that extenders or splitters may still be required.
For example if a power supply has all of its SATA power connectors coming off of a single cable, then a user who wants to connect a disc drive at the top of their PC and a hard drive or SSD at the bottom of their PC case may find the single power cable simply isn't long enough to attach both without an extender. A user may also find that they have numerous unused SATA connectors on that cable as well, which must simply be dealt with or accepted in their case.
In the author's experience semi/fully modular power supplies that have one connector per cable to alleviate these types of issue don't really exist.

7.4 Protection systems and PSU quality

There are many things that set a good, high-quality PSU apart from the lower-quality ones, and/or can be used to judge the quality of a power supply, and I will try to provide a list of what I think are the most important ones below, however be aware that this is a non-exhaustive list, and I exclude others on the grounds of: I don't deem them important enough for the user to know about, nobody cares about them, they are niche or obscure in their use, I don't know about them, or I just forgot about including them.

Things to check for on a PSU to help determine quality:

- **DC Output table**
  All PSUs will come with a chart that lists their DC output capability (reminder: PCs take AC from the wall, and convert it to DC power that the PCs components can use).

  This chart is a given for all PSUs, and any PSU that omits it should be avoided like the plague.
  Similarly, any chart that has numbers that do not add up correctly should also be avoided like the plague.
  (This sounds obvious, but many cheap or off-brand PSUs fail at this.)

  Many cheap PSUs do not list their actual Wattages (which is another warning sign to look out for), however remember that $Voltage$ x $Amps = Watts$ and you should be able to tally up all the totals and see if it matches up with the quoted amounts.

- **DC temperature**
  PSUs have output tables that list their outputs (See above), and this will be computed at a fixed temperature.
  Industrial units will do this at fifty degrees minimum, and forty degrees is a good standard to hold PSUs to, however some very cheap PSUs might list them at thirty degrees, or potentially even lower.
  Being able to supply power efficiently at higher temperatures is an advantage, so the higher the temperature the better.
  I would recommend no less than forty degrees as the rating.

- **Connectors**
  You can tell a lot about what the manufacturer expects a PSU to be capable of by the connectors that it has.
  A PSU that has no PCIe power connectors is not expected to supply a high-end system, and a PSU that has only a 4-pin CPU connector and not an 8-pin doesn't expect the CPU to draw much power either.

- **Voltage regulation**
  The voltage regulation of a system is the percentage change in the output voltage from no-load applied to full-load applied, and is thus the measure of the systems ability to

provide constant voltage over a range of load conditions.
Since the power factor determines the output voltage, the power factor does influence the voltage regulation.
A lower voltage regulation is better, and below 2% is good, with some of the best PSUs capable of under 0.5% voltage regulation.

- **Certifications and ratings**
  A lack of any certifications or ratings testing to specified standards is a sign of a low-quality PSU.
  The lack of any basic certifications, such as RoHS, should be a huge warning sign of a bad PSU.

- **Ripple, noise, and transient response**
  Low ripple, noise, and transient response are often quoted on high-quality PSUs, and are usually omitted on low-quality ones.
  Ripple is the variation in the output DC voltage of the PSU caused by incomplete suppression of the input ACs waveform.
  Noise is a sinusoidal wave overlapping the output DC voltage, and is undesirable.
  Transient response is the time taken to restore the correct output voltage after a rapid rise in current. You want the best (lowest) transient response possible.
  Some digital power supplies might not have a transient response at all.

- **Cable gauge**
  The higher the gauge of a cable, the smaller the wire diameter.
  As such, lower wire gauges can be advantageous, as they will provide a longer lifespan as thicker cables are less likely to break or wear, as well as being more able to cope with extreme amounts of power going through them.
  20AWG is about the minimum I would recommend, however 18AWG would be preferable.

- **Weight**
  The weight of a PSU is a very telling factor, and any PSU that feels very light is probably going to be low-quality.
  This is because high quality caps, heatsinks, and other components will have weight to them.
  Any PSU that would be too light to use as a paperweight should probably be avoided.

Power supplies have a variety of protection systems designed to protect against problematic scenario's.
The ATX spec requires certain protections to be in place, and many good quality PSUs include other useful ones as well:
Protection systems on PSUs can be summarised as follows:

- **OPP**
  Over-Power Protection will stop the PSU from drawing more than its rated capacity.
  The OPP trigger point is often set a little above the maximum capacity of the unit, so it can be overpowered slightly.

The ATX spec does not require OPP.

- **OCP**
  Over-current Protection ensures the output of the DC voltage rails remains below a safe limit (240VA per output rail), and is common for PSUs with multiple +12V rails, but can also protect the minor rails too.
  ATX spec requires OCP.

- **OVP**
  Over-Voltage Protection is used on the 12V, 5V, and 3.3V outputs and will shut down the PSU if the DC outputs exceed the rated levels.
  ATX specifics three levels for this protection, Minimum, nominal, and maximum.
  The minimum voltage levels required for compliance with this are 13.4V (for the +12 rail[s]), 5.74V (for the +5V rail), and 3.76V (for the 3.3V rail), and optionally 5.74V for the 5VSB rail.
  The nominal voltage levels required for compliance with this are 15V (for the +12 rail[s]), 6.3V (for the +5V rail), and 4.2V (for the 3.3V rail), and optionally 6.3V for the 5VSB rail.
  The maximum voltage levels required for compliance with this are 15.6V (for the +12 rail[s]), 7V (for the +5V rail), and 4.3V (for the 3.3V rail), and optionally 7V for the 5VSB rail.
  ATX spec requires OVP.
  Note that it is possible for components to be damaged by too much voltage before OVP kicks in.

- **UVP**
  Under-Voltage Protection protects the voltages on each rail from going below certain points.
  The ATX spec does not require UVP.

- **OTP**
  Over-Temperature Protection will shut down the PSU if the heat sensor (thermistor) in the PSU detects that the PSU is too hot.
  The ATX spec does not require OTP, however many users consider it essential.

- **SCP**
  Short-Circuit Protection will shut the PSU down if it detects an impedance of less than 0.1Ω, implying a short-circuit between the PSUs DC output rails.
  ATX spec requires SCP, and each +12 rail should have its own short circuit.

Note that UPP, UCP, UTP are not things that PSUs have, and it should be taken as an *extreme* warning sign if any PSU claims to have any of them.
The ATX specification requires OCP, SCP, and OVP. The others are usually only found on higher-quality PSUs.

There's lots of stuff I can mention to those that want to take their PSU apart and look at its quality, such as good input filtering, bridge rectifiers, good-quality caps, beefy transformers, etc, but most users are unlikely to need to know about the list above, let alone any of those, so they are omitted for the sake of brevity.

7.5 Voltage switch

The Voltage switch is a physical switch that tells the PSU what Voltage the mains input to it is providing, and should be set correctly before connecting the mains supply, and certainly before turning the PC on (or trying to).
Incorrect voltage can severely damage the PSU, and potentially any other components it is attached to.
In addition many PSUs will not operate at something other than the intended frequency at all.

PSU voltage switches will switch between 110V/127V and 220V/240V.
220-240V @50Hz is the main voltage used worldwide, and is what the majority of countries in the world use.
220-240V @60Hz is used in only a handful of countries, and is exceedingly rare.
100-127V @50Hz is extremely rare, and is only used in Northern Japan and Argentina.
100-127V @60Hz is used for North America, and numerous countries in South America.
As such, whilst not a universal rule, as a general rule assuming that a 100-127V supply will provide 50Hz, and a 220-240V supply will provide 60Hz will be correct in the majority of cases.

These days, voltage switches are usually only found on very cheap PSUs, with most decent PSUs being able to automatically adjust to either input frequency and voltage - however do note that this is not guaranteed, and if switching voltages then you should definitely check to ensure your PSU is compatible with the new frequency.

Because decent PSUs will usually be capable of automatically detecting the input frequency and adjusting accordingly, and that it is usually only very-cheap off-brand PSUs that have voltage switches, then the author recommends that any PSU that has a physical voltage switch on it should be heavily avoided.

Note also that PSUs are less efficient at 115V than 230V.
(i.e. They require more AC power input to produce the same DC power output, with the wasted power being emitted as heat.)

7.6 Form factor

PSUs can vary in their size and shape, and the specific size and shape of the PSU is called its *Form Factor*.

By-far the most common size for PSUs is the “ATX" form factor. (a.k.a. a “Full-sized", or “Full-sized ATX" sized PSU.)
Other sizes do exist, such as the rare SFX form factor that is often used for SFF cases, however the vast majority of available PSUs will be ATX sized.
Some manufacturers may also make custom sized PSUs for their own cases, which will typically come with the case. This is extremely uncommon for consumer PC cases and PSUs, however is somewhat more common with pre-built systems.

7.7 Cables

PSUs will come with a variety of cables that they use to output power to the various PC components.

Modern PSUs will come with the following:

- **24-Pin Motherboard**
  The 24-pin connector will attach directly to the motherboard, and will provide power to the motherboard and onboard components, such as RAM.

  Very old motherboards used a 20-pin connector, and modern 24-pin connectors use the same 20 pins for the same function, with 4 additional pins attached along one end.
  As such, many modern 24-pin cables have the extra 4-pins as a separate cable, or only attached on one side by a living hinge, which allows the 4 pins to be moved out of the way, allowing the 24-pin connector to adapt and be used for 20-pin motherboards. Such cables are often called *20+4*.

  Since the 24-pin is an identical 20-pin just with additional pins, there is no compatibility issues by using a 24-pin in a 20-pin slot.
  In addition, a 20-pin cable will fit a 24-pin motherboard for the same reason, however this can cause problems due to lack of available power.
  Each of the 20/24 pins/sockets will be individually notched, so the cable will only ever fit into the socket in one way, and cannot be inserted incorrectly.

  The 24-pin connector is held in place by a latching clip on one side, however often have high friction that helps hold them in place too.

- **4/8-pin CPU a.k.a *ATX/EPS 12V***
  The CPU will have a 4-pin or 8-pin cable for itself that is plugged directly into the motherboard, usually very close to the CPU socket.
  The 8-pin version is just a 4-pin connector with extra 4-pins on one side.
  This means that a motherboard that has a slot for an 8-pin can accept a 4-pin connector, however if the CPU is very power-hungry then this might cause issues with the PC as the CPU doesn't get sufficient power delivered to it. For low-powered CPUs though this setup can work fine.

  The 4-pin connector is often called an *ATX 12V* cable, and the 8-pin called an *EPS 12V*.

  Many PSUs have two separate or latching 4-pin cables, such that either the 12V ATX or 12V EPS can be used, without the extra 4 pins of the EPS 12V getting in the way should the motherboard only have the 4 pins of the ATX 12V plugged into it.
  Such cables are often called *4+4*.

  12V ATX/EPS cables are held in place by a latching clip on one side, however often have high friction that helps hold them in place too.

- **SATA Power**
  PSUs will typically have several SATA power connectors, which are thin 15-pin connectors, almost always used for modern PCS to power HDDs, SSDs, and optical drives.
  The SATA power connector has an internal “L" shaped slot/socket, meaning it can't be inserted the wrong way up.

  SATA power connectors deliver +12V, +5V, and +3.3V.
  Note though that most modern drives do not use +3.3V at all, and as such molex cables (which only provide 12V and 5V) can be used via converters to convert to SATA power without issue.

- **Molex**
  Basically all modern drives (HDD, SSD, optical drive) will use SATA power as their power input, however molex has been used for so long for drives, and is still very often used for PC expansion cards, and as such modern PSUs will usually include several molex connectors.
  Molex cables are also sometimes used for fans or fan controllers.

  Molex cables are chamfered on one side, and the socket notched appropriately, meaning that molex cables can't be inserted the wrong way up. Molex cables are friction-fit, and are often difficult to insert/remove. In addition, molex cables use pins on their cables that can be damaged, meaning some care is often recommended when trying to connect them, especially if the cable seems to require lots of force to insert.

Molex cables are usually colour coded Yellow-Black-Black-Red, which are +12V, ground, ground, and +5V, respectively.

- **PCIe Power**
  PCIe power cables are almost always used to provide additional power to discrete GPUs.
  PCIe power cables come in 2 forms: 6-pin, and 8-pin.
  The 8-pin version is the 6-pin with 2 additional pins, which are a sense-pin to detect if the connector is a 6 or 8-pin, and an additional ground pin, as well 2 from the 6-pin being changed to a +12V.
  Many PSUs will include a +12V on pin 2 for the 6-pin cables though for redundancy, making them fully compatible with the 8-pin version.
  Because the 8-pin is simply a 6-pin with 2 more pins, anything that accepts a 6-pin will also natively accept an 8-pin.
  However, because the 8-pin connector is of course physically larger than the 6-pin, the 8-pin connector might not fit into a 6-pin slot due to the extra 2 pins, and as such most PSUs will have their 8-pin connectors implemented as a 6+2-pin configuration, where the 2 additional pins are usually on a short cable and thus can be simply moved out of the way if needed.

  The 6-pin connectors are rated deliver up to 75W of power, and the 8-pin versions up to 150W.
  Note that the PCIe slot itself can provide up to 75W as well, which will be in addition to any power from the PCIe cables.
  Note also that the PCIe power connectors are usually able to carry supply more power than this, and the above is just the minimum spec.

  The pins on the side adjacent to the latch on top of the connector that holds it in place are the ground, and those on the flat, opposite side (with no latch) are the +12V pins.

Note that all PC cables are designed in such a way that they can only be correctly inserted into one slot and in one orientation only, meaning that the correct connector for any slot will be somewhat easy to visually discern upon inspection.

Please further note that PCIe 8-pin connectors and [8-pin] 12V EPS cables are both very visually similar, however have different electrical connections, and thus are incompatible. They are keyed differently, so unless extreme force is used or the socket/cable actually broken they should not fit each other anyway. If they were to somehow be connected together though, the different power outputs would damage the component they were connected to.

Lastly: in modern PCs; Black power connectors are Ground, yellow are +12V, Red are +5V, and orange are +3.3V. Some older cables might have different colour-coding though.

Chapter 8

Case

8.1 Motherboard Size / Form Factor

The most important thing about a case is the motherboard sizes that it can have mounted inside of it.
This is also sometimes called *Form Factor*, however this term is also [sometimes confusingly] used to refer to the overall size of the case itself.

Motherboards come in a range of different *form factors*.
The form factor of a motherboard will determine its physical size, as well as where all the screw-holes are for mounting it, and each form factor will usually have a fixed location for the rear-IO (such as USB sockets) on the motherboard.
The most popular current motherboard form-factors are ATX (9.6x12 inches) and micro-ATX (9.6x9.6 inches). Mini-ITX (6.7x6.7 inches) motherboards are sometimes also seen.
The larger motherboards will have more room on them for components and expansion cards, for example more slots for components such as RAM, or more PCI slots.
For ATX, MicroATX, and MiniITX the same screw-hole layout is used for securing them in place, with larger boards merely having more screw-holes, thus any case or enclosure that can fit a larger motherboard of these will also accept the smaller variations as well.
Computer cases will usually only fit up to a certain size of motherboard in them, and will not accept any larger standards.

The form factor of a case can also mean the overall size of the case itself, in which case several terms are used to describe this.
The most common term is “Full Tower" (often a.k.a. just “Tower"), which describes a full-sized upright tower system, of the approximate shape and size most people would stereotypically assume a PC case to be, and full-sized towers should be able to fit a full-sized ATX motherboard in them. Full-sized towers vary in size quite a lot, but are typically more than 50cm high.

Similarly “Mid-Tower" will refer to a tower system that is shorter than the full-sized tower, and often thinner or shallower as well, and mid-towers should be able to fit a full-sized ATX motherboard in them.
Micro-ATX cases will be able to fit micro-ATX motherboards in them, but not full-sized ATX cases, and are usually smaller than a full-sized tower case.
Mini-ITX cases are built for the mini-ITX sized motherboard, and are considerably smaller than the above types, and will typically not be able to fit Full-sized ATX or even micro-ATX in them.

Note that there is no fixed definition for what constitutes each of these sizes, and what size it is labelled as should tell you what motherboard sizes it can fit, however double checking this to be certain is still best practice.

Larger case sizes will have more room for drive bays, and will have more internal volume for cable routing. Smaller cases may also be much more limited in the maximum GPU length they can accommodate.
Thinner cases may also be more limited in GPU size as well as CPU cooler size.

Other terms exist, such as “Super Tower" (or what-not) will refer to larger versions of the full-sized tower form factor.
“Mini-tower" is also a term that is rarely used, and usually refers to a micro-ATX sized case between a mini-ITX and mid-towers in terms of size.
“SFF" is also sometimes used, and refers to extremely spatially compact cases, as often used for for example HTPCs, and can vary a lot more in their dimensions and volume than a normal PC case.
SFF cases usually accept a SFF-sized PSU as well, instead of a regular ATX PSU.

8.2 Motherboard Tray/Standoffs

The *Motherboard Tray* is the area of the case that the motherboard directly attaches to.
The term refers to the days when the tray itself was removable from the case, however for modern PC cases that is almost never the case.

Because most PC cases are metal, however, the motherboard should not touch it, or else risking electrical shorting to the case, which is bad.
Motherboards are therefore separated from the case by *Motherboard Standoffs*, which are small metal spacers that go between the PC case and the motherboard.
The motherboard will have holes in it so it can be attached to these spaces, and the PC case will have holes for them to be screwed into.
Many motherboards will have these standoffs preinstalled, and many PC cases will also list

the holes in them that the standoffs can be placed in, detailing which size of motherboard they are designed for (i.e. ATX, Micro-ATX, mini-ITX).

8.3 Fan/radiator slots

Computer cases will have certain areas in which fans can be placed. These areas will have screw-holes for a certain fan size or fan sizes, as well as a circle of grill or mesh in the case to allow air to flow through mostly unobstructed.
More fan slots can be beneficial for those who want maximum cooling for their PC, however having lots of unused fan slots can impede cooling, as there may be lots of holes for air to be pushed/pulled through, resulting in less air being moved over the critical components to be cooled. This is usually not a problem, however.

PC cases can also have mounting points for radiators. Water cooling radiators are typically rectangular in shape, as opposed to the square mounting holes of case fans, and their additional size means many cases will not fit standard watercooling radiator sizes.
Cases that have mounting-holes for watercooling radiators will often be said to have "watercooling support".
Most cases that can have watercooling radiators fitted are very limited in where those radiators can be mounted inside the case, often relegated to just the top or bottom of the case.

8.4 Dust Filters

Wherever there is an open hole in the PC case, there is a possible point of entry for dust to get into the system, which can result in clogging up of fans over time, or coating heatsinks, both of which will decrease the cooling performance of the system.
As such many PC cases come with dust filters.
Dust filters are typically either custom mesh pieces for that case specifically, or more usually; standard size mesh covers for the PCs fan slots, since fans/fan slots are the most typical entry point for dust into a system due to them moving by far the most amount of air into and out of the system, as well as the pressure currents they create being more able to move dust around.

Dust filters can also be purchased on their own and fitted to a system, typically for each of the systems fans.
Note that dust filters are always going to impede the performance of a fan, because they will have to cover the airflow path of the fan to at least some degree in order to stop dust from getting through.
Good dust filters will block the airflow as little as possible, whereas bad ones will block a

large amount of surface area that the fan could blow through.
Very bad dust filters, or dust filters fitted to fans that already have a hard time moving air due to obstructions (such as hard drive bays) may warrant an upgrade to the fan to a high static pressure model to increase airflow, however many people will not find this necessary.

8.5 Cable Management

PCs often have large quantities of cables inside to transfer data and power between the internal components, and these can get untidy and unsightly if not managed properly.
Cable management refers to the practice of securing these cables out of the way such they are unable to move about inside the case, and are neat and tidy.
Cable management is not technically necessary, and many people do not bother, however proper cable management has the advantages of allowing cable paths to be easily seen, thus allowing their start/end points to be easily determined, which can help with upgrading or troubleshooting. Cable management also prevents cables from moving about, and can thus prevent them from falling into fans should the case be knocked or moved whilst running, which can damage the fan, as well ass preventing cooling of the component if the fan is then unable to run, which can cause further issues.

Cable management usually comes in the form of cutouts in the motherboard tray and other areas of the case through which cables can be passed through, as well as loops or hooks to which cables or groups of cable ties can be cable-tied, Velcro'ed or otherwise attached to.
Good cases will also sometimes rubber coat the cable management holes, hooks, or areas to tidy them up more, as well as prevent the cables pulling on the sharper metal of the case.

8.6 Tool-less

Computer cases are [usually] big, heavy things made of metal, and thus need to be held together well for structural integrity and rigidity.
This is usually done by welding, riveting, or screwing internal components together, with side panels and drive bays being screwed in, with the user removing the screws to gain access to them or to remove them.
A *Tool-Less* design refers to one in which tools (such as the typical screwdriver) are not required for attaching or removing components.
This can be achieved by a number of methods, typically via methods such as thumb-screws or snap-on systems, however many many other types of tool-less deigns exist.
Tool-less designs are somewhat common on high-end cases, but are rare on low-mid range ones.
Tool-less designs do not necessarily make any of the components more secure, however they

can make working on the case much faster and more pleasant for the user.

8.7 Bays, 5.25", 3.25", 2.5", and expansion slots

Cases will support a certain number of *Drive Bays*, which is the number of each of the types of drives it can fit inside itself.
Typically these will be 5.25" (i.e. CD/DVD/Blu-Ray drive sized), 3.5" (i.e. standard PC hard drive sized), and/or 2.5" (i.e. SSD and laptop-hard drive sized).
Each case will list how many of each of these it has, and the total number of these is often referred to as the number of *bays* the case has.
Even if a user does not use all of the bays, then having more may still be beneficial down the line if the user wants to use them in the future, either by upgrading their current system, or for a new system altogether.
5.25" bays will almost exclusively be located on the front of the system, with an opening at their front to allow them to be accessed from outside the case, which is useful for disk drives, however it is becoming increasingly-popular to use them for fan controllers or PC readouts (such as for the cases internal temperatures).

Converters that allow smaller drives to be installed into larger bays are popular, relatively cheap, and readily available these days.

Bays are not to be confused with *expansion slots*, which refers to the number of slots open in the back of the case that can be used for expansion cards for the system, usually PCIe cards.

8.8 Front-panel I/O

Many cases will come with *Front Panel* connectivity, which is where the case will have a variety of ports on some part of it that can be connected up to the system.
This makes it drastically easier for the user to quickly connect devices to the PC, since they will not have to go around to the back of the PC case, and can do it easily and much more quickly.
Because front-panel ports are supposed to be readily accessible, they are often located next to the power and reset buttons that a PC would come with anyway.
Note that the term “front panel" is a generic term, and front-panel may refer to anywhere that is on the PC case that is not on the back, and it is not uncommon to see such I/O panels on top of certain cases.
Common examples of ports on front-panels include USB and audio (headphone and/or mi-

crophone jacks).

8.9 CPU-cooler cut-away

The CPU cooler is a large cooling system that attaches on top of the CPU of the system, and is likely the most important cooling component of any system.
The CPU is mounted on the motherboard, and is often screwed in from the back of the motherboard, meaning that to remove the CPU cooler, the motherboard must also be taken off as well.
CPU-Cooler cut-aways will therefore provide an area of the motherboard tray removed behind where the CPU would sit, which allows the CPU cooler to be removed without removing the motherboard.
For most users this is not something they will ever want to do, but for overclockers, those who like to tinker with their systems, or just when troubleshooting, it can be a useful time saver.

8.10 Noise isolation

There are a variety of terms that are used to describe these kinds of systems; "Noise-isolation", "sound-proofing", "noise-damping", etc, etc.
They all do the same thing, however; reduce the noise of the PC.
These usually work by coating the walls of the PC case with an acoustical damping material, or by rubber anti-vibration mounts for the hardware.
Noise-damped PC cases will often have support for larger fans, which can provide the same airflow as smaller ones, but at lower RPM values, which will reduce noise.

Note that most noise-damped PC cases will typically have less cooling capability than a regular PC case, however this does not necessarily make them bad in that regard.

Also note that whilst noise-damped PCs should produce less noise than an undamped PC, noise can vary in other qualities, such as pitch, and as such a noise-damped PC may not actually be more acoustically pleasing to the user than an undamped and/or louder one.

8.11 Tempered glass

It is a recent trend to try to include as many tempered glass panels in high-end cases as possible.
Tempered glass (and indeed any transparent/translucent material) will allow the inside of the case to be more easily seen, which can show any off RGB lighting inside the case, or the 'excellent' cable management, or other aesthetically-pleasing features of the case.
"Tempered" glass differs from regular glass because it is constructed differently, and is designed to break into lots of small pieces if/when broken, which will result in no large shards or small slivers of glass that could easily cut the user, and is thus a type of safety glass.
Tempered glass is also typically stronger than normal glass, as well as more resistant to thermal changes.
Users who are worried about breaking tempered glass though are in truth probably better off getting a case without any glass in at all though.

Tempered glass is almost always more expensive than regular metal panels, as it is difficult to work with, especially to bend

8.12 PSU shroud

A *PSU Shroud* is a part of the PC case that is designed to hide a length of the inside of the case, obscuring the PSU from sight, as well as a lot of the wiring from it.
PSU shrouds usually run along the top or bottom of a case, since this is where the PSU is usually mounted.
The nature of a PSU means that it needs to be connected to the system somehow, and PSU shrouds will often have holes or cutouts for the PSUs cables to feed through to actually get to the components, however for some cases they may not, and have the cables routed behind the motherboard tray or in other manners.
PSU shrouds are typically only used in cases that have transparent side panels, such that the PSU would be visible in the case if the PSU shroud were absent.
Note that PSU shrouds can negatively affect the cooling of the PSU, often substantially.

Chapter 9

Fans

Case fans are useful for keeping the components of the PC cool.
Many cases come with a fan pre-installed, but for cheap case these may be low quality, and replacing them can reduce noise and increase airflow, which can help keep the components cool.
At the low-end fans such as the *Arctic F8* are a cheap and competent option.
At the high-end fans such as those made by *Noctua*, such as the *NF-F12*, can provide very high airflow whilst producing low noise.
The general recommendation is that having at least one fan pulling air into the case, and one pushing out of it is advised, but in truth their orientation is not overly important, as the case will usually have plenty of other holes for air to move through.
Having *at least* a single case fan is *very strongly* advised, even for low-powered PCs that will not generate much heat that needs to be dissipated.
For more powerful systems that will generate more heat, such as gaming PCs, more fans can definitely be useful to prevent the components from decreasing their performance to try to avoid getting too hot.

Most PCs have protection systems that will automatically drop them down to the lowest possible power draw, or just outright shut the PC down if it gets to around 100°C or more, with 100-105°C being the maximum they will usually accept before protection systems forcibly activate, usually shutting the PC down.
As with most protection systems though, if you are getting close to tripping them, then something should probably been done some time ago to keep the system in a more optimal range.

Typically it will be the systems CPU or GPU that gets too hot, as these produce by far the most heat out of all the systems components.
Most other components in a PC will likely be relatively cool by comparison.

Try to avoid the PC from getting above 90°C, as many common components will start to

throttle at around that temperature, however for some newer components throttling may happen as early as 70°C, and as such 75-85°C is a good general maximum though before you should add more case fans or position them better, or move the case to get better airflow, or other techniques to keep it cool.
A temperature of below 65-70°C is very good though, and not to be worried about.
Whilst the above are [in the author's opinion] good general rules, if the PC is not thermal throttling then it is fine, regardless of what temperature it is at, however PCs do have lower life expectancy of their components at higher temperatures, so keeping a PC cool does still have advantages even if the PC isn't thermal throttling.

9.1 Diameter and Mounting Holes

When buying fans you will see the "Diameter" listed, which is the total length along any size of the fans case.
Spaces where fans are to be affixed generally only accept a certain size or sizes, and are designed to have fans of that diameter in them, and fan grills or grates will typically only have spaces to allow air through for fans up to the maximum size of that fan.

Larger diameter fans will move more air, and be quieter, but also more expensive.

The Diameter of a fan is not the same as its mounting size - the distance between the screw holes used to secure it.
The diameter of the screw holes used to mount fans is different from the stated size of the fan:
Note that this is either the horizontal or vertical distance between screws, *not* diagonal distance.

Stated Fan Size	Distance between Mounting Holes
40mm	32mm
50mm	40mm
60mm	50mm
70mm	60mm
80mm	72mm
920mm	83mm
120mm	105mm

9.2 RPM/speed

RPM is *Rotations per minute*, and is the metric used for how fast the fan rotates, and is thus also called *fan speed.*

All fans sold will have a maximum RPM listed, which is their max rotation speed, however they will be able operate below that.
Fans will also have a minimum rotation speed, which is the slowest they can rotate before stalling, and will not run below that stated RPM.
All else being equal, fans operating at higher RPMs will move more air, use more power, and be louder than ones at lower RPMs.
The fans RPM will determine the noise it produces, typically at a rate of roughly
$RPM^5 \sim Noise$ (i.e. Noise is proportional to the 5th power of RPM.)
Note that noise is measured on a logarithmic scale (of dB, *decibel*), whereas RPM is measured absolutely.
Thus increasing/decreasing the fan speed by around 40% will result in a perceived doubling/halving of fan noise, respectively.

9.3 Air flow

Fans rotate, and their blades move air through them, however how these blades are designed and aligned will determine how much air they can move, and at what pressure.
The *Air Flow* of a fan is the amount of air it is capable of moving, and is thus a measure of volumetric flow rate.
Higher airflow represents the ability to move more air, and thus improve cooling, however other factors, such as pressure, impedance, and fan positioning will also affect cooling as well.
Air flow is usually given in either CFM (*Cubic Feet per Minute*), or m^3/h (*Cubic Metres per Hour*), which is the SI metric.
$1m^3/h = 0.589CFM$ (or 1 $m^3/h = \frac{60}{0.3048^3}\ CFM$) [SI Metric]
$1CFM = 1.699m^3/h$ (or 1 $CFM = 60(0.3048^3)\ m^3/h$)

9.4 Pressure, and High Static Pressure

Fans use rotating blades to move air, and generate high volume throughput, typically at low pressure.

The pressure of a fan is the force that it moves air with, and thus higher pressure fans are more able to force more air past obstructions that block their air path.
This means that if there is something blocking the air flow of the fan, then a higher pressure fan will be more able to move air past the obstruction.
What this means for PC users is that High Static Pressure fans are preferable for when their air path is blocked, for example those mounted next to a rack of hard drives, where the air must be forced between the drives.

High static pressure fans are typically used in places where air paths are restricted, such as watercooling radiators.
High static pressure fans do not always have a lower airflow than normal fans, however will typically cost significantly more.
There is generally no real disadvantage to having high static pressure fans of equal airflow, save for the additional financial cost.

9.5 Fan Controller

Fans rotate at a speed determined by their voltage or PWM signal, and the point of the fans is to cool the systems components, which they do by moving more or less air, as required.
For most use cases, the systems fans will not need to rotate at their highest possible RPM to achieve sufficient cooling, and can instead spin at a lower RPM, which will decrease the noise of the system, making it more pleasant to use.
The point of a fan controller in a system is to optimize the rotation speeds of the fans in the system to maintain the system below a given temperature, whilst producing the least amount of noise possible.
Since fans are controlled via voltage or PWM signals, this is what the fan controller will use to modulate and control the fan speed.
Fan controllers are usually pieces of hardware that the fans are plugged into directly, and vary in their capabilities, maximum number of supported fans, and other factors.
All commercial fan controllers will accept 3-pin fans, however not all will have support for the 4th (PWM) pin, instead using voltage control for them.
Many fan controllers also allow custom profiles, which will change how the fan speeds vary according to the systems temperatures, allowing for systems aimed more towards extreme cooling, noise, or anywhere in between, and on some fan controllers these profiles may be set by the user instead of predefined by the manufacturer.

9.6 PWM and connector pins

Computer fans run at a voltage given to them by the wires going to them.
PC fans require a minimum of two pins connected to work, and can have up to four pins maximum.
All fans can be plugged into fan headers with less than their maximum number of pins used (e.g. plugging a 4-pin fan into a 3-pin motherboard header), however you will lose the additional functionality provided by the extra pins.

Fan pins are determined as follows:

- **2-pin**
  This is the minimum number of pins required for a fan to work at all.
  One of these pins is +12V, and the other is the ground.
  The speed of 2-pin fans is controlled exclusively by tweaking the input voltage.
  Basically all modern motherboards are capable of this, however 2-pin fans are very rare these days.
- **3-pin fans**
  3-pin fans also rely on voltage regulation for speed control, just as 2-pin fans do.
  The third pin reports the fans RPM to the motherboard/BIOS, and can allow much more precise control of the fans speed, however is not as good as PWM control [below].
- **4-pin/PWM fans**
  The 4-th pin represents a PWM signal wire, and 4-pin fans are often called “PWM fans" because of this.
  PWM (*Pulse Width Modulation*) is where the signal is not continuous, instead coming as numerous discrete parts, and the voltage is controlled by how quickly this signal is turned on and off.
  PWM fans can use less power than non-PWM fans, however fans tend to use so little power that this is unlikely to add up to a significant amount.
  PWM allows for much more control over the fans speed, which can otherwise only be controlled via voltage.
  PWM fans typically also have lower minimum RPM speeds, which can be beneficial for noise and power consumption if the extra cooling of higher RPM fans aren’t needed.

There are generally no disadvantages to having more fan pins control the fan, however for many users who don’t tax their systems the functionality of the extra pins is probably not going to be useful either, meaning that whilst fan quality is still useful, the connection method likely isn’t.
For users with more demanding workloads on their computers, the extra functionality can result in less noise or better cooling, which can be beneficial.

9.7 Bearings

Fans work by having a motor shaft rotate when given electricity, which caused the blades attached to the motor shaft to spin, thus moving air.
This motor shaft is constrained by the *bearing*, which constrains it and prevents movement in incorrect directions, allowing only rotation, as well as drastically reducing friction of the motor shaft, which can increase speed and longevity of the motor.

Fans are not all constructed identically, and will differ in their bearing mechanisms:

- **Sleeve Bearings**
  Sleeve bearings use a form of grease or oil to lubricate the motor shaft and thus reduce friction.
  Sleeve bearings can often start out quiet, however become louder as they age due to the oil drying up.
  Sleeve-bearing fans are cheap and affordable, and represent the majority of what PCs currently use.
  Sleeve fans are best mounted vertically due to their lubricant system, and will wear out quicker if mounted in other orientations.
  Sleeve-bearing fans often cannot work above a certain temperature (although this is high, say 70°C), due to less efficient lubrication preventing rotation or just causing immediate failure.
  Sleeve-bearing fans also tend to fail catastrophically and immediately, with no warning beforehand.

- **Ball Bearings**
  Ball bearings use a series of balls that are held in place around the motor shaft, and when the motor shaft rotates, these balls will rotate as well, with the smaller balls transferring the rotational load to the system.
  Because the balls are moving, they have a lower coefficient of friction than two flat surfaces moving against each other.
  Ball bearing fans are often a bit louder than sleeve-bearing ones.
  Ball bearing fans can tolerate higher temperatures than sleeve-bearing ones, and can be mounted in any orientation just fine.
  Ball bearings tend to fail slowly and over time, and this also causes noise to slowly increase over time due to failure of the grease channelling used for lubrication.

- **Fluid Bearings**
  “Fluid dynamic fans", “Fluid fans", “Hydro fans", etc, are all names for this type.
  In fluid-bearing fans the load is supported by a layer of rapidly moving pressurised oil between the surfaces, with no actual contact between the surfaces.
  Since there is no actual contact, there is no sliding friction in the system, which results in less noise and considerably longer lifespan.
  Fluid-bearing fans can also potentially move more air than other types above.
  Fluid-bearing fans are usually the most expensive type, but are also the most reliable and quietest.

- **Maglev**
  Magnetic Levitation fans use magnetic levitation to levitate the motor shaft, thus having no physical contact at all, which gives them extremely low friction whilst allowing the fastest rotation speed, as well as very low mechanical wear.
  Most maglev fans use electromagnets to stabilise the load in combination with an active control system, however passive versions with fixed magnets are also possible.

Maglev fans are typically extremely quiet, and have long lifespans and high rotation speeds , but they are still very uncommon for PC fans, and those that are available are often very expensive.

Chapter 10

VRM

10.1 About VRMs

The *Voltage Regulator Module* is a component that supplies the correct voltage to components.

GPUs will have their own VRM systems on the card for the 12V PCIe power they take, however other system components will almost always have their voltage delivery done by the motherboard.
The VRM converts the supply voltage (12V) into something the other components can use - for example many CPUs run at around 1.2V, and DDR4 typically runs on around 1.25V, whereas most PSUs tend to mostly deliver 12V, hence a conversion system is needed to turn that 12V into something the component can use, which is what the VRMs do.
(Note: This makes VRMs a type of DC-DC converter, usually a type of Buck converter.)
Note also: Whilst the PSUs 3.3V and 5V are often used for powering components in a PC, it is almost exclusively the 12V that is used to power components via VRMs.

The term VRM is almost exclusively used to refer to the CPUs VRM on the motherboard, or the GPUs VRM on the GPU itself, however they are used elsewhere, but the CPU and GPU are the components in a PC that draw the most power by far, which makes their VRMs much more important .
I'll mostly focus on the CPU VRM for this chapter, as that is where it is most commonly encountered, however all the basic terms and methodology apply to basically all PC VRMs.

VRMs are useful to increase the efficiency of the power delivered to the component, reducing power drawn when idle, and to allow for differences in voltages, as required.

The VRM can affect overclocking to a very high degree, by determining how much power can

be delivered to the components (usually the CPU when referring to manual overclocking).
For very high end motherboards designed for overclocking, the VRM becomes a very important component, and to many overclockers the CPUs VRM is *the* most important part of the motherboard.

Non-overclockers will likely not need to worry about their VRMs or their capabilities at all.
Despite this, a good VRM can still be very beneficial to them.

The main and most important VRM on a motherboard is V_{core}, which powers the CPU core.

10.2 Phases

The component will be powered by all the VRM phases at the same time, with the load distributed over all the VRMs phases simultaneously.
Thus having more phases will mean the VRM is more efficient, thus wasting less power and producing less heat, but also delivering a more clean and more stable output voltage, which helps the CPU and can help give better overclocks.
In addition, more phases will spread out the power delivery, resulting in less thermal density of the components, which can help with cooling them.
Since they are hardware components, and take up space on the board, more phases will cost more, however can be very beneficial, especially to overclockers.

VRM phases are also often quoted as “X+Y", for example “4+1", “8+2", etc.
The first number represents the number of phases the CPU power goes through (i.e. V_{core}), and the second number represents the number of other phases that are dedicated to non-CPU components, such as RAM, or the iGPU.

Simply having more phases though is not always better, as the quality of the VRM and the components used can be very important as well, and it is entirely possible for a high-phase VRM to be dramatically worse than a low-phase one, although as a general rule more phases will typically be better.
It has become common lately for motherboard manufacturer’s to quote the VRM phases, trying to instil the *“More VRM phases is better"* notion, however the quality of the VRM itself is also very important, so whilst this can be true in theory, it is often false in practice.

10.3 How it works

10.3.1 TL:DR

The VRM consists of a PWM control chip, which connects to all of the VRM phases; each VRM phase consists of a driver(which determines the operating parameters of the phase - i.e. when the MOSFETs open/close), a High-side MOSFET (Which acts as a switch to open the 12V supply to the rest of the system), a Low-side MOSFET (which prevents flyback - current spikes from the inductor when its supply voltage is cut off), an inductor (stores energy from the 12V supply to slowly output it to the component), and a capacitor (which stores excess energy should the components voltage demands drop.)

10.3.2 A more detailed explanation

The VRM phases drivers tell the high-side MOSFET to close.
The High-side MOSFET closes, allowing the 12V input to charge the inductor, which stores that energy as a magnetic field, the charging of which is what causes the voltage drop in the system from 12V input to the output voltage.
When this magnetic field in the inductor reaches the suitable output voltage, the high-side MOSFET opens, cutting off the 12V supply to the inductor, which causes the magnetic field in the inductor to collapse and as such it begins to discharge current immediately.
During the small amount of time that the high-side MOSFET is open, but before the low-side MOSFET is closed, the voltage from the inductor caused by its magnetic field collapsing discharges through the flyback diode, which prevents the current from damaging the component.
The low-side MOSFET then closes, which allows current to flow through it, effectively acting as a wire, which dramatically increases the efficiency of the system compared to using the flyback diode (since diodes are inefficient).
The inductor then slowly discharges into the component, which is where it gets its power from.
When the inductors voltage output drops sufficiently to be insufficient for the component, the high-side MOSFET will open again, repeating the cycle.

(Note that despite that last line, the VRM will generally always have exactly one high-side MOSFET open to charge an inductor at any given time, thus allowing voltage to be continually topped up to help make a more suitable output voltage.)

10.4 Other terminology and corollaries

Ripple

The voltage supplied to the component will slowly fall as the inductor drains into it, resulting in a changing (i.e. decreasing) voltage from the inductor.
The inductors voltage will start high, then fall as it drains, until it is recharged from the 12V, at which point its voltage will rapidly rise again.
Thus the output of the inductor will look like a saw-tooth wave, with rapid spikes in output, then a slow drain until the next spike.
This means that the component will receive varying voltage from an inductor - the variation in this output voltage is called the *output ripple*.

Multiple Phases

The above ripple is why systems use multiple phases - so they can have phases draining current into the component to power it, whilst charging another to the correct voltage.
Systems with multiple phases will therefore overlap their [saw-tooth] outputs with an offset, which means that the phases will help keep closer to the ideal voltage of the component at at any time as the voltage is closer to the peak output of one of the phases, and is the main reason why more phases can be better.
More phases also means that since the power delivered to the component is spread out over more components, each component can take less current, which can mean components with other beneficial properties can be chosen instead of simply smaller quantities of ones that can take high current.

For systems with multiple phases, the above cycle happens for every phase, with the output of each inductor discharging into the component simultaneously.
However there will only ever be exactly one phase with its high-side MOSFET open at any time, with all other high-side MOSFETs closed.
This means that low-side MOSFETs will be closed for considerably more time than high-side MOSFETs.
This also means that the high-side MOSFET and low-side MOSFET should always be in different states - when one is open the other should be closed, and vice versa.

Varying CPU power and VRM workload

The voltage the CPU requires can very rapidly change numerous times per second as its workload changes.
Thus the VRM will have to change its output frequently to keep up with the CPUs power

demands.
If the voltage requirement increases, then the VRM will simply charge an inductor to the correct voltage, thus supplying the correct voltage.
If the voltage requirement decreases, then the excess current will instead be taken by the phases capacitor, which prevents supplying too much current to the component.

Switching frequency and duty cycle

The VRM controller works via a negative feedback loop, monitoring voltages, currents, and temperatures, and uses the incoming voltage requests to change its PWM output signal to move towards the requests.
The PWM output carries the *switching frequency* and the *duty cycle*.

The duty cycle is $\frac{V_{out}}{V_{in}}$.
So for example to drive a 1.2V output on a 12V input would give a duty cycle of $\frac{1.2}{12} = 0.1$, or 10%.
The duty cycle represents how long the 12V supply should be connected to the inductor for, and thus what the output voltage from the inductor should be.

The switching frequency is the frequency at which the VRM switches on/off, or pulses.
For example a 300KHz switching frequency will have its phases pulse 300,000 times/second.
Higher switching frequency will mean that current will move more quickly through the VRM, and helps with transient response and decreases output ripple, but can also decreases efficiency and can increase VRM temperatures more.

10.5 VRM components

The VRM is thus made up of components are as follows:

- **The PWM Controller**
  The PWM controller/controller is what outputs the PWM signal (duty cycle and switching frequency) that the drivers work with, and the PWM signal will be output to all drivers.
  Note that whilst all of the other VRM components in the VRM will be within close proximity to each other around the component in question, the PWM control chip often isn't, and can be located elsewhere on the board.
- **Driver**
  The Driver is what determines what parameters the rest of the VRM phase it controls should function according to (i.e. determine the MOSFET on/off parameters), and

thus determines the end output voltage of the system.
Divers are used because the PWM signal itself is not sufficient for actually switching the MOSFETs.

- **MOSFET**
  A MOSFET is basically a switch, that turns on when given a voltage, and is off otherwise.
  Current can flow into a MOSFET from its input, and drain out of it to its output, controlled via voltage by an on/off control signal that tells the MOSFET when to open/close.
  Thus when the voltage is given to the MOSFET the gate closes and the voltage can flow from the source to the drain.
  The MOSFET disconnects the source and drain when the voltage given to it stops, and the MOSFET opens as a result.
  MOSFETS are often just called *"FETs"*.
  MOSFETs are often covered in heatsinks for good cooling.

- **High-Side MOSFET**
  Takes input from the 12V, and outputs to the inductor and low-side MOSFET.
  The High-side MOSFET is the switch that allows the connection between the 12V supply from the PSU and the inductor.
  If a VRM phase is on then the high-side MOSFET is also on (i.e. closed).
  Only one High-side VRM MOSFET (switch) will be closed at any given time.

- **Low-Side MOSFET**
  Takes input from the high-side MOSFET and inductor, and outputs back to the inductor.
  The low-side MOSFET closed when the high-side MOSFET opens, and acts as a wire that connects the inductor back to itself, which is much more efficient that going through the diode, which boosts efficiency massively.

- **Driver MOSFET**
  A driver MOSFET is where the driver, high-side, and low-side MOSFETs are all contained on a single chip.
  This can decrease the surface area required for the components substantially, and has thermal and power advantages.

- **Flyback Diode**
  Diodes allow current to only flow in one direction, and tend to be very inefficient.
  The flyback diode is mounted in parallel with the Low-Side MOSFET, and exists to prevent the massive voltage spike that occurs when the inductor stops receiving its 12V supply because there is no closed circuit, which would destroy the component.
  The flyback diode exists to allow current to flow through the VRM when both the low and high-side MOSFET are open, are are there to prevent damage to the component when this is the case by always allowing a closed circuit output to the inductor.

- **Inductor/Choke**
  Energy storage/filtering component.

Stores energy in a magnetic field, with the inductance value of the inductor telling you how big of a magnetic field the inductor can hold.
Inductors are used to limit how quickly the voltage can change, which prevents the VRMs output from instantly spiking to 12V, which would instantly destroy the component.
Inductors charge much faster than they discharge.
The charging and discharging of the inductor is what causes the voltage drop that concerts the 12V into the [lower] usable voltage supplied to the component.
Chokes are a type of inductor that are used for filtering.

- **Capacitor**
  Energy storage component, and helps prevent against voltage spikes.
  Capacitors store energy in an electric field.
  Capacitors take the output energy from the inductor.
  Capacitors are used to store excess current from the inductors in case the output demand of the component changes and the inductors are suddenly producing too much current - they capacitor can accept that excess current and stop it from going into the component, which might otherwise cause damage.
  Solid state capacitors have wider temperature tolerances and a longer lifespan.

10.6 Doublers

As of the time I write this, there are currently no VRM controllers available that can natively handle more than 10 phases, so any motherboard with more than 10 phases will have several phases controlled by a single output of the PWM controller.

This also leads to the term *doubler*. A doubler is a chip that takes the PWM signal from the VRM controller and splits it between two VRMs.
This can allow more phases to be used than what the VRM controller can support, which can have better output as a result, but does halve the switching frequency of the VRM.
Note that *all else being equal*, a system that had an equal amount of native phases instead of doubled ones would perform better - doubled VRM phases will never be quite as good as native ones, however for many uses they are fine.

VRMs can also have other splitting, such as the very rare quadrupler.
Quadrupling the PWM signal can split the power load over even more VRMs for better (lower) thermal density, however this is generally not needed, and is usually done to save money.
Using the term “VRM phase" for phases that have been split after a doubler is questionable, but is not overly frowned upon by the community, however many people would not consider phases split further, such as with a quadrupler, to be true phases, and they are not usually advertised as such as a result.

Chapter 11

Mice and Keyboards

11.1 USB and PS/2

Modern computer mice and keyboards will usually connect to the PC via USB, however PS/2 ports are often still available for most modern computers, and whilst few manufacturers make them, PS/2 mice and keyboards are still available.
USB has numerous advantages over PS/2, however PS/2 has several as well, in addition to not using a USB port that many people often run out of on modern PCs.
For most users, either will be fine.

Note that whilst USB to PS/2 and PS/2 to USB converters exist, these may not be guaranteed to work if the device doesn't have a controller for the new interface.

The main differences would be as follows:

- **Availability**
  USB devices are ubiquitous, and almost all mice/keyboards that are sold will come with a USB connection these days, with PS/2 being considerably rarer, and often only found of high-end keyboards.
- **Port Availability**
  Whilst all modern motherboards will come with USB slots and at least one PS/2 socket, many other devices such as laptops will not come with PS/2 inputs, so USB must be used.
- **Bus interface**
  USB mice/keyboards work over USB, of course, but USB is a general purpose interface, and many different types of device use USB. As such, a USB mouse/keyboard must identify itself as such before the computer knows what to do with it.
  This also means that anything that kills, hangs, or interferes with USB bus can interfere with all mice/keyboards attached, something that becomes more likely the more

devices and more traffic there is on the USB bus. If the USB bus hangs then all USB keyboards/mice will hang too as a result.
PS/2 has its own port and does not go through a shared bus, so the CPU will receive all inputs directly and without delay.

- **Hot swappable**
USB mice/keyboards are hot swappable, and can be plugged and unplugged at will, with the PC recognising them each time.
PS/2 is not designed to be electrically hot swappable, and unplugging a PS/2 device will usually stop it from working until a reboot, and whilst some can work after being plugged back it, this is not a guarantee.

- **Drivers**
USB drivers must load before the device can be used, and as such many BIOSs will not work or have trouble with USB devices.
PS/2 has its own dedicated port, which means the system will be able to identify it correctly before the BIOS even loads, and they should always work.

- **Input variety**
USB keyboards can send a wide variety of keystroke commands to the computer, whilst PS/2 is much more limited in what commands in can send.
This does not influence most users though, since both provide all the basic keystrokes, however those who like their hardware macro keys will likely want USB.

- **Polling and Interrupts**
USB mice/keyboards work via the same principles as other USB devices, as the CPU processes their input with the standard priority as other devices, meaning they can lag if the system is busy.
This is because USB devices have to be *polled*, in which the system has to ask what the device is doing continuously, and the input of the device is the answer to the systems polls.
PS/2 instead send hardware interrupts to the CPU only when it actually wants to send keystrokes, forcing the CPU to work on the PS/2 inputs before other processes, ensuring inputs are always recognised.

- **Key-rollover**
A keyboard can only send a maximum amount of keystrokes at a time to the system, if a keyboard tries to send too many keystrokes simultaneously, then only the first ones pushed will register. The number that will register is known as the keyboards *Key Rollover* (or KRO).
Most modern USB keyboards have 6KRO, and thus can have 6 keys pressed and registered simultaneously and any additional keys not registering.
Many high end keyboards will often have 13 or more key rollover, however USB KRO is never infinite due to how USB keyboards function.
PS/2 instead has n-key-rollover (nKRO), and can send a theoretically infinite amount of commands.

11.2 Keyboard layouts and terms

Keyboards typically have the standard array of keys that most users are aware of, including the alphabetical keys with number keys above and the function keys (CTRL, ALT, Space Bar, etc) surrounding them.
Spaced above these keys will the the Escape and function keys - usually F1-F12.
Adjacent to these alphanumerical keys, a little to the right of them will be the directional pad, and function keys above (usually, Home, END, Pg. Up, Pg. Down, and Delete and Insert.
Above the Insert/Delete group will often be three more keys of Print Screen, Scroll Lock, and Pause, which are in line with the Escape/Function keys.
Adjacent to the number pad and above keys, a little to the right of them will be the number pad, which is at minimum 10-keys, but is usually surrounded by other keys; the four mathematical operators, Enter, and Num-Lock, and a Dot/Delete key.

This amounts to 104/105/108 total keys (ANSI/ISO/JIS, or US/EU/Japan, respectively).
However many different keyboard layouts that exist that do not have all of these terms:

- **Full-Sized** keyboards will have the above keys in the above positions.
- **1800-Compact** will have the above keys, but will not have any space between the alphanumerical keys, number pad, and number pad, with the directional pad usually squished below the enter key and the 0-key of the number pad reduced to standard size.
  1800-compact is not a common size.
- **10Keyless** (a.k.a. *TKL*) Keyboards will have the layout as a full-sized keyboard, but omitting the number pad.
  TKL keyboards are also called *87%*, or *80%* keyboards, as they will have 87/88keys, which is about 80% of the normal amount.
  TKL designs are popular as they allow a more compact layout whilst retaining all basic functionality. They also allow the mouse to be held closer to the keyboard, potentially allowing more parallel arms and more comfortable arm and hand positions whilst typing.
- **75%** Keyboards will have a 10TKL layout, but no space between any keys, thus being around 75% the length of a full-sized keyboard.
  Note that some 75% keyboards remove a few keys to fit into this layout size.
  75% is not a common size.
- **60%** Keyboards have just the alphanumerical keys, missing the directional pad/function keys, number pad, and the function row, but will preserve the Escape key as their top-left key.
  60% keyboards often have a Fn (Function) key on their bottom right that allows the missing keys to be accesses via Fn key-combos.
  60% keyboards are relatively popular for high-end mechanical keyboards.

- **Tenkey** keyboards are just the number pad on their own.
  These are relatively rare, however can be very useful for data-entry, for example for economists, bankers, or scientists who do a lot of numerical data-entry work.
  They can add the functionality of the number pad to other designs, whilst being easy to move out of the way when needed.

Other sizes do exist, such as *40%*, however they are all so uncommon I don't want to add clutter and confusion by listing them here.

Ortholinear keyboards have their alphanumerical keys aligned in a grid, rather than staggered as is normally the case.

11.3 Mouse DPI

The function of a computer mouse is to accurately determine the motion of the mouse relative to the surface it is located on top of, and send that information to the computer.
Modern mice (i.e. optical and laser) have a laser and a camera that picks up differences in the surface below the mouse, translating that into relative motion information.

DPI is a measurement of the virtual pixels the mouse can register over an area of physical space.
Higher DPI therefore allows more virtual space to be covered on the PC by the mouse for a fixed area of physical surface that the mouse covers.
A DPI value that is too low can therefore mean that the users mouse will feel sluggish and doesn't move as quickly as they would like, whereas a DPI value that is too high will result in the mouse pointer on the computer moving more than the user would like for the amount the mouse was physically moved.
Modern mice always have some level of customization of their DPI, either via the mouse's software, firmware, or by the OS itself.
This allows mice to have their DPI levels adjusted according to the users preference.

Because every person is different in how they hold and operate their mouse, as well as in the output properties they expect from their mouse, there is no general one-size-fits-all "Best DPI" value, as this value will vary between users, often considerably.
As a result, higher DPI values are not necessarily better, however a DPI value that is insufficient for a user is bad.

There are many mice features that can improve the quality and performance of the mice, including higher quality switches, more accurate sensors, and better ergonomics, however provided the mouse can operate at the DPI the user wants, then having a mouse with a DPI that is simply higher is not actually any better.

11.4 Mouse Polling Rate

USB mice work by having the computer poll them - the computer repeatedly asks what the mouse is doing at fixed intervals, and the input of the mouse to the PC is what the mouse responds to those polls.
This means that polling rate is a frequency, and is measured in Hz (Hertz).

Thus having a higher polling rate can allow the mouse less delay in any outputs or changes in output that it sends to the PC, and can thus be beneficial, however in practice there is a limit to how high of a polling rate is possible before the difference becomes functionally indistinguishable to the user.

Most modern mice typically have a polling rate of between 125 and 1,000Hz.

Note that PS/2 mice do not have polling rates, this is because PS/2 mice instantly sends all data to the PC, instead of waiting for a poll to which it can respond, and only USB mice will therefore have polling rates.
This also means PS/2 mice are theoretically faster than USB mice, however in practice USB mice polling rates are often high enough that this difference is not only negligible but indistinguishable.

11.5 Mouse acceleration

Computer mice translate the physical distance the mouse moves over a surface into a discrete figure the computer can use.
Mouse acceleration changes how this figure is created.
With no mouse acceleration, the mouse will record all distance measurements based only upon the distance the mouse travelled.
This means that the speed at which the distance is covered is irrelevant - moving the mouse 10cm will register the same output from the mouse if it is done over 1 second or 10 seconds.
With mouse acceleration on, the output of the mouse changes based upon how quickly the distance is covered, such that moving the mouse over a distance quicker will result in more movement output from the mouse.
Mouse acceleration can usually be turned off.

Chapter 12

Misc terms and components

12.1 Overclocking

Overclocking is the process of increasing the performance of a component by modifying its operating parameters.
The name overclocking refers to the usual method that this is achieved: increasing the clock frequency, however other parameters will usually need to be changed too as well, such as voltage.
Overclocking has historically been a long and complex process, however for modern PCs, this is sometimes done automatically, such as a CPUs turbo frequency. However a manual overclocking process can achieve higher performance, but will take longer, often considerably so.

Overclocking has the advantage of increasing the performance of the components, and thus the system as a whole.
The disadvantages are dramatically increased power draw and heat production, as well as the high time investment required to set up the overclock. Overclocking can also cause system instability if not done correctly.

Overclocking is usually performed on a systems CPU, however GPU overclocking is also relatively common. Other components can also be overclocked, however, such as RAM, and this is often desirable to prevent limiting the other components for very high-end overclocks.
Since overclocking is usually performed on the CPU, please see the “Overclocking" section of the “CPU" section for more information on overclocking.
Overclocking is usually performed by hardcore gamers, and is very much a niche thing, with only a very very *very* small percentage of people doing it, and it is really never performed by the average PC user.
Overclocking is mostly used for increasing the performance of older systems to help keep up with modern demands, or to max out the capability of the system.

12.2 Water cooling/liquid cooling

Computers generate heat when powered on, and require that heat dissipated in order to remain cool enough to stay functional and undamaged.
How this typically works is by having heatsinks on the critical components (usually just the CPU and GPU, but potentially others for extreme cases), and having air blasted over those heatsinks.
The heatsinks thus wick heat away from the components, and that heat is dissipated from the heatsink to the air being moved over it.
Watercooling systems instead use liquid as the medium for wicking away heat from the systems components.

With watercooling, instead of heatpipes and heatsinks attached to the component that is to be cooled, instead there will be a *Cold Plate*, which will attach directly to it (with thermal compound between the two, as usual).
Heat will transfer from the hot component into the cold plate.
Behind this cold plate there will instead be moving liquid, which will take the heat from the cold plate.
The liquid in the system will be pushed around the water cooling loop in a fixed direction by the *pump* until it reaches the *radiator*.
The radiator is designed to have the liquid move around a very large surface area, and will have fans attached to it to keep lots of air moving over it to take heat, thus cooling the liquid inside.
The system will therefore need a pump to move the liquid around the system, and any water cooling system that does not have a pump will instead have the liquid remain static in the tubes, leading to very rapid overheating of the components, which will result in a system shutdown and potentially damage to them.
Water cooling systems also typically feature a *reservoir* - this does as expected, and simply holds water in the water cooling system, increasing the volume in the system.
Many water cooling systems do not need a reservoir, however it is recommended and is best practise. Whether one is needed for any individual system will be dependant on that systems cooling ability, liquid volume, and so on, and is a very case-by-case basis thing. However, if you are unsure if you need a reservoir, then it is best to have one.

Note that despite the name, the liquid used in a water cooling system does not have to be water - any liquid that is good at absorbing heat will work, and many water cooling systems sold to consumers use propylene glycol, water, or some mix thereof.
Even in systems that do not use water though, “Water Cooling" is still a common and popular term to mean any form of liquid cooling, and “Water cooling" and “Watercooling" are used interchangeably.

The fluid used is often referred to as the *coolant* of the system.
Note that water can work just fine as a coolant, however deionized water is preferable to ordinary [e.g. tap] water.

Water cooling systems will use a variety of metals in their construction, typically Copper, Aluminium, **or** Stainless Steel.
Please note the *or* above in bold. Water cooling systems should *never* use different materials in their use.
Using different types of metal in a loop is referred to as a *mixed metal* system.
Mixed metal systems are bad, as it will have galvanic corrosion present in it.
The liquid coolant will provide an electrical connection between the two metals, even if they are not physically touching, and will corrode the surface of the more active metal.
This is because the coolant effectively acts as an electrolyte, and the two metals will become an anode and a cathode.
The more stable metal will be the cathode, and the more active one the anode, with the anode being the one damaged and dissolved into the electrolyte.
(This is the same system as to how batteries generate electricity.)
This corrosion will occur faster the larger the surface area of the metals is to the electrolyte.
Metals that are near to each other in their average voltage differences will corrode slower, and mixed metal systems can work, but definitely require a lot more care, attention, and maintenance.
This simple answer to mixed metals is that *“You can, but you shouldn’t."*

12.3 AIO

An AIO is an *All In One* cooling system.
This is a term used to describe a completely self-contained cooling plate, pump, reservoir, fans to cool said reservoir, and tubes to connect it all together.
AIOs are aimed at people who want the benefits of watercooling,but without the hassle of setting up a custom loop for their system, which can be extremely time consuming and tedious, as well as potentially causing leaking if not done properly, whereas an AIO is quick, simple, and should not leak at all.
AIOs are therefore a popular choice for people who want watercooling systems in their PCs.

Do note however, that AIOs are more limited in how they work, as they used fixed radiator sizes with fixes fan mounting holes, in addition to fixed-length tubing, which mean they are more limited in how much they can cool, as well as where they can be mounted in a computers case.
This means that a proper custom water loop can often outperform an AIO, however the time, cost, and hassle of setting this up makes it undesirable, and all but extreme over-clockers are unlikely to need such cooling anyway - most non-overclockers are fine with the

stock air cooler, after all, and a good air cooler is often fine for moderate overclocking as well.

12.4 Heatpipe

Computer coolers, as found on CPUs and GPUs will usually feature *heatpipes*.
These are almost always made of Copper, but are not solid Copper.
Instead, heatpipes have a wave structure and use a liquid inside of them.
How heatpipes work is by having the liquid inside the heatpipe vaporized by the heat of the component, which then travels along the heatpipe until it reaches the cooler end, where it condenses back into a liquid, and then returns to the hot end by capillary action along the weave inside the heatpipe.
Heapipes can be extremely good conductors of heat, and can outperform even solid Copper in thermal conductivity by an order of magnitude.

Note that heatpipes are always closed-loop systems, and should never be opened. A heatpipe that has been opened will usually not function as a heatpipe any more, even if it is resealed, due to loss of the liquid inside, as well as pressure changes and potentially damage to the capillary weave.
Computer heatpipes generally can't be repaired if they are opened or damaged, and it would likely not be economical to do so anyway.

12.5 Thermal Paste

Thermal paste is what is used between a component and the heatsink that is trying to wick away heat from that component.
Components and heatsinks will basically always (at least, for consumer PC hardware, at least) have flat surfaces that make contact with each other.
Due to machining tolerances, or imperfect machining or flattening they will not be exactly flat, and might be slightly curved, or have surface imperfections, even if the user can't visibly see them.
Without thermal paste, these imperfections may mean that the heatsink does not make contact with the component at all, or may have numerous gaps between the two surfaces, which will be filled with air.
Air, however, is a *truly awful* conductor of heat.
Hence, thermal paste: thermal paste provides a high-conductivity medium between the surfaces, which will ensure that there is always a method of transfer between the two surfaces, as well as preventing air gaps, instead filling any gaps with a high-transfer material.

Thermal paste can age, and many will become hard/brittle, or decrease in thermal conductivity over time, and as such replacing old thermal paste can be a good way of decreasing temperatures of older components.

Most consumer thermal pastes are not electrically conductive, however some high end ones can be.
The surfaces that thermal paste is applied to will never be electrically conductive, so this is not usually an issue, however if the electrically-conductive paste gets onto surrounding components, or the circuit board they are atop, then it could cause serious problems.
Note: All my advice regarding thermal compound application in this section will be aimed towards non-conductive thermal compounds, as these are the most common, easiest to use, and represent the vast majority of both what is used and what is available. All my advice given in this section also goes for electrically-conductive pastes, but it is critical to ensure that none spills onto other components, circuit boards, or ... anything else at all, ideally.

Many components do not produce heat uniformly inside of them.
For example: A CPU with an iGPU may have one side of the CPU get very hot if the CPU is working hard, however the other side may be relatively cool if the GPU is not being stressed, resulting in uneven heat generation.
CPUs do have an Integrated Heat Spreader - always made of solid Copper, that help to dissipate this evenly to the CPU cooler, however basically no other computer components have this, and as such having the heatsink completely cover the component, as well as having thermal paste between the two surfaces is vital for keeping the component cool, and this advice goes for CPUs as well, ideally, however for most users is less critical than with other components.
A component that does not make contact with the hot area of a component - as may be the case if the thermal paste does not get applied there, and slight curving of the component or heatsinks surface mean there is a slight air gap - may fail to provide cooling, regardless of how good the heatsink or thermal compound is.

The question of *"How much thermal paste should you use?* is therefore relatively simple to answer: *"Not too little."*
You need enough thermal paste to *completely* cover the entirety of the surfaces that the paste is to go between. That's it, really.
Any less and you may have air gaps that may cause the component to overheat in certain areas, even if other areas are kept cool.

"Too much thermal paste" is really not a problem. Excess thermal paste will be forced out of the space between the components by them being mounted together, and will simply spread onto nearby components, which, for electrically-non-conductive thermal pastes, is really not an issue, however it can be a hassle to remove, as many components on a PC are easily damaged.
The only problem that can arise from having too much thermal paste is that extra pressure

must be applied to push the excess thermal compound out from between the surfaces that are to be cooled, and in extreme cases this pressure can damage the components, however this is extremely unlikely in practise, and easily avoided by either a modicum of care or just not applying too much thermal paste such that lots needs to be pushed out from between the components.
For electrically conductive thermal paste, however, any that spills out onto surrounding surfaces can cause damage, and more care must be taken with such compounds.

Many people recommend a pea-sized amount in the centre. Others recommend an “X" shape. Others recommend getting a spreading tool and spreading out the paste across the surface for guaranteed coverage of the entire surface.
All of these should work fine, and any method of applying thermal paste will generally work fine, provided it covers the entirety of the surface being cooled.
The easy way to test how much of the surface it covers is to take the heatsink back off again after being applied and look at where the paste spread out to - if it didn't cover the entire surface, then you need to apply more paste, or make sure the paste that is there gets spread out more.
The TL:DR for thermal paste is likely to be: “Apply a small amount, and applying too little or the paste not covering the entire surface is a problem, but too much is generally not, and any method of application is likely going to be *mostly* fine, and any method that covers the entire surface in thermal paste is *completely* fine."

12.6 USB

The *Universal Serial Bus* was introduced to standardise connecting peripherals to a PC, and is the most commonly found type of peripheral port on modern computers, with many laptops only having USB ports for peripherals.
USB is used as a general purpose connection, often being used for keyboards, mice, webcams, storage devices, and more.
USB has proven popular with consumers due to its simplicity, reliability, and backwards compatibility, and is now one of the most common connectors found on PCs.

USB has had several versions over the years, as well as several connector types.
Newer USB standards will be able to transfer more data, and provide more power to connected devices.
All USB versions are backwards compatible with older versions, however the additional power delivery and data bandwidth will not be able to be utilized by older cables or devices.

USB devices should connect to the PC automatically, requiring no setup by the user beyond a simple driver installation if necessary, making USB very quick, easy, and simple to use for the end user.

USB cables and connectors that are USB 3.0 or later will have blue inserts in the cable connector and a superspeed logo, making them easily visually distinguishable.
There are several USB connector types available, including USB-A (the “standard" type), type B, Mini (-*A* and -*B*), Micro (-*A* and -*B*), Micro-B Superspeed, and USB-C.

The USB-C standard is symmetric, meaning it can be plugged in in either orientation, whereas all other USB connectors can only be plugged in in one orientation.

12.6.1 USB Data rates

Connector	a.k.a.	Bandwidth
USB 1.X	Low Speed	1.5Mb/s
USB 1.X	Full Speed	12Mb/s
USB 2.0		480Mb/s
USB 3.0	*Superspeed*, USB 3.1 Gen 1, USB Gen 3.2 Gen 1x1	5Gb/s
USB 3.1	*Superspeed+*, USB 3.1 Gen 2	10Gb/s
USB 3.2 Gen 1x2		10Gb/s
USB 3.2 Gen 2x2		20Gb/s

The difference between USB 3.1 Gen 2 and USB 3.2 Gen 1x2 is the encoding and connectors that it can use.
USB Gen 3.2 Gen 1x2 and 2x2 both use 128/132b encoding, whereas all prior USB standards use 8/10b encoding.
USB 3.2 Gen 2x2 can also only use a USB type-C connector, and the as-yet unreleased USB4 standard is expected to only use USB-C connectors as well.

12.7 Header

A *Header* is what the plug at the end of a cable is plugged into, and is therefore the receiving end of connections.
Headers are usually referred to in the context of motherboard headers, as used for say front panel I/O (including the power/reset buttons, but also potentially things like USB or audio jacks), and fan headers, which are of course used for connecting fans.

Headers are always standard sizes and dimensions.

12.8 Standard mounting screws

PC cases tend to use standard sizes for their screws, as detailed below:

Note that imperial screw threads are defined by their thread diameter and pitch rate (threads per inch), in the format; *<Thread Diameter>-<TPI>*
So for example a 6-32 screw will be a #6 diameter, with 32 threads per inch.
The UNF and UNC specify “fine" or “course" pitch, and thus determine the pitch rate from the diameter, and this can actually be redundant, since the pitch is usually given.
As such, a 6-32 is the same as a #6 UNC, which is the same as a 6-32 UNC.

Metric screws are defined differently, and are usually referred to as *M<Diameter>x<Pitch>* screws, where the “M" represents the outer diameter of the screw in millimetres, so for example an M3 screw has a 3mm outer diameter.
Note that as per the M3 example above, the pitch rate is sometimes omitted, in which case it uses the “course" pitch rate. If the pitch is not course then it will be explicitly stated.

- **6-32 UNC**
  6-32 screws are by far the most common type used in computers, and are used for: Securing the power supply to the case, mounting 3.5" hard drives, mounting expansion card backplates to the case, as well as other tasks.
  6-32 screws used in PC cases are almost always Phillips-head, usually requiring a #2 screwdriver bit.
  It is important to note that these 6-32 screws do vary in their length, with case-screws (1/4") typically being longer than hard-drive screws (3/16" in length), however other lengths are used.
  Hard drives are not usually designed to have long screws screwed into them, as the screws can potentially damage the internal components of the hard drive if the screw is too long, so always use the shorter screws for hard drives.
- **M3**
  M3 screws are usually used for mounting 5.25" disc drives, 2.5" hard drives and SSDs, and 3.5" floppy drives.
  M3 screws used in PC cases are almost always Phillips-head, usually requiring a #1 screwdriver bit.
  M3 screws used for floppy drives and 2.5" drives are usually 4mm long M3, and those used for 5.25" are usually 5mm long M3.
- **Motherboard standoffs**
  Motherboard standoffs will be hex external and screw-thread internal.
  The external hex thread will fit a 5mm hex wrench.
  The standoff will typically mount to the PC case with a 6-32 screw thread, and internally accepts an M3 screw to secure the motherboard to the standoff, however this is not universal, and some standoffs/cases do use different standards, so always check.
  Motherboard standoffs can vary in length, however 6.5mm standoff distance is the

standard and most common.

- **KB5 Self-tapping screws**
  Self tapping screws are used to mount fans to the PC case, and this is the only place they are used in modern PCs.
  Fans use self-tapping screws such that the screw can make a thread out of the plastic fan case, meaning that the fan does not need to be tapped before being sold, instead the user doing this automatically with the screw used to secure it.
  These self-tapping screws are *KB5* specification, and are typically 10mm long.
  Self-tapping screws used in PC cases for fans are almost always Phillips-head, usually requiring a #2 screwdriver bit.

- **#4-40 UNC**
  4-40 UNC are the thumbscrew standoffs used on either side of female VGA and DVI ports, as well as many earlier D-sub ports.
  These will screw into the female port with a 14mm 4-40 screw, and accept a 5mm 4-40 screw from the cable that is to be connected to them.
  The fact that they use 4-40 for both securing them and being secured to means that they are often just referred to as “4-40".

- **Side panel thumbscrew**
  Computer side panels are usually held on by thumbscrews, of which there are typically two thumbscrews per side panel.
  These thumbscrews are typically 6-32 as well, and are usually 5mm in length, however can sometimes be longer than this.

Lastly, be aware that whilst the majority of screws used in modern computers will use the above, some cases do use different standards, so do use care, and remember that if you have to force something to fit, then it probably doesn't.