1 The central processing unit

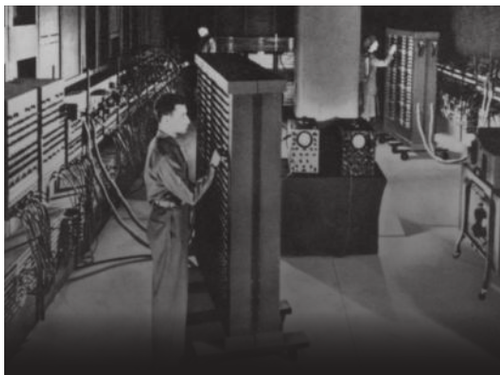
Learning outcomes

By the end of this chapter you should be able to:

- explain the purpose of the Central Processing Unit (CPU)
- describe the components of the CPU and their functions
- describe the fetch-execute cycle
- describe Von Neumann architecture
- explain how the cache size, clock speed and number of cores affect CPU performance
- explain the purpose of and characteristics of an embedded system

Why the CPU?

- The central processing unit (CPU) is responsible for executing the instructions given to it in a program.
- It follows the instructions in order to do something useful.
- The microprocessor relies on other devices:
  - to allow users to input the instructions
  - to store the instructions
  - to transfer the instructions to it so that it can carry them out
  - to carry out the commands it issues e.g. to print an essay or display an image.



ENIAC, the first computer, was 8 feet hight and 100 feet long and far less powerful than today's laptop.

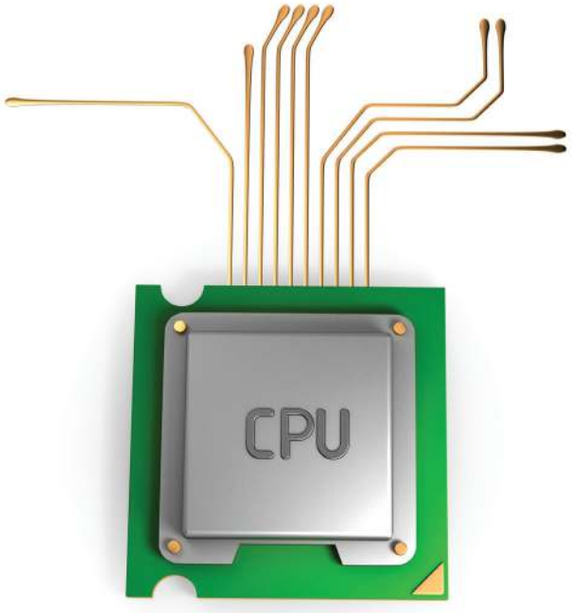
The central processing unit

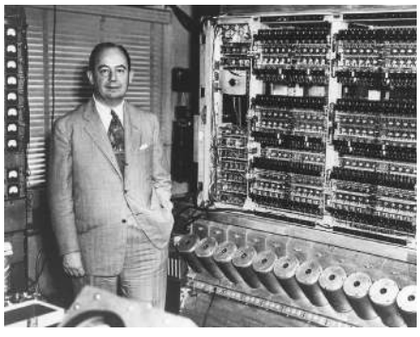
Key terms

central processing unit (CPU): this is the component of the computer that controls the other devices, executes the instructions and processes the data

execute: to run a computer program or process

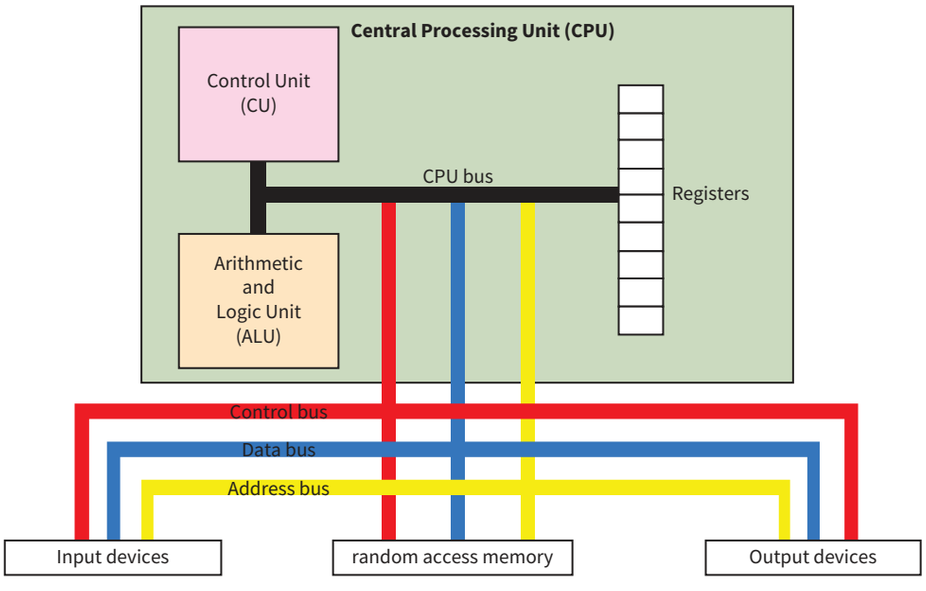
The microprocessor is the central processing unit (CPU) of the computer. It is here that the data processing takes place.





The way the CPU is designed and exetcutes (carries out) the program instructions is known as 'von Neumann architecture'. In 1945, John von Neumann (shown in the photo left) proposed his design for a 'stored program' computer where both the program and data were stored in the memory. Previously, computers had to be rebuilt for each new program that was needed!

The diagram below shows how this is put together. Take a look at the components; their functions will be described when we look at the way in which program instructions are executed.



Buses

A bus is a collection of wires that carry signals or communications between the various components of the computer system. The control bus connects the control unit (CU) with the other components of the CPU and devices in the computer system. The control unit uses it to send instructions to other components of the computer. The components use a bus to send information back to the CPU as well. The data bus is used for the transfer of data between the CPU and the RAM and the address bus for the CPU to access memory locations in the main memory.

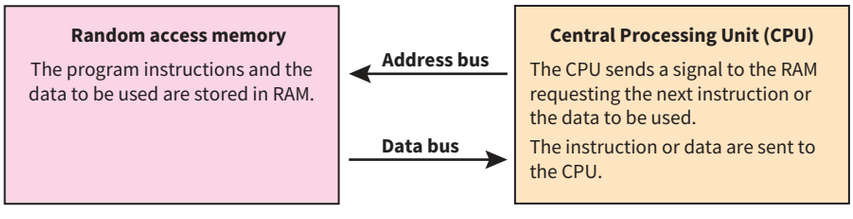
Executing the instructions: the fetch-execute cycle

The way in which the 'von Neumann architecture' executes the program instructions is through the fetch-execute cycle.

Before the cycle starts, the program instructions are copied from a storage device such as a hard disk drive or DVD to the primary storage or random access memory (RAM). (see Chapter 2, secondary storage)

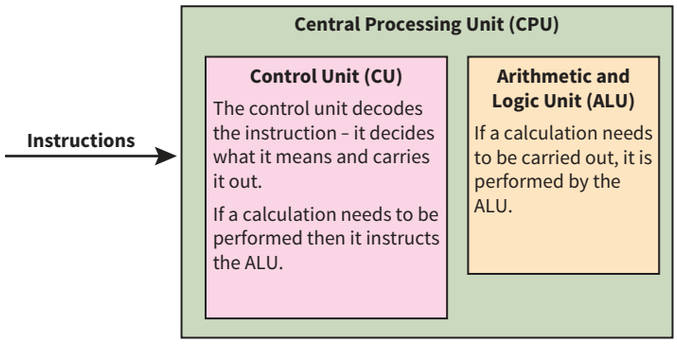
Fetch

In the fetch part of the cycle, instructions and data are moved from the random access memory to the central processing unit.



Execute

In the execute part of the cycle, the control unit decodes or interprets the instructions and decides what action to perform. These instructions are then carried out.



Key terms

bus: a bundle of wires carrying data from one component to another or a number of tracks on a printed circuit board (PCB) fulfilling the same function

RAM: (also known as random access memory) memory that can be used by computer programs to store data and instructions, but all of its data is lost when the computer is switched off

Components of the central processing unit (CPU)

The arithmetic and logic unit

The ALU performs arithmetic and logical operations. It carries out activities such as:

- addition and subtraction
- multiplication and division
- logical tests using logic gates (explained in Chapter 14)
- comparisons, such as whether one number is greater than another.

Key terms

control signals: electrical signals that are sent out to all of the devices to check their status and give them instructions

register: a storage location that is inside the CPU itself to store instructions and data that are currently been used in the fetch-execute cycle.

The control unit

The control unit coordinates the actions of the computer and controls the fetch-execute cycle by sending out control signals to the other parts of the CPU such as the ALU and registers. It also sends signals to other components of the computer system such as the input and output devices.

The two main elements of the control unit are the clock and the decoder.

The clock

Pulses are sent out to the other components to coordinate their activities and ensure instructions are carried out and completed. The timing is controlled by a vibrating quartz crystal.

One instruction can be carried out with each pulse of the clock, and therefore the higher the clock speed, the faster the CPU will be able to carry out the program instructions.

The clock speed is measured in cycles per second. 1 cycle per second is a rate of 1 Hertz. 1 megahertz (MHz) equals 1 million cycles per second and 1 gigahertz (GHz) is 1,000,000,000 cycles per second. Rates of 1 to 3 GHz are common in most home computers.

The decoder

This part of the control unit decodes the program instructions (works out what they mean) that have been brought from the memory and decides what actions should be taken. It then sends control signals to the other components to carry them out.

Registers

Key term

storage location: a place in RAM where a single piece of data can be kept until it is needed

Registers are storage locations within the CPU itself. They can be accessed even more quickly than the random access memory. The function of these registers is to store instructions and data that are currently being used in the fetch-execute cycle.

Some of the registers serve specific functions, but some of them are general purpose registers used for the quick storage of data items. Registers that serve specific functions include:

- the accumulator (A or ACC)
- the program counter (PC)
- the memory address register (MAR),
- the memory data register (MDR) or memory buffer register (MBR)

Complete Interactive Activity 1a on Cambridge Elevate

ACTIVITY 1.1

Name the parts of the CPU that perform the following functions:

a. carries out arithmetic and logical computations

b. stores data within the CPU itself

c. coordinates the activities of the CPU and computer.

Complete the Cambridge Computing Online activity www.cambridge.org/links/kose4030

Watch the fetch-execute cycle animation on Cambridge Elevate

Complete the Cambridge Computing Online activity www.cambridge.org/links/kose4029

Complete the Cambridge Computing Online activity www.cambridge.org/links/kose4031

Remember

The CPU processes data by carrying out these steps:

1. Fetch: an instruction is transferred from the memory to the CPU.
   a. The program counter supplies the address of the instruction to be fetched.
   b. The program counter is a register (also referred to as memory location) in the CPU.
2. Decode: the CPU works out what the instructions mean.
3. Execute: the control unit carries out the instructions using the ALU for instructions involving logical or mathematical operations.

Remember

1. The CPU consists of the
   a. control unit (CU)
   b. arithmetic and logic unit (ALU)
   c. registers.
2. The CU controls the activities of the CPU by sending out control signals.
3. The ALU carries out arithmetic and logical operations.
4. The registers are memory stores within the CPU.

Download Worksheet 1.1 from Cambridge Elevate

ACTIVITY 1.2

With the aid of diagrams, describe the events that take place during the fetch-execute cycle.

Tip

Overclocking is the process of increasing the clock speed to a higher level than that recommended by the manufacturer. It causes an increase in heat and can result in instability and permanent damage to the processor.

Key term

heat sink: a metal device, glued to the CPU chip with thermally conducting paste, to transfer the heat away from the chip

Watch the Von Neumann bottleneck animation on Cambridge Elevate

Key term

cache: a temporary data store so that it can be accessed very quickly when needed

CPU performance

Everyone wants their computers to work faster and faster and so manufacturers have continued to increase the speeds at which they work.

Clock speed

The rate at which instructions are processed is controlled by the clock speed. The faster the clock speed, the faster the rate of processing. Clock speeds have increased and a rate of 3 GHz is common in modern computer processors. But increasing the clock speed to increase processing speed has limitations:

- The microprocessor generates a large amount of heat and this increases with the clock speed. Although the heat is dissipated by a fan and heat sink to prevent it from malfunctioning or even melting, there are limits to the rate of cooling.
- Microprocessors with clock speeds of 9 GHz require cooling by liquid nitrogen!

ACTIVITY 1.3

In order to increase the speed of their computers, users often increase the clock frequency.

a. What is this process called?

b. A user who increased the clock frequency noticed that his computer was now far noisier.

Explain why this could be caused by increasing the clock frequency.

Cache size

Bottlenecks occur when one component cannot work as fast as other components and so hinders progress.

In the fetch-execute cycle this bottleneck is caused by the dynamic RAM (DRAM) (see Chapter 2) that is used to store the instructions and data.

Although it is far quicker to fetch the instructions from the DRAM than from a hard disk drive, the RAM is still far slower than the CPU and then the instructions have to be transferred through the data bus.

The speed of processing will be limited by the RAM which supplies the instructions regardless of how much clock speed is increased or how many cores are used.

The solution to this bottleneck problem is to use faster memory very close to, or even within, the CPU. This memory is used to store recently used data and data likely to be frequently used and is called a cache.

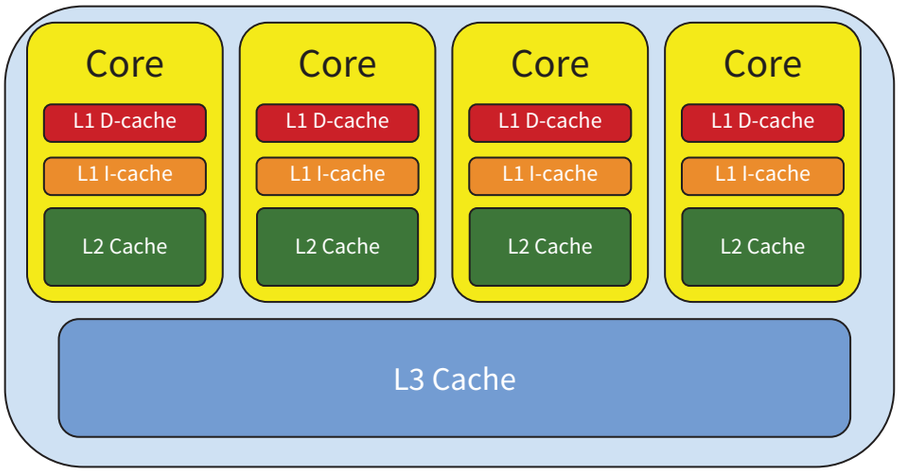
Most CPUs have independent instruction and data caches. The data caches have to be read and written to but the instruction caches just have to be read by the CPU.

The use of caches allows the CPU to check the fast cache for the data it needs. It does not have to wait for it to be fetched from the much slower DRAM. The faster SRAM is used for the cache.

The caches are located on the processor chip. The fastest is the Level 1 cache and is smaller than the Level 2 and Level 3 caches.

The L1 cache is checked first followed by the L2 and then L3 caches.

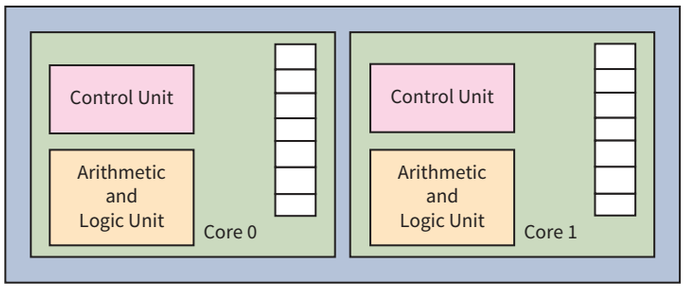
In a multi-core processor each core has its own L1 and L2 caches while the Last Level Cache is usually shared by all the cores.



The caches in a multicore processor

Number of cores

Manufacturers introduced multi-core processors in 2006 to increase processing speed. A multi-core processor has more than one CPU. The following diagram illustrates the structure of a dual-core processor.



Complete the Cambridge Computing Online activity
www.cambridge.org/links/kose4032

The following table shows the names given to processors having different numbers of cores.

Number of cores	Common name
1	Single-core
2	Dual-core
4	Quad-core
5	Penta-core
8	Octa(o)-core
10	Deca-core

The advantages of multiple core processors over single core processors are:

- the cores can work together on the same program; this is called parallel processing
- the cores can work on different programs at the same time; this is called multitasking.

However, not all programs will run at twice the speed with a dual-core processor. The tasks required might not be able to be carried out in parallel. They might be sequential so that one task requires output from a previous task and so the second task cannot start until the first has finished.

Key terms

parallel processing: when the processor cores work on different parts of the same program

multitasking: when the processor cores work on different programs at the same time

Complete the Cambridge Computing Online activity www.cambridge.org/links/kose4033

Download Worksheet 1.2 from Cambridge Elevate

ACTIVITY 1.4

A computer is advertised as being 'quad-core'.

a. Explain what is meant by 'quad-core' and how it improves the performance of the computer.

b. Explain how using cache memory also improves the performance.

Remember

1. The clock speed cannot be increased indefinitely because of the extra heat generated.
2. A processor with several CPUs is said to be multi-core.
3. The processing speed of the CPU is limited by the speed that data can be supplied by the slower RAM.
4. Cache memory stores regularly used items of data so that they can be accessed more quickly.
5. When the main memory (RAM) is full, the operating system will store data on an area of the hard disk drive (the virtual memory).

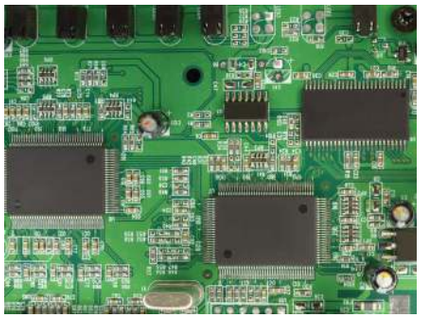
Complete Interactive Activity 11c on Cambridge Elevate

Embedded systems

An embedded system is a computer system built within a larger device such as:

- washing machines
- digital cameras

The computer systems in these devices have been built for a range of specific tasks.



Embedded devices are limited to a certain number of tasks, unlike desktop and laptop computers, which are general purpose computer systems capable of carrying out many different tasks.

All of the components in an embedded system, including microprocessor, memory and input and output interfaces, are on a single printed circuit board. The memory contains the program. The board is a component built into a larger device, hence the name 'embedded'.

In an embedded system, resources such as memory are limited and the hardware components are often unique to that system. The programs are often written in assembly language rather than in a high level language so that the hardware components can be more easily controlled.

Remember

An embedded system is built into a larger device and carries out a limited number of functions.

All of the components in an embedded system including microprocessor, memory and input and output interfaces are on a single printed circuit board.

The memory contains the program.

Key terms

printed circuit board (PCB): the base that supports the wiring and electronic components that are soldered to it or fit into sockets on the board.

hardware: the physical components making up the computer, and its peripheral devices

ACTIVITY 1.5

Many modern electronic devices contain embedded systems.

a. Explain what is meant by an 'embedded system'

b. List three devices that contain embedded systems.

Download Worksheet 1.3 from Cambridge Elevate

Practice questions

1. State the roles of the following in the fetch-execute cycle.
   a. The program counter.
   b. The address bus.
   c The data bus.
2. Catherine has bought a new laptop computer which was advertised as having a 1.6GHz dual-core central processing unit (CPU) and 512KB Level 1 cache.
   a. State the purpose of the CPU.
   b. Describe what is meant by
      i. 1.6GHz CPU
      ii. dual-core CPU
      iii. L1 cache.

Your final challenge

Your challenge is to design and code a multiple-choice quiz based on the contents of this chapter.

The design brief

- There should be at least five questions.
- Each question should have four possible answers, only one of which is correct.
- After taking the quiz the user should be
  - informed of their score
  - told the correct answers for those that they got wrong.
- You should create and then test your quiz.

Extension

- Ask the users to enter their names and then save the users and their scores in a text file.

Download Self-assessment 1 worksheet from Cambridge Elevate

What techniques did you use to remember the sequence of the fetch-execute cycle? What further research did you carry out on the fetch-execute cycle?