What is the Real RAM speed and Should I get The Most MHz?

April 25, 2017Author: Angelos Kyritsis

Hz <u>is about \$55</u>. We can also grab 8GB RAM clocked at 3400MHz <u>for \$73</u>. Getting an extra 1267MHz for \$18 seems like a good deal, and we might be tempted to buy the higher frequency RAM. However, does the higher frequency automatically mean a faster RAM? Let's see how RAM speed works.

What does the RAM frequency mean

A common misconception is that the RAM frequency is the same metric as the CPU frequency. Since both are expressed in Megahertz (MHz), it is an understandable mistake.

In fact, the RAM's frequency in MHz has absolutely nothing to do with the CPU's MHz. And, by itself, it is far from a reliable metric for the total RAM speed.

How does the MHz frequency work at a CPU

In the good (?) old days of single-core processors, the frequency in MHz was the most important metric by far.

In a nutshell, frequency expresses how many commands can a CPU process in a single second. A 200MHz frequency means that the CPU can process up to two hundred million commands per second.

When all was said and done, the CPU with the higher frequency would have the highest total performance.

Today, things are a significantly more complicated than that. The total CPU performance, and ultimately its price, also depend on the number of cores, the number of threads, cache memory size, and other parts of the CPU architecture.

Of course, everything else being equal, the cores/threads, cache memory, and general architecture, the CPU with the most GHz will always be faster. That is why overclocking is still a thing.

How the CPU is different than RAM

The function of RAM is vastly different than the function of the CPU. RAM does not perform commands or calculate anything whatsoever.

Even though the way RAM works is a complicated engineering feat, its function is relatively straightforward. The only jobs RAM will do are:

- Accepting data from the hard drive and other subsystems
- Providing part of the data to the CPU to process
- Getting the resulting data from the CPU's calculations
- Transferring the data to the hard drive, graphics card, sound card, USB controller, etc.

These are the only tasks that RAM does. Nothing more, nothing less. One could say that it is a relatively "simple" part of the computer, albeit incredibly fast in what it's doing, at about ten times quicker than the speediest SSD disk.

To have a clearer understanding of the RAM's role, let's say the processor was a sculptor. RAM would be the workers who carry the marble (data) from the warehouse (hard drive) to the sculptor to sculpt into a statue (process the data).

No matter how fast the workers carry the marble, the sculptor will process it at his own speed.

Getting faster workers will not make the sculptor work any faster. Provided, of course, that he already has enough marble to work on and is not waiting idly. Which brings us to...

The importance of bandwidth

Taking the previous metaphor a bit further, let's say that the sculptor can sculpt up to 1,000 lbs of marble per hour.

The total weight of marble that the workers can carry in one hour is the "bandwidth". If the workers can move 2,000 lbs per hour, the sculptor will always have marble to work on, and the finished statues will always be delivered on time.

However, if the workers can only carry 1,000 lbs per hour or even 1,999 lbs per hour, then we have a problem.

The sculptor will have finished with the marble he already has had to process. He will be twiddling his thumbs, waiting for the workers to deliver more. The whole system will slow down. Even if we had a faster sculptor, it would make no difference.

Back on the PC, the memory bandwidth is the max volume of data that RAM can transfer to the CPU and that the CPU can transfer back to RAM. The frequency of RAM in MHz has relatively little to do with the RAM speed and everything to do with the bandwidth.

How the RAM frequency in MHz sets the bandwidth

The maximum bandwidth can be calculated using four factors:

• The RAM's DRAM clock frequency

This is the frequency of RAM, in MHz. One Hertz (Hz) is a single clock cycle per second. 1MHz is 1,000,000Hz.

- The number of data transfers per clock cycle In all current RAM, there are two data transfers per clock cycle. This is where the name the Dual Data Rate (DDR) originated.
- The memory bus

In every motherboard from the past decade or so, the memory bus is 64 bit.

• The number of interfaces

The interfaces express how many data channels can work at the same time. If we install one RAM module, we have a single data channel. If we install two modules in Dual Channel DDR, we have two channels.

Using two memory channels, we double the maximum bandwidth. This is why it's always better to install RAM in pairs, e.g. 2x4GB instead of 1x8GB.

Some of the more expensive, enthusiast-grade motherboards support Quad Channel DDR. This means we can install four RAM modules, e.g. 4x2GB, and quadruple the memory bandwidth.

Calculating the total bandwidth

Some factors about the bandwidth remain stable, regardless of which RAM we bought. The Memory Bus will always be 64bit. If we have the proper modules installed, Dual Channel or Quad Channel will multiply the total bandwidth x2 and x4 respectively.

Also, every single RAM currently on the market supports DDR, excluding some practically ancient SDRAM modules. So, the dual data rate adds an x2 multiplier on any DDR1, DDR2, DDR3, or DDR4 RAM.

Interesting fact: this x2 multiplier is taken into account when marketing RAM modules. Which means that a RAM Module sold as 1600MHz will have a real clock frequency of about 800MHz. This is the number apps like Speccy or CPU-Z will show.

On each clock cycle, RAM transfers data twice, so 800MHz DDR is equal to 1600MHz.

So, what is the total bandwidth of a DDR3 RAM installed in Dual Channel DDR and operating at 800MHz, or at 800,000,000 cycles per second?

800.000.000 cycles x 2 transfers per cycle x 64 bit on each transfer x 2 for the dual interfaces of Dual Channel DDR.

The result is 204,8 billion bit (Gigabit) per second, or 25.6 Gigabyte per second (Gigabit/8=Gigabyte). This is the maximum volume of data that can be transferred to and from this particular RAM in a second.

In other words, it is a volume of data roughly equal to a single layer Blu-ray disk.

In the off chance that you are not a big fan of multiplication (go figure), you can spot the bandwidth of a RAM module before you even buy it. Just check the PCx XXXXX number.

PC3 or PC4 just confirms that it is a DDR3 or DDR4 RAM. The value next to it is the total bandwidth in Megabytes per second (MB/s), not including Dual Channel DDR or Quad Channel DDR.

So, these G.Skill RAM modules have a bandwidth of 12,800MB/s. Install them in Dual Channel DDR, and you will get 25,600MB/s, or the 25.6GB/s we found earlier.

The same goes for the newer DDR4 RAM. Of course, since the lowest frequency for DDR4 is 2133MHz (1066MHz x 2), the minimum bandwidth will also be higher.

Do the math, and you will see that these RAM modules will have a bandwidth of 17GB/s in single channel mode, which will double to 34GB/s with Dual Channel DDR.

1,066,000,000Hz x 2 x 64bit x 2 / 8 = 34.112GB/s

So, we now know what the memory bandwidth is, and how it is based on the RAM frequency.

However, it is crucial to understand that the bandwidth has little to do with the RAM speed.

What is the difference between the bandwidth and the RAM speed

Another metaphor that will help us visualize the RAM bandwidth is imagining a multi-lane road.

Let's say we have a two-lane road, which can support up to 1000 cars per hour.

Another road, with four lanes of traffic, can support up to 2000 cars per hour.

But you can't claim that the second road is faster than the first. It just has a larger capacity.

If 2000 cars tried to go through the two-lane road in one hour, we would get one hell of a traffic jam.

But a low number of cars could theoretically move at the same speed on both roads.

So, claiming that a 4266MHz module has twice the RAM speed of a 2133MHz module based only on the frequency, does not make any sense.

It's like arguing that a single car traveling an equal distance on empty roads would reach the destination twice as fast using the four-lane road than the two-lane road.

So, what's the use of the RAM frequency?

Ignoring the rest of the RAM specifications, which we will analyze farther down the article, the frequency by itself is only necessary to avoid bottlenecks from the rest of the PC components.

Let's say a theoretical system had a memory bandwidth of just 100MB/s, the equivalent of a narrow dirt road.

This is clearly hypothetical since even the ancient 133MHz RAM had a 1GB/s bandwidth.

On such a system, it would make no difference if we connected an HDD that could read 100MB/s or and SSD with a read speed of 500MB/s. Since RAM can only receive 100MB each second, both the hard drives would have the same actual speed.

In a traffic jam, a Ferrari moves at the same speed with the worst Datsun.

This is a simplified example because the real number would be lower.

The hard drive is not the sole component that transfers data to RAM. The memory system collects and distributes data from and to virtually every PC component, the CPU, the GPU, the sound card, the network adapter.

Having a much larger bandwidth than necessary pretty much guarantees that for most common uses there will be no bottlenecks because of RAM. It's like a 25-lane superhighway, which never gets a traffic jam.

But still, the RAM frequency by itself offers little to no information regarding the RAM speed.

How does the higher RAM frequency affect games

For most users, the most demanding applications they will ever run are games. This is why gaming PCs tend to be more expensive. And, more often than not, RAM modules with higher frequency are advertised directly to gamers.

You might know that many hardcore gamers are willing to spend massive amounts of money on their system. Even just to get 5 or 10 extra frames per second on demanding games.

So, if a gamer wants the highest possible performance, does it make sense to buy the highest RAM frequency? The answer is a resounding "no."

Benchmarks on the same equipment with just different RAM frequency, ranging from 800 to 2400MHz, have shown a difference of zero FPS, +/- 0.3.

The results are the same on benchmarks with newer games and more current equipment.

With the same GPU and GPU, the RAM frequency just doesn't make much of a difference.

When do we need higher RAM frequency?

The computer systems that require a large memory bandwidth, to avoid RAM bottlenecks, are usually the systems with multiple CPUs or CPU Cores.

In fact, it is the memory controller that defines the maximum bandwidth. Every current CPU has the memory controller integrated.

So, the top 7th gen Core i5 will work best with 2133 or 2400MHz DDR4 RAM, or 1333/1600 DDR3 RAM, depending on the motherboard.

Of course, that doesn't mean that the particular i5 won't work with higher frequency RAMs, such as 3000MHz or 4266MHz. As we will see later on, it is the motherboard and the XMP Profiles that define the maximum frequency.

However, bandwidth-wise, what the CPU says, goes. In this example, the maximum bandwidth with 2400MHz RAM on Dual Channel DDR is:

2.400.000.000 x 2 x 64 x 2 / 8 = 76,8GB/s

Even if we install RAM with higher frequency, we won't get a higher bandwidth between the CPU and RAM.

The same happens with the top 7th gen Core i7 for Socket 1151.

We need to look at CPUs for socket 2011-3 to get a larger bandwidth, such as the i7-6800K.

However, the increased bandwidth has to do with the Quad Channel DDR support, which we will find only on the 2011-3/X99 platform. The best RAM frequency remains at 2400 MHz.

2.400.000.000 x 2 x 64 x **4** / 8 = 153,6GB/s

The high-end AMD Ryzen 7 CPUs support a 2667MHz bandwidth, but only up to Dual Channel DDR.

2.667.000.000 x 2 x 64 x 2 / 8 = 85,344GB/s

CAS Latency: The real RAM speed

In the previous section, we saw that the RAM frequency by itself wouldn't affect the RAM speed and the PC's performance. Also, every CPU supports up to a particular memory bandwidth.

So, if that is the case, why are there high-frequency RAM modules? What are they good for? Is there a reason to pay a premium to get a <u>4266 Dual Channel kit</u>?

The answer to this comes from some different RAM specifications, that combined with the frequency, offer the real RAM speed.

When the memory controller commands RAM to read a specific memory column, the RAM's reaction is not instantaneous. There is a delay between the command to read the targeted part of the memory and before the first bit of the data reaches the exit pins.

This delay is named Column Address Strobe (CAS) Latency. As with any delay, the lower this number, the better.

On the SDRAM type of RAM, CAS Latency is counted in clock cycles. For example, RAM that has a CAS Latency of 9 (CL9) will need nine clock cycles between the command to read the memory and to transfer the data.

At the detailed RAM specifications, we will often see four numbers. These figures are the RAM Timings. The first one is always the CAS Latency.

The other three numbers are latencies <u>on other RAM tasks</u>. However, they are the CAS Latency plus some fixed numbers, so the CAS Latency is the main point of reference.

On motherboards that support memory overclocking, we can usually change the CAS Latency through the BIOS/UEFI.

If, however, we set a lower CAS Latency than the RAM modules can handle, we can have many different problems. From getting unexplained <u>BSODs</u> to the PC refusing to power up at all.

If the computer won't power up, there is no need to panic. We just need to clear the CMOS settings, usually through a special jumper on the motherboard. This way, the default BIOS/UEFI settings are restored, and with them the default CAS Latency.

We should not forget that lowering the CAS Latency is a form of overclocking. We might not be increasing the frequency, as when overclocking a CPU, but we are attempting to reduce the delay, and thus get a faster RAM speed.

The weird CAS Latency effect

There are some things regarding the CAS Latency that don't make sense at first. For starters, with higher RAM frequencies, we will usually get larger CAS Latency numbers.

For example, a 333MHz RAM - yes, they are still being sold - can have a CAS Latency of just 2.5.

Whereas we won't find 1600MHz RAM with a CAS Latency lower than 7.

Of course, that doesn't mean that the 333MHz model has higher RAM speed than the 1600MHz model. Remember, CAS Latency is counted in clock cycles.

On the 333MHz - which has a real frequency of 166,5MHz because of the DDR - on each second we have 166,500,000 clock cycles. That means that each clock cycle takes 0.000000006 seconds, or 6 nanoseconds (ns).

So, a 2.5 CAS Latency will take $2.5 \times 6 = 15.02$ ns. About the amount of time between the traffic light turning green and the guy behind us to start blaring the horn.

On the 1600MHz CL 7 RAM, with the true frequency of 800MHz, each clock cycle takes 1/800,000,000=0.00000000125 seconds (1,25ns).

The total CAS latency takes 7 x 1,25 = 8,75ns, which is considerably lower from the 15.02ns delay on the 333MHz model.

It is clear as day which of the two modules has the higher RAM speed. The 1600MHz model is about 71.66% faster than the 333 model

An easier way to compare two RAM modules, without doing so many calculations, is just to divide the frequency by the CAS Latency.

1600/7=228.57 333/2.5=133.2

228.57 is about 71.66% larger than 133.2. So, with these very simple divisions, we get the exact difference for the RAM speed of the two modules.

This simple fact can be a great eye opener on which RAM is truly faster. For example, a RAM with 1866MHz and CL 10 is objectively slower than a 1600MHz CL7 RAM.

1866/10=186.6 1600/7=228.57

The 1600MHz CL7 RAM is about 22.49% faster.

Does DDR3 have a faster RAM speed than DDR4?

Now that we know about the importance of the CAS Latency, it's easy to notice something weird between DDR3 and DDR4 RAM.

Currently, the faster 2400MHz DDR4 RAM has CL10, and that is the reason that it is so very expensive.

We can get a 2400MHz DDR3 RAM with CL10 at a fraction of the price.

For the same price, we can only get a 2400MHz DDR4 RAM with CL16, which theoretically should be much slower.

So, what gives? Does this mean that DDR3 RAM is faster or at a better price for the same speed?

The thing is, it's not a good idea to compare RAM of different architecture just based on the frequency and the CAS Latency.

DDR4 RAM includes new technologies, which speed up the function of the RAM even with a higher CAS Latency. An example is the <u>Command Encoding</u>, which is way too technical to get into in this guide.

Also, as we have shown earlier, 7th generation CPUs with DDR3 RAM work at a reduced bandwidth, despite the RAM's frequency. And there are very few 1151 motherboards that support DDR3 RAM.

There is no doubt that DDR4 RAM is the present and the future of RAM, at least until we get <u>DDR5 in 2020</u>.

Even for DDR4, it's a matter of time before we get affordable combinations of frequency and CAS Latency better than any DDR3. As it is currently between DDR3 and DDR2.

For the time being, we can use our newfound knowledge about the RAM speed to get the DDR4 2400MHz CL10 instead of the DDR4 2666 CL15.

In conclusion: How to get the fastest RAM speed

We have talked about many things so far, regarding bandwidth, frequencies, latency, sculptors, and roads. At the end of the day, however, getting the fastest RAM speed is relatively straightforward.

First, we need to make sure that the motherboard we have or plan to get can support modules with fast RAM speed.

For example, motherboards with the Intel chipsets H81, B85, H97 can only support 1600MHz RAM, with minimum CAS Latency 7. We can't install faster 2400MHz CL 10 DDR3 RAM, it will only run at 1600MHz.

To support more RAM frequencies and CAS Latencies, we need a Z- series chipset, such as Z97 for DDR3 and Z170/Z270 for DDR4, as well as the expensive X99.

Also, to get the highest frequency with the lowest tested CAS Latency, we need to activate the RAM's XMP Profile through the BIOS.

For AMD Ryzen, getting a motherboard is simpler. Even inexpensive motherboards will support high RAM frequencies.

Knowing our motherboard's limits, we just select the RAM modules that have the highest frequency with the lowest CAS Latency.

Is the extra RAM speed worth it?

It should be clear from the above examples that getting the highest RAM speed won't be cheap. Even RAM with relatively low frequency but very low CAS latency will be much more expensive than the average RAM price.

So, if we are on a tight budget, and have to choose between a better CPU/GPU or the faster RAM Speed, the CPU or the GPU will always be the better choice. Faster RAM speed won't be of much use is the CPU bottlenecks the whole system.

The same goes regarding the total amount of RAM. If we have to choose between 4GB RAM with faster speed and 8GB RAM with slower speed, it's better to go for the more RAM, rather than the faster RAM.