

Learn about how the CPU works

We will introduce the most general CPU architecture so that you can understand more about Intel and AMD products and the basic differences between them.

Each processor has its own very different and different internal design, but they must follow the same general principles - this is exactly what we want to introduce to you in this article. We will introduce the most general CPU architecture so that you can understand more about Intel and AMD products and the basic differences between them.

CPU (Central Processing Unit) - also called microprocessor or processor - is responsible for processing data. How it handles data will depend on the program. The program can be a spreadsheet, a word processor or a game. For CPU, no matter what program it is, it doesn't make much difference, because it doesn't understand what the program will do. It only follows "order requests" (called instructions or commands) contained within the program. These requirements could be adding 2 numbers or sending data to the video card, etc.

1. What is the CPU?

When you double-click an icon to run the program, this is what happens:

1. Programs saved inside the hard drive will be put into RAM. Here the program is a series of instructions for the CPU.
2. CPU uses hardware circuit called memory controller to load program data from RAM.
3. At that time the data inside the CPU will be processed.
4. What happens next will depend on the newly loaded program. The CPU can continue to load and execute the program or it can perform a certain task with the processed data, such as displaying certain execution results on the screen.

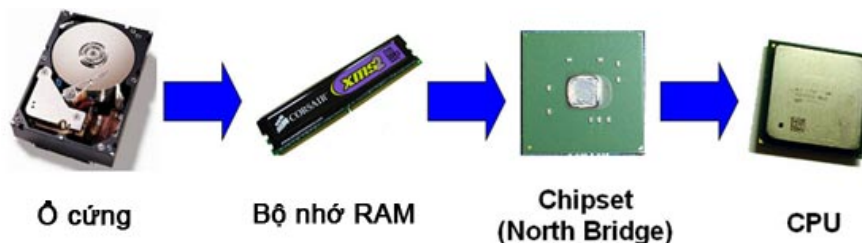


Figure 1: How saved data is transferred to the CPU

Previously, the CPU controlled the data transmission between hard disk and RAM. Since hard drives often have lower access speeds than RAM, it slows down the system, so the CPU will be very busy until all data has been transferred from the hard drive to memory. RAM. This method is called PIO, Processor I / O (or Programmed I / O). Today, the data transmission between hard disk and RAM is done without using the CPU, so it will make the system work faster. This method is called bus mastering or DMA (Direct Memory Access). To make things simpler for the drawing, we don't include the bridge chip (called the north bridge chip) between the hard drive and the RAM memory in Figure 1, but there is one chip at this connector.

AMD processors are based on sockets 754, 939 and 940 (Athlon 64, Athlon 64 X2, Athlon 64 FX, Opteron and some Sempron models) have an embedded memory controller. That means that with these microprocessors, the CPU accesses RAM directly without using north bridge chip as shown in Figure 1.

Clock

Clock is a signal used to synchronize everything inside the computer. Take a look at Figure 2, which is a typical clock: it is a square pulse that varies at '0' and '1' at a fixed fixed speed. On the drawing you can see 3 cycles of this clock pulse. The start of each cycle is when the clock signal changes from '0' to '1'; We marked it with an arrow. Clock signals are measured in units called Hertz (Hz), this is the number of clock cycles in each second of the clock. A 100MHz clock pulse means that in a second the clock has 100 million clock cycles.

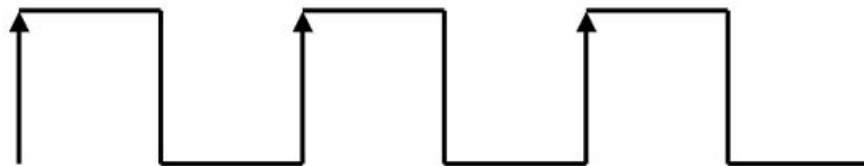


Figure 2: Clock signal

In computers, all timers are measured as clock cycles. For example, if a RAM has latency of '5', it means it will delay 5 clock cycles to start the data distribution. In the CPU, all instructions delay certain clock cycles to be executed. For example, a certain instruction may be delayed by 7 clock cycles to be executed.

For the CPU, the interesting thing is that it knows how many clock cycles each instruction needs, it knows this because it keeps a list of this information. So if it has two instructions executed and it knows that the first instruction will need 7 clock cycles to execute, it will automatically execute the next instruction on the 8th cycle. This is a common explanation for the CPU with an executable block - modern microprocessors have several execution units that work in parallel and it can execute the second instruction at the same time as the first instruction. . This is called the 'superscalar' architecture, we will talk more about this architecture in the following sections.

So what did the clock do with performance? If you think that clock and performance are the same thing, that's a completely wrong thinking about processors.

If you compare two completely different CPUs, which CPU runs at higher clock speeds will be faster. In this case, with a higher clock rate, the time between each clock cycle will be shorter, so the tasks will be executed in less time and the performance will be higher. However, when comparing two different processors, this is completely wrong.

If you take two processors with different architectures - for example, different manufacturers like Intel and AMD - the things inside these two CPUs are completely different.

As we mentioned, each instruction needs a certain number of clock cycles to be executed. Let's say that the 'A' processor takes up to 7 clock cycles to execute a certain instruction and the 'B' processor needs 5 clock cycles to perform a similar instruction. If they are running at the same clock rate, the 'B' processor will be faster, as it can handle this instruction for less time.

For modern CPUs, there are many issues that need to be considered, because CPUs have different number of execution blocks, different cache sizes, different ways of transferring data inside the CPU, how to handle instructions within the execution blocks and different clock speeds with the real world outside, etc. However, you don't need to worry about that, we will introduce them in this article.

When the microprocessor clock signal is too high, there is a problem. The motherboard, where the processor is installed, cannot work using the same clock signal. If you look at the motherboard, you will see some lines and grooves. These lines and grooves are printed circuits connecting some circuits of a computer. The problem is that with high clock speeds, these printed circuits will start to act as antennas, so the signal, instead of going to the destination at the end of the wire, disappears. , it is transmitted as radio waves.

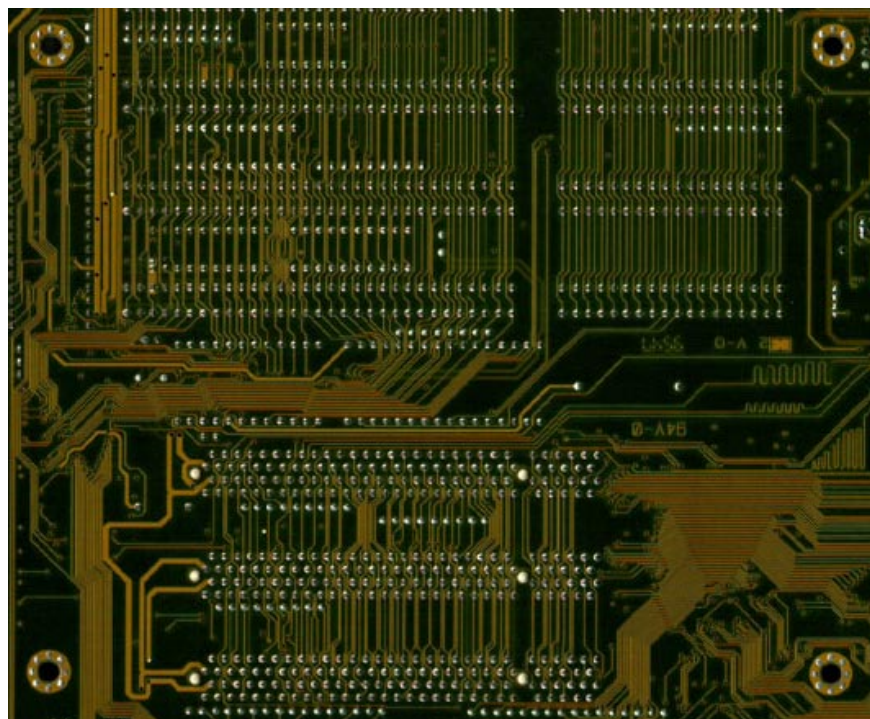


Figure 3: Printed circuit board above the motherboard can work as the antenna

External Clock

So CPU manufacturers have started using a new concept, the concept called clock multiplier, this application began to be used in 486DX2 processors. With this mechanism (used in all CPUs today), the CPU has an external clock (external clock) used when transmitting data to and from RAM (using a north bridge chip) and an internal clock. higher.

To give a real example, on Pentium 4 3.4 GHz the '3.4 GHz' number is the internal CPU clock, this clock is achieved by multiplying 17 by its external clock of 200. Simulate this example in Figure 4.

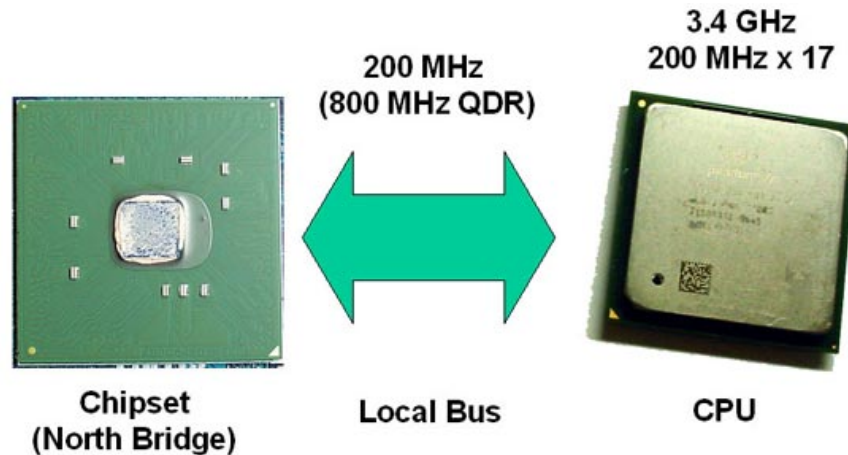


Figure 4: Internal and external clock on Pentium 4 3.4 GHz.

The big difference between the internal clock and the external clock on modern CPUs is how to overcome the magnetic weakness as mentioned above to increase computer performance. Continuing with the example of the Pentium 4 3.4 GHz above, it must reduce its speed by 17 times when performing data read from RAM! During this process, it works like a CPU at 200MHz.

Several techniques are used to minimize the effect of these differences between these clocks. One of them is to use the memory cache inside the CPU. Another method is to transmit multiple blocks of data per clock cycle. Both Intel and AMD processors use this feature, but while AMD's CPUs transmit two data per clock cycle, Intel CPUs transmit 4 data per cycle.

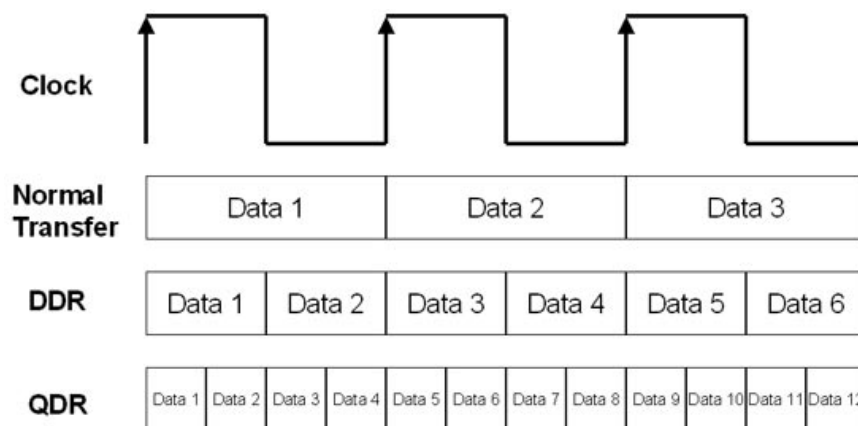


Figure 5: Transfer more data per clock cycle

Because of this, AMD's CPUs are listed as having two real external clock speeds. For example, an AMD CPU with an external clock is 200MHz listed on the CPU with an external clock of 400MHz. The same applies to Intel CPUs, with the external clock being 200MHz, its CPU will have an external clock rate of 800Mhz.

The technique of transferring two data per clock cycle is called DDR (Dual Data Rate), while the 4 data transmission per clock cycle is called QDR (Quad Data Rate).

Video about how the CPU works

You see temporarily in English, have time I will Vietnamese sub ^^.

1. Learn about how the CPU works (Part II)
2. Learn about how the CPU works (End)
3. Learn how cache works

You finished reading the article "**Learn about how the CPU works**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.