Cambridge

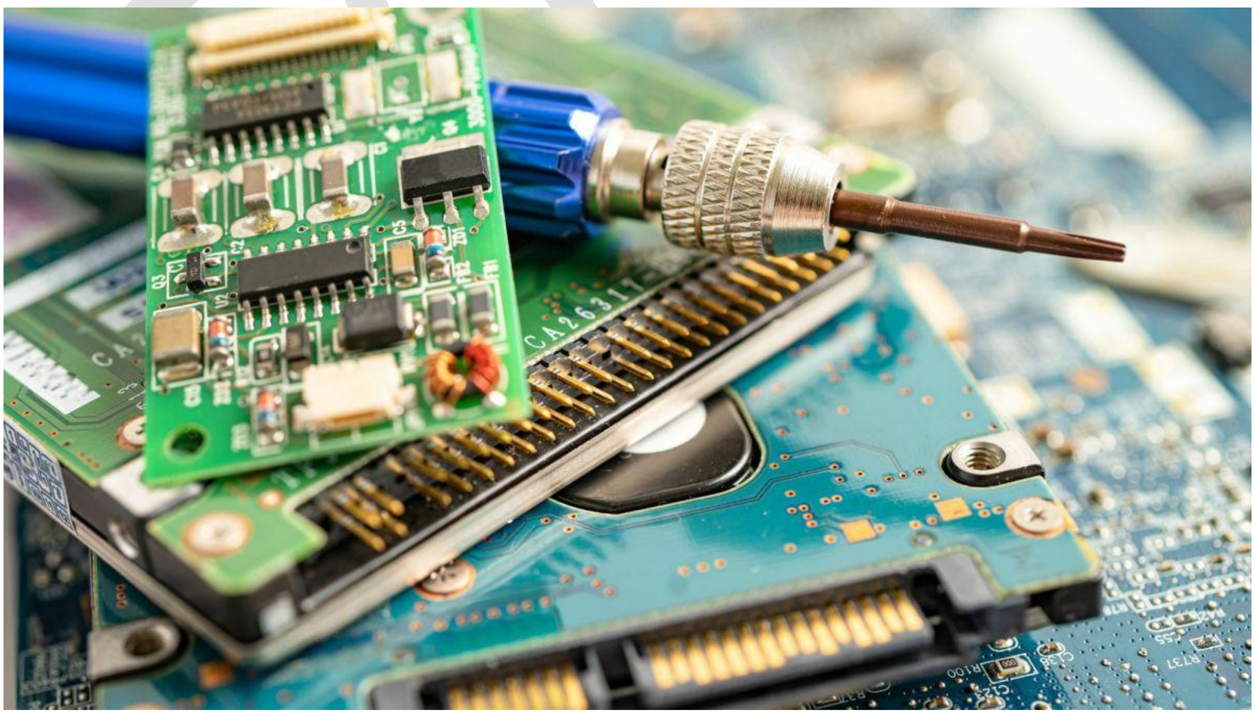
OL- IGCSE

Computer science

CODE: (0478)

Chapter 03

Hardware



3.1 Computer architecture

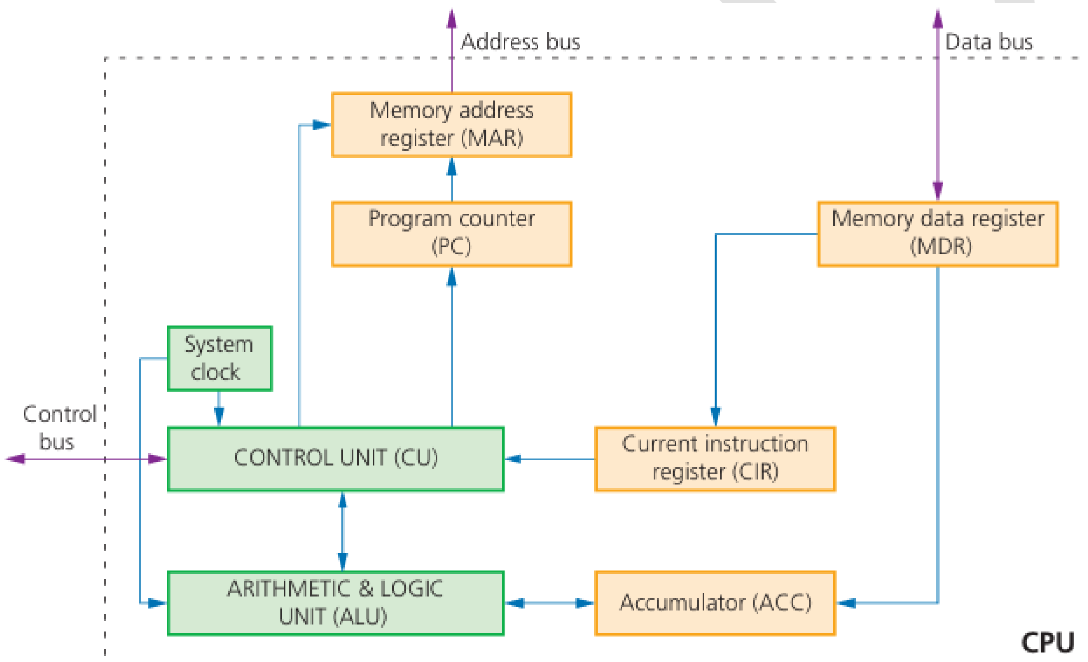
3.1.1 The central processing unit (CPU)

The **central processing unit (CPU**) (also known as a microprocessor or processor) is central to all modern computer systems (including tablets and smartphones). The CPU is very often installed as an **integrated circuit** on a single microchip. As Figure 3.1 shows, the CPU consists of: » control unit (CU) » arithmetic and logic unit (ALU) » registers and buses.

3.1.2 Von Neumann architecture

The von Neumann architecture had the following main novel features (none of which were available in computers prior to the mid-1940s):

» The concept of a central processing unit (CPU or processor)
» The CPU was able to access the memory directly
» Computer memories could store programs as well as data
» Stored programs were made up of instructions which could be executed in sequential order.



▲ **Figure 3.1** Von Neumann architecture

Components of the central processing unit (CPU)

The main components of the CPU are the Control Unit (CU), Arithmetic & Logic Unit (ALU) and system clock.

Arithmetic & Logic Unit (ALU)

The Arithmetic & Logic Unit (ALU) allows the required arithmetic (e.g. +, - and shifting) or logic (e.g. AND, OR) operations to be carried out while a program is being run; it is possible for a computer to have more than one ALU to carry out specific functions.

Control Unit (CU)

The control unit reads an instruction from memory. The address of the location where the instruction can be found is stored in the Program Counter (PC). This instruction is then interpreted using the Fetch–Decode–Execute cycle (see later in this section).

system clock is used to produce timing signals on the control bus to ensure this vital synchronisation takes place – without the clock the computer would simply crash!

The RAM holds all the data and programs needed to be accessed by the CPU. The RAM is often referred to as the **Immediate Access Store (IAS).** The CPU takes data and programs held in **backing store** (e.g. a hard disk drive) and puts them into RAM temporarily.

Registers

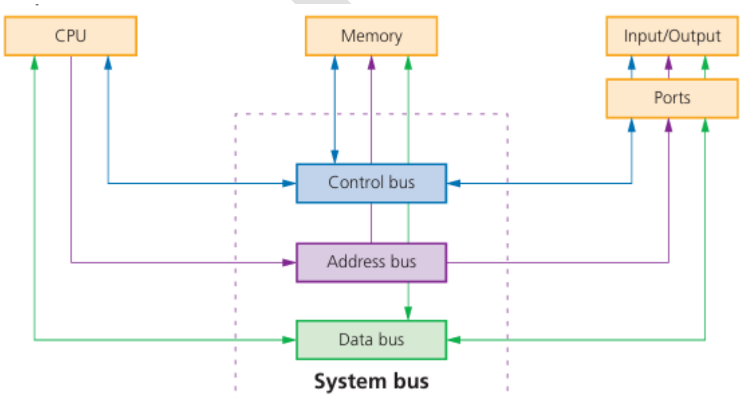
One of the most fundamental components of the von Neumann system are the registers.

▼ **Table 3.1** Specific purpose registers

Register	Abbreviation used	Function/purpose of register
current instruction register	CIR	this register stores the current instruction being decoded and executed
accumulator	ACC	this register is used when carrying out ALU calculations; it stores data temporarily during the calculations
memory address register	MAR	this register stores the address of the memory location currently being read from or written to
memory data/ buffer register	MDR	this register stores data which has just been read from memory or data which is about to be written to memory
program counter	PC	this register stores the address where the next instruction to be read can be found

System buses and memory

Earlier on, Figure 3.1 referred to some components labelled as buses. Figure 3.2 shows how these buses are used to connect the CPU to the memory and to input/ output devices.



▲ **Figure 3.2** System buses and memory

Memory

The computer memory is made up of a number of partitions. Each partition consists of an address and its contents. Table 3.2 uses 8 bits for each address and 8 bits for the content.

The address will uniquely identify every location in the memory and the contents will be the binary value stored in each **location**.

First, consider the READ operation. We will use the memory section shown in Table 3.2. Suppose we want to read the contents of memory location 1111 0001; the two registers are used as follows:

» The address of location 1111 0001 to be read from is first written into the MAR (memory address register):

MAR:	1	1	1	1	0	0	0	1

- a 'read signal' is sent to the computer memory
- the contents of memory location 1111 0001 are then put into the MDR (memory data register):

MDR:	0	1	0	1	1	0	1	1

Now let us now consider the **WRITE** operation. Again, we will use the memory section shown in Table 3.2. Suppose this time we want to show how the value 1001 0101 was written into memory location 1111 1101:

- the data to be stored is first written into the MDR (memory data register):

MDR:	1	0	0	1	0	1	0	1

- this data has to be written into location with address: 1111 1101; so this address is now written into the MAR:

MAR:	1	1	1	1	1	1	0	1

- finally, a 'write signal' is sent to the computer memory and the value 10010101 will then be written into the correct memory location.

Input and output devices

Input devices convert external data into a form the computer can understand and can then process (e.g. keyboards, touch screens and microphones). Output devices show the results of computer processing in a human understandable form (e.g. printers, monitors and loudspeakers)

(System) buses

(System) buses are used in computers as parallel transmission components; each wire in the bus transmits one bit of data.

Address bus

As the name suggests, the **address bus** carries addresses throughout the computer system. Between the CPU and memory, the address bus is **unidirectional.**

Data bus

The **data bus** is **bidirectional** (allowing data to be sent in both directions along the bus). This means data can be carried from CPU to memory (and vice versa) and to and from input/output devices.

As with the address bus, the width of the data bus is important; the wider the bus the larger the **word length** that can be transported.

Control bus

The **control bus** is also bidirectional. It carries signals from the control unit (CU) to all the other computer components.

Fetch–Decode–Execute cycle

To carry out a set of instructions, the CPU first of all **fetches** some data and instructions from memory and stores them in suitable registers. Both the address bus and data bus are used in this process. Once this is done, each instruction needs to be **decoded** before finally being **executed**. This is all known as the **Fetch–Decode–Execute cycle.**

Fetch

Both data and instruction can be stored in MDR. In the **Fetch–Decode–Execute cycle**, the next instruction is fetched from the memory address currently stored in the MAR and the instruction is stored in the MDR.

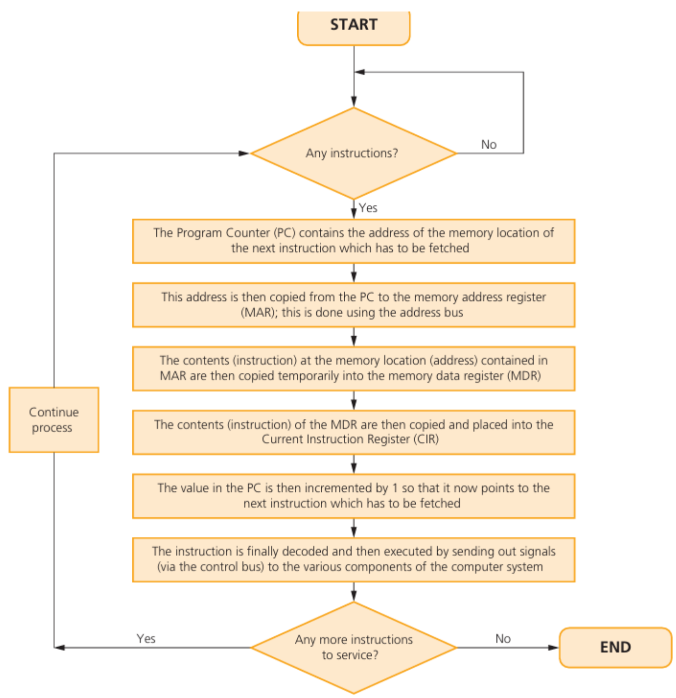
Decode

The instruction is then decoded so that it can be interpreted in the next part of the cycle.

Execute

The CPU passes the decoded instruction as a set of control signals to the appropriate components within the computer system. This allows each instruction to be carried out in its logical sequence.

Figure 3.3 shows how the Fetch–Decode–Execute cycle is carried out in the von Neumann computer model.



▲ **Figure 3.3** Fetch–Decode–Execute cycle flowchart

3.1.3 Cores, cache and internal clock

We will now consider the factors that determine the performance of a CPU. The first thing to consider is the role of the **system clock**. The clock defines the **clock cycle** that synchronises all computer operations.

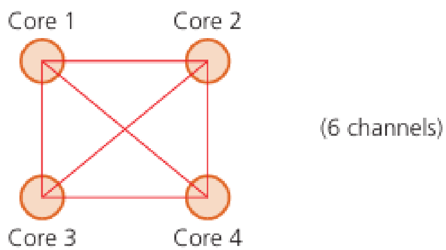
Overclocking is a factor to consider. The clock speed can be changed by accessing the **BIOS (Basic Input/Output System)** and altering the settings. However, using a clock speed higher than the computer was designed for can lead to problems,

The use of cache memories can also improve CPU performance. Unlike RAM, **cache** memory is located within the CPU itself, which means it has much faster data access times than RAM.

For example, with a **dual core** the CPU communicates with both cores using one channel reducing some of the potential increase in its performance:

▲ **Figure 3.4**

while, with a **quad core** the CPU communicates with all four cores using six channels, considerably reducing potential performance:

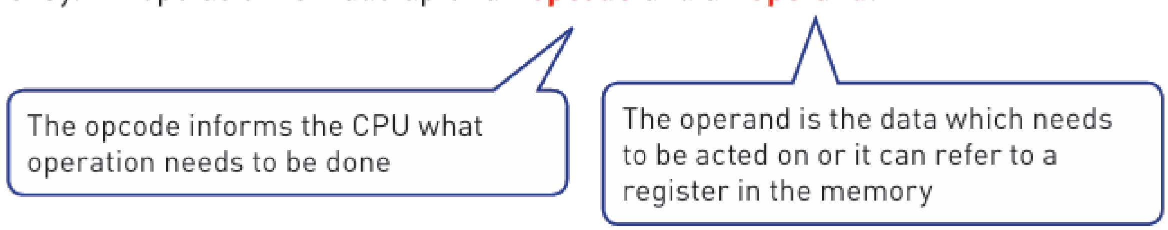


▲ **Figure 3.5**

3.1.4 Instruction set

In a computer system, instructions are a set of operations which are decoded in sequence. Each operation will instruct the ALU and CU (which are part of the CPU). An operation is made up of an **opcode** and **an operand.**

CPU). An operation is made up of an **opcode** and an **operand**.



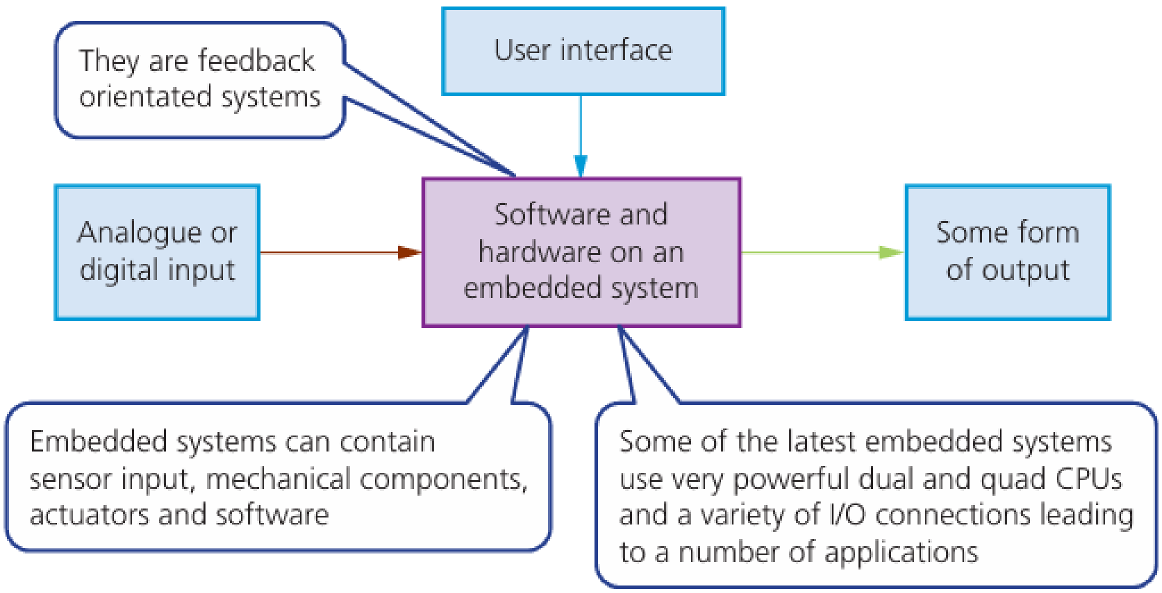
Since the computer needs to understand the operation to be carried out, there is actually a limited number of opcodes that can be used; this is known as the **instruction set.**

3.1.5 Embedded systems

An embedded system is a combination of hardware and software which is designed to carry out a specific set of functions. The hardware is electronic, electrical or electro-mechanical.

microcontrollers:	this has a CPU in addition to some RAM and ROM and other peripherals all embedded onto one single chip (together they carry out a specific task)
microprocessor:	integrated circuit which only has a CPU on the chip (there is no RAM, ROM or peripherals – these need to be added)
system on chips (SoC):	this may contain a microcontroller as one of its components (they almost always will include CPU, memory, input/output (I/O) ports and secondary storage on a single microchip)

An embedded system will have a specific set of tasks; Figure 3.6 summarises how embedded systems work in general:



▲ **Figure 3.6** Embedded systems

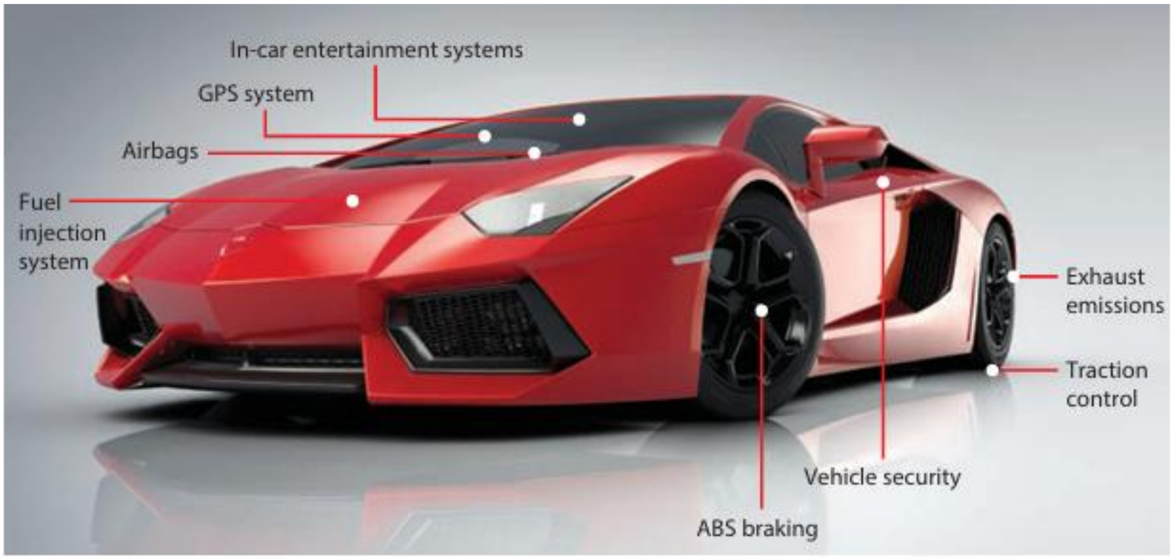
▼ **Table 3.3** Benefits and drawbacks of using embedded systems

Benefits	**Drawbacks**
they are small in size and therefore easy to fit into devices	it can be difficult to upgrade some devices to take advantage of new technology
compared to other systems, they are relatively low cost to make	troubleshooting faults in the device becomes a specialist task
they are usually dedicated to one task allowing simple interfaces and often no requirement for an operating system	although the interface can appear to be more simple (e.g. a single knob) in reality it can be more confusing (e.g. changing the time on a cooker clock can require several steps!)
they consume very little power	any device that can be accessed over the internet is also open to hackers, viruses, etc.
they can be controlled remotely using a mobile phone, for example	due to the difficulty in upgrading and fault finding, devices are often just thrown away rather than being repaired (very wasteful)
very fast reaction to changing input (operate in real time and are feedback orientated)	can lead to an increase in the 'throw away' society if devices are discarded just because they have become out-of-date
with mass production comes reliability	

Examples of the use of embedded systems

Motor vehicles

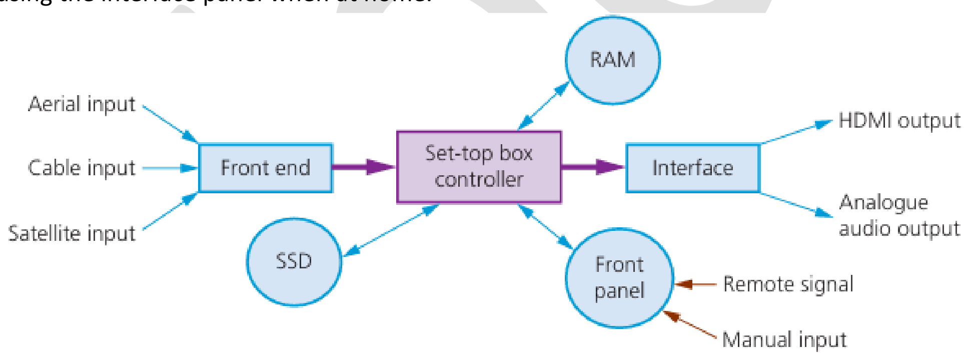
Modern cars have many parts that rely on embedded systems to function correctly. Figure 3.7 shows some of the many components that are controlled in this way.



▲ **Figure 3.7** Embedded systems found in a car

Set-top box

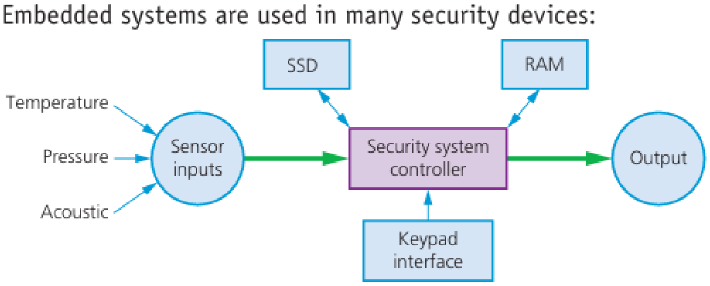
In this example, a set-top box uses an embedded system to allow, for example, recording and playback of television programmes. This can be operated remotely by the user when not at home using an internet-enabled device or by using the interface panel when at home.



▲ **Figure 3.8** Embedded system found in a set-top box

Security systems

Embedded systems are used in many security devices:



▲ **Figure 3.9** Embedded system found in a security system

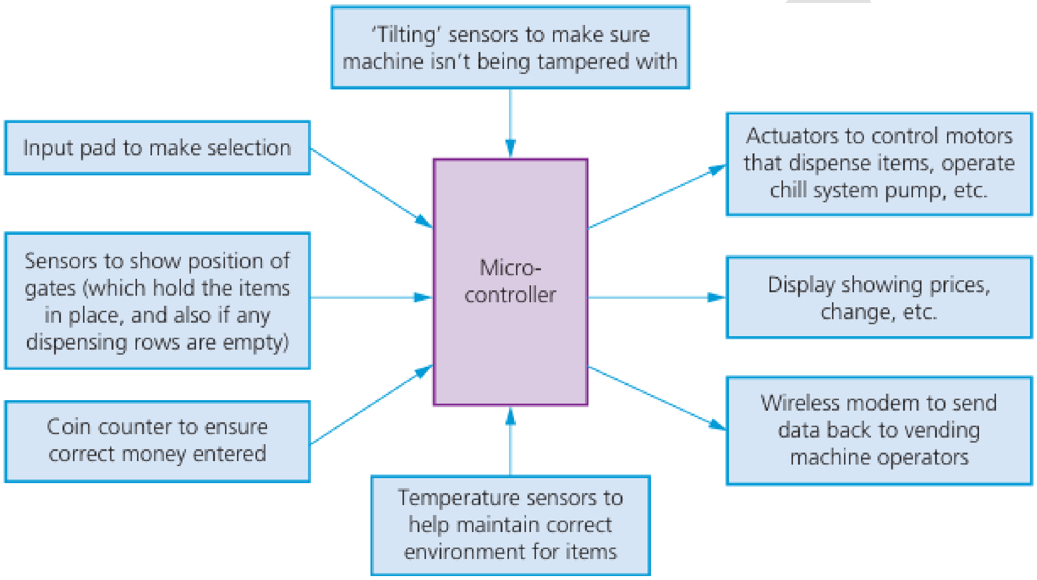
Lighting systems

Embedded systems are used in modern sophisticated lighting systems from simple home use to major architectural lighting systems. We will concentrate here on a lighting system used in a large office. The system needs to control the lighting taking into account:

» The time of day or day of the week
» Whether the room is occupied
» The brightness of the natural light.

Vending systems

Vending machines make considerable use of embedded systems. They usually use microcontrollers to control a number of functions that we all associate with vending machines:



▲ **Figure 3.11** Embedded system found in a vending machine

Inputs to this system come from the keypad (item selection) and from sensors (used to count the coins inserted by the customer, the temperature inside the machine and a 'tilt sensor' for security purposes). The outputs are:

» Actuators to operate the motors, which drive the helixes (see figure below) to give the customers their selected item(s)

» Signals to operate the cooling system if the temperature is too high

» Item description and any change due shown on an LCD display panel

» Data sent back to the vending machine company so that they can remotely check sales activity (which could include instructions to refill the machine) without the need to visit each machine.

▲ **Figure 3.12** Helix used in a typical vending machine

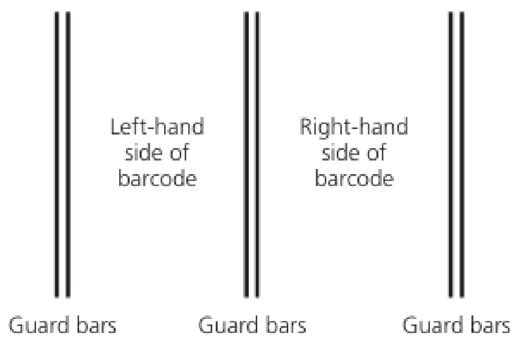
Washing machines

Many 'white goods' (such as refrigerators, washing machines, microwave ovens, and so on) are controlled by embedded systems. They all come with a keypad or dials that are used to select the temperature, wash cycle or cooking duration. This data forms the input to the embedded system, which then carries out the required task without any further human intervention.

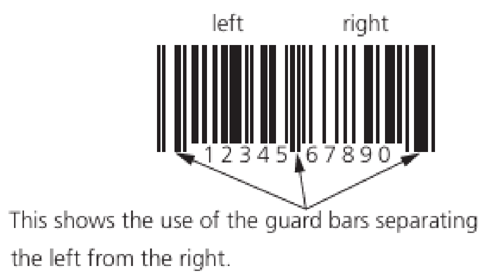
3.2 Input and output devices

3.2.1 Input devices Barcode scanners (readers)

A **barcode** is a series of dark and light parallel lines of varying thickness. The numbers 0 to 9 are each represented by a unique series of lines. Various barcode methods for representing these digits exist.

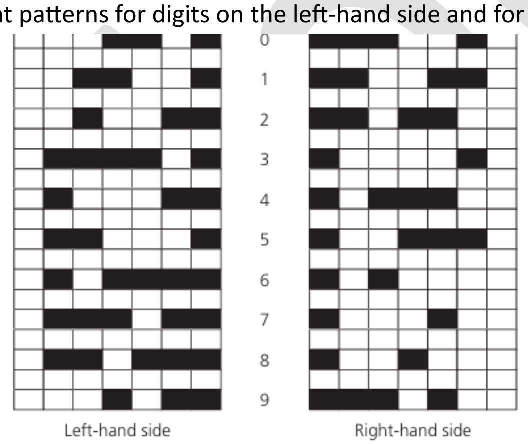


▲ **Figure 3.13** Diagram of guard bars



▲ **Figure 3.14** Sample barcode

Each digit in the barcode is represented by bars of 1 to 4 blocks thick as shown in Figure 3.15. Note there are different patterns for digits on the left-hand side and for digits on the right-hand side.



▲ **Figure 3.15** Barcode digit patterns

The section of barcode to represent the number 5 4 3 0 5 2 would therefore be:

▲ **Figure 3.16** Sample barcode section using patterns from Figure 3.15

So what happens when a barcode is scanned?

- the barcode is first of all read by a red laser or red LED (light emitting diode)
- light is reflected back off the barcode; the dark areas reflect little or no light, which allows the bars to be read
- the reflected light is read by sensors (photoelectric cells)
- as the laser or LED light is scanned across the barcode, a pattern is generated, which is converted into digital data – this allows the computer to understand the barcode
- for example: the digit '3' on the left generates the pattern: L D D D D L D (where L = light and D = dark), this has the binary equivalent of: 0 1 1 1 1 0 1 (where L = 0 and D = 1).

▼ **Table 3.4** Input and output devices at a checkout

Input/output device	How it is used
keypad	to key in the number of same items bought; to key in a weight, to key in the number under the barcode if it cannot be read by the barcode reader/scanner
screen/monitor	to show the cost of an item and other information
speaker	to make a beeping sound every time a barcode is read correctly; but also to make another sound if there is an error when reading the barcode
printer	to print out a receipt/itemised list
card reader/chip and PIN	to read the customer's credit/debit card (either using PIN or contactless)
touchscreen	to select items by touching an icon (such as fresh fruit which may be sold loose without packaging)

So the barcode has been read, then what happens?

» The barcode number is looked up in the stock database (the barcode is known as the key field in the stock item record); this key field uniquely identifies each stock item

» When the barcode number is found, the stock item record is looked up

» The price and other stock item details are sent back to the checkout (or point of sale terminal (POS))

» The number of stock items in the record is reduced by 1 each time the barcode is read

» This new value for number of stock is written back to the stock item record

» The number of stock items is compared to the re-order level; if it is less than or equal to this value, more stock items are automatically ordered.

» Once an order for more stock items is generated, a flag is added to the record to stop re-ordering every time the stock item barcode is read

» When new stock items arrive, the stock levels are updated in the database.

Advantages to the management of using barcodes

» Much easier and faster to change prices on stock items

» Much better, more up-to-date sales information/sales trends

» No need to price every stock item on the shelves (this reduces time and cost to the management)

» Allows for automatic stock control

» Possible to check customer buying habits more easily by linking barcodes to, for example, customer loyalty cards.

Advantages to the customers of using barcodes

» Faster checkout queues (staff don't need to remember/look up prices of items)

» Errors in charging customers is reduced » the customer is given an itemised bill

» Cost savings can be passed on to the customer

» Better track of 'sell by dates' so food should be fresher.

Quick response (QR) codes

Another type of barcode is the **quick response (QR) code**. This is made up of a matrix of filled-in dark squares on a light background. For example, the QR code in Figure 3.17 is a website advertising rock music merchandise. It includes a web address in the code.

Description of QR codes

» A QR code consists of a block of small squares (light and dark) known as pixels. It can presently hold up to 4296 characters (or up to 7089 digits) and also allows internet addresses to be encoded within the QR code. This compares to the 30 digits that is the maximum for a barcode. However, as more and more data is added, the structure of the QR code becomes more complex.

» The three large squares at the corners of the code function as a form of alignment; the remaining small corner square is used to ensure the correct size and correct angle of the camera shot when the QR code is read.

Advantages of QR codes compared to traditional barcodes
» They can hold much more information
» There will be fewer errors; the higher capacity of the QR code allows the use of built-in error-checking systems – normal barcodes contain almost no data redundancy (data which is duplicated) therefore it isn't possible to guard against badly printed or damaged barcodes
» QR codes are easier to read; they don't need expensive laser or LED (light emitting diode) scanners like barcodes – they can be read by the cameras on smartphones or tablets

» It is easy to transmit QR codes either as text messages or images
» It is also possible to encrypt QR codes which gives them greater protection than traditional barcodes.

Disadvantages of QR codes compared to traditional barcodes
» More than one QR format is available
» QR codes can be used to transmit malicious codes – known as attagging. Since there are a large number of free apps available to a user for generating QR codes, that means anyone can do this. It is relatively easy to write malicious code and embed this within the QR code. When the code is scanned, it is possible the creator of the malicious code could gain access to everything on the user's phone

New developments

Newer QR codes (called frame QR codes) are now being used because of the increased ability to add advertising logos (see Figure 3.19).

Digital cameras

Digital cameras have essentially replaced the more traditional camera that used film to capture the images. The film required developing and then printing before the photographer could see the result of their work.

What happens when a photograph is taken

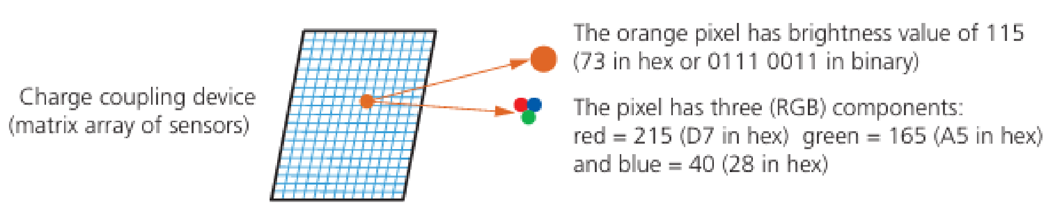
» The image is captured when light passes through the lens onto a light-sensitive cell; this cell is made up of millions of tiny sensors which are acting as photodiodes (i.e. charge couple devices (CCD) which convert light into electricity)

» Each of the sensors are often referred to as pixels (picture elements) since they are tiny components that make up the image

» The image is converted into tiny electric charges which are then passed through an analogue to digital converter (ADC) to form a digital image array

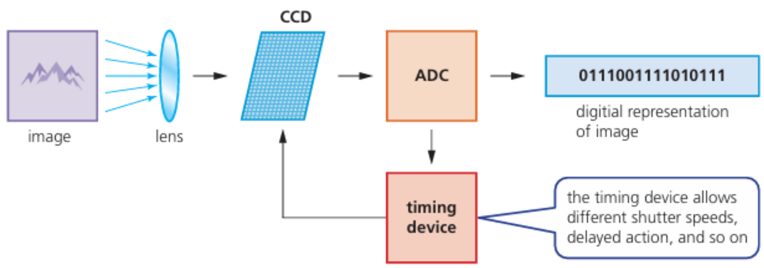
» The ADC converts the electric charges from each pixel into levels of brightness (now in a digital format); for example, an 8-bit ADC gives 28 (256) possible brightness levels per pixel (for example, brightness level 01110011)

» Apart from brightness, the sensors also measure colour which produces another binary pattern; most cameras use a 24-bit RGB system



▲ **Figure 3.21** Typical pixel brightness and colour values

Mobile phones have caught up with digital cameras as regards number of pixels. But the drawback is often inferior lens quality and limited memory for the storage of photos. But this is fast changing and, at the time of writing, many smartphones now have very sophisticated optics and photography software as standard.



▲ **Figure 3.22** Diagram of how a digital camera works

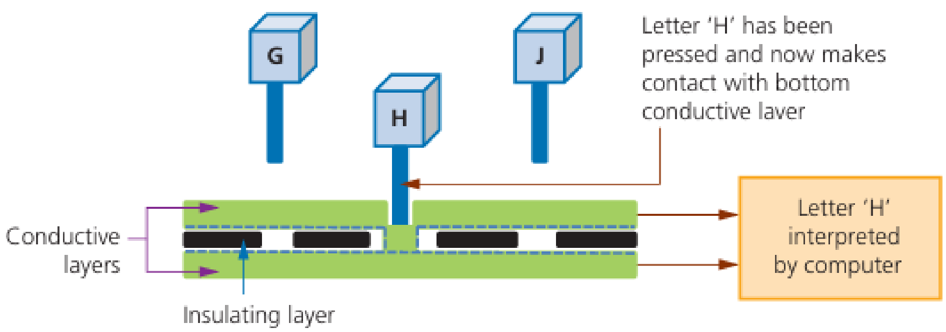
▲ **Figure 3.23** Keyboard

▲ **Figure 3.24** Ergonomic keyboard

Keyboards

The keyboard is connected to the computer either by using a USB connection or by wireless connection. In the case of tablets and mobile phones, the keyboard is often **virtual** or a type of **touch screen** technology.

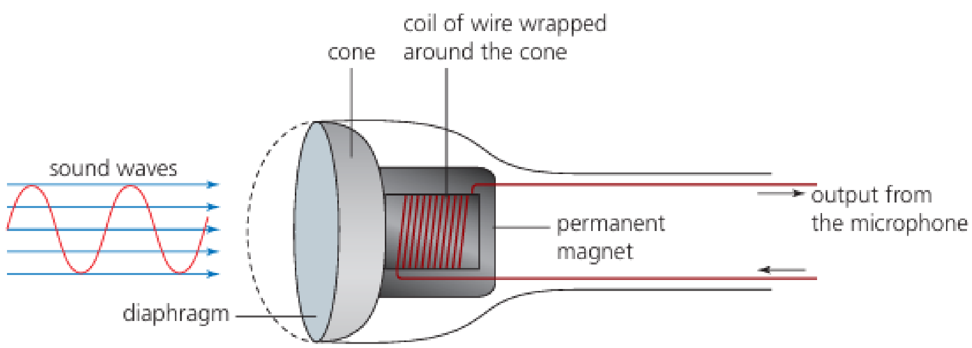
They are a relatively slow method of data entry and are also prone to errors, however keyboards are probably still the easiest way to enter text into a computer. Unfortunately, frequent use of these devices can lead to injuries, such as **repetitive strain injury (RSI)** in the hands and wrists.



▲ **Figure 3.25** Diagram of a keyboard

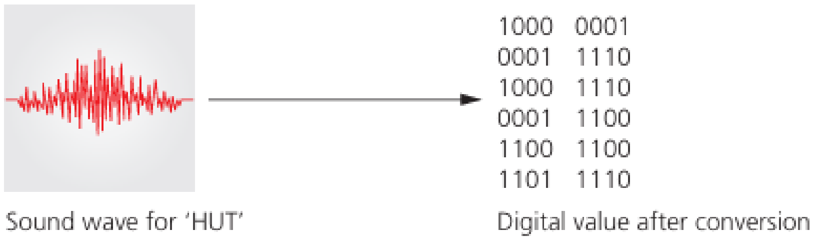
Microphones

Microphones are either built into the computer or are external devices connected through the USB port or using Bluetooth connectivity. Figure 3.26 shows how a microphone can convert sound waves into an electric current.



▲ **Figure 3.26** Diagram of how a microphone works

The electric current output from the microphone can also be sent to a computer where a sound card converts the current into a digital signal which can then be stored in the computer. The following diagram shows what happens when the word 'hut' is picked up by a microphone and is converted into digital values:

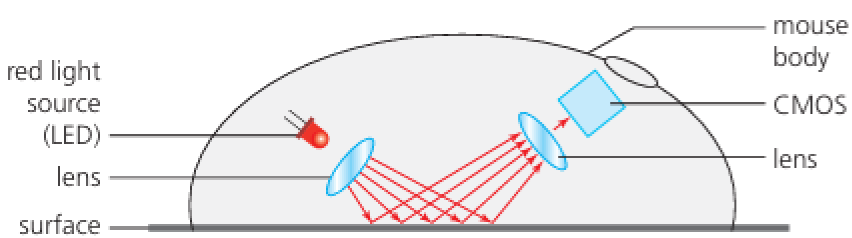


▲ **Figure 3.27** Analogue to digital conversion

Optical mouse

An **optical mouse** is an example of a **pointing device**. It uses tiny cameras to take 1500 images per second. Unlike an older mechanical mouse, the optical mouse can work on virtually any surface.

A red LED is used in the base of the mouse and the red light is bounced off the surface and the reflection is picked up by a **complementary metal oxide semiconductor (CMOS).** The CMOS generates electric pulses to represent the reflected red light and these pulses are sent to a **digital signal processor (DSP).**



▲ **Figure 3.28** Diagram of an optical mouse

Benefits of an optical mouse over a mechanical mouse

» There are no moving parts, therefore it is more reliable.
» Dirt can't get trapped in any of the mechanical components.
» There is no need to have any special surfaces.

Most optical mice use Bluetooth connectivity rather than using a USB wired connection. While this makes the mouse more versatile, a wired mouse has the following advantages:
» Not signal loss since there is a constant signal pathway (wire)
» Cheaper to operate (no need to buy new batteries or charge batteries)
» Fewer environmental issues (no need to dispose of old batteries).

2D and 3D scanners

Scanners are either two dimensional (2D) or three dimensional (3D).

2D scanners

These types of scanner are the most common form and are generally used to input hard copy (paper) documents. The image is converted into an electronic form that can be stored in a computer. A number of stages occur when scanning a document:

The cover is first raised

… then the document is placed on a glass panel …

… and then the cover is closed.

A bright light then illuminates the document – modern scanners use a type of xenon lamp or LED that produces a very bright white light.

A scan head moves across the document until the whole page has been scanned. An image of the document is produced, which is then sent to a lens using a series of mirrors.

The lens focuses the document image.

The focused image now falls onto a charge couple device (CCD), which converts light into an electric current.

Essentially the CCD is made up of thousands of light-sensitive elements (or pixels). Each element creates an electric charge when light falls on it. This means that the scanned image is now turned into an electronic form.

Software produces a digital image from the electronic form.

▲ **Figure 3.29** How a 2D scanner works

Computers equipped with **optical character recognition (OCR)** software allow the scanned text from the document to be converted into a **text file format**. This means the scanned image can now be edited and manipulated by importing it into a word processor.

3D scanners

The scanned images can be used in **computer aided design (CAD)** or, more recently, sent to a 3D printer (see Section 3.2.2) to produce a working model of the scanned image.

Application of 2D scanners at an airport

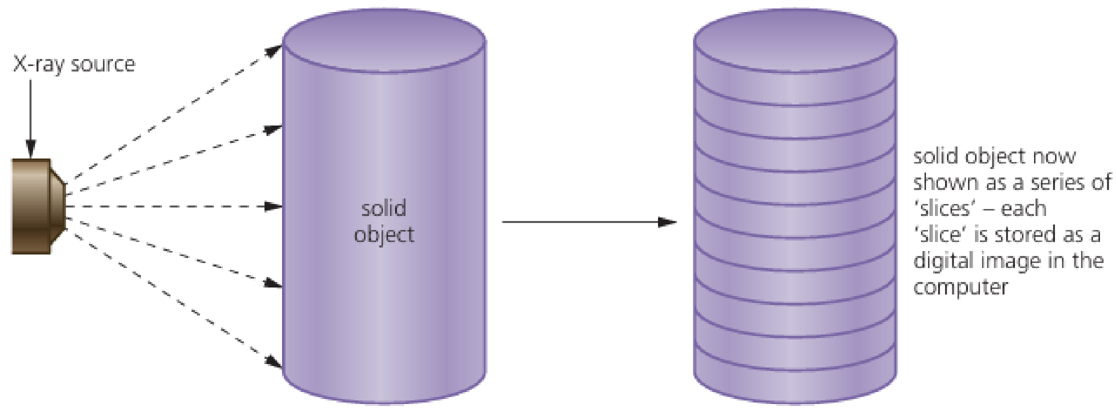
2D scanners are used at airports to read passports. They make use of OCR technology to produce digital images which represent the passport pages. Because of the OCR technology, these digital images can be manipulated in a number of ways.

Application of 3D scanning – computed tomographic (CT) scanners

Computed tomographic (CT) scanners are used to create a 3D image of a solid object. This is based on tomography technology, which basically builds up an image of the solid object through a series of very thin 'slices'. Each of these 2D 'slices' make up a representation of the 3D solid object.

Name	CT Scanner	MRI	SPECT
Stands for	computerised tomography	magnetic resonance images	single photon emission computer tomography
Uses	X-rays	radio frequencies	gamma rays

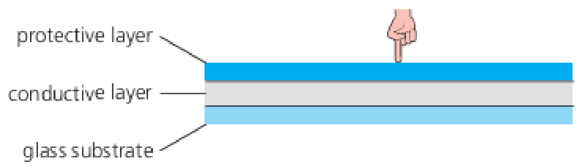
Here is a simple example of how tomography works:



▲ **Figure 3.31** Tomography

Touch screens

Touch screens are now a very common form of input device. They allow simple touch selection from a menu to launch an application (app). Touch screens allow the user to carry out the same functions as they would with a pointing device, such as a mouse.



▲ **Figure 3.32** Capacitive touch screen

There are presently two main types of capacitive touch screens:

» Surface » Projective.

With surface capacitive **screens**, sensors are placed at the corners of a screen. Small voltages are also applied at the corners of the screen creating an electric field. A finger touching the screen surface will draw current from each corner reducing the capacitance.

Projective capacitive screens work slightly differently to surface capacitive screens. The transparent conductive layer is now in the form of an X-Y matrix pattern.

Advantages compared to the other two technologies

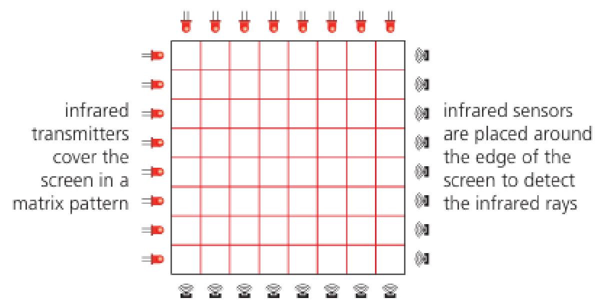
» Better image clarity than resistive screens, especially in strong sunlight
» Very durable screens that have high scratch resistance
» Projective capacitive screens allow multi-touch.

Disadvantages compared to the other two technologies

» Surface capacitive screens only work with bare fingers or a special stylus
» They are sensitive to electromagnetic radiation (such as magnetic fields or microwaves).

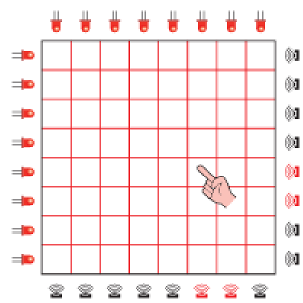
Infrared touch screens

Infrared touch screens use a glass screen with an array of sensors and infrared transmitters



▲ **Figure 3.33** Array of infrared transmitters and sensors surrounding the screen

The sensors detect the infrared radiation. If any of the infrared beams are broken (for example, with a finger touching the screen), the infrared radiation reaching the sensors is reduced. The sensor readings are sent to a microcontroller that calculates where the screen was touched:



◄ **Figure 3.34** Infrared screen touched causing sensors (shown in red) to show a reduction in infrared radiation – thus the exact position where the screen was touched can be calculated

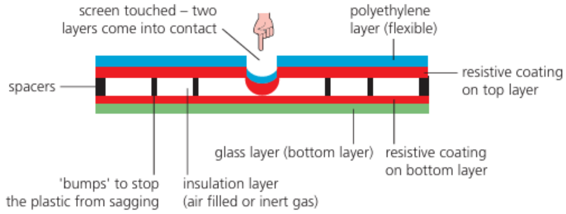
Advantages compared to the other two technologies
» Allows multi-touch facilities
» Has good screen durability
» The operability isn't affected by a scratched or cracked screen.

Disadvantages compared to the other two technologies

» The screen can be sensitive to water or moisture
» It is possible for accidental activation to take place if the infrared beams are disturbed in some way
» Sometimes sensitive to light interference.

Resistive touch screens

Resistive touch screens are made up of two layers of electrically resistive material with a voltage applied across them. The upper layer is made of flexible polyethylene (a type of polymer) with a resistive coating on one side (see Figure 3.35). The bottom layer is made of glass also with a resistive coating (usually indium tin oxide) on one side.



▲ **Figure 3.35** Resistive touch screen

A microcontroller converts the voltage (created when the two resistive layers touch) to digital data, which it then sends to the microprocessor.

Advantages compared to the other two technologies

- Good resistance to dust and water
- Can be used with bare fingers, stylus and gloved hand.

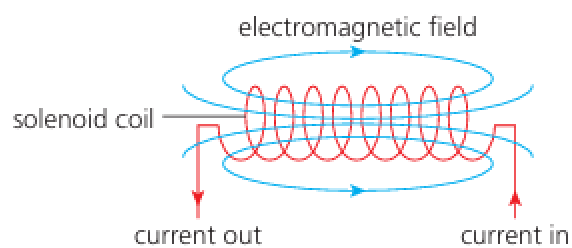
Disadvantages compared to the other two technologies

- Low touch sensitivity (sometimes have to press down harder)
- Doesn't support multi-touch facility
- Poor visibility in strong sunlight
- Vulnerable to scratches on the screen (made of polymer).

3.2.2 Output devices

Actuators

When a computer is used to control devices, such as a conveyer belt or a valve, it is usually necessary to use an actuator to, for example, start/stop the conveyer belt or open/close the valve.



▲ **Figure 3.36** A solenoid

Light projectors

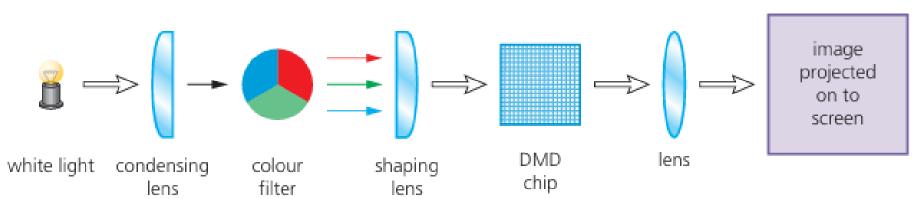
There are two common types of light projector:

» Digital light projector (DLP)

» Liquid crystal display (LCD) projector.

Digital light projectors (DLP)

The use of millions of micro mirrors on a small digital micromirror device (DMD chip) is the key to how these devices work.



▲ **Figure 3.37** A digital light projector (DLP)

Liquid crystal display (LCD) projector

These are older technology than DLP. Essentially a high-intensity beam of light passes through an LCD display and then onto a screen. How this works in principle is described below:

» A powerful beam of white light is generated from a bulb or LED inside the projector body

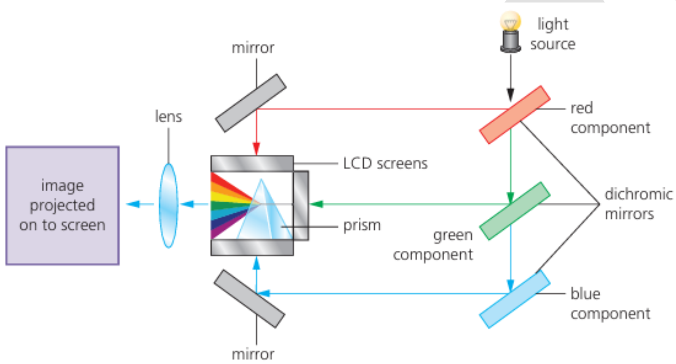
» This beam of light is then sent to a group of chromatic-coated mirrors (known as dichromic mirrors); these reflect the light back at different wavelengths

» When the white light hits these mirrors, the reflected light has wavelengths corresponding to red, green and blue light components

» these three different coloured light components pass through three LCD screens (each screen is composed of thousands of tiny pixels which can either block light or let it through; this produces a monochromatic image)...

» ... consequently, three different versions of the same image are now produced – one is the whole image in different shades of red, one is the whole image in different shades of green and one is the whole image in different shades of blue

» These images are then re-combined using a special prism to produce a full colour image » finally, the image passes through the projector lens onto a screen.



▲ **Figure 3.38** LCD projector

Advantages and disadvantages of the two types of projector

▼ **Table 3.5** Advantages and disadvantages of DLP and LCD projectors

	Advantages	**Disadvantages**
Digital light projector (DLP)	higher contrast ratios	image tends to suffer from 'shadows' when showing a moving image
	higher reliability/longevity	
	quieter running than LCD projector	DLP do not have grey components in the image
	uses a single DMD chip, which mean no issues lining up the images	the colour definition is frequently not as good as LCD projectors because the colour saturation is not as good (colour saturation is the intensity of a colour)
	smaller and lighter than LCD projector	
	they are better suited to dusty or smoky atmospheres than LCD projectors	
LCD projector	give a sharper image than DLP projectors	although improving, the contrast ratios are not as good as DLPs
	have better colour saturation than DLP projectors	LCD projectors have a limited life (that is, the longevity is not as good as DLPs)
	more efficient in their use of energy than DLP technology – consequently they generate less heat	since LCD panels are organic in nature, they tend to degrade with time (screens turn yellow and the colours are subsequently degraded over time)

Inkjet and laser printers

Inkjet printers

Inkjet printers are essentially made up of:

» A print head, which consists of nozzles that spray droplets of ink onto the paper to form characters

» An ink cartridge or cartridges; either one cartridge for each colour (blue, yellow and magenta) and a black cartridge or one single cartridge containing all three colours + black (Note: some systems use six colours)

» A stepper motor and belt, which moves the print head assembly across the page from side to side

» A paper feed, which automatically feeds the printer with pages as they are required.

The ink droplets are produced currently using two different technologies:

Thermal bubble – tiny resistors create localised heat which makes the ink vaporise. This causes the ink to form a tiny bubble; as the bubble expands, some of the ink is ejected from the print head onto the paper.

Piezoelectric – a crystal is located at the back of the ink reservoir for each nozzle. The crystal is given a tiny electric charge which makes it vibrate. This vibration forces ink to be ejected onto the paper; at the same time more ink is drawn in for further printing.

▼ **Table 3.6** Steps in inkjet printing process

Stage in process	Description of what happens
1	the data from the document is sent to a printer driver
2	the printer driver ensures that the data is in a format that the chosen printer can understand
3	a check is made by the printer driver to ensure that the chosen printer is available to print (e.g. is it busy, is it off-line, is it out of ink, and so on)
4	the data is then sent to the printer and it is stored in a temporary memory known as a printer buffer
5	a sheet of paper is then fed into the main body of the printer; a sensor detects whether paper is available in the paper feed tray – if it is out of paper (or the paper is jammed) then an error message is sent back to the computer
6	as the sheet of paper is fed through the printer, the print head moves from side to side across the paper printing the text or image; the four ink colours are sprayed in their exact amounts to produce the desired final colour
7	at the end of each full pass of the print head, the paper is advanced very slightly to allow the next line to be printed; this continues until the whole page has been printed
8	if there is more data in the printer buffer, then the whole process from stage 5 is repeated until the buffer is finally empty
9	once the printer buffer is empty, the printer sends an interrupt to the CPU in the computer; this is a request for more data to be sent to the printer; the whole process continues until the whole of the document has been printed

Laser printers

Laser printers use dry powder ink rather than liquid ink and make use of the properties of static electricity to produce the text and images. Unlike inkjet printers, laser printers print the whole page in one go. Colour laser printers use 4 toner cartridges – blue, cyan, magenta and black.

▲ **Figure 3.40** Laser printer

▼ **Table 3.7** Steps in laser printing process

Stage in process	Description of what happens
1	the data from the document is sent to a printer driver
2	the printer driver ensures that the data is in a format that the chosen printer can understand
3	a check is made by the printer driver to ensure that the chosen printer is available to print (e.g. is it busy, is it off-line, is it out of ink, and so on)
4	the data is then sent to the printer and it is stored in a temporary memory known as a printer buffer
5	the start of the printing process involves a printing drum being given a positive charge; as this drum rotates, a laser beam is scanned across it removing the positive charge in certain areas; this leaves negatively charged areas that exactly match the text/images of the page to be printed
6	the drum is then coated with positively charged toner (powdered ink); since the toner is positively charged, it only sticks to the negatively charged parts of the drum
7	a negatively charged sheet of paper is then rolled over the drum
8	the toner on the drum now sticks to the paper to produce an exact copy of the page sent to the printer
9	to prevent the paper sticking to the drum, the electric charge on the paper is removed after one rotation of the drum
10	the paper finally goes through a fuser which is a set of heated rollers; the heat melts the ink so that it fixes permanently to the paper
11	at the very end, a discharge lamp removes all the electric charge from the drum making it ready to print the next page

Applications of inkjet and laser printers

The choice of whether to use an inkjet printer or a laser printer depends on which features make it the most appropriate output device for the given application.

Inkjet printer – inkjet printers are often used for printing one-off photos or where only a few pages of good quality, colour printing is needed; the small ink cartridges or small paper trays would not be an issue with such applications.

Laser printer – these devices produce high quality printouts and are very fast when making multiple copies of a document; any application that needs high volume printing (in colour or monochrome) would choose the laser printer

3D printers

3D printers are used to produce solid objects that actually work. They are primarily based on inkjet and laser printer technology.

It was made from many layers (0.1 mm thick) of powdered metal using a technology known as **binder 3D printing.**

The following information describes some of the features of 3D printing:

» Various types of 3D printers exist; they range from the size of a microwave oven up to the size of a small car.

» 3D printers use additive manufacturing (i.e. the object is built up layer by layer); this is in sharp contrast to the more traditional method of subtractive manufacturing (i.e. removal of material to make the object).

» **Direct 3D printing** uses inkjet technology; a print head can move left to right as in a normal printer. However, the print head can also move up and down to build up the layers of an object.

» **Binder 3D printing** is similar to direct 3D printing. However, this method uses two passes for each of the layers; the first pass sprays dry powder and then on the second pass a binder (a type of glue) is sprayed to form a solid layer.

» Newer technologies are using lasers and UV light to harden liquid polymers; this further increases the diversity of products which can be made.

How to create a solid object using 3D printers

There are a number of steps in the process of producing an object using 3D printers. The steps are summarised below:

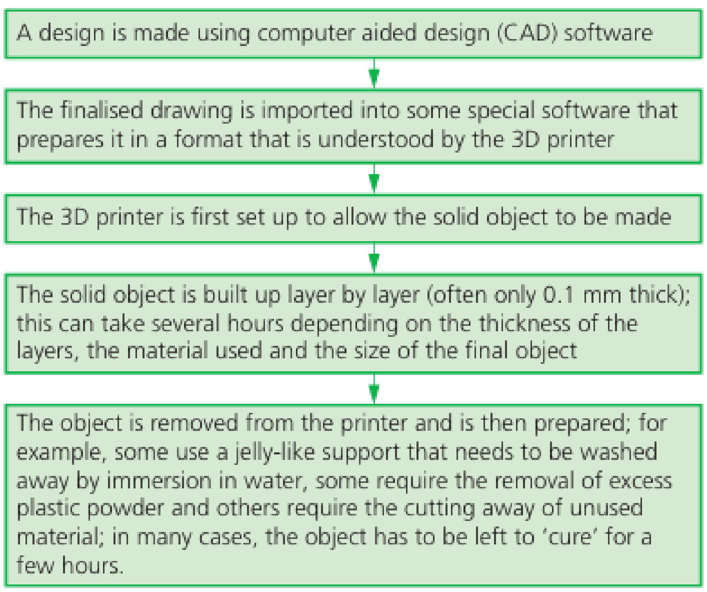


Figure 3.43 How to create an object using a 3D printer

Uses of 3D printing

3D printing is regarded as being possibly the next 'industrial revolution' since it will change the manufacturing methods in many industries.

LED and LCD screens

LED screens

An LED screen is made up of tiny light emitting diodes (LEDs). Each LED is either red, green or blue in colour. By varying the electric current sent to each LED, its brightness can be controlled, producing a vast range of colours.

The reader needs to be very careful here. Many television screens are advertised as LED when in fact they are LCD screens which are backlit using LEDs.

LCD screens

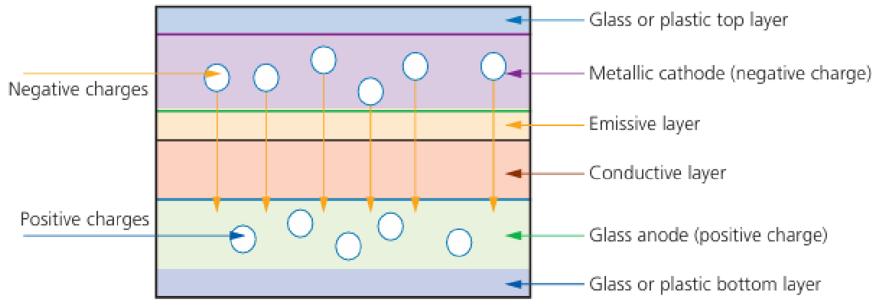
LCD screens are made up of tiny liquid crystals. These tiny crystals make up an array of pixels that are affected by changes in applied electric fields. How this works is outside the scope of this book.

LEDs have become increasingly more popular, as the method of back lighting, due to a number of advantages over older CCFL technology:

» LEDs reach their maximum brightness almost immediately (there is no need to 'warm up' before reaching full efficiency)
» LEDs give a whiter light that sharpens the image and makes the colours appear more vivid; CCFL had a slightly yellowish tint
» LEDs produce a brighter light that improves the colour definition » monitors using LED technology are much thinner than monitors using CCFL technology
» LEDs last indefinitely; this makes the technology more reliable and makes for a more consistent product
» LEDs consume very little power which means they produce less heat as well as using less energy.

Organic light emitting diodes (OLED)

Newer LED technology is making use of **organic light emitting diodes (OLEDs**). These use organic materials (made up of carbon compounds) to create semi-conductors that are very flexible. Organic films are sandwiched between two charged electrodes (one is a metallic **cathode** and the other a glass **anode**).



▲ **Figure 3.44** How an OLED screen works

Advantages of using OLED compared to existing LEDs and LCDs:

» The plastic, organic layers of an OLED are thinner, lighter and more flexible than the crystal structures used in LEDs or LCDs.

» The light-emitting layers of an OLED are lighter; OLED layers can be made from plastic rather than the glass as used in LED and LCD screens.

» OLEDs give a brighter light than LEDs. » OLEDs do not require backlighting like LCD screens – OLEDs generate their own light.

» Since OLEDs require no backlighting, they use much less power than LCD screens (most of the LCD power is used to do the backlighting); this is very important in battery-operated devices such as mobile phones.

» Since OLEDs are essentially plastics, they can be made into large, thin sheets (this means they could be used on large advertising boards in airports, subways, and so on).

» OLEDs have a very large field of view, about 170 degrees, which makes them ideal for use in television sets and for advertising screens.

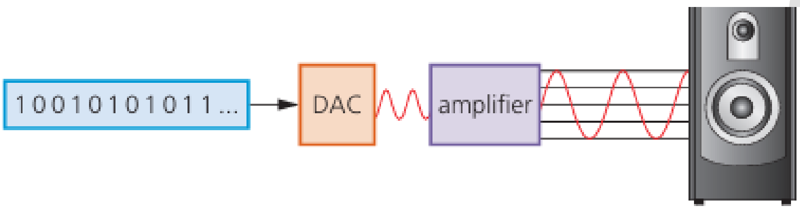
(Loud) speakers

Loudspeakers are output devices that produce sound. When connected to a computer system, digitised sound stored on a file needs to be converted into sound as follows:

» The digital data is first passed through a digital to analogue converter (DAC) where it is changed into an electric current.

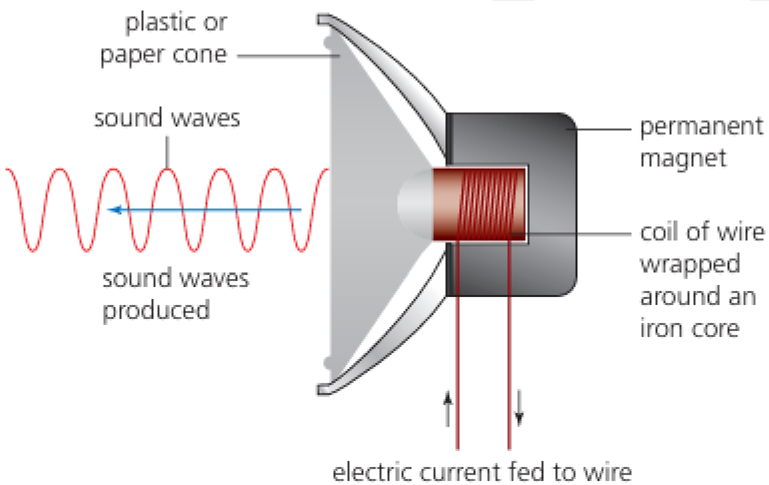
» This is then passed through an amplifier (since the current generated by the DAC will be very small); this creates a current large enough to drive a loudspeaker.

» This electric current is then fed to a loudspeaker where it is converted into sound. The following schematic shows how this is done:



▲ **Figure 3.46** Digital to analogue conversion

As Figure 3.46 shows, if the sound is stored in a computer file, it must pass through a **digital to analogue converter (DAC**) to convert binary (digital) data into an analogue form (electric current) that can then drive the loudspeaker.

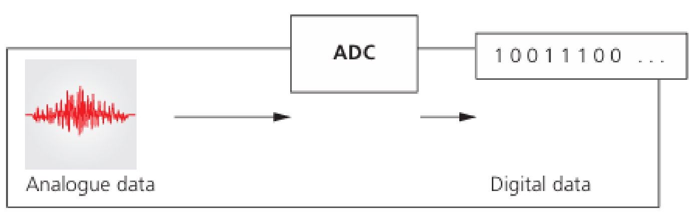


▲ **Figure 3.47** Diagram showing how a loudspeaker works

3.2.3 Sensors

Sensors are input devices which read or measure physical properties from their surroundings.

However, computers cannot make any sense of these physical quantities so the data needs to be converted into a digital format. This is usually achieved by an **analogue to digital converter (ADC).** This device converts physical values into discrete digital values.

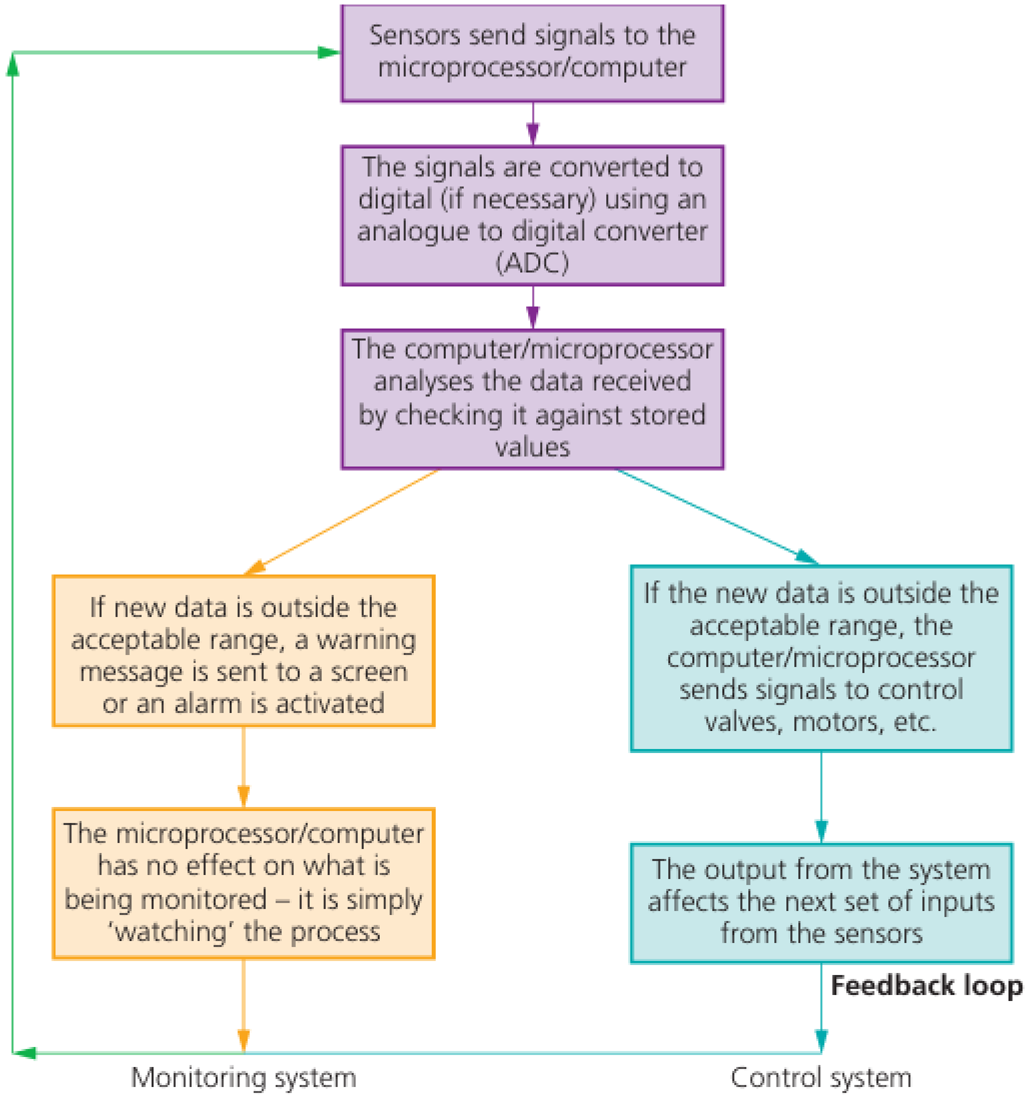


▲ **Figure 3.49** ADC

▼ **Table 3.8** Sensors

Sensor	Description of sensor	Example applications
Temperature	measures temperature of the surroundings by sending signals; these signals will change as the temperature changes	• control of a central heating system • control/monitor a chemical process • control/monitor temperature in a greenhouse
Moisture	measures water levels in, for example, soil (it is based on the electrical resistance of the sample being monitored)	• control/monitor moisture levels in soil in a greenhouse • monitor the moisture levels in a food processing factory
Humidity	this is slightly different to moisture; this measures the amount of water vapour in, for example, a sample of air (based on the fact that the conductivity of air will change depending on the amount of water present)	• monitor humidity levels in a building • monitor humidity levels in a factory manufacturing microchips • monitor/control humidity levels in the air in a greenhouse
Light	these use photoelectric cells that produce an output (in the form of an electric current) depending on the brightness of the light	• switching street lights on or off depending on light levels • switch on car headlights automatically when it gets dark
Infrared (active)	these use an invisible beam of infrared radiation picked up by a detector; if the beam is broken, then there will be a change in the amount of infrared radiation reaching the detector (sensor)	• turn on car windscreen wipers automatically when it detects rain on the windscreen • security alarm system (intruder breaks the infra-red beam)
Infrared (passive)	these sensors measure the heat radiation given off by an object, for example, the temperature of an intruder or the temperature in a fridge	• security alarm system (detects body heat) • monitor the temperature inside an industrial freezer or chiller unit
Pressure	a pressure sensor is a transducer and generates different electric currents depending on the pressure applied	• weighing of lorries at a weighing station • measure the gas pressure in a nuclear reactor
Acoustic/sound	these are basically microphones that convert detected sound into electric signals/pulses	• pick up the noise of footsteps in a security system • detect the sound of liquids dripping at a faulty pipe joint
Gas	most common ones are oxygen or carbon dioxide sensors; they use various methods to detect the gas being monitored and produce outputs that vary with the oxygen or carbon dioxide levels present	• monitor pollution levels in the air at an airport • monitor oxygen and carbon dioxide levels in a greenhouse • monitor oxygen levels in a car exhaust
pH	these measure acidity through changes in voltages in, for example, soil	• monitor/control acidity levels in the soil in a greenhouse • control acidity levels in a chemical process
Magnetic field	these sensors measure changes in magnetic fields – the signal output will depend on how the magnetic field changes	• detect magnetic field changes (for example, in mobile phones and CD players) • used in anti-lock braking systems in cars
Accelerometer	these are sensors that measure acceleration and motion of an application, i.e. the change in velocity (a piezoelectric cell is used whose output varies according to the change in velocity)	• used in cars to measure rapid deceleration and apply air bags in a crash • used by mobile phones to change between portrait and landscape mode
Proximity	these sensors detect the presence of a nearby object	• detect when a face is close to a mobile phone screen and switches off screen when held to the ear
Flow (rate)	these sensors measure the flow rate of a moving liquid or gas and produce an output based on the amount of liquid or gas passing over the sensor	• used in respiratory devices and inhalers in hospitals • measure gas flows in pipes (for example, natural gas)
Level	these sensors use ultrasonics (to detect changing liquid levels in, for example, a tank) or capacitance/conductivity (to measure static levels (for example, height of water in a river) – note, level sensors can also be optical or mechanical in nature	• monitor levels in a petrol tank in a car • in a pharmaceutical process where powder levels in tablet production need to be monitored • leak detection in refrigerant (air conditioning)

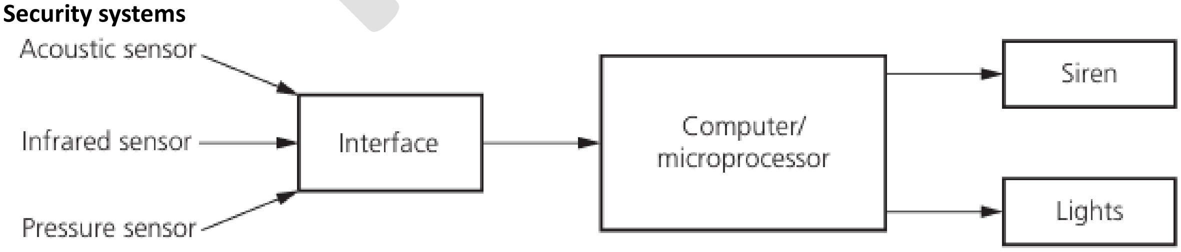
Sensors are used in both monitoring and control applications. There is a subtle difference between how these two methods work (the flowchart is a simplification of the process):



▲ **Figure 3.50** Monitoring and control systems using sensors

Monitoring applications

Security systems



▲ **Figure 3.51** Security system

The security monitoring system will carry out the following actions:
» The system is activated by keying in a password on a keypad
» The **infrared sensor** will pick up the movement of an intruder in the building
» The **acoustic sensor** will pick up sounds such as footsteps or breaking glass
» The **pressure sensor** will pick up the weight of an intruder coming through a door or through a window

Control applications

Control of street lighting

This next sequence shows how a microprocessor is used to control the operation of a street lamp. The lamp is fitted with a light sensor which constantly sends data to the microprocessor.

Anti-lock braking systems (on cars)

Anti-lock braking systems (ABS) on cars use magnetic field sensors to stop the wheels locking up on the car if the brakes have been applied too sharply:

» when one of the car wheels rotates too slowly (i.e. it is locking up), a magnetic field sensor sends data to a microprocessor

» the microprocessor checks the rotation speed of the other three wheels » if they are different (i.e. rotating faster), the microprocessor sends a signal to the braking system …

» … and the braking pressure to the affected wheel is reduced …

» … the wheel's rotational speed is then increased to match the other wheels

» the checking of the rotational speed using these magnetic field sensors is done several times a second …

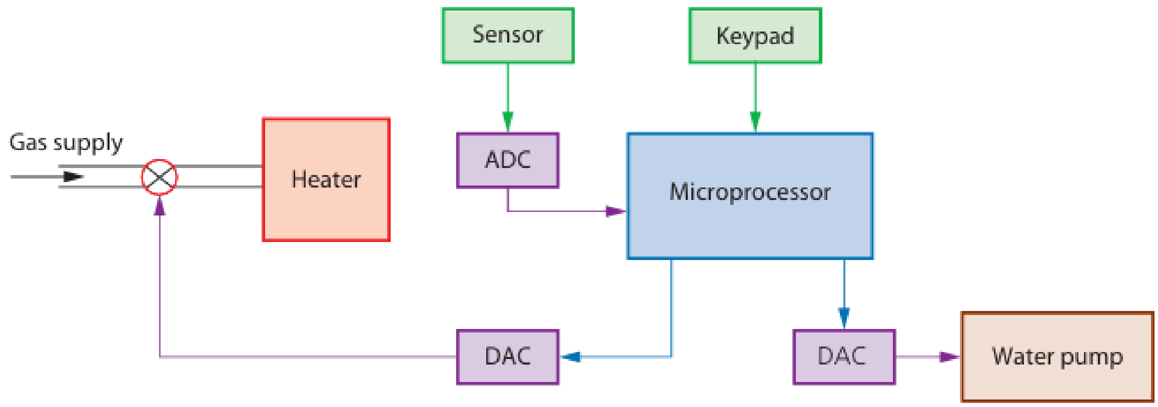
» … and the braking pressure to all the wheels can be constantly changing to prevent any of the wheels locking up under heavy braking ...

» … this is felt as a 'judder' on the brake pedal as the braking system is constantly switched off and on to equalise the rotational speed of all four wheels

» if one of the wheels is rotating too quickly, braking pressure is increased to that wheel until it matches the other three.

Central heating systems

In this example, a gas supply is used to heat water using a heater. A valve on the gas supply is controlled by a microprocessor and is opened if the heating levels need to be increased.



▲ **Figure 3.53** Controlling a central heating system

Chemical process control

The following description shows how the sensors and computer are used to control this process:

» **Temperature and pH sensors** read data from the chemical process

» This data is converted to digital using an ADC and is then sent to the computer » the computer compares the incoming data with pre-set values stored in memory

» … if the temperature < 70°C, a signal is sent to switch on the heater

» … if the temperature >= 70°C, a signal is sent to switch off the heaters

» … if the pH > 3.5, then a signal is sent to open a valve and acid is added

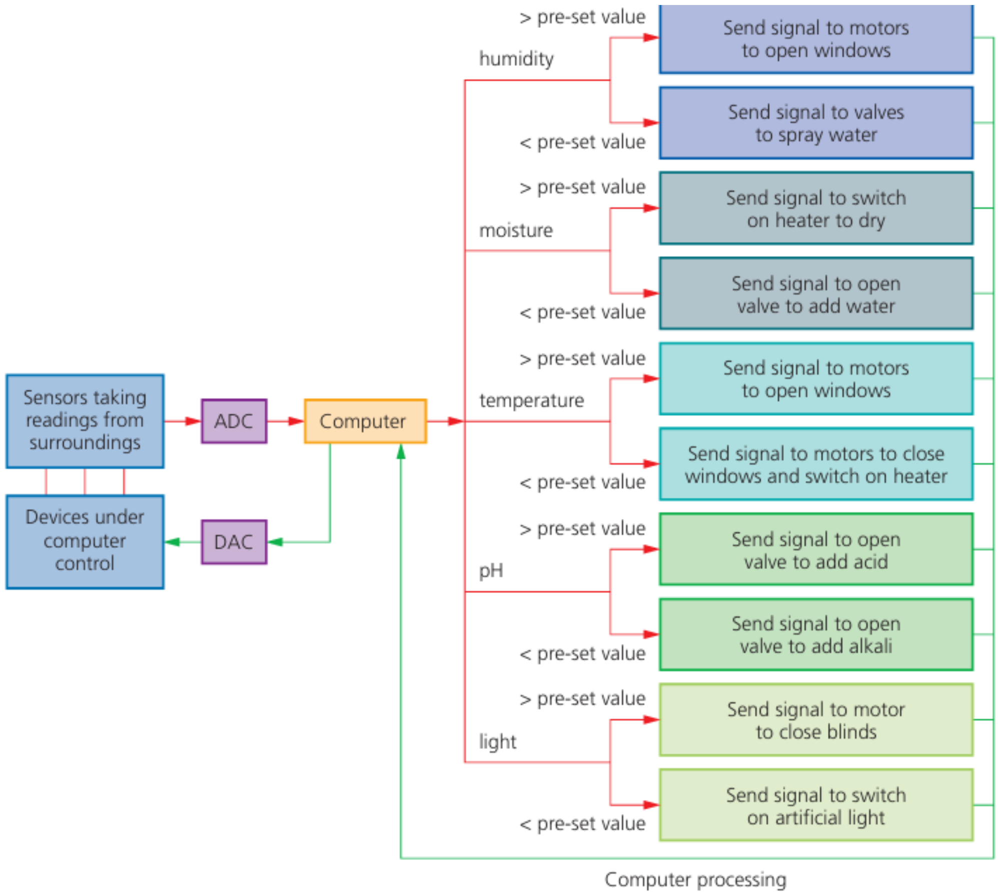
» … if the pH <= 3.5, then a signal is sent to close this valve

» The computer signals will be changed into analogue signals using a DAC so that it can control the heaters and valves

» This continues as long as the computer system is activated.

Greenhouse environment control

Five different sensors could be used here to control the greenhouse environment, namely: **humidity, moisture**, **temperature, pH and light**. To simplify this problem the control mechanisms are shown in Figure 3.54.



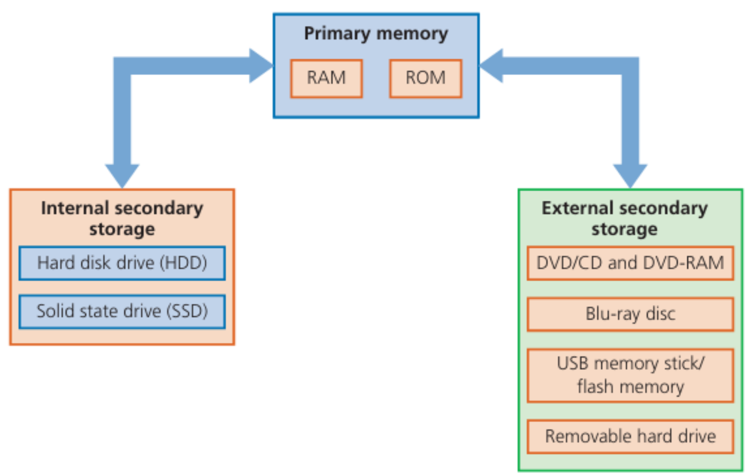
▲ **Figure 3.54** Control of greenhouse environment

3.3 Data storage

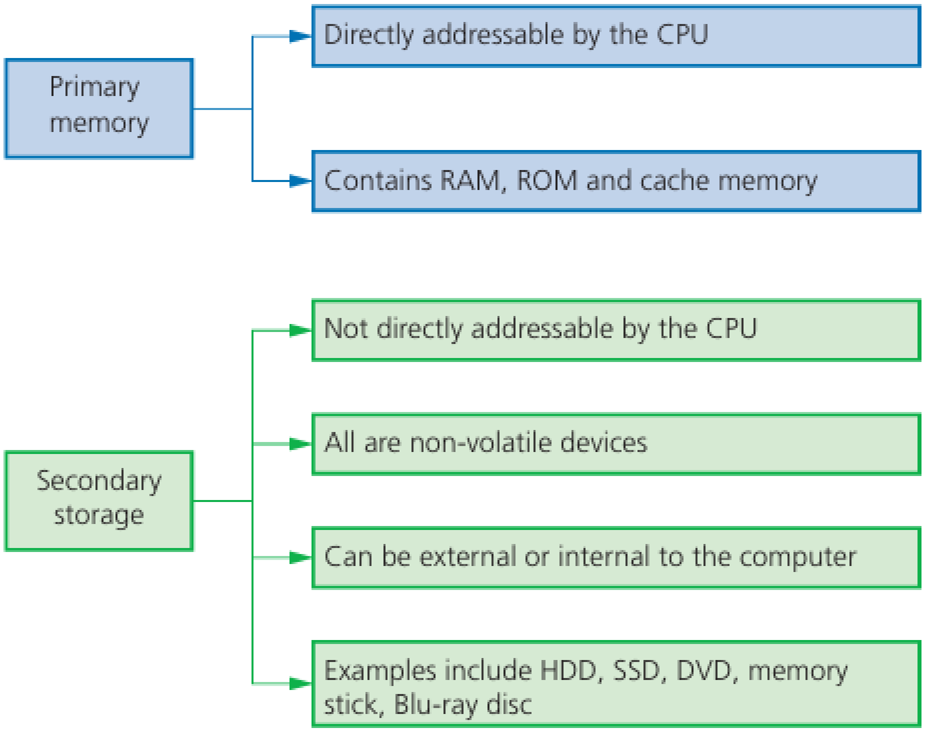
All computers require some form of memory and storage. Memory is usually referred to as the internal devices used to store data that the computer can access directly. This is also known as primary memory.

Memory and storage devices can be split up into two distinct groups:

» Primary memory » Secondary storage.



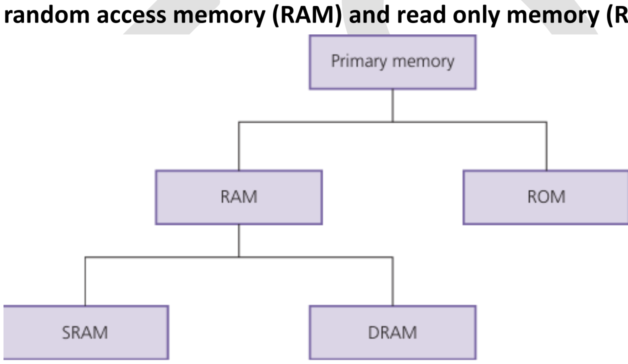
▲ **Figure 3.57** Typical memory and storage devices



▲ **Figure 3.58** Summary of primary, secondary and off-line devices

3.3.1 Primary memory

Primary memory is the part of the computer memory which can be accessed directly from the CPU; this includes **random access memory (RAM) and read only memory (ROM) memory chips.**



▲ **Figure 3.59** Primary memory

Random access memory (RAM)

Features of RAM include:

» Can be written to or read from, and the data can be changed by the user or the computer (i.e. it is a temporary memory)

» Used to store data, files, part of an application or part of the operating system currently in use

» It is volatile, which means memory contents are lost when powering off the computer.

There are currently two types of RAM technology:
» Dynamic RAM (DRAM) » Static RAM (SRAM).

Dynamic RAM (DRAM)

▲ **Figure 3.60** DRAM

Each DRAM chip consists of transistors and capacitors. Each of these parts is tiny since a single RAM chip will contain millions of transistors and capacitors. The function of each part is:

» Capacitor – this holds the bits of information (0 or 1)

» Transistor – this acts like a switch; it allows the chip control circuitry to read the capacitor or change the capacitor's value.

This type of RAM needs to be constantly **refreshed** (that is, the capacitor needs to be re-charged every 15 microseconds otherwise it would lose its value).

Static RAM (SRAM)

It makes use of **flip flops**, which hold each bit of memory. SRAM is much faster than DRAM when it comes to data access (typically, access time for SRAM is 25 nanoseconds and for DRAM is 60 nanoseconds).

▼ **Table 3.9** Differences between DRAM and SRAM

DRAM	SRAM
consists of a number of transistors and capacitors	uses flip flops to hold each bit of memory
needs to be constantly refreshed	doesn't need to be constantly refreshed
less expensive to manufacture than SRAM	has a faster data access time than DRAM
has a higher memory capacity than SRAM	
main memory is constructed from DRAM	CPU memory cache makes use of SRAM
consumes less power than SRAM	

Read-only memory (ROM)

ROM chips have the following features:
» They are non-volatile (the contents are not lost after powering off the computer)
» They are permanent memories (the contents cannot be changed or written to by the user, the computer or any application/program)
» The contents can only be read » they are often used to store data that the computer needs to access when powering up for the first time (the basic input/output system (BIOS)); these are known as the start-up instructions (or bootstrap)

▼ **Table 3.10** RAM and ROM features

RAM	ROM
temporary memory device	permanent memory device
volatile memory	non-volatile memory device
can be written to and read from	data stored cannot be altered
used to store data, files, programs, part of OS **currently** in use	always used to store BIOS and other data needed at start up
can be increased in size to improve operational speed of a computer	

We will consider the function of each type of memory independently:

ROM

» Storing the factory settings such as remote-control frequencies

» Storing the 'start-up' routines when the toy car is first switched on

» Storing of the set routines; for example, how the buttons on the hand-held device control turning left, acceleration, stopping, and so on.

RAM

» The user may wish to program in their own routines; these new instructions would be stored in the RAM chip » the RAM chip will store the data/instructions received from the remote-control unit.

3.3.2 Secondary and off-line storage Secondary (and off-line)

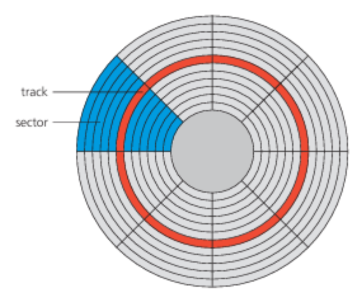
storage includes storage devices that are not directly addressable by the CPU. They are non-volatile devices that allow data to be stored as long as required by the user. This type of storage can store more data than primary memory, but data access time is considerably longer than with RAM or ROM.

3.3.3 Magnetic, optical and solid-state storage

Secondary (and off-line) storage falls into three categories according to the technology used:
» Magnetic » Solid state » Optical.

Magnetic storage

Hard Disk Drives (HDD)
Hard disk drives (HDD) are still one of the most common methods used to store data on a computer.



▲ **Figure 3.63** Tracks and sectors

The effects of **latency** then become very significant. Latency is defined as the time it takes for a specific block of data on a data track to rotate around to the read-write head.

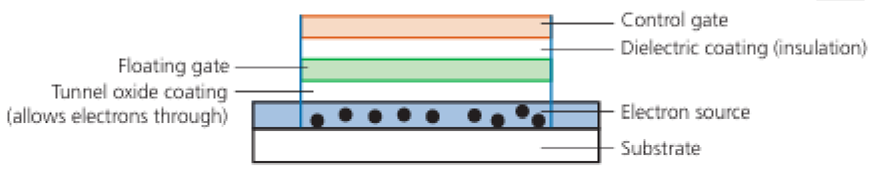
Through time, the HDD will undergo numerous deletions and editing which leads to sectors becoming increasingly **fragmented** resulting in a gradual deterioration of the HDD performance (in other words, it takes longer and longer to access data).

Removable hard disk drives are essentially HDDs external to the computer that can be connected to the computer using one of the USB ports. In this way, they can be used as a back-up device or another way of transferring files between computers.

Solid state drives (SSD)

Latency is an issue in HDDs as described earlier. Solid state drives (SSD) remove this issue considerably since they have no moving parts and all data is retrieved at the same rate. They don't rely on magnetic properties; the most common type of solid state storage devices store data by controlling the movement of electrons within NAND or NOR chips.

intersection on the matrix there is a floating gate and a control gate arranged as follows.



▲ **Figure 3.64** Flash memory

The main benefits of this newer solid-state technology over hard disk drives are:

» They are more reliable (no moving parts to go wrong)

» They are considerably lighter (which makes them suitable for laptops)

» They don't have to 'get up to speed' before they work properly

» They have a lower power consumption

» They run much cooler than HDDs (both these points again make them very suitable for laptop computers)

» Because of no moving parts, they are very thin » data access is considerably faster than HDD.

The main drawback of SSD is still the longevity of the technology (although this is becoming less of an issue). Most solid state storage devices are conservatively rated at only 20 GB of write operations per day over a three year period – this is known as **SSD endurance.**

However, the durability of these solid state systems is being improved by a number of manufacturers and they are rapidly becoming more common in applications such as servers and **cloud storage** devices.

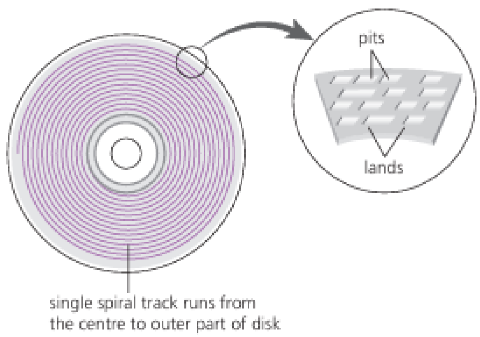
Memory sticks/flash memories

Memory sticks/flash memories (also known as pen drives) use solid state technology.

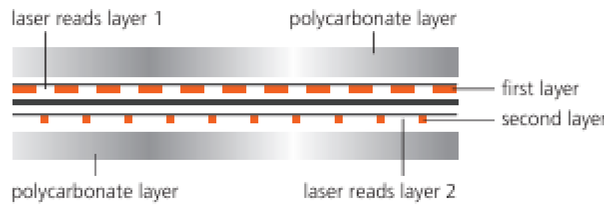
Optical media

CD/DVD disks CDs and DVDs are described as optical storage devices. Laser light is used to read and write data to and from the surface of the disk.

DVD technology is slightly different to that used in CDs. One of the main differences is the potential for **dual layering**, which considerably increases the storage capacity.



▲ **Figure 3.65** Optical media



▲ **Figure 3.66** Dual-layering on a DVD

Blu-ray discs

Blu-ray discs are another example of optical storage media. However, they are fundamentally different to DVDs in their construction and in the way they carry out read-write operations.

The main differences between DVD and Blu-ray are:

» A blue laser, rather than a red laser, is used to carry out read and write operations; the wavelength of blue light is only 405 nanometres (compared to 650 nm for red light)

» Using blue laser light means that the 'pits' and 'lands' can be much smaller; consequently, Blu-ray can store up to five times more data than normal DVD

» Single-layer Blu-ray discs use a 1.2 mm thick polycarbonate disk; however, dual layer Blu-ray and normal DVDs both use a sandwich of two 0.6 mm thick disks (i.e. 1.2 mm thick)

» Blu-ray disks automatically come with a secure encryption system that helps to prevent piracy and copyright infringement

» The data transfer rate for a DVD is 10 Mbps and for a Blu-ray disc it is 36 Mbps (this equates to 1.5 hours to transfer 25 GiB of data).

Comparison of the capacity and interactivity of DVDs and Blu-ray discs

» A standard dual-layer DVD has a storage capacity of 4.7 GB (enough to store a 2-hour standard definition movie)

» A single-layer Blu-ray disc has a storage capacity of 27 GB (enough to store a 2-hour high definition movie or 13 hours of standard definition movies)

» A dual-layer Blu-ray disc has a storage capacity of 50 GB (enough to store 4.5 hours of high definition movies or 20 hours of standard definition movies).

▼ **Table 3.11** Comparison of CD, DVD and Blu-ray (Note: nm = 10^{-9} metres and μm = 10^{-6} metres)

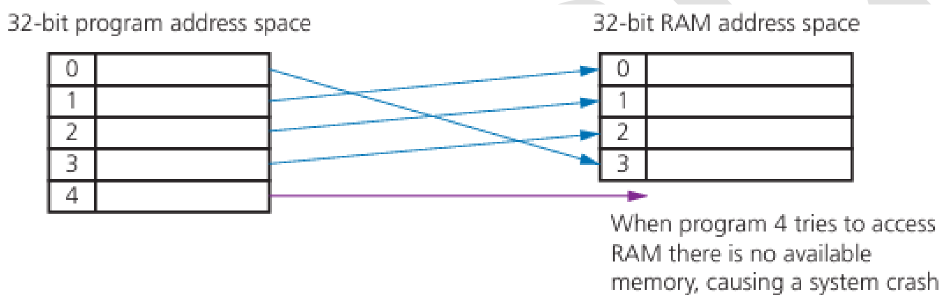
Disk type	Laser colour	Wavelength of laser light	Disk construction	Track pitch (distance between tracks)
CD	Red	780 nm	single 1.2 mm polycarbonate layer	1.60 μm
DVD (dual-layer)	Red	650 nm	two 0.6 mm polycarbonate layers	0.74 μm
Blu-ray (single layer)	Blue	405 mm	single 1.2 mm polycarbonate layer	0.30 μm
Blu-ray (dual-layer)	Blue	405 nm	two 0.6 mm polycarbonate layers	0.30 μm

3.3.4 Virtual memory

This is the basis behind **virtual memory**. Essentially RAM is the **physical memory**, while **virtual memory** is RAM + swap space on the hard disk or SSD.

Without virtual memory

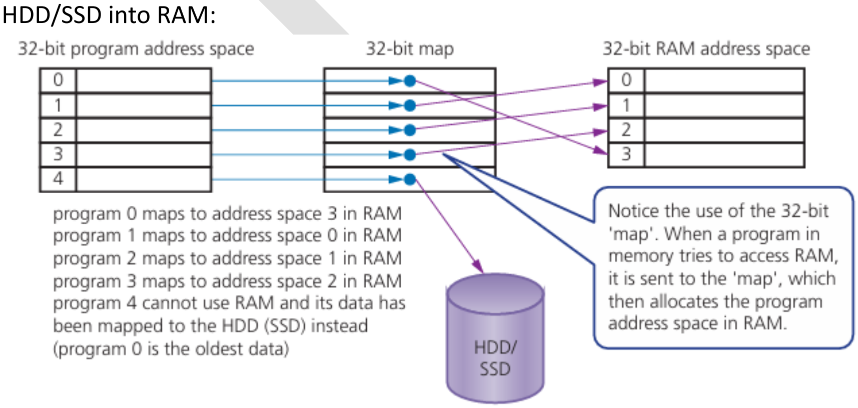
Suppose we have five programs (numbered 0 to 4) that are in memory, all requiring access to RAM. The first diagram shows what would happen without virtual memory being used (i.e. the computer would run out of RAM memory space):



▲ **Figure 3.68** Normal memory management

With virtual memory

We will now consider what happens if the CPU uses virtual memory to allow all five programs to access RAM as required. This will require moving data out of RAM into HDD/SSD and then allowing other data to be moved out of HDD/SSD into RAM:

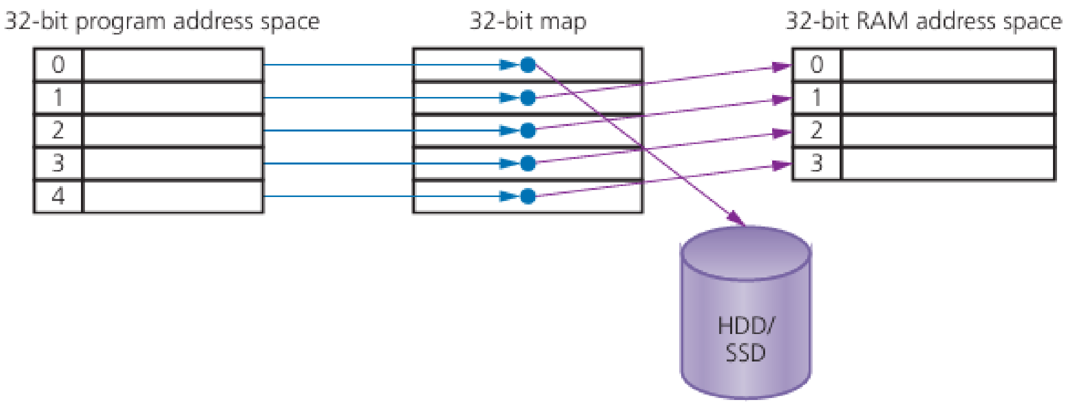


▲ **Figure 3.69** Status just before program 4 is given RAM space

Virtual memory now moves the oldest data out of RAM into the HDD/SSD to allow program 4 to gain access to RAM. The 32-bit 'map' is now updated to reflect this new situation:

» Data from program 0 (which was using RAM address space 3 – the oldest data) is now mapped to the HDD/SSD instead, leaving address space 3 free for use by program 4
» Program 4 now maps to address space 3 in RAM, which means program 4 now has access to RAM.

Our diagram now changes to:



▲ **Figure 3.70** Status with program 0 now mapped to HDD and program 4 has access to RAM

In computer operating systems, **paging** is used by memory management to store and retrieve data from HDD/SSD and copy it into RAM. A **page** is a fixed-length consecutive (or contiguous) block of data utilised in virtual memory systems.

This is a key part of how virtual memory works allowing **data blocks** (pages) to be moved in and out of a HDD/SSD.

The main benefits of virtual memory are:
» Programs can be larger than physical memory and still be executed
» There is no need to waste memory with data that isn't being used (e.g. during error handling)
» It reduces the need to buy and install more expensive RAM memory (although as mentioned earlier there are limits to the value of doing this).

When using HDD for virtual memory the main drawback is **disk thrashing**. As main memory fills, more and more data needs to be swapped in and out of virtual memory leading to a very high rate of hard disk read/write head movements; this is known as disk thrashing.

A point can be reached when the execution of a process comes to a halt since the system is so busy moving data in and out of memory rather than doing any actual execution – this is known as the **thrash point.**

3.3.5 Cloud storage

Public and private cloud computing

Cloud storage is a method of data storage where data is stored on remote servers. The same data is stored on more than one server in case of maintenance or repair, allowing clients to access data at any time. This is known as data **redundancy**.

There are three common systems:

» Public cloud – this is a storage environment where the customer/client and cloud storage provider are different companies

» Private cloud – this is storage provided by a dedicated environment behind a company firewall; customer/client and cloud storage provider are integrated and operate as a single entity

» Hybrid cloud – this is a combination of the two above environments; some data resides in the private cloud and less sensitive/less commercial data can be accessed from a public cloud storage provider.

▼ **Table 3.12** Benefits and drawbacks of cloud storage

Benefits of using cloud storage	Drawbacks of using cloud storage
customer/client files stored on the cloud can be accessed at any time from any device anywhere in the world provided internet access is available	if the customer/client has a slow or unstable internet connection, they would have many problems accessing or downloading their data/files
there is no need for a customer/client to carry an external storage device with them, or even use the same computer to store and retrieve information	costs can be high if large storage capacity is required; it can also be expensive to pay for high download/upload data transfer limits with the customer/client internet service provider (ISP)
the cloud provides the user with remote back-up of data with obvious benefits to alleviate data loss/disaster recovery	the potential failure of the cloud storage company is always possible – this poses a risk of loss of all back-up data
if a customer/client has a failure of their hard disk or back-up device, cloud storage will allow recovery of their data	
the cloud system offers almost unlimited storage capacity	

Data security when using cloud storage

Companies that transfer vast amounts of confidential data from their own systems to a cloud service provider are effectively relinquishing control of their own data security. This raises a number of questions:

» What physical security exists regarding the building where the data is housed?
» How good is the cloud service provider's resistance to natural disasters or power cuts?
» What safeguards exist regarding personnel who work for the cloud service company; can they use their authorisation codes to access confidential data for monetary purposes?

Potential data loss when using cloud storage

The following breaches of security involving some of the largest cloud service providers suggest why some people are nervous of using cloud storage for important files:

» The XEN security threat, which forced several cloud operators to reboot all their cloud servers, was caused by a problem in the XEN hypervisor

» A large cloud service provider permanently lost data during a routine back-up procedure.

» The celebrity photos cloud hacking scandal, in which more than 100 private photos of celebrities were leaked. Hackers had gained access to a number of cloud accounts, which enabled them to publish the photos on social networks and sell them to publishing companies.

» In 2016, the National Electoral Institute of Mexico suffered a cloud security breach in which 93 million voter registrations, stored on a central database, were compromised and became publicly available to everyone. To make matters worse, much of the information on this database also linked to a cloud server outside Mexico

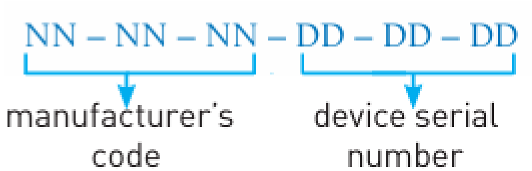
3.4 Network hardware

3.4.1 Network interface card (NIC)

A **network interface card (NIC**) is needed to allow a device to connect to a network (such as the internet). It is usually part of the device hardware and contains the **Media Access Control (MAC**) address generated at the manufacturing stage.

3.4.2 Media Access Control (MAC)

A MAC address is made up of 48 bits which are shown as six groups of hexadecimal digits with the general format:



Types of MAC address

It should finally be pointed out that there are two types of MAC address: the **Universally Administered MAC Address (UAA) and the Locally Administered MAC Address (LAA).**

There are a few reasons why the MAC address needs to be changed using LAA:

» Certain software used on mainframe systems need all the MAC addresses of devices to fall into a strict format; because of this, it may be necessary to change the MAC address of some devices to ensure they follow the correct format

» It may be necessary to bypass a MAC address filter on a router or a firewall; only MAC addresses with a certain format are allowed through, otherwise the devices will be blocked if their MAC address doesn't adhere to the correct format

» To get past certain types of network restrictions it may be necessary to emulate unrestricted MAC addresses; hence it may require the MAC address to be changed on certain devices connected to the network.

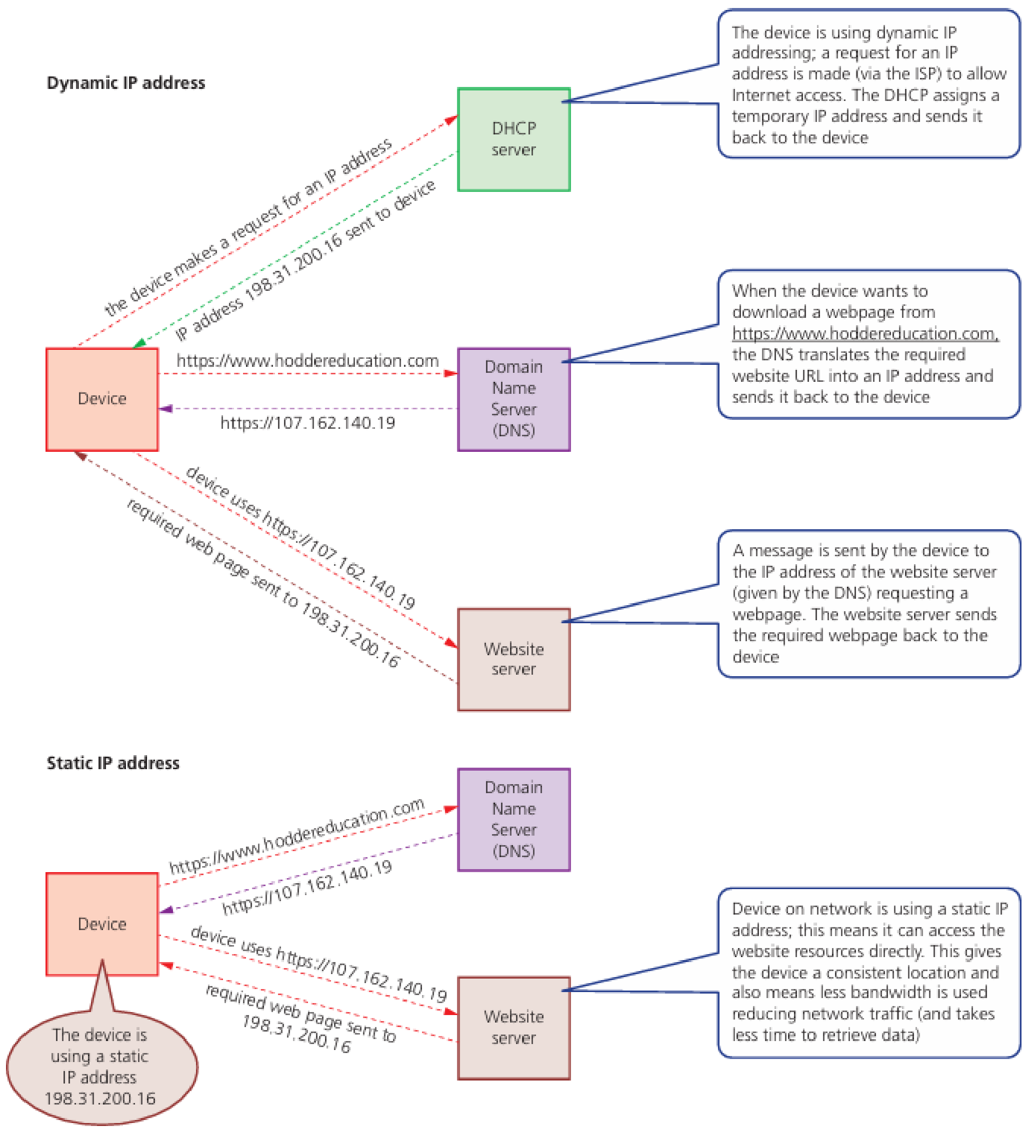
3.4.3 Internet protocol (IP) address

When a device connects to a private network, a router assigns a private IP address to it. That IP address is unique on that network, but might be the same as an IP address on a separate network. However, when a router connects to the internet it is given a unique public IP address. This is usually supplied by the internet service provider (ISP).

The main advantages of IPv6 compared to IPv4 are:
» Removes the risk of IP address collisions
» Has built-in authentication checks
» Allows for more efficient packet routes.

▼ **Table 3.13** MAC addresses and IP addresses

MAC addresses	IP addresses
identifies the physical address of a device on the network	identifies the global address on the internet
unique for device on the network	may not necessarily be unique
assigned by the manufacturer of the device and is part of the NIC	dynamic IP addresses are assigned by ISP using DHCP each time the device connects to the internet (see later)
they can be universal or local	dynamic IP addresses change every time a device connects to the internet; static IP addresses don't change
when a packet of data is sent and received, the MAC address is used to identify the sender's and recipient's devices	used in routing operations as they specifically identify where the device is connected to the internet
use 48 bits	use either 32 bits (IPv4) or 128 bits (IPv6)
can be UAA or LAA	can be static or dynamic



▲ **Figure 3.71** Comparison of dynamic and static IP addressing

Static and dynamic IP addresses

IP addresses can be either static (don't change) or dynamic (change every time a device connects to the internet).

Static

Static IP addresses are permanently assigned to a device by the internet service provider (ISP); they don't change each time a device logs onto the internet. Static IP addresses are usually assigned to:

» Remote servers which are hosting a website

» An online database

» A File Transfer Protocol (FTP) server. FTP servers are used when files need to be transferred to various computers throughout the network.

Dynamic

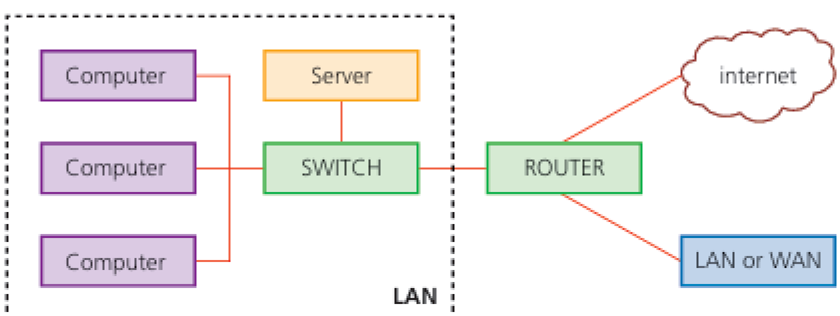
Dynamic IP addresses are assigned by the ISP each time a device logs onto the internet. This is done using **Dynamic Host Configuration Protocol (DHCP).**

▼ **Table 3.14** Dynamic and static IP addresses

Dynamic IP addresses	Static IP addresses
greater privacy since they change each time a user logs on	since static IP addresses don't change, they allow each device to be fully traceable
dynamic IP addresses can be an issue when using, for example, VoIP since this type of addressing is less reliable as it can disconnect and change the IP address causing the VoIP connection to fail	allow for faster upload and download speeds (see Figure 3.71)
	more expensive to maintain since the device must be constantly running so that information is always available

3.4.4 Routers

Routers enable data packets to be routed between different networks, for example, to join a LAN to a WAN. The router takes data transmitted in one format from a network (which is using a particular protocol) and converts the data to a protocol and format understood by another network, thereby allowing them to communicate.



▲ **Figure 3.72** Router flow diagram

Key terms used throughout this chapter

central processing unit (CPU) – responsible for the execution or processing of all the instructions and data in a computer

integrated circuit – usually a chip made from a semi-conductor material which carries out the same tasks as a larger circuit made from individual components

von Neumann architecture – a type of computer architecture which introduced the concept of the stored program in the 1940s

Arithmetic & Logic Unit (ALU) – the component of the CPU that carries out all arithmetic and logical operations

accumulator (ACC) – temporary general-purpose register that stores numerical values at any part of a given operation

memory address register (MAR) – a register that stores the address of the memory location currently being read from or written to

current instruction register (CIR) – a register that stores the current instruction being decoded and executed

memory data register (MDR) – a register that stores data that has just been read from memory or data that is about to be written to memory

program counter (PC) – a register that stores the address where the next instruction to be read can be found

control unit – the component of a computer's CPU that ensures synchronisation of data flow and programs throughout the computer by sending out control signals along the control bus

system clock – produces timing signals on the control bus to ensure synchronisation takes place

clock cycle – clock speeds are measured in terms of GHz; this is the vibrational frequency of the system clock which sends out pulses along the control bus; for example, a 3.5GHZ clock cycle means 3.5 billion clock cycles a second

immediate access store (IAS) – memory that holds all data and programs needed to be accessed by the control unit

backing store – a secondary storage device (such as HDD or SSD) used to store data permanently even when the computer is powered down

cache – is temporary memory using static RAM to hold frequently used data/instructions by the CPU thereby increasing CPU performance. More generally, cache means any area of storage used to quickly access frequently-used data - other examples include web cache, database cache, DNS cache

register – a temporary component in the CPU which can be general or specific in its use; it holds data or instructions as part of the Fetch–Decode–Execute cycle

address – a label for a memory location used by the CPU to track data

memory location – a numbered place in memory where values can be stored

system buses – a connection between major components in a computer that can carry data, addresses or control signals

address bus – the system bus that carries the addresses throughout the computer system

data bus – the system bus that allows data to be carried from CPU to memory (and vice versa) or to and from input/output devices

control bus – the system bus that carries signals from control unit to all other computer components

unidirectional – can travel in one direction only; used to describe data

bidirectional – can travel in both directions; used to describe data

word – a group of bits used by a computer to represent a single unit; for example, modern computers often use 64-bit word lengths

overclocking – changing the clock speed of a system clock to a value higher than the factory/recommended setting

core – a unit on a CPU made up of an ALU, control unit and registers; a CPU may contain a number of cores

dual core – a CPU containing two cores

quad core – a CPU containing four cores

Fetch–Execute–Decode – a cycle in which instructions and data are fetched from memory, decoded and finally executed

Basic Input/Output System (BIOS) – a suite of programs on firmware that are used to perform the initialisation of a computer system during the boot-up process

opcode – part of a machine code instruction that identifies what action the CPU has to perform

operand – part of a machine code instruction that identifies what data is to be used

instruction set – the complete set of machine code instructions used a particular microprocessor

embedded system – a combination of hardware and software designed to carry out a specific set of functions

barcode – a series of dark and light lines of varying thickness used to represent data; the code has to be scanned using laser or LED light source

key field – the field that uniquely identifies a record in a file

quick response (QR) code – a matrix of dark and light squares which represent data; the pattern can be read and interpreted using a smartphone camera and QR app

frame QR code – a type of QR code that includes a space for advertising

DAC (digital to analogue converter) – device that converts digital data into electric currents that can drive motors, actuators and relays, for example

ADC (analogue to digital converter) – a device that converts analogue data (for example, data read from sensors) into a form understood by a computer

charge couple device (CCD) – a light sensitive cell made up of millions of tiny sensors acting as photodiodes

virtual keyboard – an onscreen keyboard which uses the features of the touch screen to emulate a physical keyboard

touch screen – a screen that allows the user to select or manipulate a screen image using the touch of a finger or stylus; touch screens most frequently use capacitive, infra-red or resistive technology

repetitive strain injury (RSI) – pain felt in the muscles, nerves and tendons caused by a repetitive action (for example, excessive clicking of a mouse button over a period of time)

optical mouse – a pointing device that uses a red LED to track the movement of the device and then relays its coordinates to a computer

pointing device – an input device that allows the user to control the movement of an onscreen cursor or to allow onscreen selection by clicking a button on the device

complementary metal oxide semi-conductor (CMOS) – a chip that generates an electric current (or pulses) when light falls on its surface

digital signal processor (DSP) – a processor that calculates, for example, the coordinates of a pointing device based on the pulses of electricity received

optical character recognition – technology that can convert hard copy text or images into a digital format to be stored in a computer memory

computer sided design (CAD) – software used to create drawings (for example, to send to a 3D printer or to produce blue-prints of a microprocessor design)

computed tomographic (CT) scanner – technology that can create a 3D image of a solid object by slicing up the object into thin layers (tomography)

capacitive touch screen – a type of touch screen that uses the change in the screen's capacitance (the ability to store an electrical charge) when it is touched by a finger or stylus

infra-red touch screen – a type of touch screen that uses infra-red beams and sensors to detect where the screen has been touched

resistive touch screen – a type of touch screen that uses two conductive layers which make contact where the screen has been touched

actuator – an output device that converts electrical energy into mechanical movement

digital micromirror device (DMD) – a chip that uses millions of tiny mirrors on its surface to create a video display

thermal bubble – inkjet printer technology whereby tiny resistors create heat and form an ink bubble which is ejected onto paper in an inkjet printer

piezoelectric crystal – a crystal located in an ink reservoir within an inkjet printer; the crystal vibrates and forces ink out onto paper

direct 3D printing – a 3D printing technique in which the print head moves in the x, y and z directions

binder 3D printing – a 3D printing method that uses a two-stage pass; the first stage uses dry powder and second stage uses a binding agent

cathode – a negative electrode

snode – a positive electrode

organic LED (OLED) – a light-emitting diode that uses the movement of electrons between a cathode and an anode to produce an on-screen image; it generates its own light so no backlighting is required

loudspeaker – an output device that converts electric current into sound

memory – the devices within the computer that are directly accessible by the CPU; there are two types of memory – RAM and ROM; memory is different to hard disk drives, for example, which are known as storage devices

random access memory (RAM) – primary memory that can be written to or read from

read only memory (ROM) – primary memory that cannot be written to (changed) and can only be read

dynamic RAM (DRAM) – a type of RAM chip that needs to be constantly refreshed

static RAM (SRAM) – a type of RAM chip that uses flip flops and doesn't need to be constantly refreshed

volatile – describes memory that loses its contents when the power is turned off

refresh – recharge every few seconds in order to maintain charge; for example with a device such as a capacitor

flip flop – electronic circuit with only two stable conditions

latency – the lag in a system; for example, the time it takes to find a track on a hard disk, which depends on the time it takes for the disk to rotate around to its read-write head

SSD endurance – the total guaranteed number of times data can be written to or read from a solid state drive (SSD) in its usable life cycle

optical storage – a type of storage that uses laser light to read and write data, and includes CDs, DVDs and Blu-ray discs

dual layering – using two recording layers in storage media such as DVDs and some Blu-rays

virtual memory – a memory management system that makes use of secondary storage and software to enable a computer to compensate for the shortage of actual physical RAM memory

disk thrashing (HDD) – a problem in a hard disk drive (HDD) caused by excessive swapping in and out of data causing a high rate of head movements during virtual memory operations

thrash point – the point at which the execution of a program comes to a halt because the system is so busy moving data in and out of memory rather than actually executing the program

data redundancy – the unnecessary storing of the same data on several storage devices at the same time

cloud storage – a method of data storage where data is stored on offsite servers; the physical storage may be on hundreds of servers in many locations

network interface card (NIC) – a hardware component (circuit board or chip) that is required to allow a device to connect to a network, such as the internet

router – a device that enables data packets to be moved between different networks, for example, to join a LAN to a WAN

static IP address – an IP address that doesn't change

MAC address – a unique identifier which acts as a network address for a device; it takes the form NN-NN-NN-DD-DD-DD, where NN is the manufacturer code and DD is the device code

dynamic IP address – a temporary IP address assigned to a device each time it logs onto a network

dynamic host configuration protocol (DHCP) – a server that automatically provides and assigns an IP address

Revision questions

1. An alarm clock is controlled by a microprocessor. It uses the 24-hour clock. The hour is represented by an 8-bit register, A, and the number of minutes is represented by another 8-bit register, B

(a) Identify what time is represented by the following two 8-bit registers. A

A									B							
128	64	32	16	8	4	2	1		128	64	32	16	8	4	2	1
0	0	0	1	0	0	1	0	:	0	0	1	1	0	1	0	1

Hours .... Minutes ....

b) An alarm has been set for 07:30. Two 8-bit registers, C and D, are used to represent the hours and minutes of the alarm time. Show how 07:30 would be represented by these two registers:

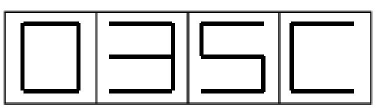
C									D							
								:								
Hours									Minutes							

(c) Describe how the microprocessor can determine when to sound the clock alarm.

(d) The LCD (liquid crystal display) on the clock face is back-lit using blue LEDs (light emitting diodes). The brightness of the clock face is determined by the level of light in the room. The amount of light given out by the LEDs is controlled by a control circuit. Describe how the sensor, microprocessor and LEDs are used to maintain the correct brightness of the clock face.

(e) Modern LCD monitors and televisions use LED back-lit technology. Give two advantages of using this new technology compared to the older cold cathode fluorescent lamp (CCFL) method.

2. Each seat on a flight is uniquely identified on an LCD above the seat. For example, seat 035C is shown as:



The first three characters are digits that represent the row. The fourth character is the seat position in that row. This is a single letter, A to F, that is stored as a hexadecimal value.

Each of the four display characters can be stored in a 4-bit register. For example, 0 and C would be represented as:

	8	4	2	1
0:	0	0	0	0
C:	1	1	0	0

(a) Show how the 4-bit registers would store the remaining two characters, 3 and 5.

3				
5				

(b) Identify which seat is stored in the following 4-bit registers.

0	0	0	1	→ ....
1	0	0	1	→ ....
0	1	0	0	→ ....
1	1	1	0	→ ....

3. (a) Street lighting is controlled automatically. A light sensor and a microprocessor are used to decide when to switch each street light on or off. Describe how the sensor, microprocessor and light interact to switch the street light on or off. Include in your answer how the microprocessor stops the street lights being frequently switched on and off due to brief changes in the light intensity.

(b) Name three different sensors (other than light and pH) and describe an application for each of these sensors.

4. A cold store is kept at a constant low temperature using a sensor, a microprocessor and a cooling unit. Explain how the sensor and microprocessor will maintain a constant low temperature

5. A passenger logs onto an airline website and types in the reference number for their flight. Once the passenger accesses their account they can choose their seat and also print out a boarding pass which contains a unique barcode. This barcode is scanned at the airport check-in desk. Name one input and one output device found at the check-in desk and give a reason for your choice.

6. Four input devices are shown in the table below. Give an application which makes use of each device and state a reason why the device is appropriate for that application. Your application must be different in each case.

Input device	Application and reason
Light sensor	Application .... Reason ....
Keyboard	Application .... Reason ....
Barcode reader	Application .... Reason ....
Touch screen	Application .... Reason ....

7. A sports stadium uses a pressure sensor and a microprocessor to monitor the number of people entering the sports stadium. For the counter to increment the weight on the pressure sensor must exceed 5 kg. Explain how the system uses the pressure sensor and the microprocessor to monitor the number of people entering

8. Personal computers (PCs) use an operating system. Explain why this type of computer needs an operating system.

9. (a) A manufacturer of aeroplane engines assigns a denary identification number (ID) to each engine. One engine has the ID: 0431 (i) Convert this denary number to a 12-bit binary format

(ii) Show how this number would be represented in hexadecimal.

(b) The current status of the engine is sent to a computer in the aeroplane. Each piece of data collected is 8 bytes in size. Data collection occurs every 30 seconds. Calculate the number of kilobytes that would be needed to store the data collected during a 10-hour flight. Show your working.

(c) At the end of the flight, all of the data are sent to the aeroplane engine manufacturer using the Internet. The computer in the aeroplane has a MAC address and an IP address. State what is meant by these two terms.

(d) When sending this data, security is very important. Data are sent over the Internet using Transport Layer Security (TLS) protocol. Name the two layers that make up TLS.

10. Six statements are given about touch screen technology. Tick (3) to show if the statement applies to Capacitive or Resistive touch screen technology

Statement	Capacitive (✓)	Resis (✓)
Needs pressure to be applied to create a circuit		
May not register a touch if the user is wearing gloves		
More commonly used in smartphones		
More responsive to a touch		
Needs an electrical field to be changed to register a touch		
Cheaper to manufacture		