UNIT 5:

I/O and Memory organization

Input Devices

Input device enables the user to send data, information, or control signals to a computer. The Central Processing Unit (CPU) of a computer receives the input and processes it to produce the output.

Some of the popular input devices are:

1. Keyboard
2. Mouse
3. Scanner
4. Joystick
5. Light Pen
6. Digitizer
7. Microphone
8. Magnetic Ink Character Recognition (MICR)
9. Optical Character Reader (OCR)
10. Digital Camera
11. Paddle
12. Steering Wheel
13. Gesture recognition devices
14. Light Gun
15. Touch Pad
16. Remote
17. Touch screen

18. VR
19. Webcam
20. Biometric Devices

---

1) Keyboard

The keyboard is a basic input device that is used to enter data into a computer or any other electronic device by pressing keys. It has different sets of keys for letters, numbers, characters, and functions. Keyboards are connected to a computer through USB or a Bluetooth device for wireless communication.

Types of keyboards: There can be different types of keyboards based on the region and language used. Some of the common types of keyboards are as follows:

i) QWERTY Keyboard:

It is the most commonly used keyboard with computers in modern times. It is named after the first six letters of the top row of buttons and is even popular in countries that do not use Latin-based alphabet. It is so popular that some people think that it is the only type of keyboard to use with computers as an input device.

ii) AZERTY Keyboard:

It is considered the standard French keyboard. It is developed in France as an alternative layout to the QWERTY layout and is mainly used in France and other European countries. Some countries have manufactured their own versions of AZERTY.

Its name is derived from the first six letters that appear on the top left row of the keyboard. The Q and W keys in AZERTY keyboard are interchanged with A and Z keys in QWERTY keyboard. Furthermore, in AZERTY keyboard M key is located to the left of the L key.

AZERTY keyboard differs from QWERTY keyboard not only in the placement of letters but also in many other ways, e.g., it gives emphasis on accents, which is required for writing European languages like French.

iii) DVORAK Keyboard:

This type of keyboard layout was developed to increase the typing speed by reducing the finger movement while typing. The most frequently used letters are kept in a home row to improve typing.

---

2) Mouse

The mouse is a hand-held input device which is used to move cursor or pointer across the screen. It is designed to be used on a flat surface and generally has left and right button and a scroll wheel between them. Laptop computers come with a touchpad that works as a mouse. It lets you control the movement of cursor or pointer by moving your finger over the touchpad. Some mouse comes with integrated features such as extra buttons to perform different buttons.

The mouse was invented by Douglas C. Engelbart in 1963. Early mouse had a roller ball integrated as a movement sensor underneath the device. Modern mouse devices come with optical technology that controls cursor movements by a visible or invisible light beam. A mouse is connected to a computer through different ports depending on the type of computer and type of a mouse.

Common types of the mouse:

i) Trackball Mouse:

It is a stationary input device that has ball mechanism to move the pointer or cursor on the screen. The ball is half inserted in the device and can be easily rolled with finger, thumb or the palm to move the pointer on the screen. The device has sensor to detect the rotation of ball. It remains stationary; you don't need to move it on the operating surface. So, it is an ideal device if you have limited desk space as you don't need to move it like a mouse.

ii) Mechanical Mouse:

It has a system of a ball and several rollers to track its movement. It is a corded type of mouse. A mechanical mouse can be used for high performance. The drawback is that they tend to get dust into the mechanics and thus require regular cleaning.

iii) Optical Mouse:

An optical mouse uses optical electronics to track its movement. It is more reliable than a mechanical mouse and also requires less maintenance. However, its performance is affected by the surface on which it is operated. Plain non-glossy mouse mat should be used for best results. The rough surface may cause problems for the optical recognition system, and the glossy surface may reflect the light wrongly and thus may cause tracking issues.

iv) Cordless or Wireless Mouse:

As the name suggests, this type of mouse lacks cable and uses wireless technology such as IrDA (infrared) or radio (Bluetooth or Wi-Fi) to control the movement of the cursor. It is used to improve the experience of using a mouse. It uses batteries for its power supply.

---

3) Scanner

The scanner uses the pictures and pages of text as input. It scans the picture or a document. The scanned picture or document then converted into a digital format or file and is displayed on the screen as an output. It uses optical character recognition techniques to convert images into digital ones. Some of the common types of scanners are as follows:

Types of Scanner:

i) Flatbed Scanner:

It has a glass pane and a moving optical CIS or CCD array. The light illuminates the pane, and then the image is placed on the glass pane. The light moves across the glass pane and scans the document and thus produces its digital copy. You will need a transparency adapter while scanning transparent slides.

ii) Handheld Scanner:

It is a small manual scanning device which is held by hand and is rolled over a flat image that is to be scanned. The drawback in using this device is that the hand should be steady while scanning; otherwise, it may distort the image. One of the commonly used handheld scanners is the barcode scanner which you would have seen in shopping stores.

iii) Sheetfed Scanner:

In this scanner, the document is inserted into the slot provided in the scanner. The main components of this scanner include the sheet-feeder, scanning module, and calibration sheet. The light does not move in this scanner. Instead, the document moves through the scanner. It is suitable for scanning single page documents, not for thick objects like books, magazines, etc.

iv) Drum Scanner:

Drum scanner has a photomultiplier tube (PMT) to scan images. It does not have a charge-coupled device like a flatbed scanner. The photomultiplier tube is extremely sensitive to light. The image is placed on a glass tube, and the light moves across the image, which produces a reflection of the image which is captured by the PMT and processed. These scanners have high resolution and are suitable for detailed scans.

v) Photo Scanner:

It is designed to scan photographs. It has high resolution and color depth, which are required for scanning photographs. Some photo scanners come with in-built software for cleaning and restoring old photographs.

---

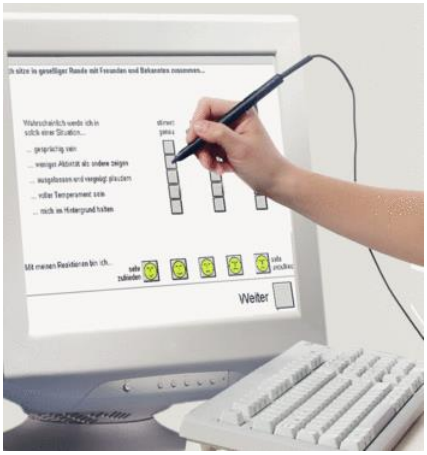
4) Joystick

A joystick is also a pointing input device like a mouse. It is made up of a stick with a spherical base. The base is fitted in a socket that allows free movement of the stick. The movement of stick controls the cursor or pointer on the screen.

The frist joystick was invented by C. B. Mirick at the U.S. Naval Research Laboratory. A joystick can be of different types such as displacement joysticks, finger-operated joysticks, hand operated, isometric joystick, and more. In joystick, the cursor keeps moving in the direction of the joystick unless it is upright, whereas, in mouse, the cursor moves only when the mouse moves.

---

5) Light Pen



A light pen is a computer input device that looks like a pen. The tip of the light pen contains a light-sensitive detector that enables the user to point to or select objects on

the display screen. Its light sensitive tip detects the object location and sends the corresponding signals to the CPU. It is not compatible with LCD screens, so it is not in use today. It also helps you draw on the screen if needed. The first light pen was invented around 1955 as a part of the Whirlwind project at the Massachusetts Institute of Technology (MIT).

---

6) Digitizer

Digitizer is a computer input device that has a flat surface and usually comes with a stylus. It enables the user to draw images and graphics using the stylus as we draw on paper with a pencil. The images or graphics drawn on the digitizer appear on the computer monitor or display screen. The software converts the touch inputs into lines and can also convert handwritten text to typewritten words.

It can be used to capture handwritten signatures and data or images from taped papers. Furthermore, it is also used to receive information in the form of drawings and send output to a CAD (Computer-aided design) application and software like AutoCAD. Thus, it allows you to convert hand-drawn images into a format suitable for computer processing.

---

7) Microphone

The microphone is a computer input device that is used to input the sound. It receives the sound vibrations and converts them into audio signals or sends to a recording medium. The audio signals are converted into digital data and stored in the computer. The microphone also enables the user to telecommunicate with others. It is also used to add sound to presentations and with webcams for video conferencing. A microphone can capture audio waves in different ways; accordingly the three most common types are described below:

i) Dynamic:

It is the most commonly used microphone with a simple design. It has a magnet which is wrapped by a metal coil and a thin sheet on the front end of the magnet. The sheet transfers vibrations from sound waves to the coil and from coil to electric wires which transmit the sound like an electrical signal.

ii) Condenser:

It is designed for audio recording and has a very sensitive and flat frequency response. It has a front plate called diaphragm and a back plate parallel to the front plate. When sound hits the diaphragm, it vibrates the diaphragm and alters the distance between the two plates. The changes in distance are transmitted as electric signals.

iii) Ribbon:

It is known for its reliability. It has a thin ribbon made of aluminum, duraluminum, or nanofilm suspended in a magnetic field. The sound waves cause vibrations in the ribbon, which generate a voltage proportional to the velocity of the vibration. The voltage is transmitted as an electrical signal. Early ribbon microphones had a transformer to increase the output voltage, but modern ribbon microphones come with advanced magnets to produce a strong signal.

---

8) Magnetic Ink Character Recognition (MICR)

MICR computer input device is designed to read the text printed with magnetic ink. MICR is a character recognition technology that makes use of special magnetized ink which is sensitive to magnetic fields. It is widely used in banks to process the cheques and other organizations where security is a major concern. It can process three hundred cheques in a minute with hundred-percent accuracy. The details on the bottom of the cheque (MICR No.) are written with magnetic ink. A laser printer with MICR toner can be used to print the magnetic ink.

The device reads the details and sends to a computer for processing. A document printed in magnetic ink is required to pass through a machine which magnetizes the ink, and the magnetic information is then translated into characters.

---

9) Optical Character Reader (OCR)

OCR computer input device is designed to convert the scanned images of handwritten, typed or printed text into digital text. It is widely used in offices and libraries to convert documents and books into electronic files.

It processes and copies the physical form of a document using a scanner. After copying the documents, the OCR software converts the documents into a two-color (black and white), version called bitmap. Then it is analyzed for light and dark areas, where the dark areas are selected as characters, and the light area is identified as background. It is widely used to convert hard copy legal or historic documents into PDFs. The converted documents can be edited if required like we edit documents created in ms word.

---

10) Digital camera:

It is a digital device as it captures images and records videos digitally and then stores them on a memory card. It is provided with an image sensor chip to capture images, as opposed to film used by traditional cameras. Besides this, a camera that is connected to your computer can also be called a digital camera.

It has photosensors to record light that enters the camera through the lens. When the light strikes the photosensors, each of the sensors returns the electrical current, which is used to create the images.

---

11) Paddle:

It is a simple input device that is widely used in games. It is a wheel that is held by hand and looks like a volume knob on a stereo that is used to increase or decrease the volume. Paddle moves or controls cursor or any other objects in the game in a back-and-forth motion. It is widely used as an alternative to the joystick. Besides this, the term paddle also refers to many handheld devices designed to control a function in an electronic device, computer, etc.

12) Steering wheel:

It is used as an input device in racing video games such as car racing games or in driving programs as virtual simulators to steer a vehicle. It works like the real steering wheel by allowing you to take a right or left turn. A steering wheel may be provided with acceleration and brake pedal devices and a mechanism for shifting gears. Thus, it makes racing games more adventurous and entertaining.

---

13) Gesture recognition devices:

These devices take human gestures as input. There are many such devices that respond to gestures. For example, Kinect is one such device that observes the movement of a player's body and interprets these movements as inputs to video games. This feature is also available in certain tablets and smartphones where you can perform certain tasks such as taking pictures using finger gestures such as swiping, pinching, etc.

14) Light Gun:

As the name suggests, it is a pointing input device that is designed to point at and shoot the targets on the screen in a video game, or arcade, etc. The light gun was used for the first time on the MIT Whirwind computer. When the gun is pointed at the target on the screen and the trigger is pulled, the screen goes blank for a fraction of a second. During this moment, the photodiode, which is present in the barrel, determines where the gun is pointed. For example, shooting ducks in a duck hunt game.

15) Touchpad:

It is usually found in laptops as a substitute for the mouse. It allows you to move or control the cursor on the screen using your finger. Just like a mouse, it also has two buttons for right and left click. Using the touchpad, you can perform all the tasks that you do with a mouse, such as selecting an object on the screen, copy, paste, delete, open a file or folder, and more.

16) Remote:

It is a hardware device designed to control the functioning of a device, e.g., a TV remote that can be used to change channels, increase or decrease the volume, from a distance without leaving the seat. The first cordless TV remote was invented by Dr. Robert Adler of Zenith in 1956. The remote sends the electromagnetic waves to communicate with the device. These waves can be infrared rays, radio waves, etc.

---

17) Touch screen:

It is the display screen of a device such as a smartphone, tablet, etc., that allows users to interact or provide inputs to the device by using their finger. Today, most of the electronic devices come with touchscreen as an alternative to a mouse for navigating a graphical user interface. For example, by touching, you can unlock your phone, open emails, open files, play videos, etc. Besides this, it is used in lots of devices such as Camera, Car GPS, Fitness machine, etc.

The concept of the touch screen was first introduced and published by E.A. Johnson in 1965. The first touch screen was developed at the beginning of the 1970s by CERN engineers Frank Beck and Bent Stumpe.

---

18) VR:

VR stands for virtual reality. It is an artificial or virtual environment which is generated by computers. A person can interact with virtual objects of this artificial environment using some input devices such as headsets, gloves, headphones, etc. For example, he or she can find himself or herself walking on a beach, watching a football match, walking in the sky, etc., without actually doing all this.

---

19) Webcam:

Any camera which is connected to a computer is called a webcam. The in-built camera provided on a computer can also be considered a webcam. It is an input device as it can take pictures, and can be used to record videos if required. The pictures and videos are stored in the computer memory and can be displayed on the screen if required. Although

it works almost the same as the digital camera, it is different from a digital camera, as it is designed to take compact digital photos that can be uploaded easily on the webpages and shared with others through the internet.

---

20) Biometric Devices:

Biometrics refers to a process in which a person is identified through his or her biological features such as fingerprints, eye cornea, face structure, etc. It is done by using biometric devices, which can be of different types based on their scanning features and abilities, such as:

i) Face Scanner:

It is designed to identify a person by scanning his or her face. It takes the face measurements of a person. For example, the distance between eyes, nose, and mouth, etc., accordingly, it confirms the identity of a person. Besides this, it is smart enough to differentiate between a person's picture and the real person.

ii) Hand Scanner:

The hand of a person can also be used to verify his or her identity as every person has a unique pattern of veins in the palm, just like fingerprints. This device takes advantage of this feature; it identifies a person by scanning the palm of his hand. It uses infrared light to scan veins' patterns and blood flowing in them. Palm is even more unique than fingerprints.

iii) Fingerprint Scanner:

It scans the fingerprints to identify people or for biometric authentication. This device is developed, keeping in mind the fact that no two persons in the world can have the same fingerprints. It is widely used in companies as a fingerprint attendance system to mark the attendance of employees. This type of scanners captures the pattern of valleys and ridges found on a finger and store it in the memory or database. When you press your finger on the given space, it verifies the identity by using its pattern-matching software.

iv) Retina or Iris Scanner:

It scans the retina or iris of a person's eye to confirm the identity. This device is more secure than others as it is next to impossible to copy the retina or iris. It works by mapping

the retina's blood vessel patterns of the eye. The blood vessels of retina absorb light more easily as well as can be identified with appropriate lighting.

In this scan, a beam of low-energy infrared light falls on the retina through the scanner's eyepiece. Then, the software captures the network of blood vessels in the retina and uses it to verify a person's identity.

v) Voice Scanner:

It records the voice of a person and digitizes it to create a distinctive voice print or template. The voiceprints are stored in the database, and are used to verify the voice of a person to confirm his or her identity. The person is required to speak in the normal or same voice that was used to create a voice template. It is not much reliable as it can be misused using a tape recording.

Difference between Synchronous and Asynchronous Transmission

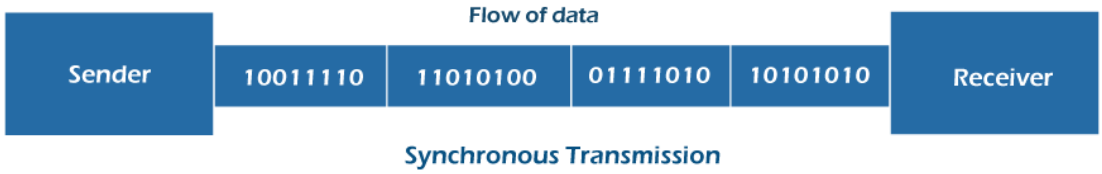
Before starting the topic difference between ***synchronous*** and ***asynchronous transmission***, you must know about the transmission. The action of transferring data or anything from one place to other is referred to as transmission. It is a method of sharing data between two devices linked by a network, also known as communication mode. Synchronous and asynchronous transmissions are the two main types of transmission used in computer networking.

In both synchronous and asynchronous transmission, data is sent between the transmitter and the receiver based on a clock pulse utilized for synchronization. These serial data transmission techniques are both known as synchronous transmission.

In this article, you will learn about the difference between ***Synchronous*** and ***Asynchronous transmission***. But before discussing the differences, you must know about Synchronous and Asynchronous transmission with their advantages and disadvantages.

What is Synchronous Transmission?

Synchronous transmission is an effective and dependable method of sending huge amounts of data. The data travels in a full-duplex method in the type of frames or blocks in Synchronous Transmission. The transmitter and receiver must be synced so that the sender knows where to start the new byte. As a result, every data block is marked with synchronization characters, and the receiving device obtains the data until a certain ending character is found.



Synchronous Transmission

It also allows connected devices to interact in real time. Synchronous transmission can be seen in chat rooms, video conferencing, telephonic talks, and face-to-face interactions. It utilizes the broad-band and voice band channels because they enable quicker speeds of up to ***1200 bps*** and meet the objective of high data transfer speed.

Advantages and Disadvantages of Synchronous Transmission

There are various advantages and disadvantages of synchronous transmission. Some advantages and disadvantages of synchronous transmission are as follows:

Advantages

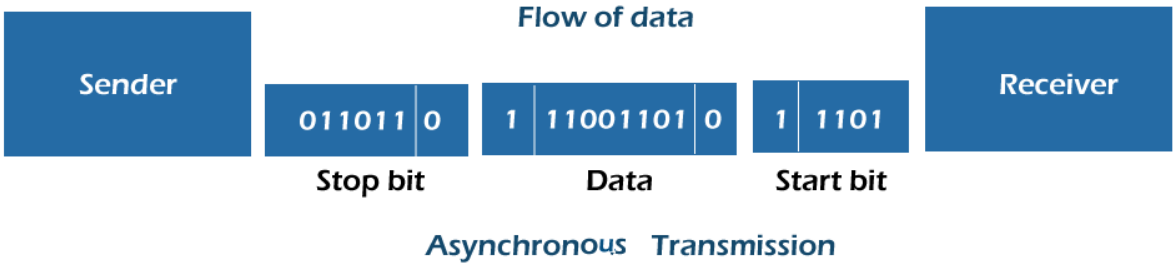
1. It aids the user in transferring a huge amount of data.
2. Every byte is sent without a pause before the next.
3. It also helps to reduce timing errors.
4. It allows connected devices to communicate in real-time.

Disadvantages

1. The sender and receiver must operate at the same clock frequency simultaneously.
2. The accuracy of the received data is determined by the receiver's capacity to count the received bits precisely.

What is Asynchronous Transmission?

Asynchronous transmission is also referred to as start and stop transmission. It sends data from the transmitter to the receiver using the flow control approach and synchronizes data between the source and destination without utilizing a clock.



Asynchronous Transmission

This transmission technique sends ***8 bits*** or one letter at a time. In this system, each character transmits the start bit before the transmission process begins, and it also transmits the stop bit when the character is sent. The total number of bits is ***10***, including the character, start, and stop bits.

It employs character-based synchronization for the receiving terminal to synchronize with receiving data on a character. It is easy, quick, and inexpensive and doesn't need two-way communication. Asynchronous transmission is demonstrated via letters, televisions, emails, forums, and radios.

Asynchronous transmission makes use of voice-band channels that are narrow and operate at a slower speed. In this case, the transmitting device operates manually or intermittently.

Advantages and Disadvantages of Asynchronous Transmission

There are various advantages and disadvantages of Asynchronous transmission. Some advantages and disadvantages of Asynchronous transmission are as follows:

Advantages

1. It doesn't require synchronizing the receiver and transmitter.
2. It is a very flexible technique of data transmission.
3. This kind of transmission is simple to implement.
4. It allows users to send signals from sources with varying bit rates.
5. When the data byte transmission is complete, the data transmission may be resumed.

Disadvantages

1. The timing errors may occur because synchronization is difficult to determine.
2. These bits could be mistakenly recognized due to the noise on the channel.
3. The start and stop bits are extra bits that must be utilized in asynchronous transmission.
4. It transmits information at a slower rate.

Key differences between Synchronous and Asynchronous Transmission

Here, you will learn about the key differences between ***Synchronous*** and ***Asynchronous Transmission***. Some of the main differences between Synchronous and Asynchronous Transmission are as follows:

1. A synchronous transmission is a form of transmission that enables synchronized communication by sharing a common clock pulse between the sender and the receiver. In contrast, an asynchronous transmission is a form of transmission in which the transmitter and receiver have their internal clocks and hence don't require an external common clock pulse.
2. Data is transmitted as frames in synchronous transmission. In contrast, asynchronous transmission transmits data one byte at a time.
3. In synchronous transmission, the amount of time between two successive broadcasts remains constant. In contrast, the time gap between two successive transmissions is random in asynchronous transmission.
4. The data transfer rate of synchronous transmission is fast. In contrast, the data transfer rate of asynchronous transmission is slow.

5. Synchronous transmission is complicated and costly. In contrast, asynchronous transmission is simple and cost-effective.
6. Synchronous transmission is simple to design. In contrast, asynchronous transmission is both complex in nature and design.
7. There is no gap between data in Synchronous transmission due to the common clock pulse. Whereas there is a gap between the data bytes in asynchronous transmission. It has start and end bits between which actual data is present.
8. Local storage is not necessary for synchronous transmission at the terminal end. In contrast, local buffer storages are needed to construct blocks at both ends of the line in asynchronous transmission.
9. In Synchronous Transmission, the voice-band and broad-band channels are primarily utilized. In contrast, asynchronous transfer is employed with voice-band channels that have a limited type.

Head-to-head comparison between Synchronous and Asynchronous Transmission

Here, you will learn the head-to-head comparisons between Synchronous and Asynchronous Transmission. The main differences between Synchronous and Asynchronous Transmission are as follows:

Features	Synchronous Transmission	Asynchronous Transmission
Definition	It is a type of transmission that enables synchronized communication by sharing a common clock pulse between the transmitter and the receiver.	It is a form of transmission in which the transmitter and receiver have their own internal clocks and hence don't require an external common clock pulse.
Basic	The transmission begins with the block header, which contains a bit sequence.	It employs the start and stops bits to precede and follow a character.

Data Unit	Data is transmitted as frames in synchronous transmission.	It transmits data one byte at a specific time.
Storage	It doesn't require any storage at the terminal end.	The local buffer storages are needed to construct blocks at both ends of the line in asynchronous transmission.
Transmission speed	The data transfer rate of synchronous transmission is fast.	The data transfer rate is slow.
Cost	It is complicated and costly.	It is simple and cost-effective.
Gap between the data	There is no gap between data in Synchronous transmission due to the common clock pulse.	There is a gap between the data bytes, and it has start and end bits between which actual data is present.
Implementation	It is implemented by hardware and software.	It is only implemented by hardware.
Time interval	The time delay between two transmissions is constant.	The time delay between two transmissions is random.
Bits	The start and stop bits are not utilized in data transmission.	The start and stop bits are used to transmit data with additional overhead.
Synchronized clocks	It doesn't require any synchronized clocks.	It needs synchronized clocks at both ends.
Complexity	It is simple and easy to design.	It is complex to design.
Band Channels	It mainly uses both voice-band and broad-band channels.	It mainly uses voice-band channels that have a limited type.

Examples	Some examples of synchronous transmission are Video Conferencing, Chat Rooms, and Telephonic Conversations.	Some examples of asynchronous transmission are emails, letters, forums, etc.

Modes of I/O Data Transfer

Data transfer between the central unit and I/O devices can be handled in generally three types of modes which are given below:

1. Programmed I/O
2. Interrupt Initiated I/O
3. Direct Memory Access

Programmed I/O

Programmed I/O instructions are the result of I/O instructions written in computer program. Each data item transfer is initiated by the instruction in the program.

Usually the program controls data transfer to and from CPU and peripheral. Transferring data under programmed I/O requires constant monitoring of the peripherals by the CPU.

Interrupt Initiated I/O

In the programmed I/O method the CPU stays in the program loop until the I/O unit indicates that it is ready for data transfer. This is time consuming process because it keeps the processor busy needlessly.

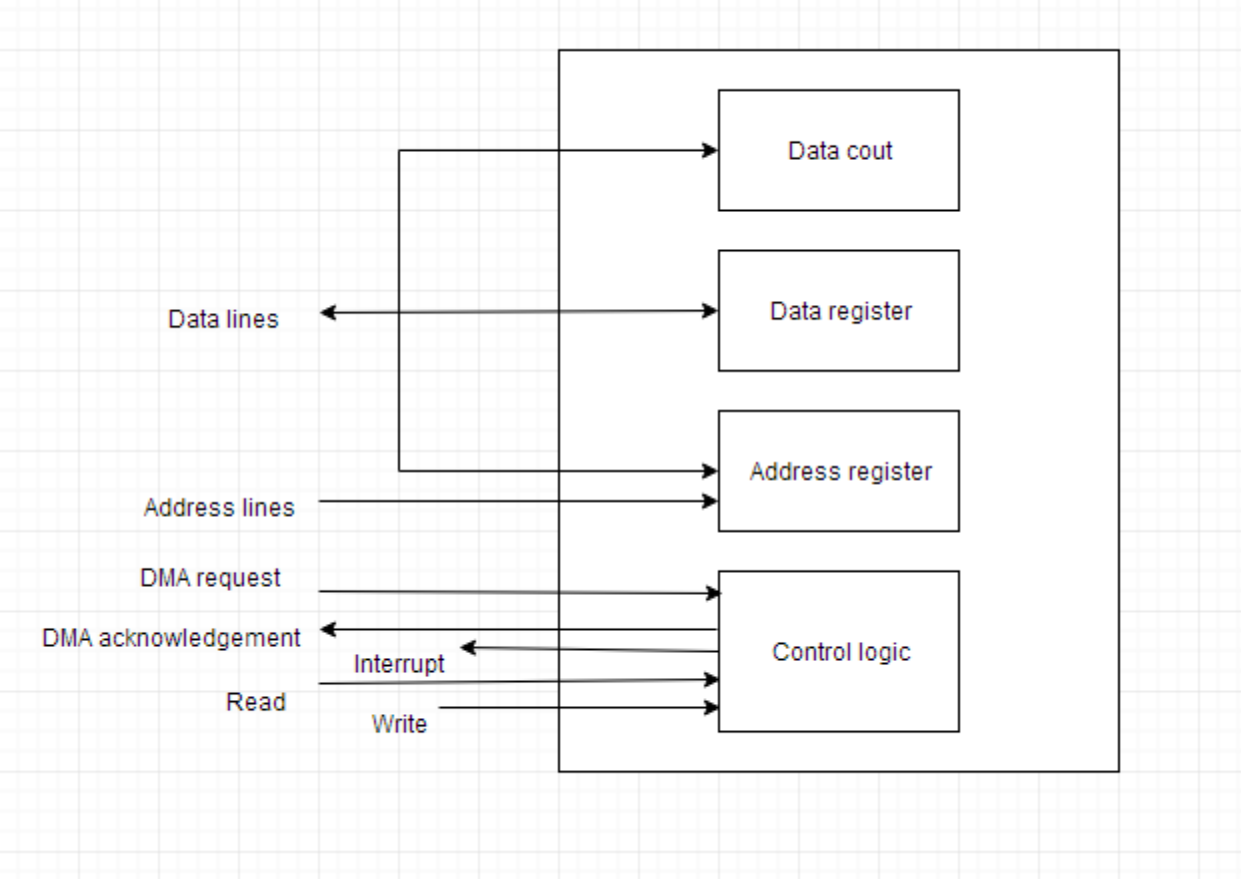
This problem can be overcome by using **interrupt initiated I/O**. In this when the interface determines that the peripheral is ready for data transfer, it generates an interrupt. After receiving the interrupt signal, the CPU stops the task which it is processing and service the I/O transfer and then returns back to its previous processing task.

Direct Memory Access

Removing the CPU from the path and letting the peripheral device manage the memory buses directly would improve the speed of transfer. This technique is known as **DMA**.

In this, the interface transfer data to and from the memory through memory bus. A DMA controller manages to transfer data between peripherals and memory unit.

Many hardware systems use DMA such as disk drive controllers, graphic cards, network cards and sound cards etc. It is also used for intra chip data transfer in multicore processors. In DMA, CPU would initiate the transfer, do other operations while the transfer is in progress and receive an interrupt from the DMA controller when the transfer has been completed.



Above figure shows block diagram of DMA

Direct Memory Access

Transferring data processes involves processors, memory, and input-output devices. It is the processor which initiates the transfer process, and it ends when the data gets stored. Because of this, the processor's load increases, and it has to stay ideal, resulting in decreasing efficiency. The **Direct Memory Access**(DMA) controller performs the role of the station master and helps in speeding up the process of transferring data between input-output devices and memory. The **Direct Memory Access**(DMA) controller tries to decrease the processor's involvement in the entire process. In this blog, we will learn about the concept of **Direct Memory Access**.

What is DMA Controller

The hardware device used for this purpose(Direct Memory Access) is known as a DMA controller. This controller is present in input-output devices as part of its interface circuit. It performs the role of the station master and helps speed up the process of transferring data between Input-output devices and memory. And it tries to decrease the processor's involvement in the entire process.

Block Diagram

DMA block diagram

Source: https://binaryterms.com/

Whenever a processor is asked to read or write a data block, that is, to transfer a data block, it instructs the DMA controller by sending the following information:-

1. The first information is that the data should be read or the data should be written in memory. This information is transferred through reading or writing control lines between the processor and DMA controls to control the mental unit.
2. The processor also provides the first address/data block in memory, where the memory block in memory should be read or where the data block should be written in memory. The DMA controller keeps this in the address register. They are also called first-time address registers.
3. The processor also provides the starting address of the data block in memory, where the data in memory should be read or where the data block should be written in memory. The DMA controller keeps this in the address register. They are also called starting-time address registers.
4. The processor also sends the number of words or say word count, how many words to read or write. This information gets stored in a data count or a word count register.
5. Essential is the address of the I / O device that is looking for reading or writing data. This information gets stored in the data register.

Working

DMA Controller Data Transfer

Source: https://binaryterms.com/

The following points describe the working of the **Direct Memory Access**(DMA) controller as per the given diagram:-

1. Whenever an input-output device wants to transfer some data to the memory or from memory, it sends a DMA Request that can be referred to as DRQ to the DMA controller.
The DMA controller accepts this (DMA Request)DRQ and asks the CPU to capture or hold a few clock cycles by requesting a Hold request(HLD).
2. The CPU receives a hold request (HLD) from the DMA controller, releases the bus, and sends the hold acknowledgment (HLDA) to the DMA controller.
3. After obtaining a hold acknowledgment (HLDA), the DMA controller approves the input-output device (DACK) to perform data transfer. The DMA controller takes power from the system bus and transfers data to or from memory.
4. Once the data transfer completes, the DMA raises an interrupt to notify the processor about completing the data transfer function. The processor can control the bus again and begin processing where it left off.

The DMA controller can now be a separate unit assigned to various I / O devices, or it can also be part of the I / O device interface.

Modes of DMA controller for transferring data

The alternative methods are not effective for large block data transfer. And the DMA controller does this work at a faster rate and effectively for large blocks.

The modes in which the **Direct Memory Access**(DMA) controller transfers the data are:

Burst Mode

In this mode, when the DMA controller gets the charge of the system bus, then after that, only after completion of the data transfer does it get to release the system bus, and till then, the processor has to wait for those system buses.

Cycle Stealing Mode

When the DMA controller is in this mode, it forces the processor to stop the operation and leave control over the bus for a short time to the DMA controller. After each byte transfer, the DMA controller releases the bus and re-requests the system bus. In this way, the DMA controller steals the clock cycle to transfer all bytes.

Transparent Mode

Here, the DMA controller only takes over the system bus if the processor does not require a system bus.

Memory Hierarchy

A memory unit is an essential component in any digital computer since it is needed for storing programs and data.

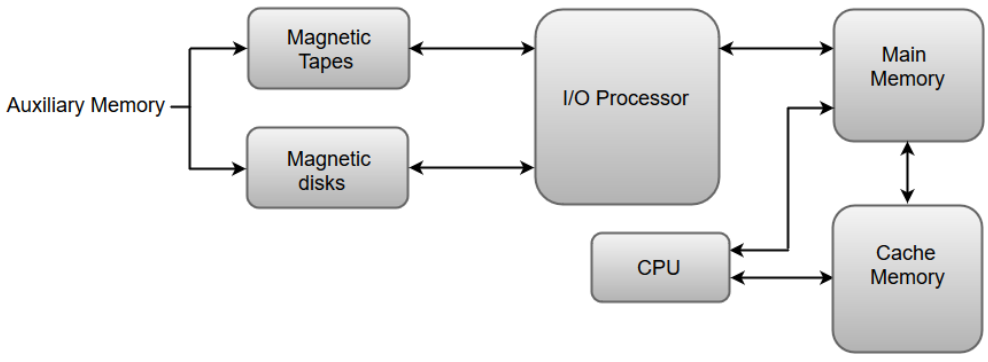
Typically, a memory unit can be classified into two categories:

1. The memory unit that establishes direct communication with the CPU is called **Main Memory**. The main memory is often referred to as RAM (Random Access Memory).
2. The memory units that provide backup storage are called **Auxiliary Memory**. For instance, magnetic disks and magnetic tapes are the most commonly used auxiliary memories.

Apart from the basic classifications of a memory unit, the memory hierarchy consists all of the storage devices available in a computer system ranging from the slow but high-capacity auxiliary memory to relatively faster main memory.

The following image illustrates the components in a typical memory hierarchy.

Memory Hierarchy in a Computer System:



Auxiliary Memory

Auxiliary memory is known as the lowest-cost, highest-capacity and slowest-access storage in a computer system. Auxiliary memory provides storage for programs and data that are kept for long-term storage or when not in immediate use. The most common examples of auxiliary memories are magnetic tapes and magnetic disks.

A magnetic disk is a digital computer memory that uses a magnetization process to write, rewrite and access data. For example, hard drives, zip disks, and floppy disks.

Magnetic tape is a storage medium that allows for data archiving, collection, and backup for different kinds of data.

Main Memory

The main memory in a computer system is often referred to as **Random Access Memory (RAM)**. This memory unit communicates directly with the CPU and with auxiliary memory devices through an I/O processor.

The programs that are not currently required in the main memory are transferred into auxiliary memory to provide space for currently used programs and data.

I/O Processor

The primary function of an I/O Processor is to manage the data transfers between auxiliary memories and the main memory.

Cache Memory

The data or contents of the main memory that are used frequently by CPU are stored in the cache memory so that the processor can easily access that data in a shorter time. Whenever the CPU requires accessing memory, it first checks the required data into the cache memory. If the data is found in the cache memory, it is read from the fast memory. Otherwise, the CPU moves onto the main memory for the required data.

Difference Between Cache Memory and Virtual Memory

Computer Memory is just like a human brain used to store data and instructions either temporarily or permanently. It is a physical device capable of storing information temporarily like RAM (Random Access Memory), or permanently, like ROM (Read Only Memory). The main memory refers to physical memory, and it is known as RAM. In computer memory, we can edit or update only the data that is in the main memory. We can say that when we want to access the secondary storage media or any file that must be loaded into the main memory from the secondary device.

Cache Memory

Cache Memory is a faster memory used by the central processing unit (CPU). It is a memory that helps to reduce the access time for files or data that is recently used by the main memory. It is smaller in size, high-speed memory, and located near a processor core that stores the copies of the information or instruction frequently used by the main memory locations.

Furthermore, it behaves like a buffer between the CPU and the main memory to hold those data or programs most frequently called by the CPU. For example, whenever we execute a program by the processor, it fetches data from the main memory and fetched data to be copied to the cache memory. When the program's copy is already available to the cache memory, it directly calls the processor to execute it; otherwise, the program/files are fetched from memory. Hence, it reduces the access time of the data from the main memory.

Advantages of Cache Memory

1. The access time of files or instruction in the cache memory is less than the main memory.
2. It stores frequently used data by the main memory.

3. It is the faster computer memory as compared to the main memory.
4. Store the program in a cache memory that is executed within a short time.

Disadvantages of Cache Memory

1. It has limited space to store the data.
2. It is very costly as it is a fast memory to access the data.

Virtual Memory

Virtual Memory is used in the computer memory to increase the storage capacity of the main memory. It is a logical storage unit of a computer that creates an illusion to execute a large program that may not be completely placed in the main memory. Furthermore, it allows the user to load or store the data program or files larger than the size of the main memory.

Advantages of Virtual Memory

1. Virtual Memory allows the users to run more than one application at once.
2. It enhances the degree of multiprogramming in the virtual memory.
3. Virtual Memory is a logical unit of computer memory that increases the main memory capacity by storing or executing a large size program than the main memory.
4. It does not require any fixed limit on the degree of multiprogramming.
5. It increases the CPU utilization in the virtual memory.
6. It is required whenever the system does not have much space to store any big programs or files.

Disadvantages of Virtual Memory

1. Virtual Memory can slow the process of application in the system.
2. It may take more time to switch between the applications.
3. It reduces the stability of the system.
4. It allows the user to lesser hard disk space for its use in the system.

Difference Between Cache Memory and Virtual Memory

S. N.	Parameter Difference	Cache Memory	Virtual Memory
1.	**Definition**	Cache Memory is the high speed of computer memory that reduces the access time of files or documents from the main memory.	Virtual Memory is a logical unit of computer memory that increases the capacity of main memory by storing or executing programs of larger size than the main memory in the computer system.
2.	**Memory Unit**	Cache Memory is defined as a memory unit in a computer system.	Virtual Memory is not defined as a memory unit.
3.	**Size**	Its size is very small as compared to Virtual Memory.	Its size is very large as compared to the Cache Memory.
4.	**Speed**	It is a high-speed memory as compared to Virtual Memory.	It is not a high-speed memory as compared to the Cache Memory.
5.	**Operation**	Generally, it stores frequently used data in the cache memory to reduce the access time of files.	The virtual memory keeps those data or programs that may not completely be placed in the main memory.
6.	**Management**	Cache Memory is controlled by the hardware of a system.	Whereas the virtual memory is control by the Operating System (OS).
7.	**Mapping**	It does not require a mapping structure to access the files in Cache Memory.	It requires a mapping structure to map the virtual address with a physical address.