ANATOMY OF A CPU

Htoo Myat Soe Paing/ 110102722

COMP 2660 Technical Article

Reem Al-Saidi

I. INTRODUCTION

The Central Processing Unit, also known as (CPU), is the most integral component of the computing system. A CPU is a single-chip digital machine that is made up of millions or billions of tiny transistors. As comparable to the heart of the human body, the CPU serves as the final piece of the puzzle of a computer. Since the invention of the CPU, it has completely revolutionized technological advancements and reshaped the digital future of our existence. CPUs have seamlessly integrated into our daily lives, from mobile devices to laptops, computers, and beyond. They have not just been integral in our daily lives; they have become inseparable, shaping our routines and making it so much more accessible and easier to use these devices. The following article touches deeply on the importance of a CPU and will discuss the components contributing to the anatomy of the CPU.

II. IMPORTANCE

While the CPU itself is crucial, the significance of the supporting components is just as crucial. These components tie in to make sure the computer runs smoothly. The harmony between the CPU and the supporting components is vital for the whole system to function effectively.

III. TECHNOLOGICAL DETAILS

A. HISTORY OF THE CPU

The CPU has had a fascinating history, which dates back to 1971 when Intel released the first microprocessor. Federico Faggin, in particular, led the project and is credited with the design and implementation of the 4004 [5]. The Intel 4004 marked a significant milestone in the history of computing. It was a 4-bit microprocessor designed for specific tasks, and its success paved the way for the development of subsequent microprocessors that formed the foundation of modern CPUs. It had a clock speed of 108 KHz and held 2300 transistors and 10µm [4]. It was capable of executing 92,000 instructions per second [6]. While it is nothing compared to today's powerful microprocessors, it started the revolution of superior microprocessors that were soon to follow.

Fig 1. The first microprocessor intel 4004 processors held up by Federico Faggin, Ted Hoff, and Stanley Mazor [17].

In the following year, the Intel 8008, the first 8-bit microprocessor, hit the market [5]. In 1974, Motorola introduced the MC6800 with a 1-2 MHz clock frequency [7].

In the same year, Intel unveiled an improved version of its chip called the Intel 8080. Other companies soon entered the competition with their processors, including Panafacom with the MN1610, the first single-chip 16-bit processor, and Zilog with the Z80. [7]. Companies like AMD, Cyrix, and Acorn Computers also made their processors aiming to surpass the capabilities of the previous processors. It marked an era of intense competition between these companies and sparked continuous advancements in microprocessors.

The 1980s saw a surge of personal computers (PCs) with advancing microprocessor technology. Intel's 32-bit 80286 and 80386 with 134000 and 27500 transistors respectively [8]. It was now 100 times faster than the first Intel chip and allowed multi-tasking, which allowed multiple programs to run at the same time [8]. The rise of PCs was directly linked to the continuous enhancements and adaptability of CPUs during this era.

In the 1990s, Intel introduced the Pentium series, enabling higher operating speeds by doubling the clock rate and executing multiple instructions per cycle [7]. This era is also known as the "megahertz wars" as manufacturers attempt to one-up one another to produce faster processors. Notable releases include Cyrix with a Cx5x86 processor, AMD with the K6 processor, Intel with the Pentium Pro and Pentium II and Pentium III [7].

Now in the modern era of CPUs, processors have multi-core which significantly increase production and performance while reducing energy consumption. The present-day processors now equip a 64-bit chip with expanded memory and faster clock speeds. Most modern-day processors have a speed of 1GHz to 4GHz. Companies like Intel and AMD lead the way for modern processors.

In addition, modern CPUs are now equipped with technologies such as hyper-threading, dividing tasks for seamless multitasking, beneficial for professionals and gamers alike handling intensive programs [9]. These innovations reflect a bright present and a brighter future, all thanks to the remarkable progress in tiny chip processors.

Fig 2. A modern CPU [18].

B. The Control Unit

The primary role of the Control Unit is to execute instructions, often coming from the system programs of the device. These series of instructions are firstly fetched from the computer's memory, decode the binary representation, and then execute the following instruction. This process is

also known as the fetch-decode-execute cycle. The fetch stage involves the Control Unit to fetch the next instruction from the computer's memory. The memory address of the next instruction is often stored in a register called The Program Counter(PC) [1]. Once the instruction has been fetched, the Control Unit moves on to the decode stage. In this stage, it evaluates the binary code of the instruction. Each instruction typically contains an opcode(operation code), which is an operation code indicating the specific task for the CPU to perform and directs the components within the CPU to execute the instruction [2]. This stage, known as the execute stage, involves directing other components like the Arithmetic Logic Unit (ALU) to perform mathematical or logical operations. It could also involve other components of the CPU to access certain memory parts or control other parts of the system [2]. The Control Unit then makes sure the instructions are executed properly and are in the correct sequence. Finally, the Program Counter is then updated to point to the next instruction in memory and then the continuous cycle repeats. The Control Unit is integral to the CPU, serving as its main hub for executing instructions. Similar to a receptionist at an information desk, it forms the backbone of the CPU, carrying out computations and operations swiftly and effectively.
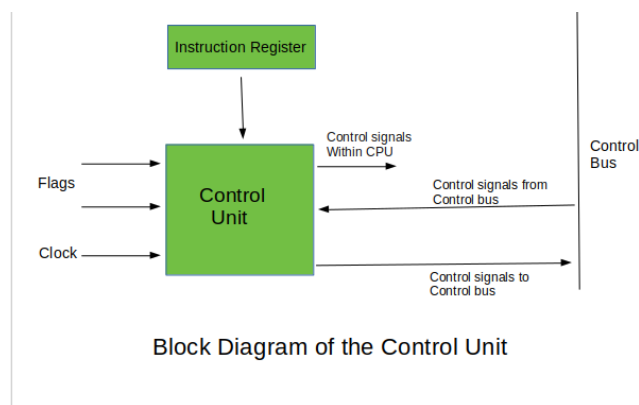

Diagram 1. A diagram of the Control Unit [19]

C. Arithmetic Logic Unit

The Arithmetic Logic Unit (ALU) is an important component of the Central Processing Unit (CPU) that is responsible for carrying out calculations and logical operations on data. Within the system itself, the ALU is specialized in two units, the Arithmetic Unit (AU) and Logic Unit (LU), which handle different tasks. The primary role of the AU is to execute mathematical operations such as addition, subtraction, multiplication, and division. LU executes logical operations including OR, AND, and NOT, also carrying out both signed and unsigned operations [2]. The ALU operates on binary numbers, executes binary calculations, and carries out more complex mathematical requests [3]. In logical operations, the ALU checks conditions and produces outcomes based on logical rules, such as comparing values for greater-than, less-than, or equal-to conditions [2]. The ALU receives inputs, including arithmetic or logical operations, from registers within the CPU to perform tasks. The ALU then performs an output based on the instruction provided through the signals from the Control Unit (CU). Integral to the Fetch-Decode-Execute cycle, the ALU plays a pivotal role in the 'execute' stage, where it performs specific arithmetic or logical operations detailed by the instruction [1]. In the modern CPUs, the inclusion of multiple ALUs enhances overall speed and efficiency.

D. Registers

Registers are small, but high-speed storage locations within the Central Processing Unit (CPU) that temporarily store data and instructions during the execution of the program. Depending on the type of the CPU, the types of registers can vary from one processor to another. Here are some common types. [11].

- The Program Counter: Acts like a GPS that keeps track of the memory address of the next instruction to be fetched. It guides the CPU through the execution of instructions [11].

- The Instruction Register: Temporarily holds the current instruction being executed and allows the Control Unit to decode the actions and instructions within the CPU.

- General Purpose Registers: Used to store data temporarily during various operations such as arithmetic, logical, and data operations.

- Accumulator: A register specialized for arithmetic and logic operations. It then stores the results and outcomes within the CPU [2].

- Address registers: Hold memory addresses for data access or are used to transfer data across memory locations [11].

- Data Registers: A temporary storage for operands and results during arithmetic and logic operations.

- Control registers: Manage several control signals within the CPU. The common tasks executed by the Control registers are interrupt control, memory management, and coordination of different operations [11].

- Stack Pointer: Manages an area of memory used for temporary data storage [2]. It maintains the memory address of the stack's top and keeps track of it using push and pop operations [12].

- Status register or Flag register: Flags the status of the CPU after an operation, indicating conditions such as Carry Flag (CF), Sign Flag (SF), Overflow Flag (OF), Negative Flag (NF), Zero Flag (ZF) and other conditional flags. The flags are useful for conditional branching, decision-making, and conditional execution [13].

Each Register Type contributes to the CPU's functionality in their way, working together to perform and execute instructions effortlessly.

E. Cache Memory

Cache memory is a high-speed computer memory that provides fast and quick data access to a processor and stores frequently used computer applications, web searches, programs, and data algorithms. Cache memory is approximately 10 to 100 times faster than the RAM, making it the fastest in computer memory [14]. This is because the Cache memory itself is located directly to the CPU chip and its small size makes it a quick access for the CPU. The CPU Cache memory is divided into 3 different cache levels (L1, L2, and L3), according to the speed of the cache levels [1].

Level 1 (L1) Cache memory is the fastest and the smallest of the three cache levels, also known as the "Primary cache". L1 cache typically ranges from 16 to 128 kilobytes (KB) [15]. The L1 cache memory is divided into two parts; the instruction cache, which stores and deals with the instructions for the CPU to perform, and the data cache, which stores the program data [14]. The primary role of the L1 cache is to store routinely used data and instructions that the CPU needs to use immediately when completing a task.

Level 2 (L2) Cache memory is the second level of the memory system. L2 is bigger than the L1 but is slower due to its location being on a separate chip near the CPU. L2 cache typically ranges from 128 kilobytes (KB) to 8 megabytes (MB), and the size ranges from one CPU to another [16]. L2 cache's main purpose is to serve as a mediator between the CPU and the main memory, by storing frequently used data, it helps reduce delays and latency with fetching data from the main memory, helping the system improve performance and speed [16]. Since L2 is much bigger, it can fetch more data and instructions from the CPU.

Level 3 (L3) Cache memory is the third and last level of the memory system, having the largest size, but is the slowest of the three Cache levels. L3 Cache is typically located on a separate chip and is typically built on the motherboard. L3 ranges from a few megabytes to and can even exceed 128 MB [2]. The primary role of the L3 cache is to facilitate resource sharing among cores, exchanging data with L1 and L2 caches [2].

Cache memory provides rapid access to recently used data, enhancing CPU utilization, reducing latency and delays, and contributing to the overall efficiency of the entire system.

F. System Clock

The System Clock is the clock speed of all the systems in the CPU and acts as a timekeeping device, like a clock that synchronizes the operations of various components within the system. The clock speed is measured in Hertz (Hz), which represents the number of clock cycles per second. A higher clock speed usually equals a faster CPU because it allows more instructions to be executed during a certain time frame.

G. Bus

A bus is a communication system that transfers data, and addresses between components within a computer or through different computers. They are a bunch of microscopic wires that transfer data between components. A bus is represented in megahertz (MHz) and ranges from 66 MHz to and even exceeds 800MHz. A higher bus speed means faster transfer of data. A bus resembles a vehicle that takes people to different locations. These are the common types of buses.

Data Bus transmits data between the CPU, and memory and can transfer to other computers. The width of the data bus indicates how many bits of data can be transferred continuously, also known as bandwidth [2]. The speed of data exchange of the data bus helps impact the performance of the system.

The address bus is responsible for transmitting memory addresses from the CPU to the memory [2].

The Control Bus is a communication channel that carries control signals and manages between the CPU and other hardware components [2]. The control bus transmits command signals like read and write commands, interrupt signals, and request and clock signals [2]. The control bus plays an important role in neutralizing the components to synchronize and helps keep the execution of instruction as smoothly as possible.

The Front Side Bus (FBS) is a bus that connects the CPU to the memory system and other system components. It is a crucial bus that determines the overall speed and efficiency of the system [2].

Understanding which bus does which, is vital for comprehending how data flows and works in a computer system, and determining the speed and efficiency of the functionality.
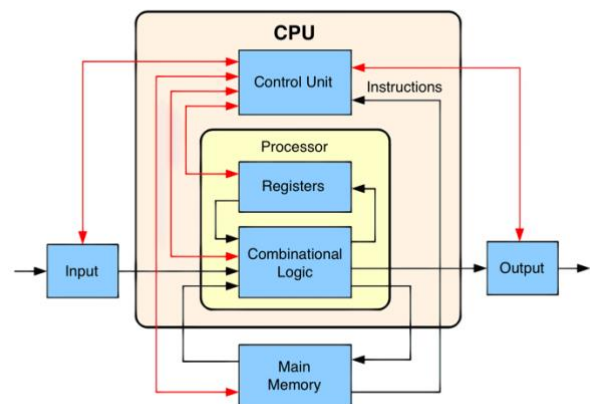

Fig 3. Diagram of a CPU [1].

IV.   DISCUSSION

As discussed, the CPU is a complicated system that requires other components to perform as well. CPUs without the main components would not function properly and not be cohesive. The synergy between the components with each other is just as crucial for the whole system to run smoothly.

A. CU and ALU synergy

The CU fetches instructions from memory and sends them over to the ALU. The ALU decodes the instructions and performs either arithmetic or logic operations depending on what the instruction has provided. The Control Unit then does the final execution for the provided instruction. The role of the ALU is to perform arithmetic operations and the role of the CU is to direct communication between the ALU and the memory unit [2]. These two components are essential and are inseparable as one needs another to successfully execute an instruction.

B. The trend in CPU design

In recent times, we have seen the evolution of microprocessors dating back to 1971 when the first chip came out. There has been an ongoing trend of microprocessors slightly becoming smaller than the previous. Measured in nanometers (nM), the lower the value of nM is, the more transistors can be placed in the chip, increasing more powerful processing power and efficiency of the CPU [1]. The greater the number of transistors correlates to a more powerful CPU. So, by decreasing the size of the transistors, more transistors can fit in it. Not only that, smaller CPUs are proven to be faster than the slightly bigger ones. A small CPU means the signals inside it will travel faster making it process faster [2]. However, while this is an ongoing trend with smaller CPUs, we must know the limitations of the processor, such as overheating, overclocking, or crashing.

V. CONCLUSION

In conclusion, The Central Processing Unit (CPU) is the most important part of a computer system, as well as the components that make up the CPU. Just like a human body, a heart alone can't function without the rest of the parts' contribution. A CPU most resembles it, it is crucial and the other components wouldn't work without it. The Control Unit (CU) and The Arithmetic Logic Unit (ALU) handle and process instructions. Registers are crucial as they provide data storage to the other components. Cache memory optimizes data levels, reduces latency and lagginess, and improves efficiency in all components. Buses also play a major role as they are responsible for transferring data between the CPU and other memory systems. The System Clock keeps track of and synchronizes the operations of various components within the system. As technology continues to emerge, the CPU will see major improvements and tweaks in the coming year. However, the synergy between the CPU and the components contributing to it will be the backbone of the entire system and will distinguish a good-performing CPU and a bad-performing CPU.

REFERENCES

1. Gayde, William. "Anatomy of a CPU." TechSpot, TechSpot, 6 Apr. 2020, www.techspot.com/article/2000-anatomy-cpu/.

2. "CPU Components." StudySmarter UK, www.studysmarter.co.uk/explanations/computer-science/computer-organisation-and-architecture/cpu-components/. Accessed 16 Oct. 2023.

3. Online CS Modules: The Central Processing Unit, courses.cs.vt.edu/~csonline/MachineArchitecture/Lessons/CPU/index.html#:~:text=The%20CPU%20is%20composed%20of,with%2016%20bytes%20of%20RAM. Accessed 16 Oct. 2023.

4. "The Story of the Intel® 4004." Intel, www.intel.com/content/www/us/en/history/museum-story-of-intel-4004.htm. Accessed 16 Oct. 2023.

5. "History of Microprocessor." Online Tutorials, Courses, and eBooks Library, www.tutorialspoint.com/history-of-microprocessor. Accessed 16 Oct. 2023.

6. "4004 - Intel." WikiChip, en.wikichip.org/wiki/intel/mcs-4/4004#cite_note-2. Accessed 16 Oct. 2023.

7. "Computer Processor History." Computer Hope, 12 Sept. 2023, www.computerhope.com/history/processor.htm.

8. Backgrounder - Intel, www.intel.com/pressroom/kits/45nm/IntelTransistor_Bkgndr_FINAL.pdf. Accessed 16 Oct. 2023.

9. "What Is Hyper-Threading?" Intel, www.intel.com/content/www/us/en/gaming/resources/hyper-threading.html. Accessed 16 Oct. 2023.

10. "Control Unit." Encyclopædia Britannica, Encyclopædia Britannica, inc., www.britannica.com/technology/control-unit. Accessed 16 Oct. 2023.

11. "What Is a Register in a CPU and How Does It Work?" Total Phase Blog, 1 June 2023, www.totalphase.com/blog/2023/05/what-is-register-in-cpu-how-does-it-work/#:~:text=Registers%20are%20a%20type%20of,bit%20sequence%20or%20individual%20characters).

12. Embedded System Design - Baylor University, web.ecs.baylor.edu/faculty/dong/elc4438_spring2016/CortexMArch1.pdf. Accessed 16 Oct. 2023.

13. "Processor Status Register." C64, www.c64-wiki.com/wiki/Processor_Status_Register. Accessed 16 Oct. 2023.

14. Phillips, Gavin. "How Does CPU Cache Work? What Are L1, L2, and L3 Cache?" MUO, 30 Jan. 2023, www.makeuseof.com/tag/what-is-cpu-cache/.

15. What Is Level 1 Cache (L1 Cache)? - Definition from Techopedia, www.techopedia.com/definition/8048/level-1-cache-l1-cache. Accessed 16 Oct. 2023.

16. "Arc HS L2 Cache Memory Option." Synopsys, www.synopsys.com/dw/ipdir.php?ds=arc-l2cache#:~:text=The%20L2%20cache%20is%20user,16%20or%2032%20memory%20banks. Accessed 16 Oct. 2023.

17. Ing, Graeme. "The Intel 4004-the World's First Microprocessor." *Medium*, Medium, 5 Mar. 2023, graemeing.medium.com/the-intel-4004-the-worlds-first-microprocessor-84e25269c797.

18. "Pearson IT Certification." *Processor Overview | On the Motherboard | Pearson IT Certification*, www.pearsonitcertification.com/articles/article.aspx?p=2731934&seqNum=2. Accessed 17 Oct. 2023.

19. "Introduction of Control Unit and Its Design." *GeeksforGeeks*, GeeksforGeeks, 31 Aug. 2023, www.geeksforgeeks.org/introduction-of-control-unit-and-its-design/.