

MODULE 2

INPUT/OUTPUT ORGANISATION

One of the features of a computer is its ability to exchange data with other devices. A human uses a keyboard and display screen to process text and graphics. Here discuss various ways for I/O operations are performed.

- In computing, input/output, or I/O, refers to the communication between computers and the outside world.
- Inputs are the signals or data received by the system, and outputs are the signals or data sent from it.
- I/O devices are used by a person to communicate with a computer.
- Some of the input devices are keyboard, mouse, track ball, joystick, touch screen, digital camera, webcam, image scanner, fingerprint scanner, barcode reader, microphone and so on.
- Some of the output devices are speakers, headphones, monitors and printers.
- Devices for communication between computers, such as modems and network cards, typically serve for both input and output.

ACCESSING I/O DEVICE

- ☐ A simple arrangement to connect I/O devices to a computer is to use a single bus structure. It consists of three sets of lines to carry
 - Address
 - Data
 - Control Signals.
- ☐ When the processor places a particular address on address lines, the devices that recognize this address respond to the command issued on the control lines
- ☐ The processor request either a read or write operation and the requested data are transferred over the data lines.

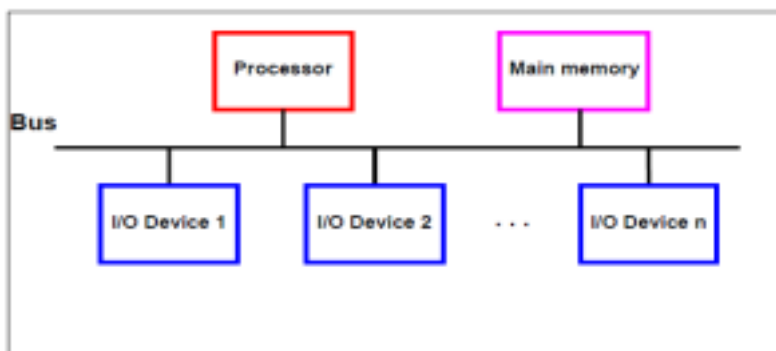


Figure 1 A Single bus structure

MEMORY-MAPPED I/O

- The arrangement of I/O devices and the memory share the same address space is called memory-mapped I/O.
- With memory-mapped I/O, any machine instruction that can access memory can be used to transfer data to or from an I/O device.
- For example, if DATAIN is the address of the input buffer associated with the keyboard, the instruction
 - ❑ **Move DATAIN, R₀** - Reads the data from DATAIN then into processor register R₀.
 - ❑ **Move R₀, DATAOUT** - Send the contents of register R₀ to location DATAOUT.
 - ❑ **DATAIN** - Input buffer associated with keyboard.
 - ❑ **DATAOUT** - Output data buffer of a display unit / printer.

Figure 5.2 illustrates the hardware required to connect an I/O device to the bus.

- The **address decoder** enables the device to recognize its address when this address appears on the address lines.
- The **data register** holds the data being transferred to or from the processor.
- The **status register** contains information relevant to the operation of the I/O device.
- Both the data and status registers are connected to the data bus and assigned unique addresses.
- The address decoder, the data and status registers, and the control circuitry required to coordinate I/O transfers constitute the **device's interface circuit**.

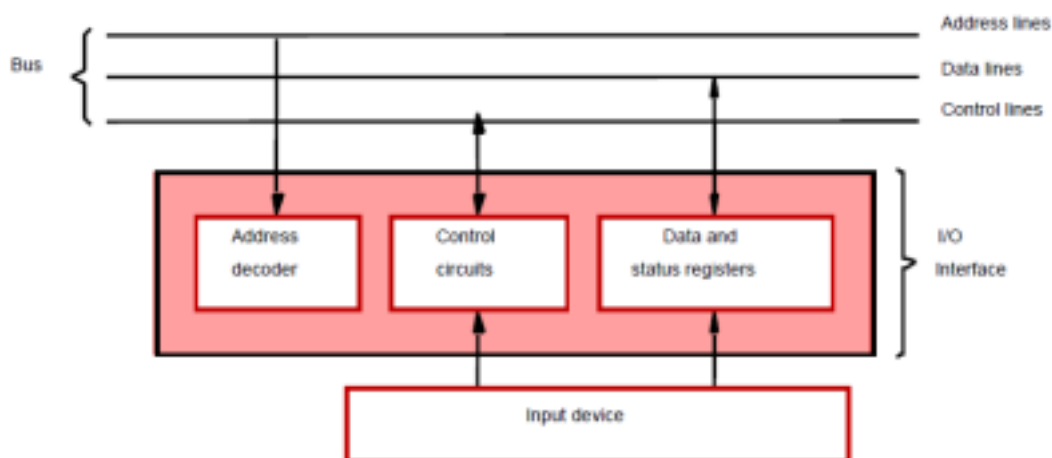


Figure 5.2 I/O interface for an input device

- For an input device such as a keyboard, a status flag, SIN, is included in the interface circuit as part of the status register.
- SIN = 1, when a character is entered at the keyboard, SIN = 0, once the char is read by processor.

- For an output device, SOUT status flag is used.

PROGRAMMED I/O

The processor executes a program that gives it direct control of the I/O operation, including sensing device status, sending a read or write command, and transferring the data. When the processor issues a command to the I/O module, it must wait until the I/O operation is complete. If the processor is faster than the I/O module, this is wasteful of processor time. The overall operation of the programmed I/O can be summarized as follows:

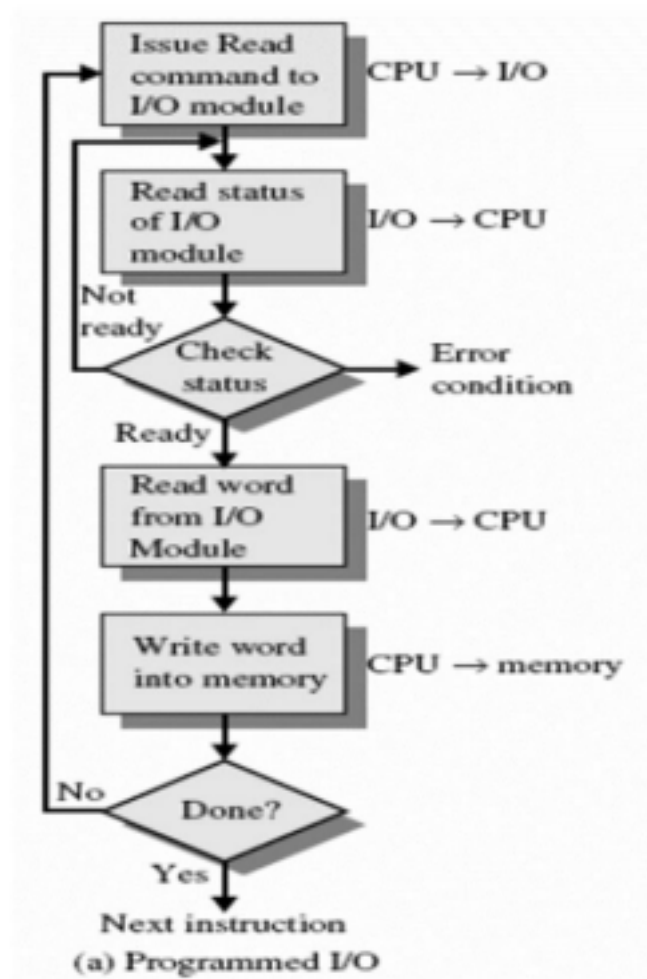
- 1. The processor is executing a program and encounters an instruction relating to I/O operation.*
- 2. The processor then executes that instruction by issuing a command to the appropriate I/O module.*
- 3. The I/O module will perform the requested action based on the I/O command issued by the processor (READ/WRITE) and set the appropriate bits in the I/O status register.*
- 4. The processor will periodically check the status of the I/O module until it finds that the operation is complete.*

Advantages

- simple to implement
- very little hardware support

Disadvantages

- busy waiting
- ties up CPU for long period with no useful work



Mechanisms to handle I/O operations

Polling-the processor repeatedly checks a status flag to achieve the required synchronization between Processor & I/O device. (ie) the processor polls the device.

Interrupt - Synchronization is achieved by having I/O device send special signal over the bus where it is ready for data transfer operation.

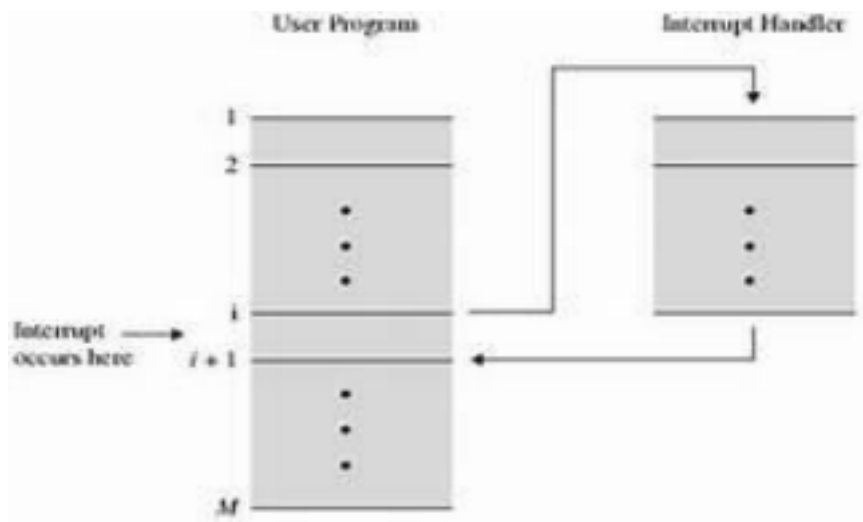
DMA - It is a technique used for high speed I/O device. Here, the device interface transfer data directly to or from the memory without continuous involvement by the processor.

INTERRUPTS

When a program enters a wait loop, it will repeatedly check the device status. During this period, the processor will not perform any function. There are many situations where other tasks can be performed while waiting for an I/O device to become ready. To allow this to happen, we can arrange for the I/O device to alert the processor when it becomes ready.

It can do so by sending a hardware signal called an **interrupt** to the processor. At least one of the bus control lines called an **interrupt request line** is usually dedicated for this purpose.

Since the processor is no longer required to continuously check the status of external devices, it can use the waiting period to perform other useful functions. Indeed by using interrupts such waiting periods can ideally be eliminated.



Above figure shows what happens when an interrupt occurs while executing the i th instruction of the user program.

- The processor first completes the execution of instruction i . Then it loads the PC (Program Counter) with the address of the first instruction of the ISR.
- After the execution of ISR, the processor has to come back to instruction $i + 1$.
- Therefore, when an interrupt occurs, the current contents of PC which point to $i+1$ is put in temporary storage in a known location.
- A return from interrupt instruction at the end of ISR reloads the PC from that temporary storage location, causing the execution to resume at instruction $i+1$.
- When the processor is handling the interrupts, it must inform the device that its request has been recognized so that it remove its interrupt requests signal.
- This may be accomplished by a special control signal called the **interrupt acknowledge signal**.
- The task of saving and restoring the information can be done automatically by the processor.

The interrupt cycle might performs the following tasks:

- Suspend execution of current program
- Save context
- Set PC to start address of interrupt handler routine

- Process interrupt
- Restore context and continue interrupted program

Types of Interrupts:

1. Hardware Interrupts: If the signal for the processor is from external device or hardware is called hardware interrupts.

2. Software Interrupts: Software interrupt can also divided in to two types. They are

Normal Interrupts and Exception

INTERRUPTS HARDWARE

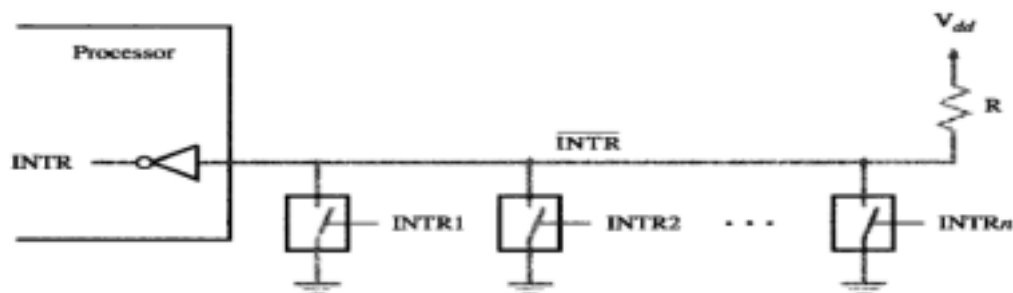


Figure 4.6 An equivalent circuit for an open-drain bus used to implement a common interrupt-request line.

A

- A single interrupt request line may be used to serve n devices. All devices are connected to the line via switches to ground.
- To **request** an interrupt, a device **closes** its associated switch, the voltage on INTR line drops to 0(**zero**). • If all the interrupt request signals (INTR1 to INTRn) are **inactive**, all switches are open and the voltage on INTR line is equal to **Vdd**..
- When a device requests an interrupt by closing its switch, the voltage on the line drops to 0, causing INTR request signal received by the processor to go to 1.
- Since closing one or more switches will cause line voltage to drop to 0, the value of INTR is the logical OR of the requests from individual devices. ie;

$$\text{INTR} = \text{INTR1} + \dots + \text{INTRn}$$

HANDLING MULTIPLE INTERRUPTS

When more than one device raises an interrupt request signal, then additional information is needed to decide which device to be considered first. The following methods are used to decide which device to select: Polling, Vectored Interrupts, and Interrupt Nesting. These are explained as following below.

• Polling:

In polling, the first device encountered with with IRQ bit set is the device that is to be serviced first. Appropriate ISR is called to service the same. It is easy to implement but a lot of time is wasted by interrogating the IRQ bit of all devices. **Daisy chain:** A

hardware poll.

- **Vectored Interrupts:**

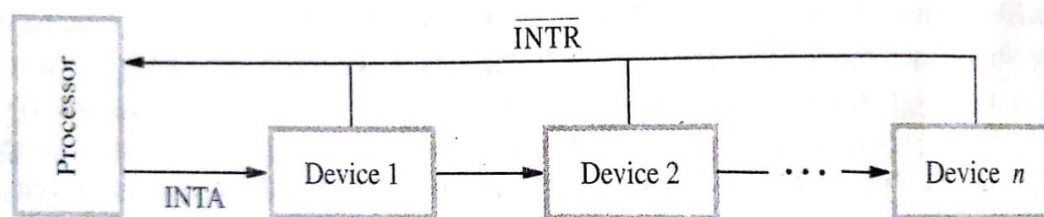
In vectored interrupts, a device requesting an interrupt identifies itself directly by sending a special code to the processor over the bus. This enables the processor to identify the device that generated the interrupt. The special code can be the starting address of the ISR or where the ISR is located in memory, and is called the interrupt vector.

- **Interrupt Nesting:** In this method, I/O device is organized in a priority structure. Therefore, interrupt request from a higher priority device is recognized where as request from a lower priority device is not. Processor accepts interrupts only from devices/processes having priority more than it.

- **Simultaneous requests:** The processor must have some mechanisms to decide which request to service when simultaneous requests arrive. Here, daisy chain and arrangement of priority groups as the interrupt priority schemes are discussed. Priority based simultaneous requests are considered in many organizations.

Daisy chain

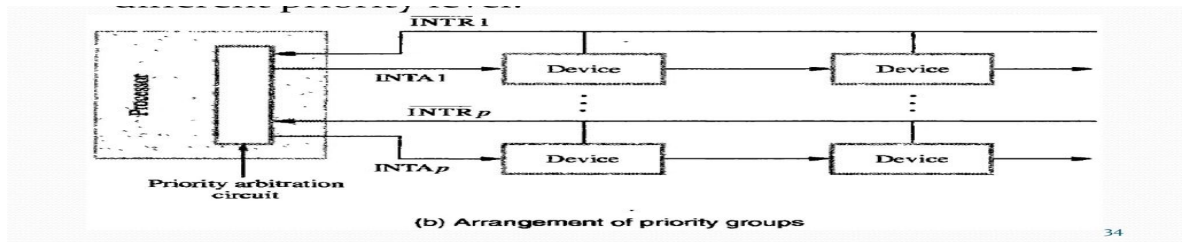
- The interrupt request line INTR is common to all devices. The interrupt acknowledge line INTA is connected in a daisy chain fashion such that INTA signal propagates serially through the devices.
- When several devices raise an interrupt request, the INTR is activated & the processor responds by setting INTA line to 1. this signal is received by device.
- Device1 passes the signal on to device2 only if it does not require any service.
- If device1 has a pending request for interrupt blocks that INTA signal & proceeds to put its identification code on the data lines.
- Therefore, the device that is electrically closest to the processor has the highest priority.



(a) Daisy chain

Arrangement of Priority Groups:

- Here the devices are organized in groups & each group is connected at a different priority level.
- Within a group, devices are connected in a daisy chain.



DIRECT MEMORY ACCESS

- It is a technique used for high speed I/O device.
- Here, the device interface transfer data directly to or from the memory without continuous involvement by the processor
- A special control unit may be provided to allow the transfer of large block of data at high speed directly between the external device and main memory, without continuous intervention by the processor. This approach is called **DMA**.
- DMA transfers are performed by a control circuit called the **DMA Controller**.

To initiate the transfer of a block of words , the processor sends,

Starting address

Number of words in the block

Direction of transfer.

- When a block of data is transferred , the DMA controller increments the memory address for successive words and keeps track of the number of words and it also informs the processor by raising an interrupt signal.
- While DMA control is taking place, the program requested the transfer cannot continue and the processor can be used to execute another program.
- After the DMA transfer is completed, the processor returns to the program that requested the transfer.

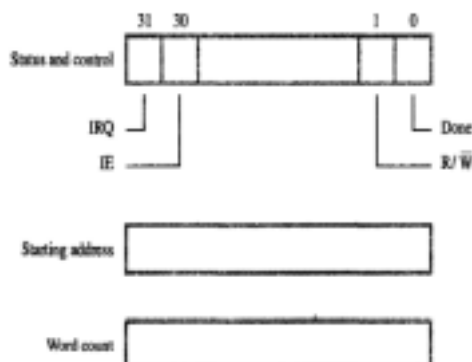


Figure 4.18 Registers in a DMA interface.

R/W - Determines the direction of transfer.

When

R/W =1, DMA controller read data from memory to I/O device.

R/W =0, DMA controller perform write operation.

Done Flag=1, the controller has completed transferring a block of data and is ready to receive another command.

IE=1, it causes the controller to raise an interrupt (interrupt Enabled) after it has completed transferring the block of data.

IRQ=1, it indicates that the controller has requested an interrupt.

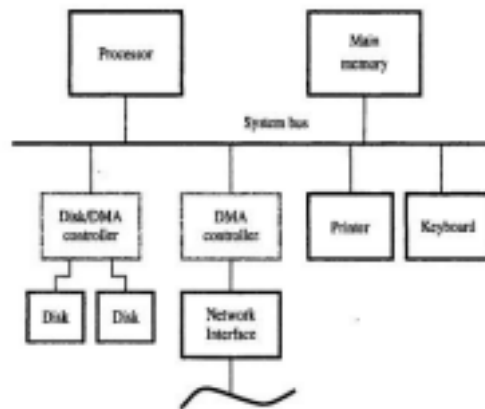
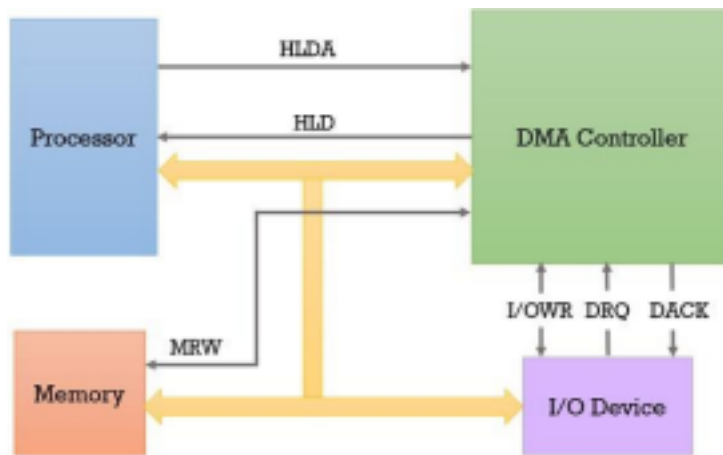


Figure 4.19 Use of DMA controllers in a computer system.

DMA controller transfers the block of data to and from memory in three modes burst mode, cycle steal mode and transparent mode.

1. **Burst Mode:** Here, once the DMA controller gains the charge of the system bus, then it releases the system bus only after completion of data transfer. Till then the CPU has to wait for the system buses.
2. **Cycle Stealing Mode:** In this mode, the DMA controller forces the CPU to stop its operation and relinquish the control over the bus for a short term to DMA controller. After the transfer of every byte, the DMA controller releases the bus and then again requests for the system bus. In this way, the DMA controller steals the clock cycle for transferring every byte.
3. **Transparent Mode:** Here, the DMA controller takes the charge of system bus only if the processor does not require the system bus.

Here, we will discuss the working of the DMA controller. Below we have the diagram of DMA controller that explains its working:



DMA Controller Data Transfer

1. Whenever an I/O device wants to transfer the data to or from memory, it sends the DMA request (DRQ) to the DMA controller. DMA controller accepts this DRQ and asks the CPU to hold for a few clock cycles by sending it the Hold request (HLD).
2. CPU receives the Hold request (HLD) from DMA controller and relinquishes the bus and sends the Hold acknowledgement (HLDA) to DMA controller.
3. After receiving the Hold acknowledgement (HLDA), DMA controller acknowledges I/O device (DACK) that the data transfer can be performed and DMA controller takes the charge of the system bus and transfers the data to or from memory.
4. When the data transfer is accomplished, the DMA raise an interrupt to let know the processor that the task of data transfer is finished and the processor can take control over the bus again and start processing where it has left.

STANDARD I/O INTERFACES

A personal computer, include a printed circuit board called motherboard. The board houses the processor chip, the main memory, and some I/O interfaces. The processor bus is the bus defined by signals on the processor chip itself. It is not possible to define a uniform standard for the processor bus. Some organisation such as IEEE (**Institute of Electrical and Electronics Engineers**), ANSI (**American National Standards Institute**), ISO (**International Organization for Standardization**).

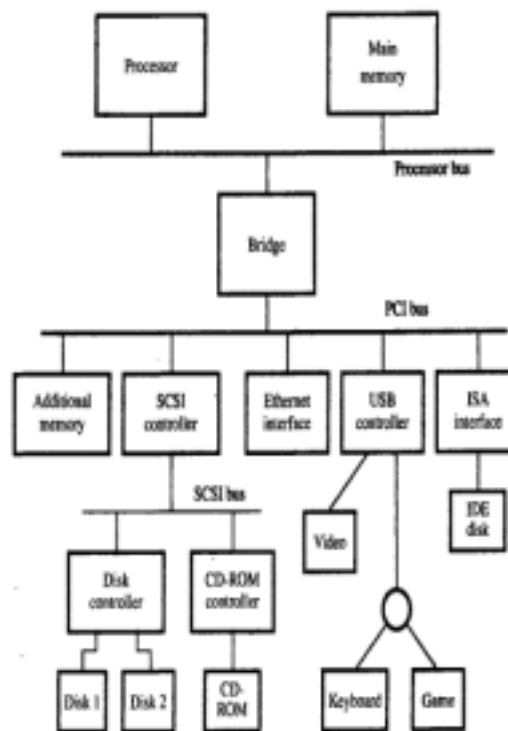


Figure 4.38 An example of a computer system using different interface standards.

Consider a computer system using different interface standards. Let us look in to Processor bus and Peripheral Component Interconnect (PCI) bus. These two buses are interconnected by a circuit called **bridge**. It is a bridge between processor bus and PCI bus. An example of a computer system using different interface standards is shown in figure 4.38. The three major standard I/O interfaces discussed here are:

- PCI (Peripheral Component Interconnect)
- SCSI (Small Computer System Interface)
- USB (Universal Serial Bus)

Peripheral Component Interconnect (PCI)

- The PCI was developed as a low-cost bus that is truly processor independent. It has growing demand for bus bandwidth to support high-speed disks and graphic and video devices, as well as the specialized needs of multiprocessor systems.
- An important feature of PCI is a plug-and-play capability for connecting I/O devices. To connect a new device, the user simply connects the device interface board to the bus. The software takes care of the rest. The PCI bus is a good example of a system bus.

Data Transfer: Most memory transfers involve a burst of data rather than just one word. The reason is that modern processors include a cache memory. The PCI is designed primarily to support this mode of operation. TtT

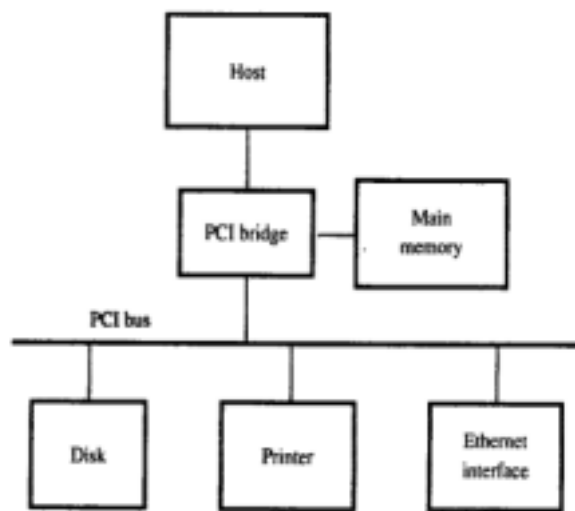


Figure 4.39 Use of a PCI bus in a computer system.

The bus supports three independent address spaces: memory, I/O and configuration

Device Configuration

- When an I/O device is connected to a computer, several actions are needed to configure both the device and the software that communicates with it.
- The PCI simplifies this process by incorporating in each I/O device interface a small configuration ROM memory that stores information about that device. The configuration ROMs of all devices is accessible in the configuration address space.

SCSI Bus

The acronym SCSI stands for Small Computer System Interface.. It's a fast bus that can connect lots of devices to a computer at the same time, including harddrives, scanners, CD-ROM, printers and tape drives. Other technologies, like serial-ATA (SATA), have largely replaced it in new systems, but SCSI is still in use.

A controller connected to a SCSI bus is an initiator or a target. The processor sends a command to the SCSI controller, which causes the following sequence of events to take place:

- The SCSI controller contends for control of the bus (initiator).
- When the initiator wins the arbitration process, it selects the target controller and hands over control of the bus to it.
- The target starts an output operation. The initiator sends a command specifying the required read operation.

- The target sends a message to the initiator indicating that it will temporarily suspend the connection between them. Then it releases the bus.
- The SCSI bus is connected to the processor bus through a SCSI controller. This controller uses DMA to transfer data packets from the main memory to the device, or vice versa.
- A packet may contain a block of data, commands from the processor to the device, or status information about the device.
- Communication with a disk drive differs substantially from communication with the main memory.
- A controller connected to a SCSI bus is one of two types. An initiator or a target. An initiator has the ability to select a particular target and to send commands specifying the operations to be performed.
- Clearly, the controller on the processor side, such as the SCSI controller, must be able to operate as an initiator. The disk controller operates as a target. It carries out the commands it receives from the initiator.
- The initiator establishes a logical connection with the intended target. Once this connection has been established, it can be suspended and restored as needed to transfer commands and bursts of data.
- While a particular connection is suspended, other device can use the bus to transfer information. This ability to overlap data transfer requests is one of the key features of the SCSI bus that leads to its high performance.
- Data transfers on the SCSI bus are always controlled by the target controller. To send a command to a target, an initiator requests control of the bus and, after winning arbitration, selects the controller it wants to communicate with and hands control of the bus over to it. Then the controller starts a data transfer operation to receive a command.

Universal Serial Bus (USB)

The USB has been designed to meet several key objectives such as:

- Provide a simple, low-cost and easy to use interconnection system that overcomes the difficulties due to the limited number of I/O ports available on a computer
- Accommodate a wide range of data transfer characteristics for I/O devices, including telephone and Internet connections
- Enhance user convenience through a “plug-and-play” mode of operation

Port Limitation: To add new ports, a user must open the computer box to gain access to the internal expansion bus and install a new interface card. The user may also need to know how to configure the device and the software. And also it is to make it possible to add many devices to a computer system at any time, without opening the computer box.

USB architecture: To accommodate a large number of devices that can be added or removed at any time, the USB has the tree structure. Each node has a device called a hub. Root hub, functions, split bus operations – high speed (HS) and Full/Low speed (F/LS).

COMPUTER PERIPHERALS

Computer peripherals are external devices that enhance machine's functionality through better performance, design, sound quality and more. There are input and output peripherals.

Input Devices

transfers information from outside to the computer. An input device converts incoming data and instructions into electrical signals in binary code and sends them into the CPU. Commonly used input devices are: Mouse, Keyboard, Scanner, Barcode reader, Digitizer, Biometric devices, Touch Screen, Gamepads and Joysticks, Multimedia input devices.

Keyboard

A **computer keyboard** is an input device used to enter characters and functions into the computer system by pressing buttons, or **keys**. It is the primary device used to enter text. A keyboard typically contains keys for individual letters, numbers and special characters, as well as keys for specific functions. A keyboard is connected to a computer system using a cable or a wireless connection.

Keyboard is available in two ways. One type consisting of an array of mechanical switches mounted on a printed circuit board. These switches are organised in rows and columns and connect to microcontroller on the board. When a switch is pressed the micro controller identify correct row and column, then determine the key.

The second type uses a flat structure consist of three layers. Top layer is a plasticized material with key position marked on top surface. The middle layer made of rubber with holes at key position. The bottom layer is metallic with raised bumps at key position. When pressure applied on top layer come in contact corresponding to the bumps in bottom layer thus make a electrical circuit in the same way as mechanical switch. The current flow send to microcontroller. This is the low cost keyboard.

Keyboards are of two sizes 84 keys or 101/102 keys, but now keyboards with 104 keys or 108 keys are also available.

Keys & Description

• Typing Keys

These keys include the letter keys (A-Z) and digit keys (0-9) which generally give the same layout as that of typewriters.

• Numeric Keypad

It is used to enter the numeric data or cursor movement. Generally, it consists of a set of 17 keys that are laid out in the same configuration used by most adding machines and calculators.

• Function Keys

The twelve function keys are present on the keyboard which are arranged in a row at the top of the keyboard. Each function key has a unique meaning and is used for some specific purpose.

• Control keys

These keys provide cursor and screen control. It includes four directional arrow keys. Control keys also include Home, End, Insert, Delete, Page Up, Page Down, Control(Ctrl), Alternate(Alt), Escape(Esc).

- **Special Purpose Keys**

Keyboard also contains some special purpose keys such as Enter, Shift, Caps Lock, Num Lock, Space bar, Tab, and Print Screen.

Mouse

- Mouse is the most popular pointing device. It is a very famous cursor-control device having a small palm size box with a round ball at its base, which senses the movement of the mouse and sends corresponding signals to the CPU when the mouse buttons are pressed.
- Generally, it has two buttons called the left and the right button and a wheel is present between the buttons. A mouse can be used to control the position of the cursor on the screen, but it cannot be used to enter text into the computer. Many variations of the mouse exist, including trackballs, tablets, touchpads, and pointing sticks.
- Mechanical and optical(using light emitting diode) mice are available.

Scanner

It scans images, printed text, handwriting or an object and converts it to a digital image. In scanners the document placed on a glass window for scanning. They use light to reflect off a surface and measure the relative reflections of the different dots that make up the scanner is able to detect. Convert the light they receive into electrical impulses and is forwarded to the software producing the scan. The most common type of scanner is flatbed scanner.

Another type of scanner, called document scanner, uses Optical Character Recognition (OCR) software that can recognize text.



Output Devices

An output device is used to send data out of the system. The user gets the result after processing of data by the computer through output devices. Examples of some output devices are: monitor, Printer etc. The main output devices are Display Devices, Printers, Speaker.

video display

video displays are **electronic device that represents information in visual form.**

Monitors, commonly called as **Visual Display Unit (VDU)**, are the main output device of a computer. It forms images from tiny dots, called pixels that are arranged in a rectangular form. The sharpness of the image depends upon the number of pixels.

There are two kinds of monitors.

- Cathode-Ray Tube (CRT)
- Flat-Panel Display

The CRT display is made up of small picture elements called pixels. The smaller the pixels, the better the image clarity or resolution.

Flat-Panel Display Monitor

The flat-panel display refers to a class of video devices that have reduced volume, weight and power requirement in comparison to the CRT. You can hang them on walls or wear them on your wrists. Current uses of flat-panel displays include calculators, video games, monitors, laptop computer, and graphics display.

The flat-panel display is divided into two categories –

- **Emissive Displays** – Emissive displays are devices that convert electrical energy into light. For example, plasma panel and LED (Light-Emitting Diodes).
- **Non-Emissive Displays** – Non-emissive displays use optical effects to convert sunlight or light from some other source into graphics patterns. For example, LCD (Liquid-Crystal Device).

Printers

- Printer is an external hardware device responsible for taking computer data and generating a hard copy of that data.
- Printers are one of the most used peripherals on computers and are commonly used to print text, images, and photos.
- Printer is an output device, which is used to print information on paper. There are two types of printers –

1. Impact Printers
2. Non-Impact Printers

Impact Printers

Impact printers print the characters by striking them on the ribbon, which is then pressed on the paper. Characteristics of Impact Printers are the following –

- Very low consumable costs
- Very noisy
- Useful for bulk printing due to low cost
- There is physical contact with the paper to produce an image
- Impact printers are of two types –

1. Character printers
2. Line printers

Non-Impact Printers

Non-impact printers print the characters without using the ribbon. These printers print a complete page at a time, thus they are also called as Page Printers.

- These printers are of two types –
- Laser Printers
- Inkjet Printers

Characteristics of Non-impact Printers

- Faster than impact printers

- They are not noisy
- High quality
- Supports many fonts and different character size

Graphics Accelerator

Graphics Accelerator(s) Graphics accelerators are **hardware that is optimized for doing the computations for three-dimensional computer graphics**. In desktop systems, the graphics accelerator is usually in the form of a graphics card, but sometimes the graphics accelerator is part of the motherboard.