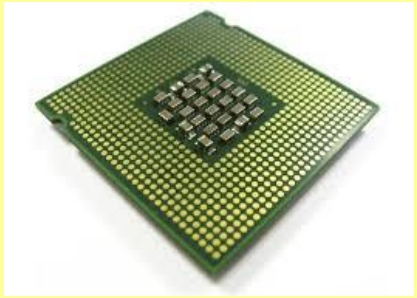
National 5



Computer Architecture

What we need to know!

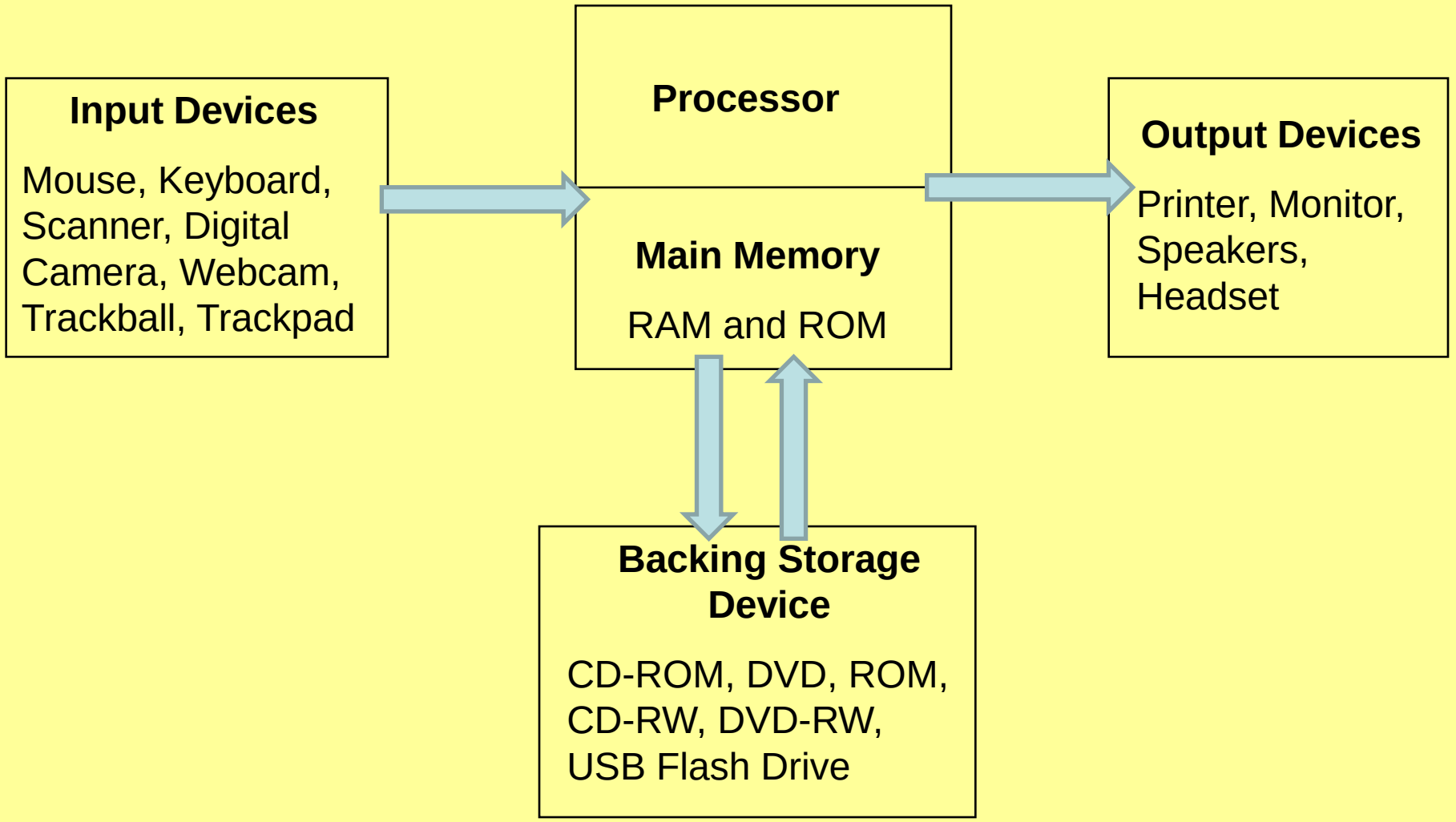
The Processor – ALU, Control Unit, Registers

Memory – Main Memory, RAM, ROM, Memory Sizes

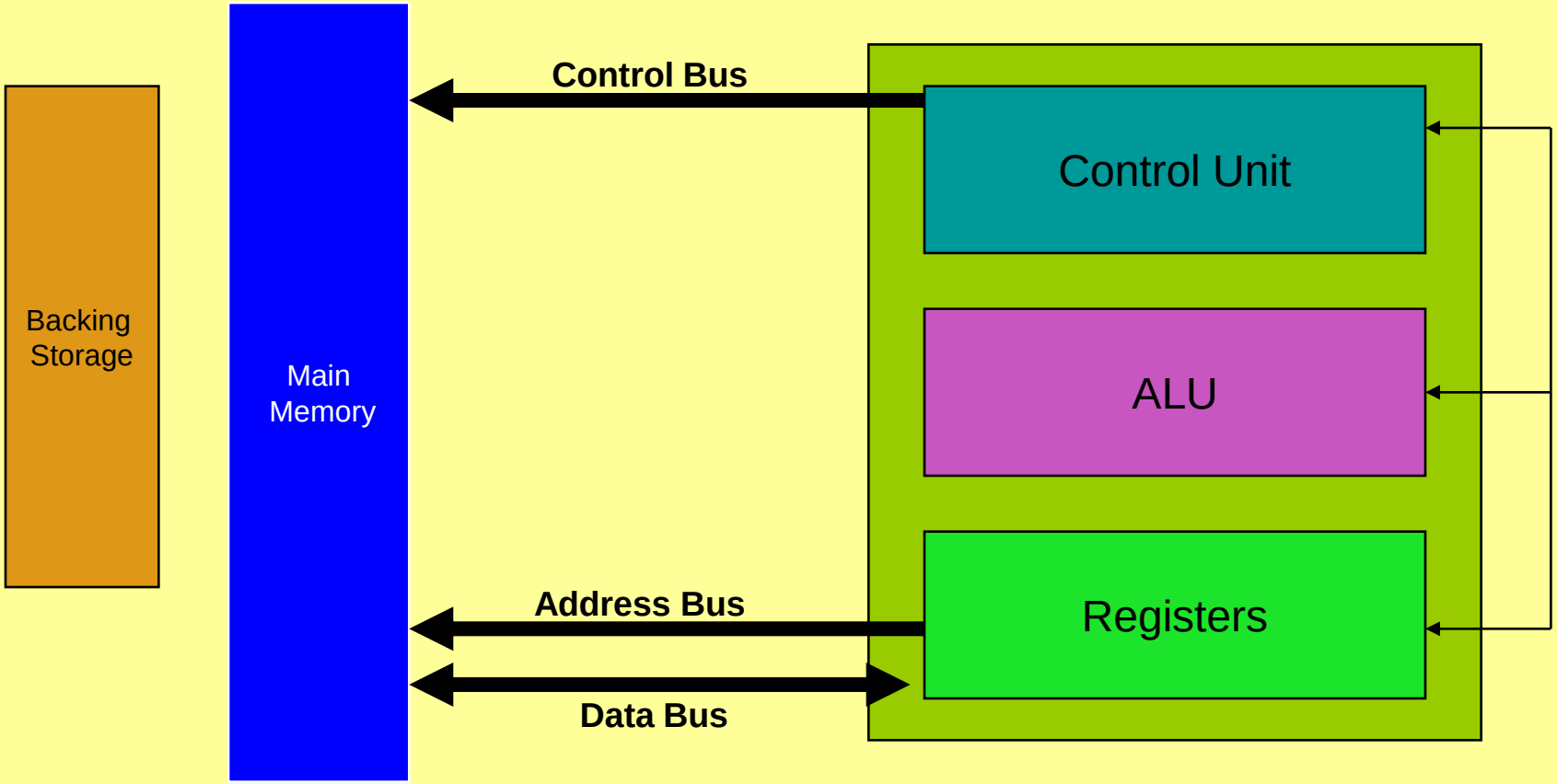
Buses – Address, Data, Control

Interfaces – Data Conversion, Speed of Operation, Temporary Storage, Different Voltage Levels

4 Box Diagram



Basic Computer Structure



The purpose of the ALU

- The Arithmetic and Logic Unit (ALU) is the part of the Central Processing Unit (CPU) where the following take place:

  - Calculations (+, -, /, *)
  - Logical Decisions (AND, OR, NOT)
  - Comparisons (<,=, >, <=, >=)

The purpose of the Control Unit

- The main functions of the control unit are:
  - to control the timing of operations within the processor
  - to send out signals that fetch instructions from the main memory
  - to decode these instructions
  - to carry out these decoded instructions

The purpose of the Registers

- These are very fast temporary storage locations on the processor, they provide temporary storage places for data being manipulated. This data could be an address in memory, data being read or written, or instructions waiting to be decoded.

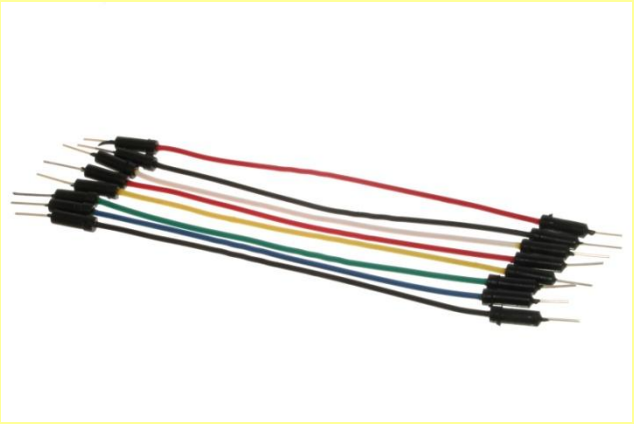
- They are used to:
  - Hold data that is being transferred to or from memory.
  - Hold the address of the location in memory which the processor is accessing to read or write data.
  - Hold the instructions that are being carried out.

Complete Exercise 1 and get your teacher to check the answers.

Processor Buses

- A BUS is a channel along which data/control signals flow.

- These channels are electronic pathways (sets of wires).



Data Bus

- The main pathway along which data flows between the Processor and Memory
- The Data Bus is a bi-directional (two way) bus as data can travel from the processor to memory or from memory to the processor.
- The width of the data bus is measured by the number of lines on the bus. Each line can carry one bit.
- The wider the Data bus the more data can travel along it in one clock cycle e.g. A 32-bit data bus can transfer 32 bits at a time.
- The width of the data bus is known as the **word length**.

Address Bus

- All addresses for read or write operations are generated by the processor and are placed on the address bus.
- This means the address bus holds the address of the memory location being accessed.
- The Address Bus is unidirectional (one way) only - *always from the processor*
- The wider the Address bus the more memory that can be addressed.

Control Bus

- The Control Bus is of a set of control lines which work independently, controlling all processor operations.

Control Line	Function
Read Line	controls a read from memory operation
Write Line	controls a write to memory operation
Clock	constant pulse that controls the timing of all processor operations
Reset	resets control lines and processor registers
Interrupt	signals the processor that an Input or Output device or a systems software routine requires the processors time

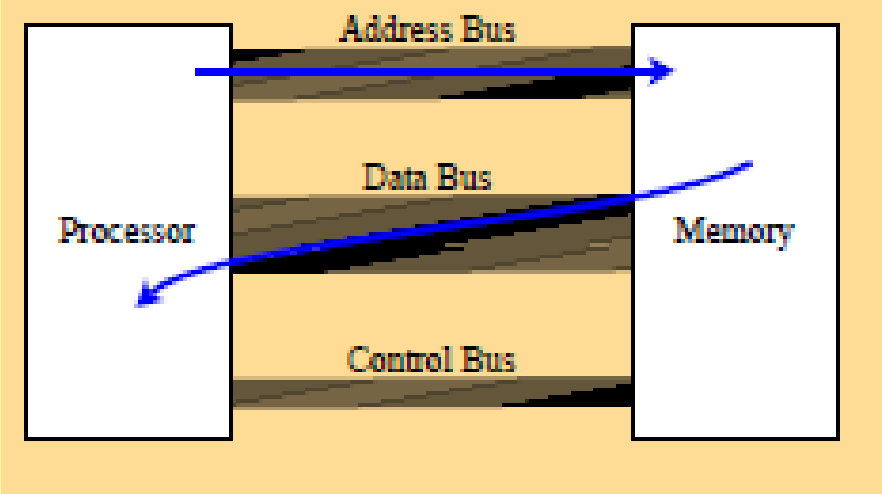
Complete Exercise 2 and get your teacher to check the answers.

The Role of Buses

When instructions and data are transferred from the memory to the processor the following steps are carried out.

Memory Read

1. The address bus is used to select the address of the desired memory location.
2. The control bus activates the Read Line.
3. The machine code instruction (or data) in the selected location is sent along the data bus to the processor.

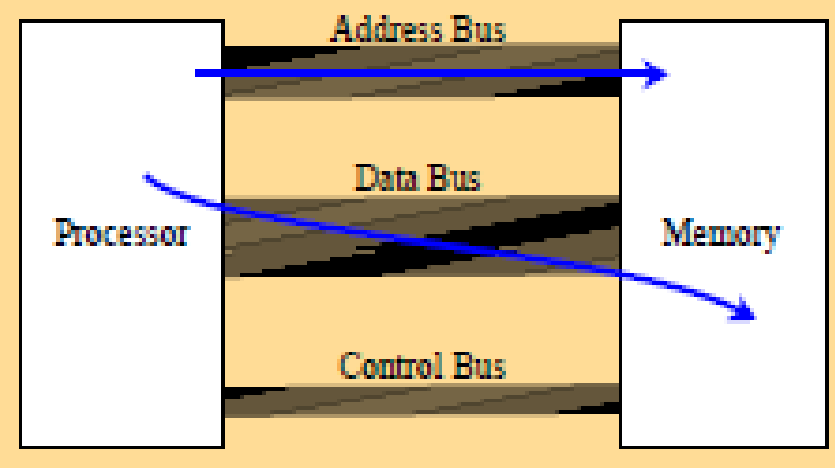


The Role of Buses cont…

When data is transferred from the processor back to the memory the following steps are carried out.

Memory Write

1. The address bus is used to select the desired memory location.
2. The control bus activates the Write Line.
3. The data is sent along the data bus to the selected location.



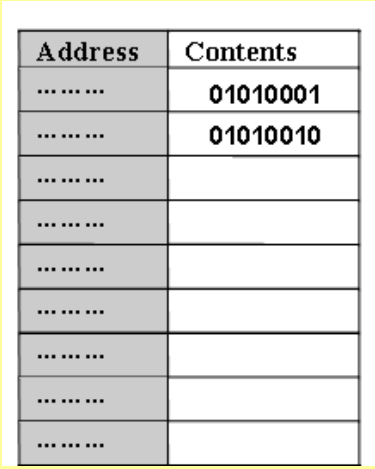
- Note - The address bus is one directional as it only ever sends address information from the processor to the memory. The data bus is bi-directional as data can travel to and from the processor.

Main Memory

- Main memory stores programs and data that the user is currently using. Instructions have to pass into the processor to be carried out. The processor can't store all the instructions it needs inside it so main memory is used.

- Memory is organised into consecutive memory locations, each having its own unique address, and each capable of storing one byte of information (the amount needed to store a single letter eg A). Main memory is just a collection of memory chips.

- It is divided into **ROM** and **RAM** sections.



Address	Contents
……	01010001
……	01010010
……	
……	
……	
……	
……	
……	
……	

Main Memory (ROM)

- Read Only Memory (ROM) is used to store a small part of the operating system called the Bootstrap Loader. When your computer is switched on, the bootstrap loader examines the backing storage devices to find the operating system. Once found it is loaded into RAM.

ROM has the following features:

- data in ROM is permanently stored onto a microchip
- ROM is read-only so it cannot be changed
- data on ROM is not lost when the computer is switched off.

Main Memory (RAM)

- Random Access Memory (RAM) is the largest part of the main memory. This is where the operating system is stored; it also holds all programs and data. You can purchase additional RAM chips and install them in your desktop/laptop computer.

- RAM chips - to improve system performance they can be added to your computer, increasing the ram from, say, 2 GB to 4 GB.

- RAM has the following features:
  - the data in RAM is read/write so it can be changed
  - all data stored in RAM is lost when then computer is switched off
  - RAM stores the programs and associated data of the programs currently in use

Complete Exercise 3 and get your teacher to check the answers.

Storage

- A computer is called a *two-state* system as it uses only two digits for all processing and storage. These are the digits 0 and 1 - called BINARY. Think about binary as being a light bulb that is either ON or OFF, ie, it has two states!

Bulb On
1

Bulb Off
0

Storage refers to the *media and methods* used to permanently store data, this enables data to be loaded in the future.

Storage

- The smallest form of storage in a computer is a bit, this is a **Binary Digit**, either 1 or 0. This table shows you how bytes are organized:

Bit	= binary digit: 0 or 1
1 byte	= 8 bits
1 KiloByte	= 1024 bytes
1 MegaByte	= 1024 KB
1 GigaByte	= 1024 MB
1 TeraByte	= 1024 GB
1 Petabyte	= 1024 TB

To remember the order just remember the rhyme:

Peter, **T**he **G**iant **M**onster, **K**ills **B**oy

Storage

- To help you to understand the concept of the amount of storage, study this table:

Unit of Storage	Composed of		Typical files
1 bit	Can be 1 or 0		
1 byte	8 bits		1 character
1 Kilobyte (KB)	2^{10} = 1024 bytes	1024 bytes	Half an A4 page
1 Megabyte (MB)	2^{20} = 1,048,576 bytes	1024 KB	500 A4 pages
1 Gigabyte (GB)	2^{30} = 1,073,741,824 bytes	1024 MB	7 minutes of HD-TV video
1 Terabyte (TB)	2^{40} = 1,099,511,627,776 bytes	1024 GB	472 hours of HD-TV video
1 Petabyte (PB)	2^{50}= 1,125,899,906,842,624 bytes	1024 TB	13.3 years of HD-TV video

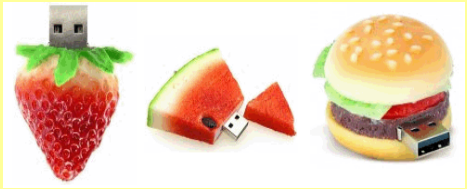
Complete Exercise 4 and get your teacher to check the answers

Types of Storage

- There are three main types of storage used in devices:

- Solid State Storage
- Optical Storage
- Magnetic Storage

Solid State Storage



- Modern computer systems, such as mobile phones and MP3 players, all make use of solid state storage devices.

- These have no moving parts, which means they can:
  - Be smaller - devices they are in can be smaller too.
  - Use less battery power - devices they are in stay charged for longer.
  - Be more robust - a magnetic disk could be ruined by dropping it, because its moving parts may be damaged.

- Examples include:
  - SD cards, used in mobile phones, smartphones and tablets
  - USB flash drives

Optical Storage

- This is the use of laser technology to store data.

- Examples include:
  - CD / DVD-ROM (Read Only Memory)
  - CD / DVD-R (Recordable)
  - CD / DVD-RW (ReWriteable)

Magnetic Storage

- This is the use of a magnetisable coating. Magnetised dots are made on the surface of the material. These dots are created, read and erased by very small electromagnets.

-

- Examples include:
  - Magnetic Tape
  - Hard Disk Drive

Interfaces

- An interface is the combination of hardware and software allow the processor to communicate with an external or internal device such as a printer, modem or hard drive.

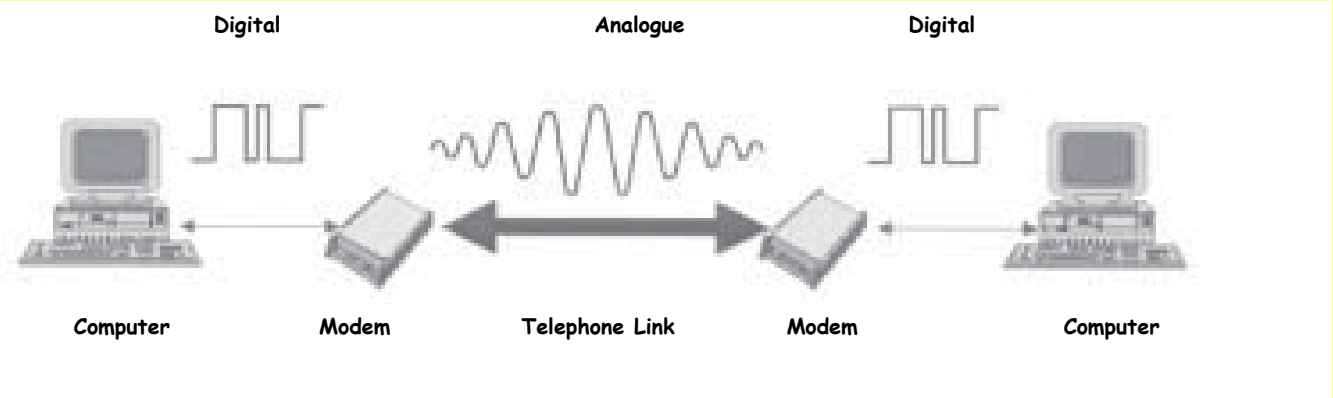
- Sometimes the interface is a board in the computer and sometimes it is a connection to a port.

- The reason that an interface is required is that there are differences in characteristics between the peripheral device and the processor. Those characteristics include:
  - Data Conversion
  - Speed of Operation
  - Temporary Storage of Data
  - Different Voltages

- This means that independent systems can communicate, despite their differing characteristics.

Data Conversion

- The common example of data conversion is when the peripheral accepts an analogue signal that must be converted into digital for the processor to understand it. A modem is typical of this, as shown below.

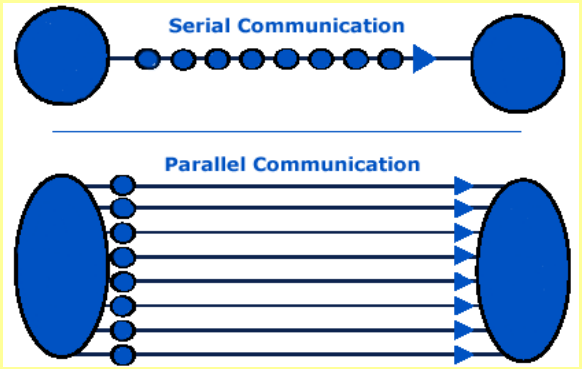


Data Conversion – Serial & Parallel

- In a computer systems data is either transferred in **Serial** or in **Parallel** between devices, eg transferring a key-press on a keyboard into the processor.

- Serial communication means a stream of bits travel along a single channel (one bit after another in single file). Think of this as 8 swimmers swimming a length of a pool but in one only lane, each swimmer follows the next into the pool.

- Parallel communication means a byte(s) of data travel along several channels at the same time. Think of this as 8 swimmers each having their own lane to swim in. All swimmers begin their swim at the same time.



Speed of Operation

- The speed of operation of peripheral devices tends to be in terms of pages per minutes, frames per second or megabytes per second; however, the processor works at a rate in line with its internal clock, which is much faster.

- The speed of the internal operations is measured in gigahertz (Ghz) and a processor will typically work at 2.8 GHz, i.e. 2800,000,000 cycles per second.

- This difference in the speed of operation between the processor and devices requires an interface between the two devices as the processor can deliver data much faster than the peripheral device can handle.

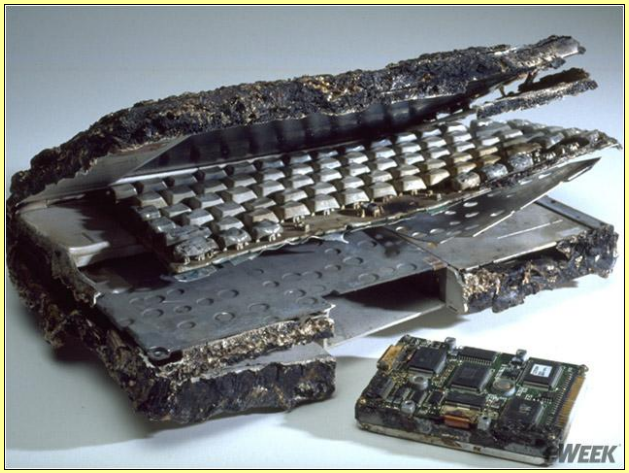
Temporary Storage of Data

- In older computer systems the processor would stand inactive while the printer was finishing a print job.
- One way around this problem is to have the data held temporarily in transit between the processor and the printer.
- Interfaces are used to hold this data, thus releasing the processor; the data is held in a 'buffer'. Another term for buffer is 'spooling'.
- Keyboard characters entered by the user are stored in the keyboard buffer while they are being processed.

Different Voltage Levels

- Some peripherals may use different voltage levels than the processor, for example a keyboard uses a higher voltage than the processor

- The interface must manage these differences in voltage so the hardware does not get damaged



Basic Computer Structure