

Syllabus Content:

1.3.1 Input, output and storage devices



identify hardware devices used for input, output, secondary storage



show understanding of the basic internal operation of the following specific types of device:

- keyboard
- speakers
- tracker ball mouse
- hard disk
- optical mouse
- solid state (flash) memory
- scanner
- optical discs
- inkjet printer
- microphone
- laser printer
- touchscreen
- 3D printer



show understanding of the need for secondary (including removable) storage

Input, output and storage devices

Computer System:

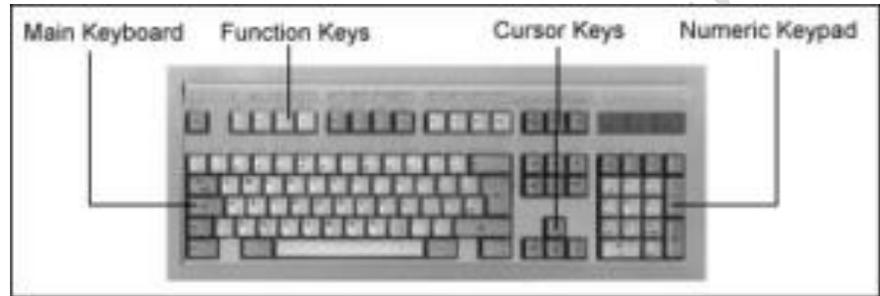
A computer is an electronic machine that accepts data, stores and processes data into information. The computer is able to work because there are instructions in its memory directing it, instructions that direct the computer are called **software** or **computer program**. The physical parts of the computer that you can see and touch, such as the keyboard, monitor and the mouse are called **hardware**. There are four major categories of computer hardware:

1. **Input devices:** used to enter data into the computer
2. **Processing devices:** manipulate the data.
3. **Storage devices:** store data and program.
4. **Output device:** show results of the process.

Input devices:

An input device is any hardware component that allows you the user to enter data or instruction into the computer. There are many manual/automatic input devices. Most widely used input devices are:

-  Keyboard
-  Pointing devices
-  Tracker ball mouse
-  Laser mouse
-  2D, 3D Scanners



Keyboards:

The keyboard is the most common and widely used input device. It is made up of buttons called 'keys'. The keys are arranged into sections:

- Alphabet keys
- Function or F keys (F1, F2, F3)
- Numeric keys
- Arrow keys
- Command keys (insert, delete, home, end, page up/down)

Most keyboards are called '**QWERTY**' keyboards. This name comes from the first six letters on the top row of the alphabet keys.

QWERTY



DVORAK



Uses: - Keyboards are used to enter characters and data into computers. A QWERTY keyboard uses the standard layout for English-speaking countries.

The **Dvorak** keyboard layout is an alternative and

was designed to avoid awkward key combinations.

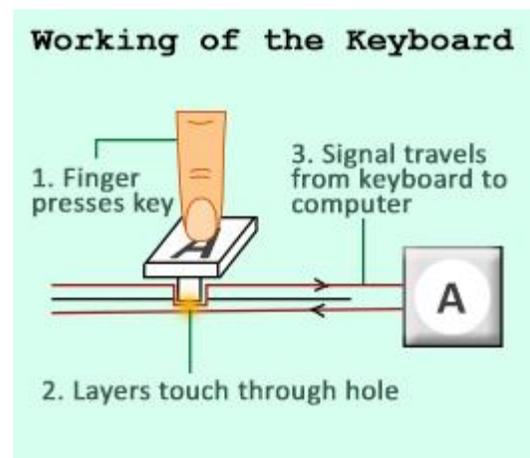
Did You Know?

The current keyboard layout, or the QWERTY layout, which is based on the layout of the typewriter, was designed not to increase the speed of typing, but to slow it down in order to avoid typewriters from jamming.



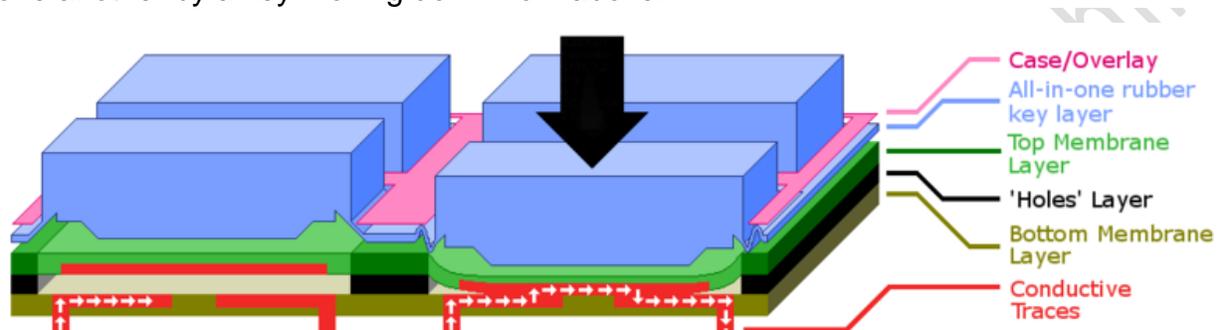
The working of a computer keyboard can be compared to a miniature computer.

Inside the keyboard, there are metallic plate, circuit board. The keyboard has its own processor and circuitry, a majority of which forms an important component called (key matrix) and processor, which are responsible for transferring information from the keyboard to the computer.



Internal operation of a keyboard:

What actually happens behind the scenes is that, there are three separate layers of plastic. Two of them are covered in electrically conducting metal tracks, and there's an insulating layer between them with holes in it. You can see dots at the places where the keys press the two conducting layers together. There are lines, which are electric connections allowing tiny electric currents to flow when the layers are pressed tight to one another by a key moving down from above.



1. When you type or press any key, a switch is pressed, which completes the circuit (closes an open circuit) and allows a tiny amount of current to flow.
2. A processor (on keyboard) analyses the position of the keys pressed and sends this information to the computer, where it is sent to something called the 'keyboard controller'.
3. This controller processes the information that is sent by the keyboard's processor, and, in turn, sends it to the operating system (OS).
4. The OS then checks this data to analyze if it contains any system level commands, like **Ctrl+Shift+Esc**, which is the keypress to bring up the Task Manager. If such system level commands are present, the computer executes them; if not, it forwards the information to the current application.
5. The application then checks if the keypresses relate to commands in the application, like Ctrl+P, which is the keypress for the print command. Again, if there are such commands, they are executed first, and if not, then these keypresses are accepted as content or data.
6. All this happens in a fraction of a second, so even if you press many keys, there is no lag in the system.

Character Mapping

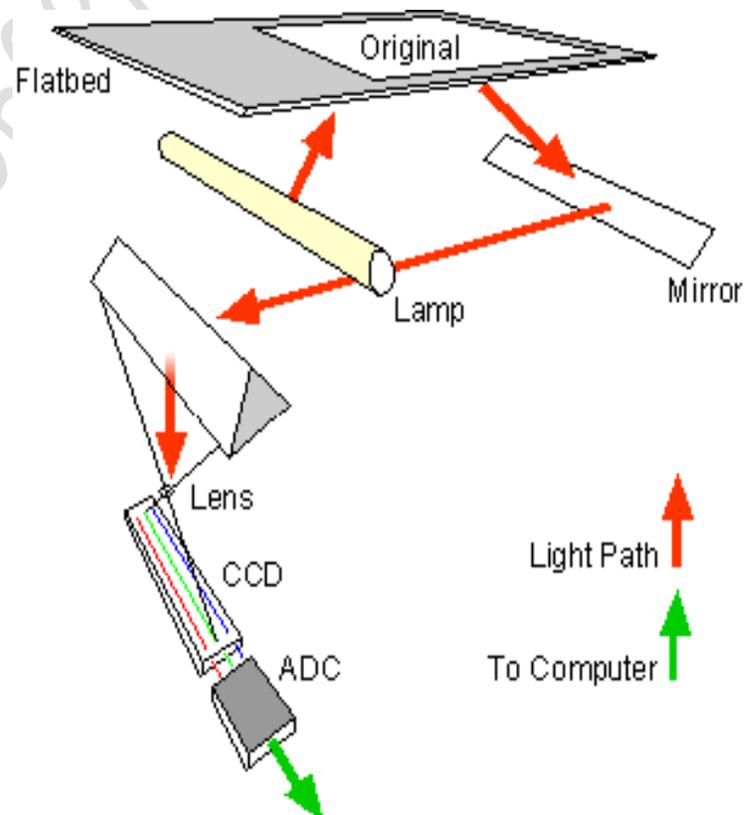
The key matrix has a corresponding chart or character map that is stored in the read-only memory (ROM) of the computer. When you press a key, the processor looks up the position of the circuit that was closed, or completed, with the character map, and determines which key was pressed. All the keys are mapped and stored in the memory.

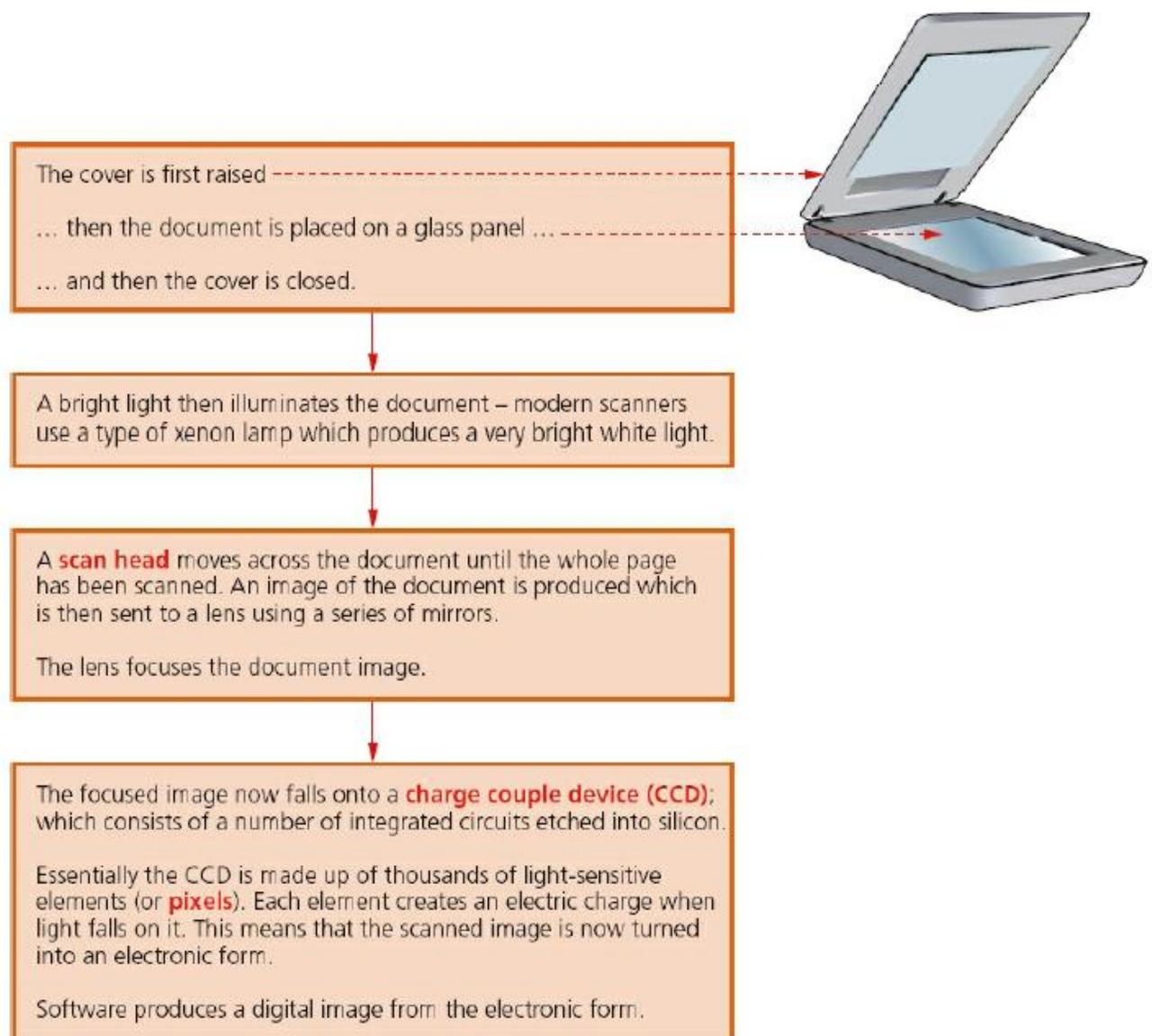
To put it simply, keyboards use switches and circuits to change keystrokes to a format the computer understands. Every keyboard contains a processor that does the work of translating the keystrokes, or the keys pressed, to the computer.

2D-Scanner:

A scanner creates a digital photograph of a paper document. It scans the illuminated surface of the document with a single row of hundreds of light sensors. Each sensor produces an analogue signal that depends on the **intensity of the light it receives**.

The scanner's embedded computer repeatedly scans the signals from the sensors as they move across the document. The embedded computer then digitizes, processes them and sends them to the computer.

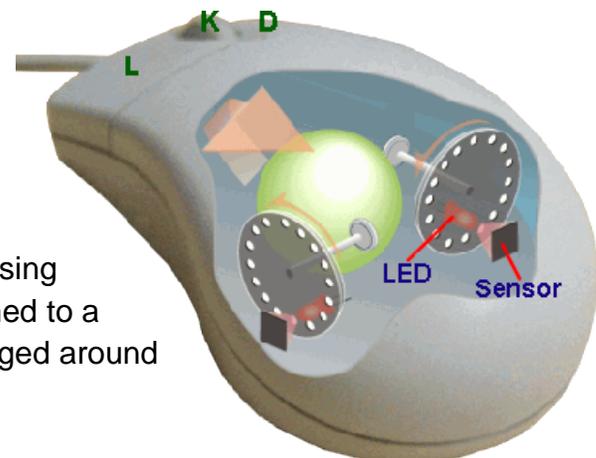




Mouse:

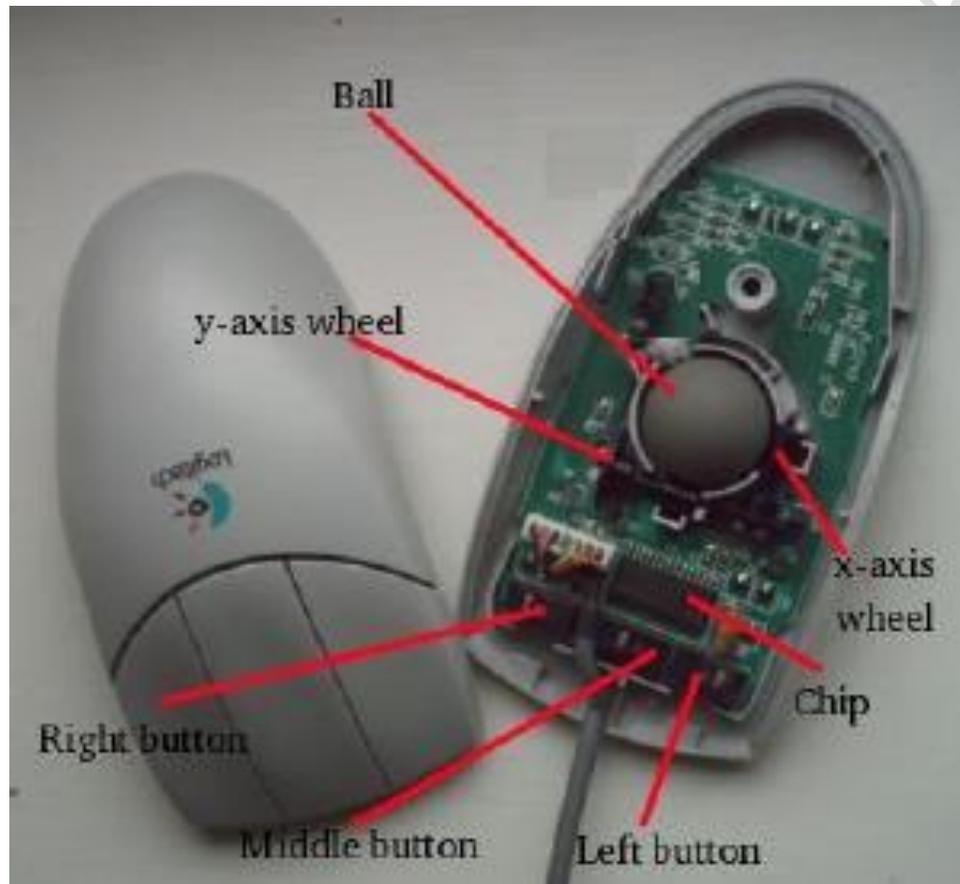
The computer mouse introduced initially contained a rubber ball held in contact with two rollers set perpendicularly to each other.

As the mouse moves the rubber ball rotates causing one or both rollers to rotate. Each roller is attached to a spindle on which there is a disc with holes arranged around the outer edge.



A light beam and detector are arranged so that the intermittent transmission of the light through the holes in the disc is recorded as the disc rotates and the circuitry attached to the pair of detectors then sends the appropriate data to the computer to activate movement of the screen cursor.

Figure below has the components of a mechanical mouse

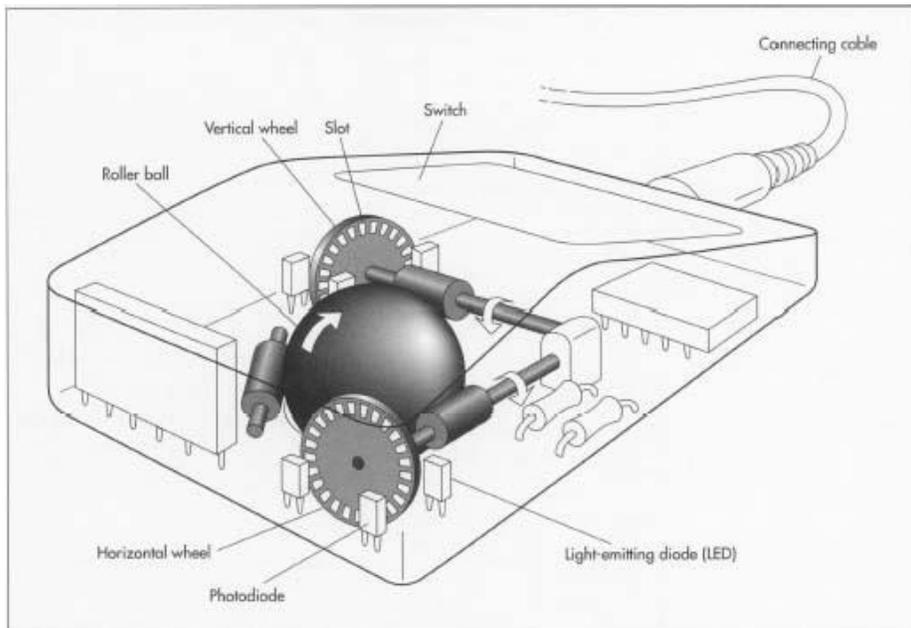


More recently the tracker ball mouse was phased out and the optical mouse was introduced. This technology dispenses with the mechanical aspects associated with the movement of a rubber

Tracker ball Mouse:

A trackball is a pointing device consisting of a ball held by a socket containing sensors to detect a rotation of the ball about two axes—like an upside-down mouse with an exposed protruding ball. The user rolls the ball with the thumb, fingers, or the palm of the hand to move a pointer.



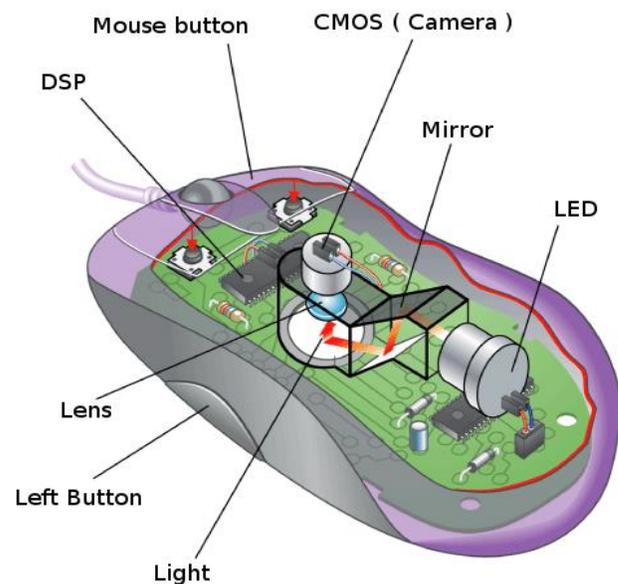


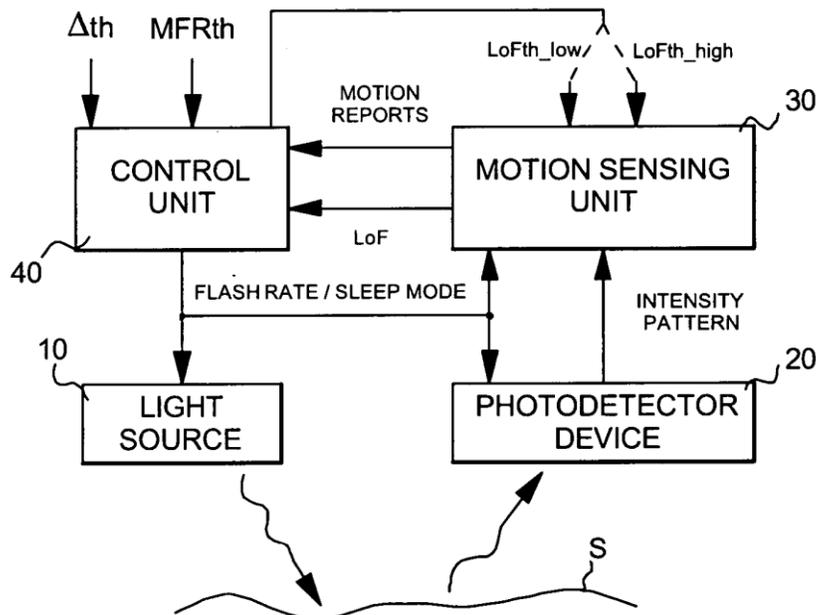
Compared with a mouse, a trackball has no limits on effective travel; at times, a mouse can reach an edge of its working area while the operator still wishes to move the screen pointer farther. With a trackball, the operator just continues rolling, whereas a mouse would have to be lifted and re-positioned. Some trackballs, such as Logitech's optical-pickoff types, have notably low friction, as well as being dense (glass), so they can be spun to make them coast. The trackball's buttons may be situated to that of a mouse or to a unique style that suits the user.

Laser Mouse:

A mouse allows the user to point by moving the cursor in graphical user interface on a PC's screen.

The optical mouse actually uses a tiny camera to take 1,500 pictures every second. Able to work on almost any surface, the mouse has a small, red light-emitting diode (LED) that bounces light off that surface onto a complementary metal-oxide semiconductor (CMOS) sensor.





The CMOS sensor sends each image to a digital signal processor (DSP) for analysis. The DSP, operating at 18 MIPS (million instructions per second), is able to detect patterns in the images and see how those patterns have moved since the previous image.

Based on the change in patterns over a sequence of images, the DSP determines how far the mouse has moved and sends the corresponding coordinates to the computer. The computer moves the cursor on the screen based on the coordinates received from the mouse. This happens hundreds of times each second, making the cursor appear to move very smoothly.

Touch Screen

A touch screen is an alternative to a separate pointing device. With a touch screen the user selects items on the screen by touching the surface. This makes touch screen systems very intuitive and simple to use.

Often used for information terminals in public places e.g. libraries or



museums where mice or keyboards may be stolen or damaged.

They allow simple touch to launch an application or to carry out many of the functions of pointing devices such as a mouse. One of the main uses of touchscreen technology is in mobile phones.

At present, there are three major types of touchscreen technologies applied to mobile phone screens:

-  Capacitive
-  Infra-red
-  Resistive.

We will consider how each of these technologies work and what their main benefits and drawbacks are.

INFRA-RED (heat and optical)		CAPACITIVE
<p>Heat-sensitive</p> <ul style="list-style-type: none"> - Uses glass as the screen material. - Needs a warm object (e.g. fingers) to carry out an input operation. 	<p>Optical</p> <ul style="list-style-type: none"> - Uses glass as the screen material. - Uses an array of sensors (in the form of a grid); the point of contact is based on which grid coordinate is touched. 	<ul style="list-style-type: none"> • This is made up of many layers of glass that act like a capacitor, creating electric fields between the glass plates in layers. • When the top glass layer is touched, the electric current changes and the coordinates where the screen was touched is determined by an on-board microprocessor.
Benefits		Benefits
<ul style="list-style-type: none"> • Both systems allow multi-touch capabilities. • The optical system allows the use of bare fingers, gloved fingers or a stylus for input. • Both systems have good screen durability; it takes a major impact to break the glass. 		<ul style="list-style-type: none"> • This is a medium cost technology. • Screen visibility is good even in strong sunlight. • It permits multi-touch capability. • The screen is very durable; it takes a major impact to break the glass.
Drawbacks		Drawbacks
<ul style="list-style-type: none"> • It is relatively expensive technology. • Heat-sensitive system only allows bare fingers to be used for input (gloved fingers or stylus don't work). • Both systems (optical and heat-sensitive) have fairly good screen visibility in strong sunlight. 		<ul style="list-style-type: none"> • Allows only the use of bare fingers as the form of input; although the latest screens permit a special stylus to be used.

RESISTIVE

- This makes use of an upper layer of polyester (a form of plastic) and a bottom layer of glass.
- When the top polyester layer is touched, the top layer and the bottom layer complete a circuit.
- Signals are then sent out which are interpreted by a microprocessor, the calculations determine the coordinates of where the screen was touched.

Benefits

- It is relatively inexpensive technology.
- It is possible to use bare fingers, gloved fingers or a stylus to carry out an input operation.

Drawbacks

- Screen visibility is poor in strong sunlight.
- It doesn't permit multi-touch capability.
- The screen durability is only fair; it is vulnerable to scratches and the screen wears out through time.



Microphone

An input device that converts **sound** into a signal that can be fed into a computer.

The signal from a microphone is usually **analogue** so, before it can be processed by a computer, it must be converted into digital data. An [Analogue-to-Digital Converter \(ADC\)](#) is used for this (usually built into the computer's sound card)

Many headphones now come with microphones to allow them to be used with chat and phone applications

3D- Scanners:

A **3D scanner** is a device that analyses a real-world object or environment to collect data on its shape and possibly its appearance (e.g. colour). The collected data can then be used to construct digital **three-dimensional models** using **3D-printers**.

Many different technologies can be used to build these 3D-scanning devices; each technology comes with its own limitations, advantages and costs. Many limitations in the kind of objects that can be digitised are still present, for example, optical technologies encounter many difficulties with shiny, mirroring or transparent objects.

Types of 3D-Scanners: There are a variety of technologies for digitally acquiring the shape of a 3D object. A well established classification divides them into two types:



contact scanners



non-contact scanners.

There are a variety of technologies that fall under each of these categories.

Contact 3D scanners:

Contact 3D scanners probe the subject through physical touch, while the object is in contact with or resting on a **precision flat surface plate**, ground and polished to a specific maximum of surface roughness. Where the object to be scanned is not flat or cannot rest stably on a flat surface, it is supported and held firmly in place by a fixture.

The scanner mechanism may have three different forms:

- A carriage system with rigid arms held tightly in perpendicular relationship and each axis gliding along a track. Such systems work best with flat profile shapes.
- An articulated arm with rigid bones and high precision angular sensors. The location of the end of the arm involves complex math calculating the wrist rotation angle and hinge angle of each joint.
- A combination of both methods may be used, such as an articulated arm suspended from a traveling carriage, for mapping large objects with interior cavities or overlapping surfaces.



Non-contact 3D-scanners:

Active scanners emit some kind of radiation or light and detect its reflection or radiation passing through object in order to probe an object or environment.

Possible types of emissions used include **light, ultrasound** or **x-ray**.

Hand-held laser scanners

Hand-held laser scanners create a 3D image through the triangulation mechanism described above: a laser dot or line is projected onto an object from a hand-held device and a sensor (typically a **charge-coupled device** or **position sensitive device**) measures the distance to the surface.

Data is collected in relation to an internal coordinate system and therefore to collect data where the scanner is in motion the position of the scanner must be determined.

External tracking often takes the form of a **laser tracker** (to provide the sensor position) with **integrated camera (to determine the orientation of the scanner)**.

Both techniques tend to use **infra-red** or **Light-emitting diodes** attached to the scanner which are seen by the camera(s) through filters providing resilience to ambient lighting.

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'. Together these 2D 'slices' make up a representation of the 3D solid object.

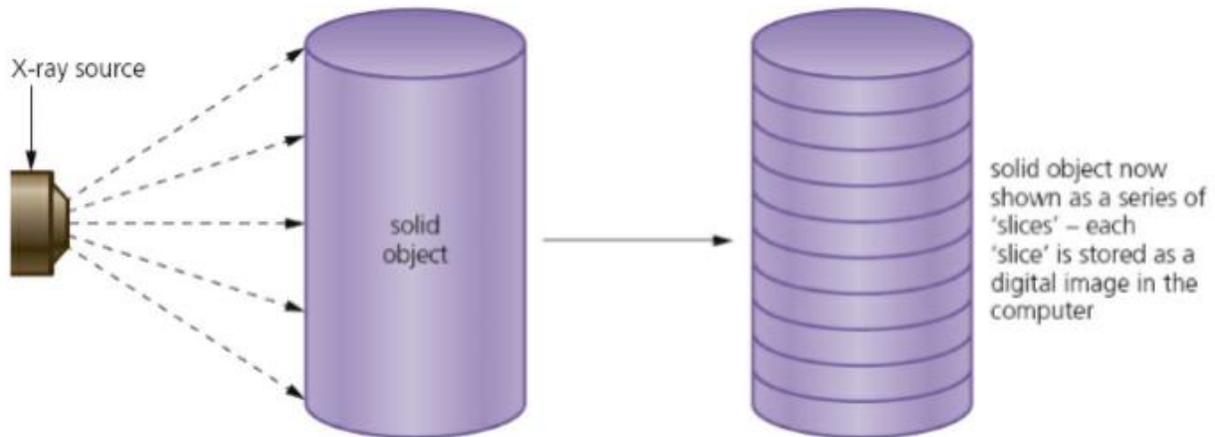
Each slice is built up by use of X-rays, radio frequencies or gamma imaging; although a number of other methods exist. Each 'slice' is then stored as a digital image in the computer memory.

The whole of the solid object is represented digitally in the computer memory. Depending on how the image is formed, the type of tomographic scanner can have different names. For example:



 X-rays	CT scanners	computerised tomography
 radio	radio frequencies	MRI magnetic resonance imaging
 gamma rays	SPECT	single photon emission computed tomography

Figure below shows a simple example of how tomography works.



Output devices

An output device is a piece of hardware that is used to display or output data which has been processed or has been stored on the computer.

There are many different kind of output devices such as inkjet, laser and 3D printers; 2D and 3D cutters; speakers and head phones; actuators; flat panel display screens including Liquid Crystal Display (LCD) and Light-Emitting Diodes (LED); LCD projectors and Digital Light Projectors (DLP)

Printer:

Printer is an output device that prints character and graphics on paper or other materials.

Laser Printer:

LASER PRINTERS differ greatly from inkjet printers in the way they print pages. They 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 (inkjet printers print the page line by line).

The following steps describes briefly the stages that occur when a document is printed using a laser printer:

1. Laser printer uses a laser scanning a drum to print with powdered ink, known as toner.
2. The printer places an even, negative, static charge on a photoconductive drum.
3. It scans a very narrow laser beam across the surface of the rotating drum.
4. The laser beam causes the negative charge to leak away wherever it shines on the drum.
5. The drum revolves past a supply of toner which is also charged negatively. The toner is attracted onto those regions of the drums surface where no charge remains.
6. Toner particles are repelled by those regions that remain charged because **they were not lit by the laser's beam**.
7. **The printer rapidly switches the beam on and off** to draw the required pattern of output.
8. A roller presses a sheet of paper against the rotating drum and the toner particles are attracted to the **positively charged paper with greater force** due to laws of physics.
9. Another roller **presses the paper against a heated 'fuser' roller. The heated toner melts and bonds to the paper**, producing a printed copy.
10. If there are four drums with four different colors of toner the printer can print in color.

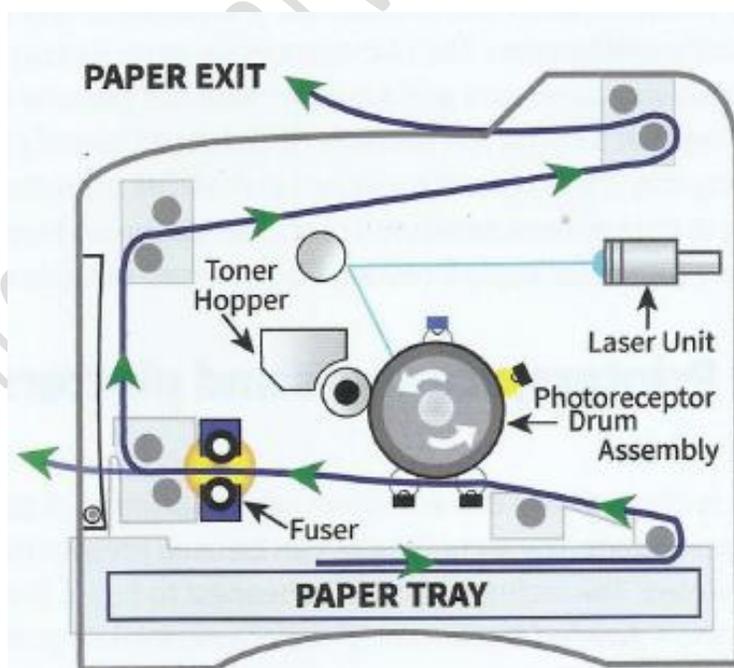


Figure 3.06 A schematic diagram of a laser printer



Inkjet printer

An inkjet printer uses a print head to propel differently sized droplets of ink, measuring just Pico liters, from a number of fine nozzles onto a sheet of paper. The printhead scans from side to side on stabilizer bar and rollers feed the paper forward.

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 (see Section 4.3)
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

Speakers:

Speakers are one of the most common output devices used with computer systems. The purpose of speakers is to produce audio output that can be heard by the listener.

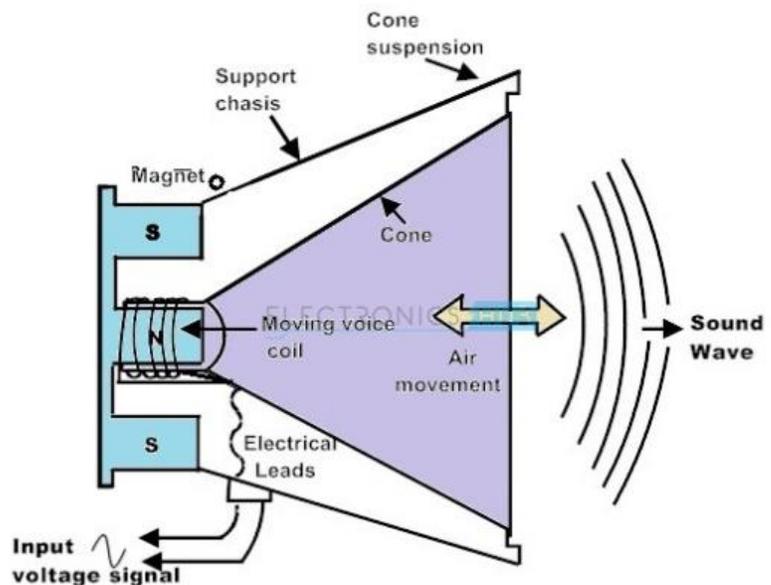
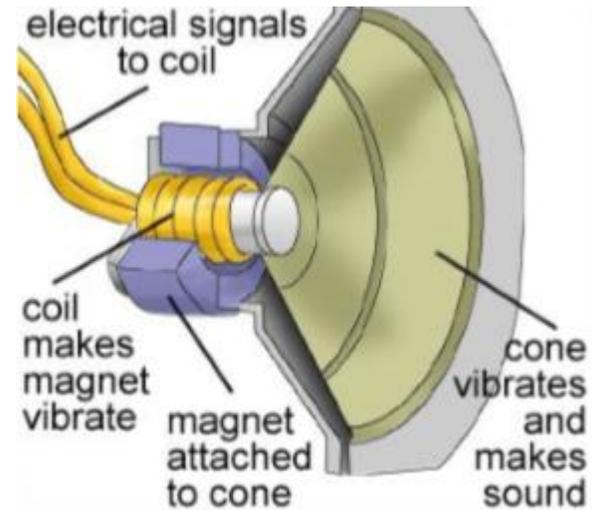
Speakers are transducers that convert electromagnetic waves into sound waves. Speakers use magnets to convert electricity into sound waves. This is a basic principle of physics.

Sound is made when an object makes the particles around it vibrate. These vibrations travel through the air, and reach your ears. Our brain interprets this motion as sound. High frequencies of sound are made when the wavelength of the vibrations are close together. Low frequencies occur when they are farther apart. The amplitude of the vibrations causes the level of volume you hear.

To make these vibrations, speakers have a set of magnets. One of them is called the permanent magnet. It **doesn't move or change polarity and is made of a magnetic metal like iron. The other magnet is an electromagnet.**

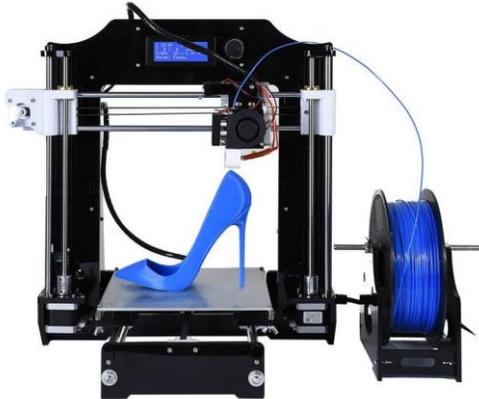
It is a coil of metal wire like copper or aluminum. When an electric current is sent through the electromagnet, it is either attracted to or repelled away from the permanent magnet.

The polarity of the coil can be reversed depending on the current. This back and forth movement causes the diaphragm or cone to vibrate, because it is connected to the magnetic coil. This is the sound that you hear.



3D- Printers:

3D PRINTERS are primarily used in COMPUTER-AIDED DESIGN (CAD) applications. They can produce solid objects which actually work. The solid object is built up layer by layer using materials such as powdered resin, powdered metal, paper or ceramic powder.



The motorcycle was made using an industrial 3D printer.



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

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

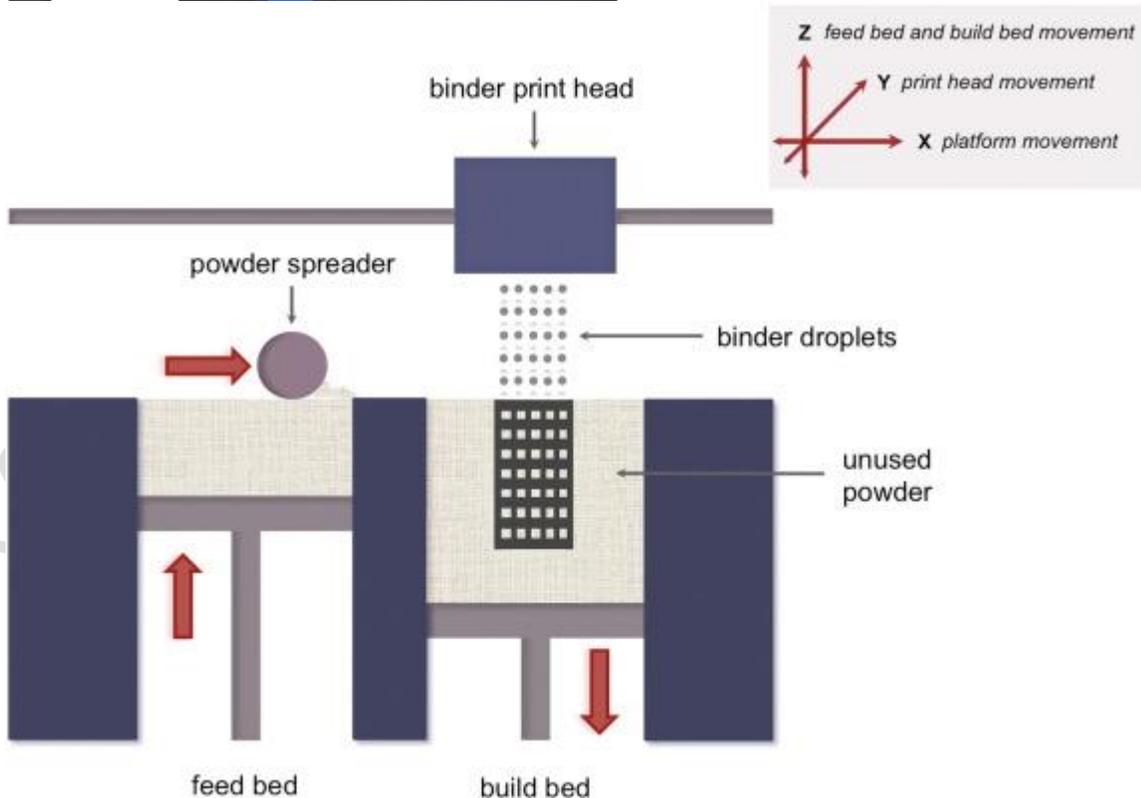
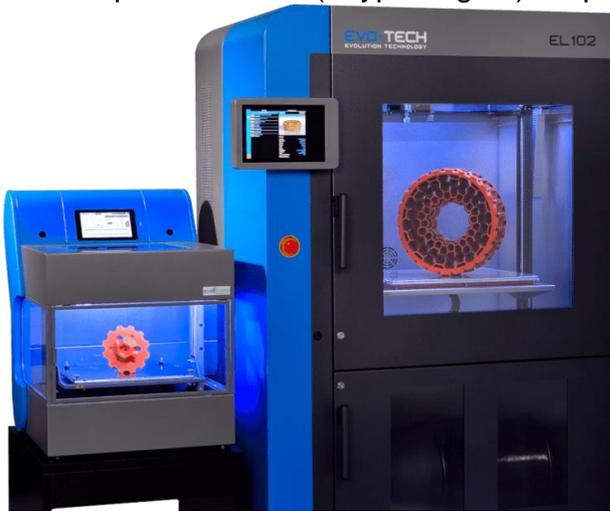
Various types of 3D printers exist; they range from the size of a microwave oven up to the size of a small car.



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). For example, making a statue using a 3D printer would involve building it up layer by layer using powdered stone until the final object was formed.

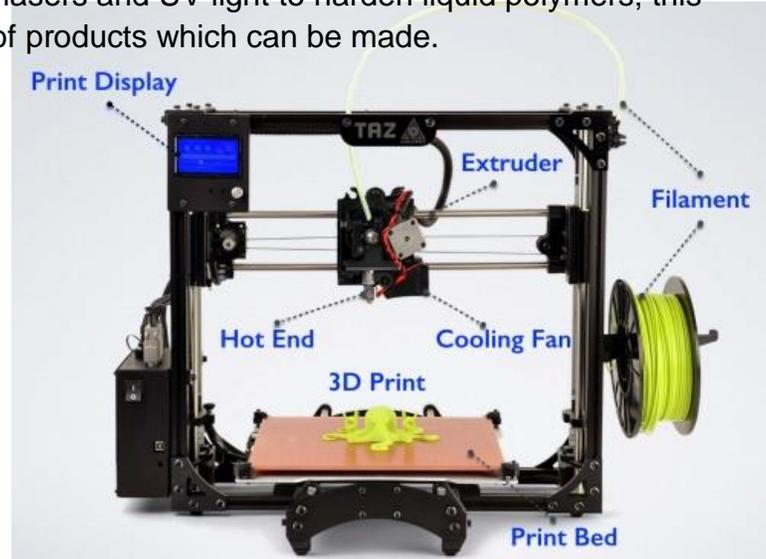
The subtractive method would involve carving the statue out of solid stone (i.e. removing the stone not required) until the final item was produced. Similarly, CNC machining removes metal to form an object; 3

-  3D printing would produce the same item by building up the object from layers of powdered metal.
-  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.

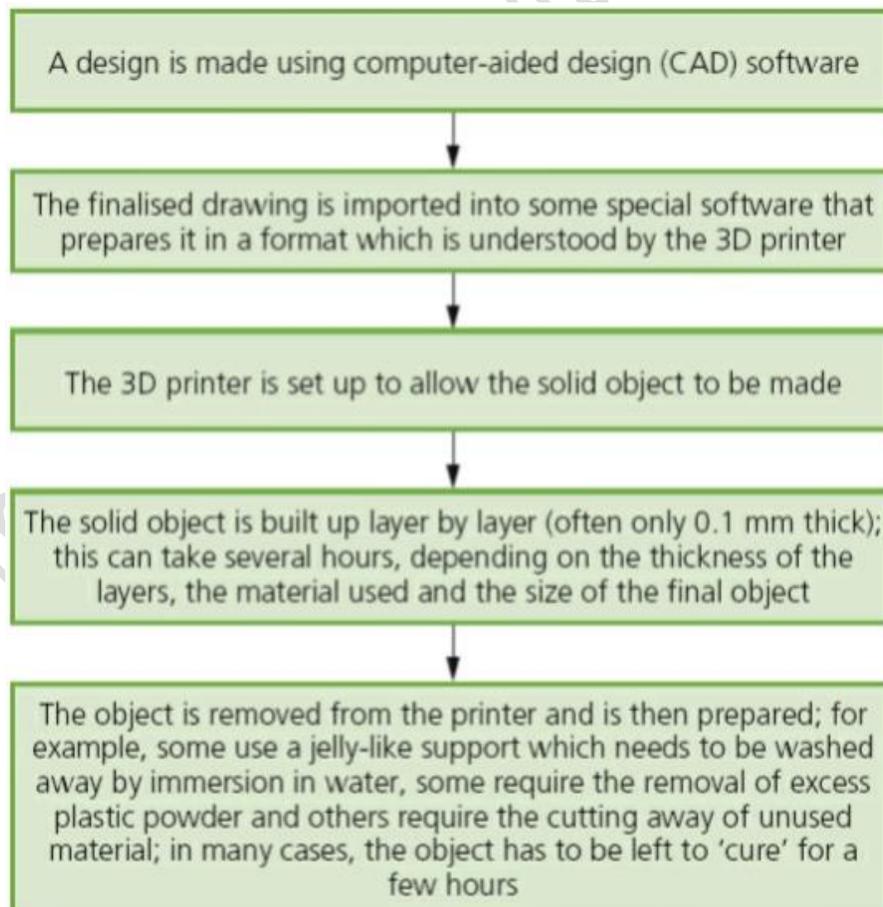




Newer technologies are using lasers and UV light to harden liquid polymers; this further increases the diversity of products which can be made.



How to create a solid object using a 3D printer There are a number of steps in the process of producing an object using these 3D printers. The steps are summarised below



Backing storage:

Backing storage (also called auxiliary storage) stores programs and data for future use. In order to store data while the electricity is switched off or unavailable storage must be non-volatile. Access to backing storage is slower than internal memory. Operating systems and program files are loaded into RAM from backing storage when required for execution.

It is important to distinguish between a storage device and storage medium. The storage device is the machine that stores data; the storage medium is the material on which the device stores data. There are three different types of backing storage device:

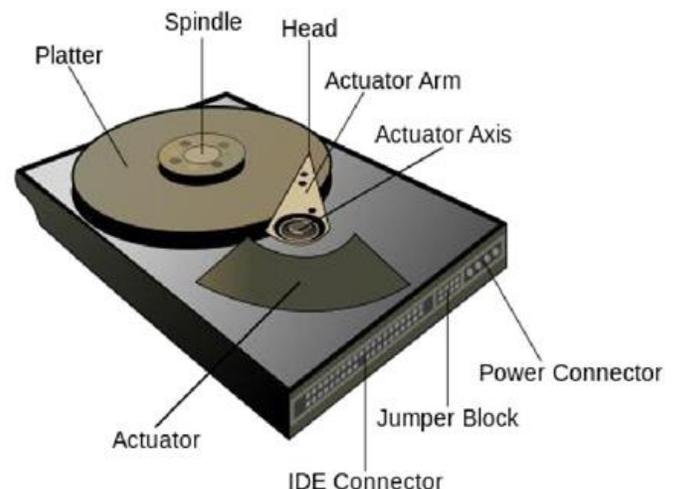
1. Magnetic storage device
2. Optical storage device
3. Solid state storage device.

Magnetic Storage device:

Magnetic media stores data by assigning a magnetic charge to metal. This metal is then processed by a read head, which converts the charges into ones and zeros. Historically, magnetic media has been very popular for storing programs, data, and making backups. It looks set to continue in this role for some time. However, solid state technology is starting to be used more and more, storing programs and data on new devices such as mobile phones and cameras.

Hard Disk Drive (HDD)

Hard disks are usually found inside computers to store programs and data. They are increasingly cheap and more and more companies are using them to back things up. Hard disks can vary in physical size with some disks getting as small as your thumb. The capacity of a commercial disk is currently up to about 4 terabytes allowing users to read and write to them. They are constructed from several key components:





Platter - Metallic disks where One or both sides of the platter are magnetized, allowing data to be stored. The platter spins thousands of times a second around the spindle. There may be several platters, with data stored across them

Head - The head reads magnetic data from the platter. For a drive with several platters there may two heads per platter allowing data to be read from top and bottom of each

Actuator Arm - used to move the read heads in and out of the disk, so that data can be read and written to particular locations and you can access data in a Random fashion, you don't need to read your way through the entire disk to fetch a particular bit of information, you can jump right there. Seek time is very low.

Power connector - provides electricity to spin the platters, move the read head and run the electronics **IDE connector** - allows for data transfer from and to the platters

Jumper block - used to get the disk working in specific ways such as RAID

For the exam you must be able to explain how a hard disk works:

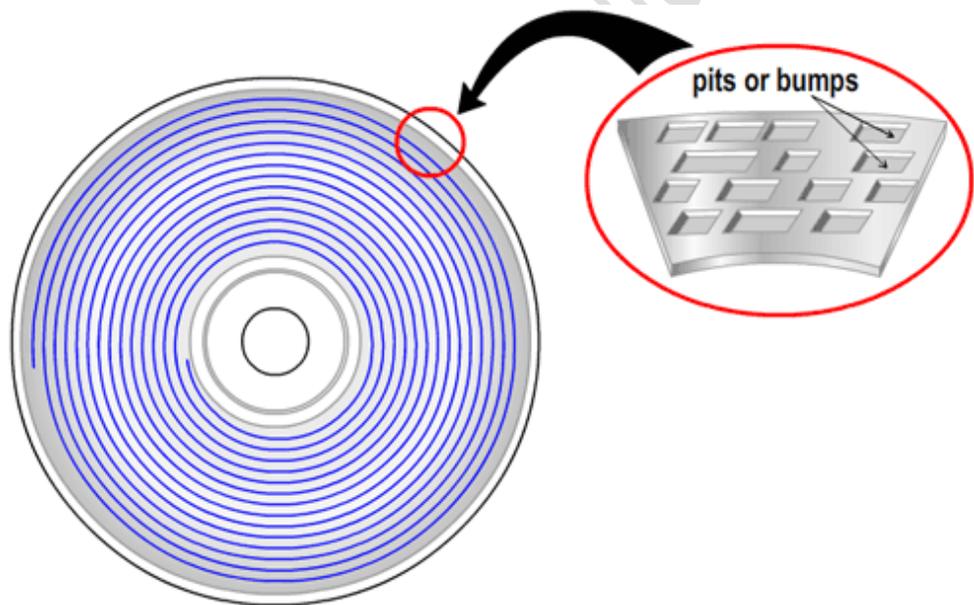
1. The platters spin around the spindle
2. data is requested to be read from a particular area of a platter
3. the actuator arm moves the read head to that track
4. Once the data sector that is required has spun around and under the read head, data is read
5. Read data is sent from the IDE connector to main memory

Writing data is very similar:

1. The platters spin around the spindle
2. data is sent to the hard disk using the IDE connector
3. the actuator arm moves the write head to the track that will be written to
4. Once the data sector that is required has spun around and under the write head, data is written to the platter

Optical media

Optical media works by creating a disc with a pitted metallic surface. There are several different types of disk out there ranging from 650 MB to 128 GB, with the pits and lands getting closer together for higher volume disks. The principle behind how each of them works is the same.

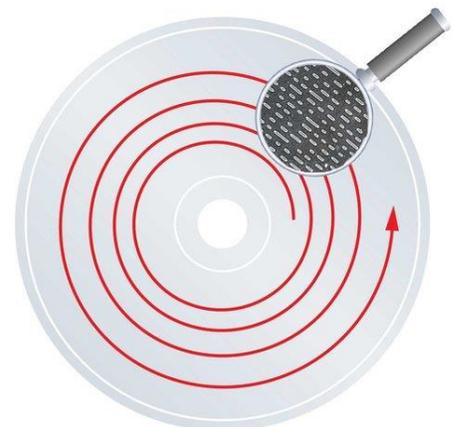


CD/ DVD disks:

CDS and DVDS are described as OPTICAL STORAGE DEVICES.

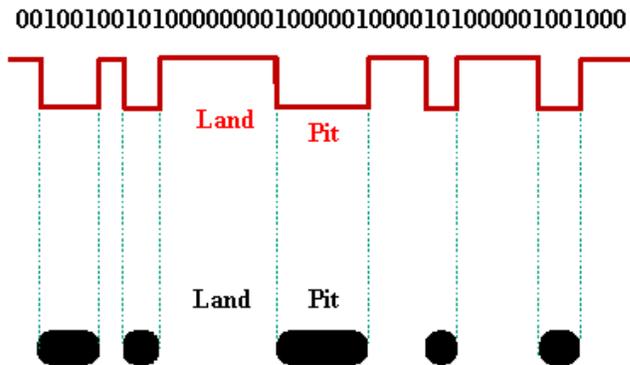
Laser light is used to read data and to write data in the surface of the disk.

Both CDs and DVDs use a thin layer of metal alloy or light-sensitive organic dye to store the data. As can be



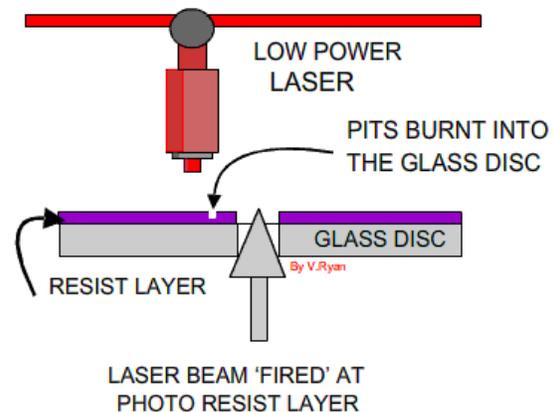
seen in Figure, both systems use a single, spiral track which runs from the centre of the disk to the edge.

The data is stored in 'pits' and 'bumps' on the spiral track. A red laser is used to read and write the data. CDs and DVDs can be designated R (write once only) or RW (can be written to or read from many times).



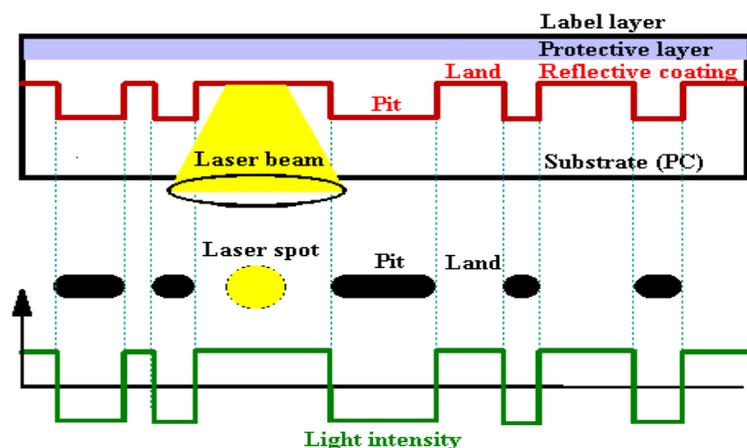
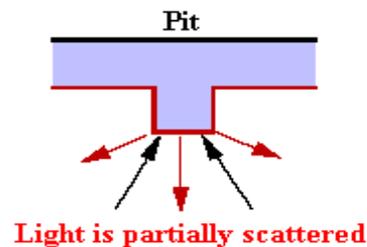
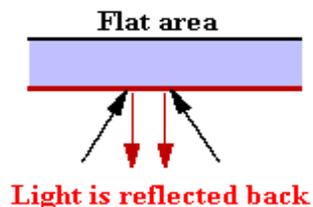
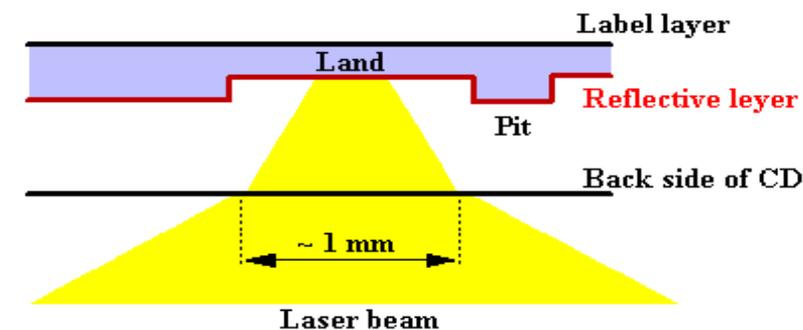
Writing to a CD/ DVD

1. A single track runs in a spiral pattern from the center of the disc to the outside, this track is made of pits and lands to represent the ones and zeroes of binary data
2. A low-powered laser is shone onto the CD-ROM, burning pits into the metal
3. The disc spins and the laser follows the track, putting the binary data onto the CD in a spiral track
4. The data has been written

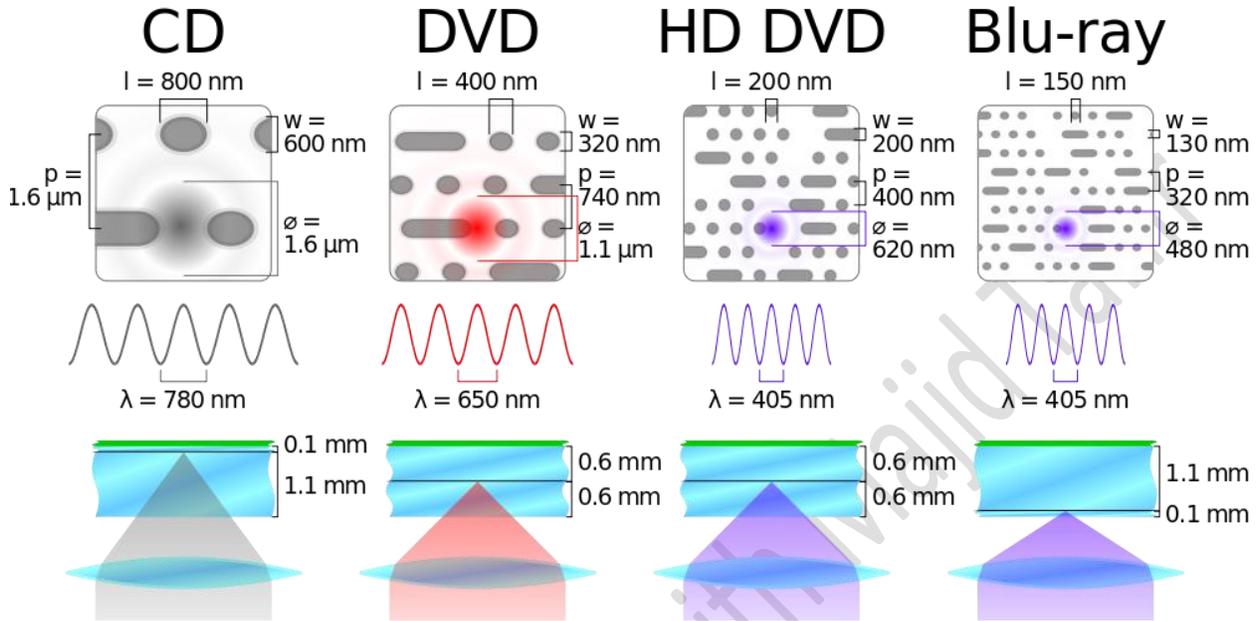


Reading from a CD/DVD

1. A single track runs in a spiral pattern from the center of the disc to the outside, this track is made of pits and lands to represent the ones and zeroes of binary data
2. A low-powered laser is shone on the metallic surface and the reflection is captured in a photodiode sensor, the lands reflect differently to the pits, meaning it can tell the difference between a 1 and a 0
3. The disc spins and the laser follows the track
4. The binary data (the 1s and 0s) are put together and the CD-ROM has been read



Difference between CD, DVD, HD-DVD and Blu-Ray disks:



The main differences are:

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	red	650 nm	two 0.6 mm polycarbonate layers	0.74 μm
Blu-ray	blue	405 nm	single 1.1 mm polycarbonate layer	0.30 μm

- 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 'bumps' can be much smaller; consequently, Blu-ray can store up to five times more data than normal DVD
- Blu-ray uses a single 1.1 mm thick polycarbonate disk; normal DVDs use a sandwich of two 0.6 mm thick disks
- using two sandwiched layers can cause BIREFRINGENCE (light is refracted into two separate beams causing reading errors); because Blu-ray uses only one layer, the disks don't suffer from birefringence
- Blu-ray disks automatically come with a secure encryption system which helps to prevent piracy and copyright infringement.

Device	Type	Size	Image
<input type="checkbox"/> CD-ROM <hr/> <input type="checkbox"/> CD-R <input type="checkbox"/> CD-RW	<input type="checkbox"/> Read Only <input type="checkbox"/> Write once then Read only <input type="checkbox"/> re-Writable	650 - 900 MB	
<input type="checkbox"/> DVD-ROM <input type="checkbox"/> DVD-R <input type="checkbox"/> DVD-RW <input type="checkbox"/> DVD-RAM	<input type="checkbox"/> Read Only <input type="checkbox"/> Write once then Read only <input type="checkbox"/> re-Writable <input type="checkbox"/> re-Writable	4.7 - 9.4 GB	
<input type="checkbox"/> Blu-ray (BD) Disc <input type="checkbox"/> HD DVD (obsolete)	Re-Writable and Read Only versions available. Uses a blue laser, that is able to recognise smaller pits and lands, which allows for the pits and lands to be more closely packed, and so store more data	25 - 128 GB	

Solid State Storage Media:

What is a Solid State Drive?

A Solid State Drive is a storage device which emulates a rotating hard disk drive but is made from semiconductor components.



Typically SSD are made using NAND flash memory, but a DRAM based device would also be considered a solid state drive. The 2.5" form factor is most popular, but flash cards and modules such as CompactFlash, SD Card, mSATA all the way up to PCIe boards are also SSD.

Since a SSD has no moving parts, it handles shock and vibration much better than a rotating disk drive. Without the spindle motor needed to rotate platters and an actuator to move the heads to the correct location, a SSD generally uses much less power than its mechanical HDD counterpart.



Smaller Physical Size – which could also be removable in the case of CompactFlash, SD Card, etc.



Lower Power Usage

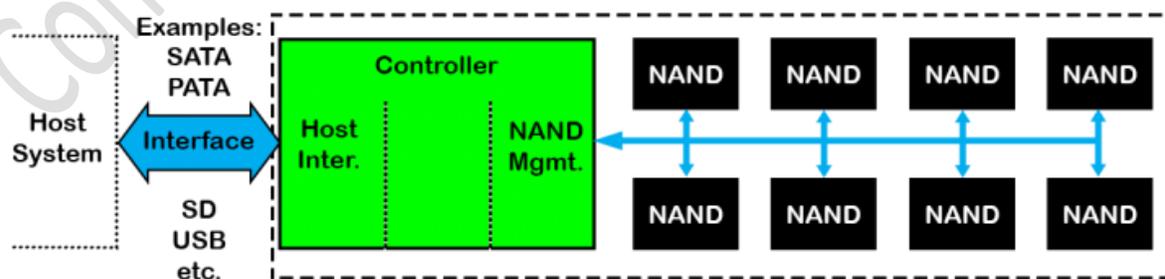


Higher Performance – especially small block transfers used for transactional processing



Lower Absolute Cost – if only a small amount of memory is needed

How do Solid State Drives Work?



Basic Solid State Drive (SSD) Architecture

A Solid State Drive emulates a hard disk drive in a host computer or embedded system. It does this with a specially designed controller which has the same electrical interface as a rotating hard disk drive. In many cases it has the exact same connector and physical form factor as the hard drive as well.

Other circuitry of the controller manages the NAND flash memory which stores the data of the SSD. While the controller function appears relatively straight forward, there are many behind the scenes issues with NAND flash management and it requires a great deal of planning to create a reliable SSD controller.



How long do Solid State Drives last?

The life of a SSD depends on the application which it is being used. Unlike rotating hard disk drives which SSD were designed to replace, the SSD memory has a finite number of write/erase cycles. This makes the management of data written to the SSD extremely important.

As an example, a consumer 2.5" SATA SSD could last for decades in low intensity application, but may only last a few weeks in a high intensity data logging application.

Solid state storage devices are electronic and made as integrated circuits or chip. The currently predominant technology is flash memory, which like ROM holds data that are non-volatile but can be erased and rewritten in large blocks. We often refer to this as non-volatile memory.

Memory cards

Digital cameras use a slightly different form of solid-state memory, known as **XD (eXtreme Digital)** or **SD (Secure Digital)** cards. The technology is essentially the same as memory sticks.

These memory cards allow photos to be transferred from camera to computer via the USB port. Many printers and computers also have card slots allowing the device to read the memory card directly.

Each memory card is made up of NAND chips (see Section 6.4.2) and, as with all solid-state memories, there are no moving parts. All solid-state memories need to be treated with some care.

If they are removed from a device before being made safe (for example, while still in the middle of a read–write operation), data on the memory card or stick can be corrupted. In extreme cases, the memory card or stick become totally unusable. It is essential to wait for the message: 'it is now safe to remove your device' if plugged into a computer. With mobile phones or digital cameras, it is usually advisable to switch them off before removal of the card.

How flash memory works—the simple explanation

Flash works using an entirely different kind of transistor that stays switched on (or switched off) even when the power is turned off. A normal transistor has three connections (wires that control it) called the source, drain, and gate.



Think of a transistor as a pipe through which electricity can flow as though it's water.



One end of the pipe (where the water flows in) is called the source—think of that as a tap or faucet. The other end of the pipe is called the drain—where the water drains out and flows away.



In between the source and drain, blocking the pipe, there's a gate. When the gate is closed, the pipe is shut off, no electricity can flow and the transistor is off. In this state, the transistor stores a zero.



When the gate is opened, electricity flows, the transistor is on, and it stores a one. But when the power is turned off, the transistor switches off too. When you switch the power back on, the transistor is still off, and since you can't know whether it was on or off before the power was removed, you can see why we say it "forgets" any information it stores.

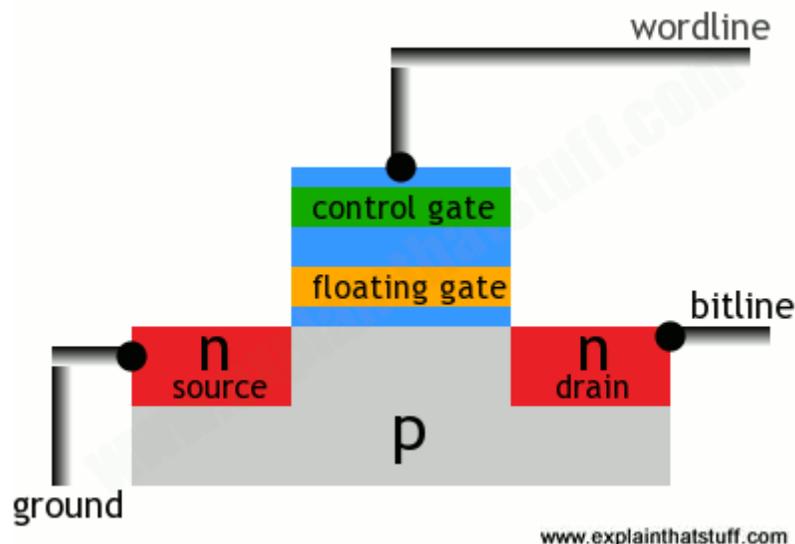


A flash transistor is different because it has a second gate above the first one. When the gate opens, some electricity leaks up the first gate and stays there, in between the first gate and the second one, recording a number one. Even if the power is turned off, the electricity is still there between the two gates. That's how the transistor stores its information whether the power is on or off. The information can be erased by making the "trapped electricity" drain back down again.

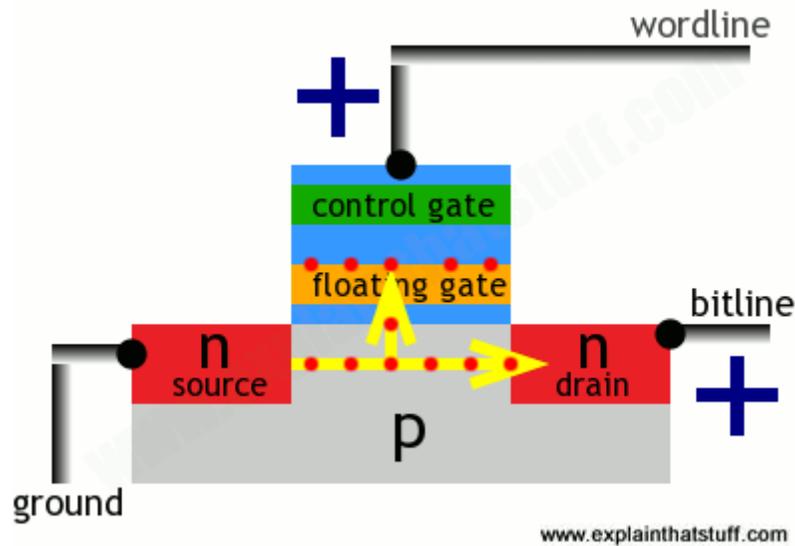
How flash memory works—a more detailed explanation

That's a very glossed over, highly simplified explanation of something that's extremely complex. If you want more detail, it helps if you read our article about [transistors](#) first, especially the bit at the bottom about MOSFETs—and then read on.

The transistors in flash memory are like MOSFETs only they have two gates on top instead of one. This is what a flash transistor looks like inside. You can see it's an n-p-n sandwich with two gates on top, one called a control gate and one called a floating gate. The two gates are separated by oxide layers through which current cannot normally pass:



In this state, the transistor is switched off—and effectively storing a zero. How do we switch it on? Both the source and the drain regions are rich in electrons (because they're made of n-type silicon), but electrons cannot flow from source to drain because of the electron deficient, p-type material between them. But if we apply a positive voltage to the transistor's two contacts, called the bitline and the wordline, electrons get pulled in a rush from source to drain. A few also manage to wriggle through the oxide layer by a process called tunneling and get stuck on the floating gate:



The presence of electrons on the floating gate is how a flash transistor stores a one. The electrons will stay there indefinitely, even when the positive voltages are removed and whether there is power supplied to the circuit or not. The electrons can be flushed out by putting a negative voltage on the wordline—which repels the electrons back the way they came, clearing the floating gate and making the transistor store a zero again.

Not an easy process to understand, but that's how flash memory works its magic!

References:

www.wikipedia.com

<https://www.explainthatstuff.com/flashmemory.html>

https://courses.cs.washington.edu/courses/cse370/01au/minirproject/BAH/bah_cdpage.html

Cambridge International AS & A level by Silvia Langfield and Dave Duddell.

IGCSE Computer Science by Hodder Education.