Higher
Computing Science

Computer Systems

Computer Structure

Name:_____

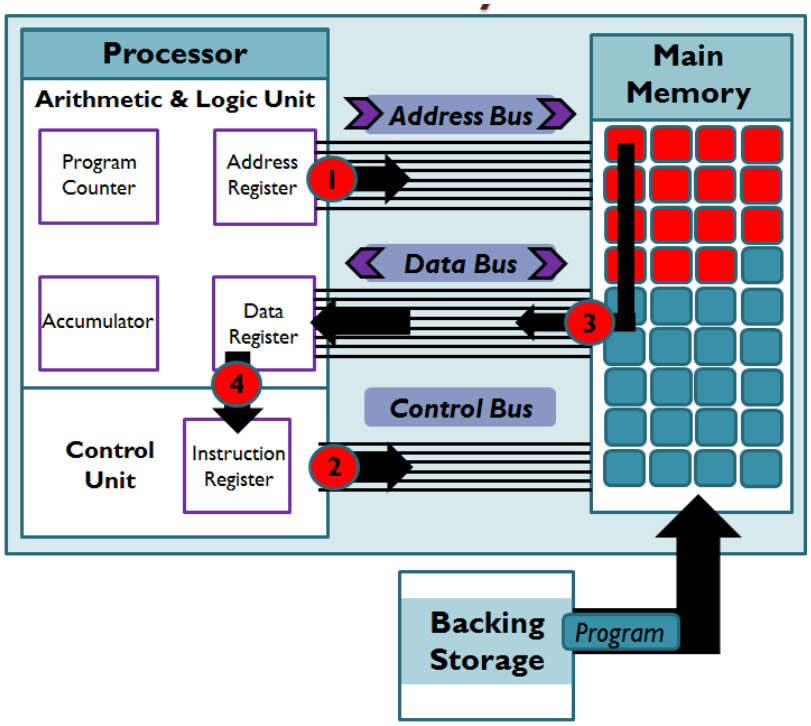
Contents

Fetch Execute Cycle

When a program is started, machine code **instructions** are **loaded** from backing storage into **main memory.** Each instruction for the program is located in a different **unique address location** in main memory

Fetch-Execute

The **fetch-execute cycle** is the process that retrieves the program instructions, one at a time, from RAM to the CPU where they are executed.



1. The address bus is set up (by the MAR) with the memory address location of next instruction to be fetched.

2. The **Read line** is activated on the **Control Bus** (to inform RAM that an instruction or data is to be transferred to the CPU).

3. **Instruction** at specified memory location is loaded onto the *Data Bus* which sends it to the Data Register.

4. Instruction is passed from Data Register to Instruction Register to be **decoded and executed.**

When the CPU needs to write data back to memory (e.g. an updated variable value) then a **memory write** operation must take place.

Memory Write

1. The address bus is set up (by the MAR) with the memory address location of next instruction to be fetched.

2. The data bus is set up (by the MDR) with the data to be stored in RAM.

3. The **Write line** is activated on the **Control Bus** (to inform RAM that an instruction or data is to be transferred from the CPU).

4. Data on the **data bus** is stored in the memory location specified by the **address bus**.

Factors affecting Computer System Performance

The fetch-execute cycle can be speeded up by altering some of the components that are involved in the cycle itself.

- CPU clock speed
- Number of CPUs
- Width of the data bus
- Increase in Cache memory

Clock Speed

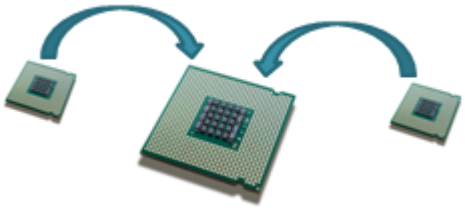
The CPU **clock** sends out a constant, steady pulse. **One** action is carried out by the processor on **each** clock pulse.

Clock speed (frequency) is measured in Hertz (Hz), meaning per second.	Typical computer is 3GHz That's 3000 million pulses per second

The faster the **clock speed**, the more instructions can be executed in a fixed time.

Number of Processors (Cores)

In the past, each CPU chip had a **single core**. This means that it contains one processor and can do **one** thing at a time.



Nowadays, CPU chips contain **dual core** (two cores) **or more** meaning they can do **more than one** thing simultaneously.

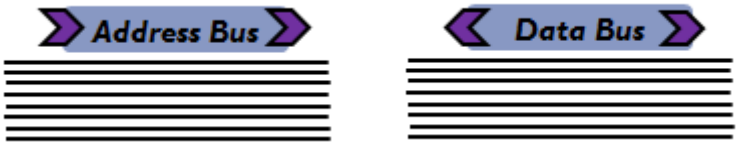
The more cores a CPU has, the more tasks it **can be** programmed to carry out at the same time, improving the performance of the computer.

If a piece of software isn't programmed to make use of these extra cores then they will not make any difference to the system performance.

Data Bus Width

Buses connect the CPU to main memory allowing them to communicate with each other.

The main CPU buses are:



Address Bus

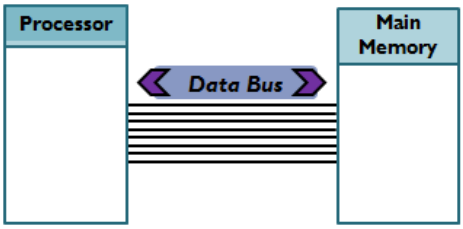
The Address Bus is **unidirectional** meaning it only carries address information in **one direction**.

The Address Bus is used by the CPU to tell main memory which address location to open.

Data Bus

The Data Bus is **bi-directional** meaning that it carries data in **both directions** between the CPU and memory.

As well as data, the Data Bus carries **instructions** from memory to the CPU where they are executed.



The number of lines on the data bus is equivalent to the size of each memory location. This is known as the **word size.**

8 lines	= **1** byte word size
16 lines	= **2** byte word size
24 lines	= **3** byte word size

Increasing the width of the data bus improves CPU performance by reducing the number of fetches to memory**.**

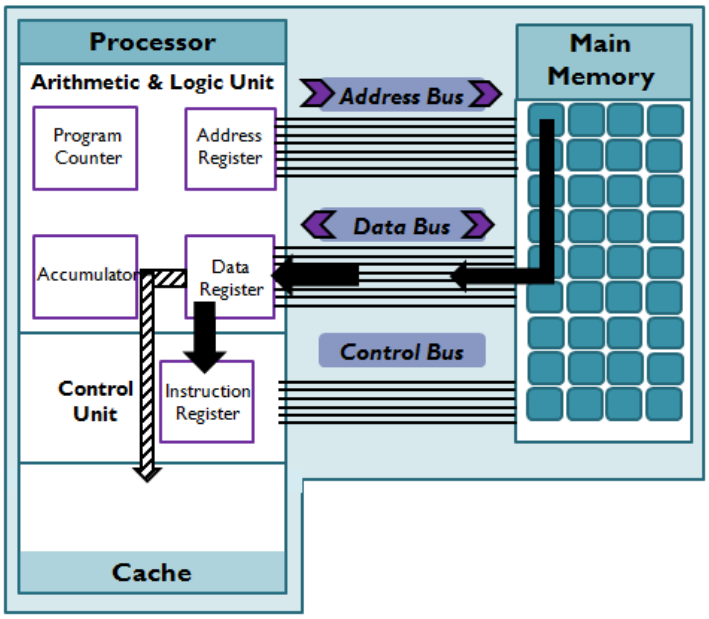
This is because it will take fewer fetches to transfer data if the data bus is larger.

Cache Memory

CPU overall performance is compromised by buses and main memory.

The CPU is much faster than these components so it has to wait on instructions and data being fetched.

The use of **cache memory**, located on the CPU, improves the CPU performance.



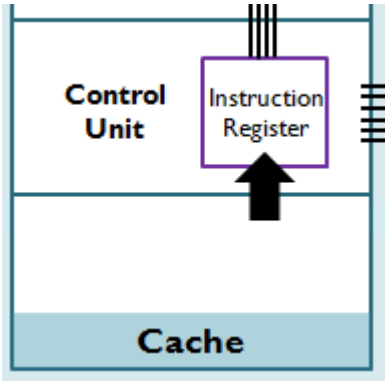
The CPU has to **wait** on the buses and Main Memory **fetching instructions** and data which is a **slow** process**.**

CPU **performance can be improved** by adding cache memory to the CPU

Cache memory is located on or close to the CPU itself which makes it **faster** to access data.

Cache stores copies of data from frequently used memory locations, or **reads ahead** to load the next instruction into cache

The CPU will firstly check cache for the instruction it needs saving it having to carry out a slow fetch to main memory.



Cache uses S-RAM chips which are **more expensive** (per MB) than main memory D-RAM chips, but **cache access time is faster.**

Summary: Improving computer system performance

Increasing Clock Speed

- Fetch-Execute cycle is carried out more quickly
- CPU can execute more instructions per second.
- Power consumption/heat increases so cooling is more difficult (super-cooled circuitry)

Increasing number of processors (cores)

- Several instructions can be carried out simultaneously (1 per core)
- Programs can execute more quickly
- Programs must be written specially to make use of multiple cores.

Increased Data Bus Width

- Reduces the number of fetches to main memory by transferring more data in a single fetch
- Larger instructions can be transferred in fewer fetches.
- Performance improved by making less use of slow buses.

Increased Cache Memory

- Reduces the number of fetches to memory by storing frequently accessed instructions and data in cache (close to the CPU)
- Prevents a slow fetch using the buses for every instruction
- CPU can obtain instructions and data from cache allowing for less use of slow buses.

Revision Questions – Computer Structure

1. Explain why a processor with larger cache would outperform an identical processor with smaller cache.

2. Two computer systems have the same number of processor cores, the same width of data bus and the same clock speed.

   (a) State one other factor that could account for one computer system performing better than the other, when tested for processing speed.

   (b) Explain why increasing the width of the data bus will improve the system performance.

3. An instruction to be executed is in memory address 3412. Complete the missing steps of the fetch-execute cycle shown below.
   Step 1 The processor sets up the address bus with the address 3412.

   Step 2 _____

   Step 3 _____

   Step 4 The instruction in the instruction register is then interpreted by the decoder and carried out.