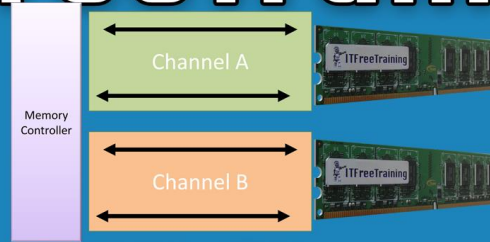


ITFreeTraining

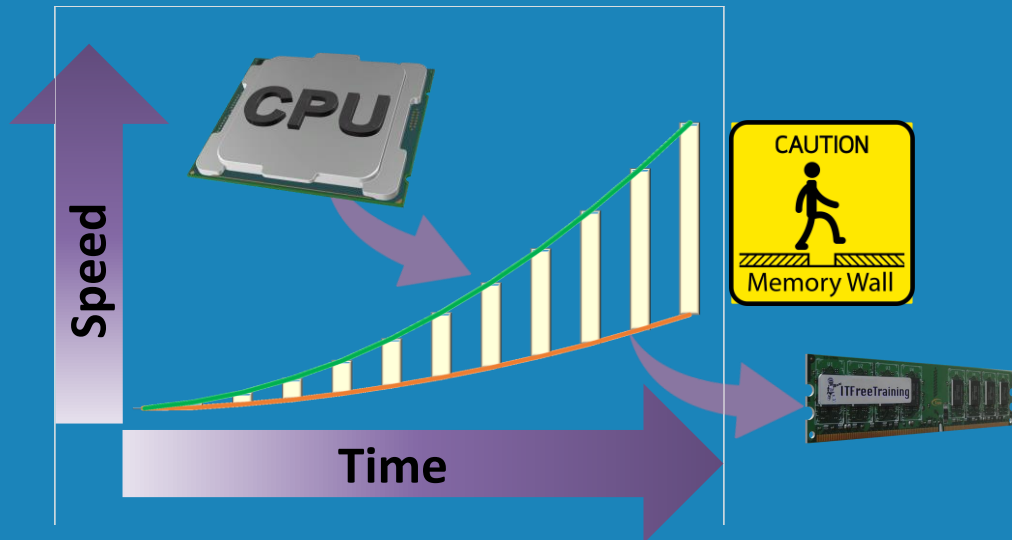


Multi-Channel Memory Architecture

For the free video please see
<http://itfreetraining.com/ap/3a25>

Welcome to the ITFreeTraining video on multi-channel memory architecture. Multiple channel is a technology that increases the amount of data that can be transferred between memory and the memory controller by grouping memory modules together. By the end of this video, you will have a better understanding of this technology, how to implement it and what performance effects you can expect to receive.

Memory Wall/Bandwidth Wall

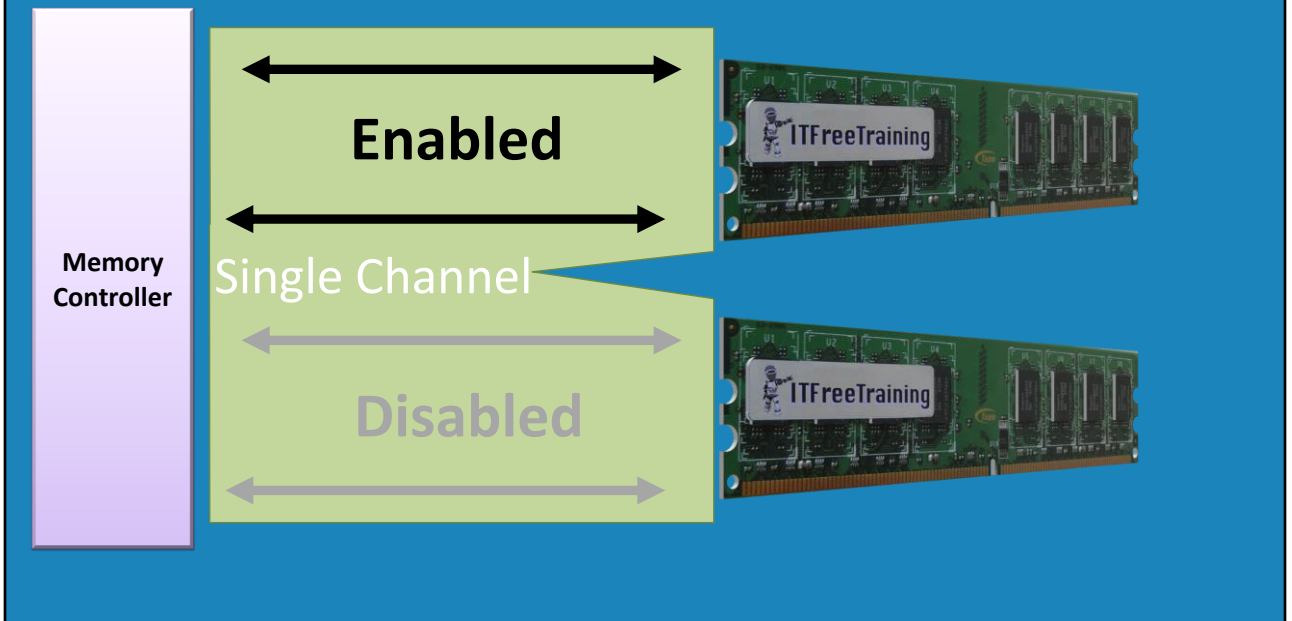


0:24 Before I look at how multi-channels work, I will first look at the memory wall (also referred to as the bandwidth wall). This will give you a better understanding of why multi-channel memory was developed.

In the early days of computing, CPU speed increased very quickly. Over some years, CPU speed doubled each year. The problem was memory speed was not increasing at the same rate as CPU speed. This created a gap between CPU performance and memory performance.

This gap in performance is referred to as the memory wall or bandwidth wall. This was more of a problem in the old days of computing as CPU speed kept increasing much faster than memory speed. However, nowadays CPU speed does not increase anywhere near as fast as in the old days, but performance can be improved by adding features like additional cores to a CPU. The slow increase of CPU speed has allowed memory speed to increase and catch up with CPU speed.

No Multi-Channel (Single)

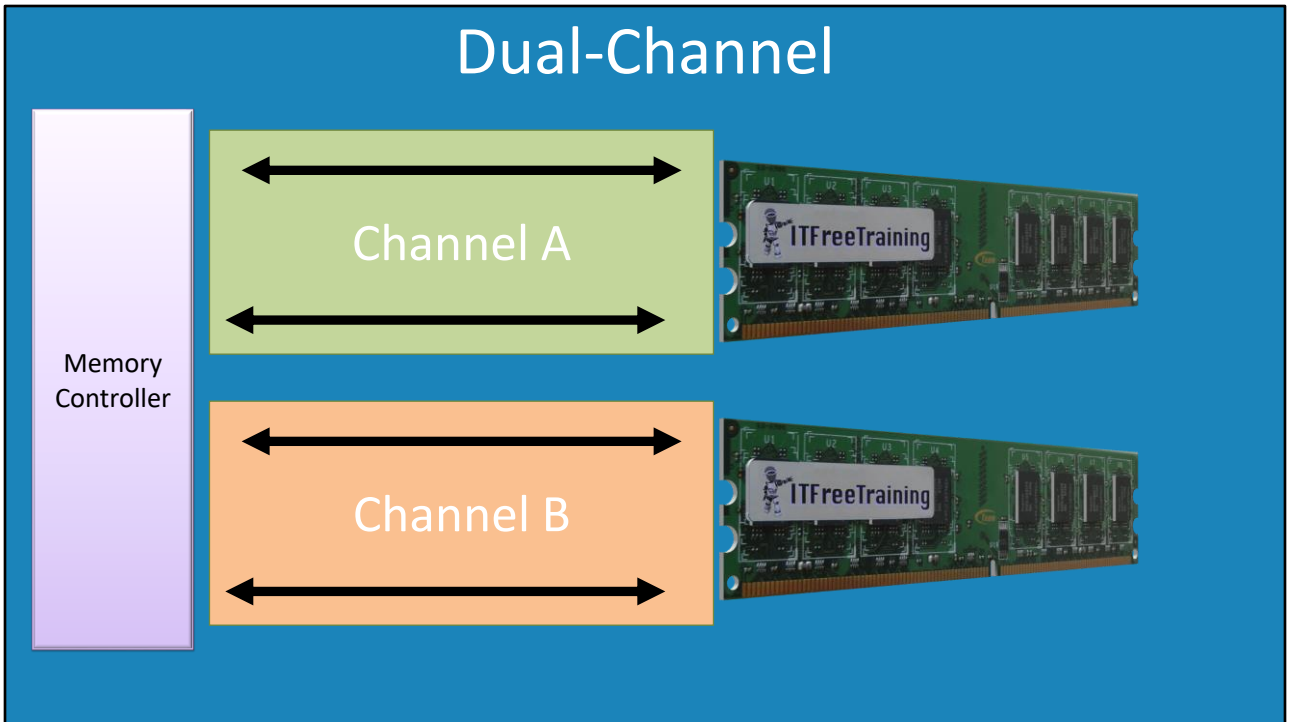


1:24 To understand how multi-channel works, I will first look at what occurs when it is not used. Consider that you have a memory controller, either inside the CPU or on its own chip. Inside the computer, there are two memory modules.

The memory controller needs to access the memory modules. With two memory modules, dual-channel would be used. Without it enabled, which is the case in this example, the memory controller shares the access path from the memory controller with the two memory modules.

This essentially means that it can only access one memory module at a time. While the first memory module is being accessed, the second memory module cannot be accessed. So what is the difference between this and dual-channel access?

Dual-Channel



2:11 When dual-channel is enabled, the memory controller is able to access both memory modules at the same time. By being able to access two memory modules at the same time, this increases the amount of data that can either be read or written to the memory modules at once.

I will look at triple and quad-channel shortly, but first I will look at the requirements needed to run multi-channel.

Requirements



- Same DIMM configuration
 - Memory size, number of chips etc. (*)
- Not required (Caution advised)
 - Brand/Timing/Speed
 - Speed determined by slowest memory module

* Some newer motherboards/CPU's allow different sizes

2:32 In order to use multi-channel, first the memory modules must have the same DIMM configuration. This essentially means that both need to be of the same size and have the same number of chips on them. Traditionally, you won't be able to mix and match, for example a 4GB memory module with an 8GB memory module. If the memory modules have a different number of chips, most likely they will operate differently. For example, how they access and transfer data will differ - so they will not work together.

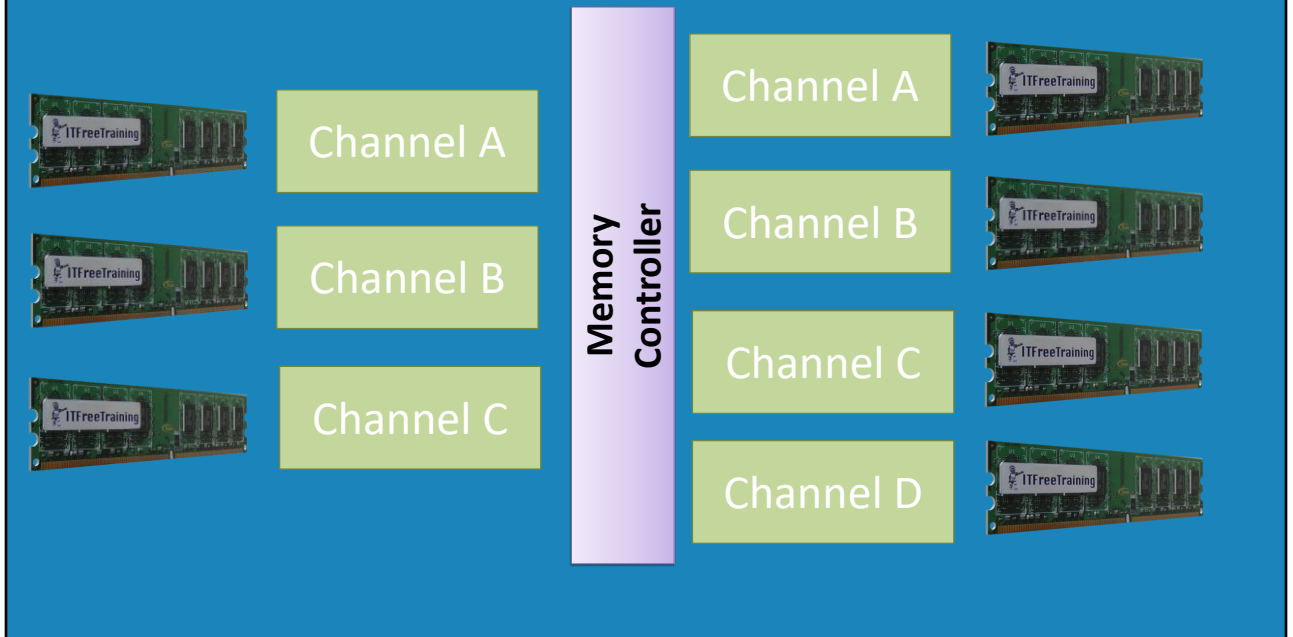
Having said all this, some of the newer motherboards and CPUs will allow differently sized memory to be configured using multi-channel. They achieve this by mapping the matching memory amount to the multi-channel. Any extra memory would be mapped to its own channel. This essentially means you would get a speed improvement for the lower memory amount that is mapped directly to another memory module and no speed increase for the memory that cannot be mapped, since there is not enough memory in the other memory module to do this.

When you have been shopping for RAM, you may have noticed memory modules sold as a kit. For example, this one shown. Manufacturers will often sell two memory modules bundled together to ensure that there are no problems working in multi-channel mode.

If you do have different memory modules, this does not necessary mean they will not work together; however I would exercise some caution if you attempt this. Generally speaking, memory modules do not have to be of the same brand, timing or speed. If you put two

memory modules in that are of different speeds, the motherboard should reduce the speed to match the slowest memory module. I do emphasize it should; if you are going to mix and match memory modules, test first - as you may experience unexpected problems.

Triple-Channel/Quad-Channel



4:28 We understand what dual-channel is, so what is triple or quad-channel. Triple-channel is essentially when the memory controller uses three memory modules at the same time. Quad-channel is when the memory controller can use four memory modules at the same time.

The same rules apply, memory modules need to be of the same memory size and specifications, however speed and brand can vary. Dual-channel is not a recent invention and its use dates back to the very early days of computing. By contrast, triple and quad-channel are relatively new technology. If you are planning to use either of these, check that your motherboard supports it. Also nowadays, since a lot of CPUs have the memory controller in the CPU, you should also check that the CPU supports it.

Quad Motherboard Example

- Memory slots are generally color coded



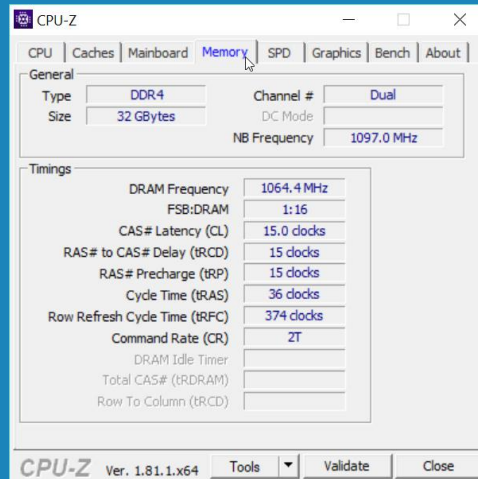
5:15 When installing memory modules, you need to ensure that the memory modules are put in the correct slots. In most motherboards nowadays the slots are color coded to make it easier. For example, on this motherboard the slots are colored gray and black.

Ensure you check the documentation that is supplied with the motherboard. Often, a motherboard requires a single memory module installed in a particular slot before it will even start up. In the case of this motherboard, the first memory module should be installed in the first gray slot below the CPU. If you want to enable dual-channel, the next memory module must be installed in the lowest gray slot. For triple-channel the next memory module will be installed in the gray slot above the CPU. Lastly, for quad-channel, the last memory module will need to be installed in the second gray slot above the CPU.

On some motherboards, if you do not have a memory module in the required slot the computer will not start. Other than this, if you don't put the memory modules in the correct slots, multi-channel will not work, but the computer may still be able to access the memory. In short, check the motherboard documentation to ensure memory is placed in the correct slots.

CPU-Z

- <https://www.cpuid.com/software/cpu-z.html>



6:27 Once you have your memory installed, you may want to check that multi-channel is working. Sometimes BIOS will tell you if multi-channel is working. CPU ID also provides a free tool called CPU-Z which can be used to check.

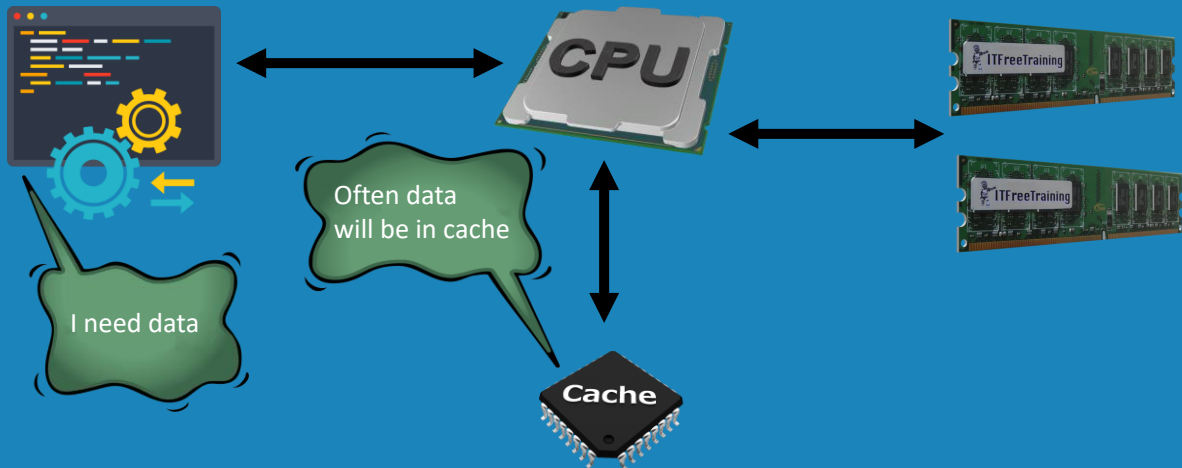
Shown here, is CPU-Z. The first tab of CPU-Z will give you a lot of information about the CPU and what features it supports. There are a lot of tabs with useful information, however the tab I am interested in is Memory.

At the top right of the tab, notice the channel is shown. In this case, the channel is shown as dual. If you are having some problems getting multi-channel to work, you may also want to look at the tab called SPD.

SPD stands for Serial Presence Detect. SPD provides information about the memory. This information includes timing information about each memory module. Selecting the slot from the pull-down will display information about that memory module. If you are having problems with matching different types of memory, this may give you an idea about what is causing the problem.

In the Real World

- Testing shows varying results



7:30 So in the real world, does configuring multi-channel make much of a difference? To understand this, consider how memory gets used by a computer system. An application running on a computer requires memory to run.

The application will request memory via the CPU. The CPU will first check its cache to see if it has the required memory in the cache. Often, the data will be in the cache. This is because the CPU will transfer a block of memory at once into the cache. When requesting memory, applications will often request memory that is near to where the last memory was requested.

Think of it like this, if you were reading a book, more than likely after reading a word in a book you would next read the following word in the sentence. You could flip through the book looking for something, but when you found it you would most likely read at least a few words or sentences before looking somewhere else. Computer programs follow this behavior. Often a number of requests will be in the same part of memory before moving to another part of memory. For this reason, most memory requests are fulfilled by the CPU getting the required memory out of its cache.

So what happens if the request cannot be fulfilled by the cache? In this case the CPU will transfer data from the memory modules into its cache. Multi-channel allows two or more memory modules to transfer data at once to and from the CPU.

This may sound like performance should be doubled, however consider that accessing the

memory modules is only a small part of the overall process. If an application is very memory intensive, it will require more memory to be transferred via memory modules to the CPU. This may also sound like multi-channel would make a big difference; however, the application may also only use memory that is spatially close together at one time. If this is the case, most application requests can be fulfilled by the cache.

Testing of multi-channel verses single-channel shows that, with some applications, there is significant improvement, whereas with others not much at all. For example, gaming showed little improvement while data-intensive graphic applications like Adobe products showed a higher improvement. Other testing found a 70% improvement in a quad-channel system, but it all depends on what software you are running.

In the real world, don't expect too much improvement in a dual-channel system and testing has shown varying results, but it won't make a system run slower. On server-based or performance workstation systems it is generally a good idea to run multi-channel. This is because these systems tend to run under high-load for long periods and even a small increase in performance of a few percent makes a difference. Consider a server that converts video all day. A performance increase of only 3% per hour in a 24-hour period amounts to just over an hour of performance improvement. On a home system running a lot of applications and not under heavy load, a small improvement like this may not even be noticeable.

The main point to remember is, you won't get double performance for running dual-channel and results will vary depending on what applications you are running.

This concludes this video on multi-channel. I hope you have found this video useful and I hope to see you in other videos from us. Until the next video, I would like to thank you for watching.

References

"Random-access memory"

"Multi-channel memory architecture" https://en.wikipedia.org/wiki/Multi-channel_memory_architecture

"Clock rate" https://en.wikipedia.org/wiki/Clock_rate

"RAM Performance Benchmark: Single-Channel vs. Dual-Channel - Does It Matter?" <https://www.gamersnexus.net/guides/1349-ram-how-dual-channel-works-vs-single-channel?showall=1>

Credits

Trainer: Austin Mason <http://ITFreeTraining.com>

Voice Talent: HP Lewis <http://hplewis.com>

Quality Assurance: Brett Batson <http://www.pbb-proofreading.uk>