2.0 INTRODUCTION

Till now we have discussed about various components and memory system for a computer. We have also introduced one interconnection structure called the Bus. In this unit we will discuss more about the input/output organisation and newer technologies in computer hardware. We will discuss briefly about input/output devices, two popular device interfaces and the concepts of parallel organisation. We will also discuss about the concepts of pipelining, vector processing and the concepts relating to Reduced Instruction Set Computers (RISC) architecture.

2.1 OBJECTIVES

At the end of this unit you should be able to:

- identify some of the Input/output devices
- identify the serial and parallel interfaces
- define the Flynn's classification of computers
- list the characteristics of multiprocessor systems
- define pipelining and vector processing
- define the concepts relating to RISC

2.2 INPUT/OUTPUT DEVICES (PERIPHERALS)

The computer will be of no use if it is not communicating with the external world. Thus, a computer must have a system to receive information from outside world and must be able to communicate results to external world. Thus, a computer consists of an I/O system. This system includes two basic components, one is the I/O devices and another called I/O module, which not only connects an I/O device with the system bus, but plays a very crucial role in between. A device, which is connected to an I/O module of computer, is called a peripheral. Let us discuss few Input/Output devices in this section and we will define the I/O module in the subsequent sections.

2.2.1 Input Devices

Input technologies are rapidly developing. These, as the name suggests, are used for transferring user command or choice or data to the computer.

Keyboard

The keyboard is one of the most common input devices for computers. The layout of the keyboard is like that of the traditional QWERTY typewriter, although there are some extra commands and function keys provided for. Substantial development has taken place in the ergonomics of keyboard design to ensure that operator strain is minimal.

Pointing Devices

The keyboard provides facility to input data and commands to computer in text form. We find that, while working with a display based packages, we are mostly pointing to some area in the display to select an option and move across on the screen to select subsequent options. For such cases pointing devices are very useful. There are several pointing devices, some of them are:

(a) **Mouse:** Mouse is a handy device which, can be moved on a smooth surface to simulate the movement of cursor that is desired on the display screen. Mouse could be optical; offering quite and reliable operation, or mechanical, which is cheaper but noisier. User can move the mouse, stop it at a point where the pointer is to be located and, with the help of buttons, make selection of choices.

(b) **Light Pen:** This is a pen shaped device allowing natural movement on the screen. The pen contains the light receptor and is activated by pressing the pen against the display screen. Receptor is the scanning beam which, helps in locating the pen's position. Suitable system software is provided to initiate necessary action when we locate an area on the display surface with the help of the light pen.

There are a few other pointing devices known as track balls and joysticks, which are used more on entertainment usage, like games. We will not discuss them in this section.

Voice/Speech Input

One of the most exciting areas of research is in recognising human voices/speech so that this could from input to computer directly. This approach will eliminate the need for keying in data and will facilitate casual users to use the computer very easily. There are several problem areas for research since speech recognition system should be able to identify who is speaking and what the message is. Voice recognition techniques along with several other techniques to convert the voice signals to appropriate words and derive the correct meaning of words are required for a commercially viable comprehensive speech recognition system. We have found limited success in this area and today devices are available commercially to recognise and interpret human voices within limited scope of operation.

Scanners

Scanners facilitate capturing of the information and storing them in graphic format for displaying back on the graphical screen. Scanner consists of two components, the first one to illuminate the page so that the optical image can be captured and the other to convert the optical image into digital format for storage by computer.

The graphic image scanned can now be seen and processed directly by the computer. Substantial research work is going on to establish methods by which the scanned image can be automatically converted into equivalent text for further processing.

Source Data Automation

Most recent trends for data input are towards source data automation. The equipments used for source data automation capture data as a by-product of a, business activity thereby completely eliminating manual input of data. Some examples are:

1. Magnetic Ink Character Recognition (MICR) devices are generally used by the banking industry to read the account numbers on cheques directly and does the necessary processing.

2. Optical Mark Recognition (OMR) devices can sense marks on computer readable papers. This kind of device is used by academic and testing institutions to grade aptitude tests where candidates mark the correct alternatives on a special sheet of paper. The optical mark recognition devices then directly read these answer sheets and the information sent to a computer for processing.

3. Optical Bar Code Reader (OBR) scans a set of vertical bars of different widths for specific data and are used to read tags and merchandise in stores, medical records, library books, etc. These are available as hand held devices.

2.2.2 Output Devices

The output normally can be produced in two ways; either on a display unit/device or on a paper. Other kinds of output such as speech output, mechanical output is also being used in certain applications. In this section, we will discuss only the display and printing output devices.

Display Devices

One of the most important peripherals in computer is the graphic display device. Conventional computer display terminals knows as alphanumeric terminals, display characters (images) from a multi-dot array (normally 5 x 7 or 7 x 9). These are used to read text information displayed on the screen. However, there are increasing demand for display of graphics, diagrams and pictures to make the visual presentation of information more effective for user interaction and decision making.

Graphic display is made up of a series of dots called 'pixels' (picture elements) whose pattern produces the image. Each dot on the screen is defined as a separate unit, which can be directly addressed. Since each dot can, be controlled individually there is much greater flexibility in drawing pictures. There are three categories of display screen technology:

1. Cathode Ray Tube (CRT)

2. Liquid Crystal Display (LCD)

3. Projection Displays

CRT Displays

The main components of a CRT terminals are the electron gun, the electron beam controlled by an electromagnetic field and a phosphor coated display screen.

The electromagnetic field in order to create an image (Refer figure 5). There are two types of CRT displays

(a) Vector CRT displays in which the electron beam is directed only to the places where the image is to be created.

(b) Raster displays in which the image is projected on to the screen by directing the electron beam across each row of picture elements from the top to the bottom of the screen. This type of display provides a high dynamic capability since image is continuously refreshed and it allows for continual user input and output. It offers full colour display at a relatively low cost and is becoming very popular.

Liquid Crystal Displays (LCD)

First introduced in watches and clocks in 1970s, LCD is now applied to display terminals. The major advantage of LCD is the low energy consumption. The CRT (Cathode Ray Tube) is replaced by liquid crystal to produce the image. This also have colour capability but the image quality is relatively poor. These are commonly used in portable devices because of compactness and low energy requirements.

Projection Displays

The personal size screen of the previous displays is replaced by a large screen upon, which images are normally used for large group presentation. These systems can be connected to computer and whatever appears on the computer terminal gets enlarged and projected on a large screen. Another popular method is to connect computer to a LCD flat screen. and to project the LCD image using Overhead Projector. These are popularly used for seminars, classrooms, marketing presentations, etc.

Printers

Printers are used for producing output on paper. There are a large variety of printing devices, which can be classified according to the print quality and the printing speeds. Current estimates indicate that about 1500 types of printers are commercially available conforming to about 15 different printing technologies. The following categories of printers are identified.

(a) Printing Technology - impact printers vs. non-impact printers

Impact printers use variations of standards typewriter printing mechanism where a hammer strikes Paper impact printers use variations of standards typewriter printing mechanism through inked ribbon. Non-impact printer uses chemical, heat or electrical signals to etch or induce symbols on paper. Many of these require special coated or treated paper.

(b) Character forms-fully formed characters vs. dot matrix

Fully formed characters are constructed from solid lines and curves like the characters of typewriter whereas a dot matrix character is made up from a carefully arranged sequence of dots packed very close to each other. Obviously print quality of a dot matrix printer will be poorer compared to that from fully formed characters.

(c) Printing sequence - serial vs. line vs. page

This indicates the amount of information of printer can output within a single cycle of operation. Serial printing is done character by character whereas line printing forms an entire line and prints a line at a time whereas a page printer outputs a whole page of character and images simultaneously during one cycle. Clearly the speed of output will depend upon the printing sequence incorporated in the printing device. We will now look at three of the most popular printers:

Dot Matrix Printers

This is one of the most popular printers used for personal computing systems. These printers are relatively cheaper compared to other technologies. This uses impact technology and a print head containing banks of wires moving at high 'speeds against inked ribbon and paper. Characters ire produced in matrix format. The speeds range from 40 characters per second (CPS) to about 1,000 cps. A disadvantage of this technology is that the print quality is low.

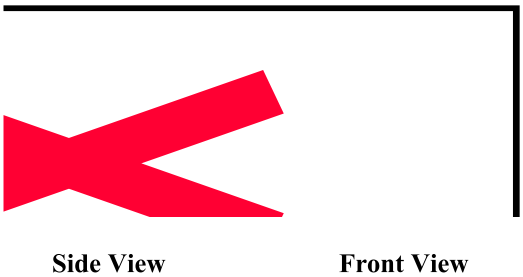


Figure 1: Dot Matrix Printer

Ink Jet Printers

These print by spraying a controlled stream of tiny ink droplets- accurately on the pa er dot matrix or solid characters. These are non-impact and hence relatively silent and high quality p ical speeds range from 50 cps to above 300 cps and this technology has been used well for production g and elaborate graphics.

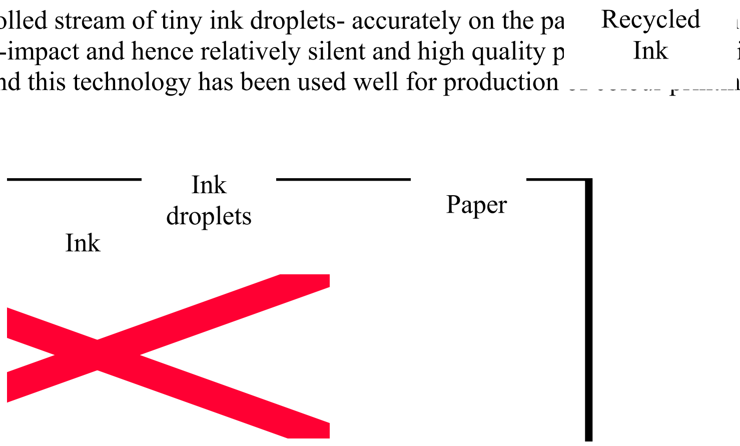


Figure 2: Inkjet Printing

Laser Printers

This is a high quality, high speed, and high volume technology, which works in non-impact fashion on plain paper or pre-printed forms. Printing is achieved by deflecting laser beam on to the photosensitive surface of a drum and the latent image attracts the toner to the image areas. The toner is then electrostatically transferred to the paper and fixed into a permanent image. Speeds can range from 10 pages a minute to about 200 pages per minute. This technology is relatively expensive but is becoming very popular because of the quality, speed and noiseless operations.

These three printers are compared in figure 3.

Plotters

To produce graphical output on paper the plotters are used.

After discussing so much about the I/O devices, let us come back to one of the basic question: How does I/O devices are connected to computers? We will try to answer the question in the next section.

Style of Printing	Print quality
Dot matrix	Low quality but steadily improving. Very useful for high-volume work (low quality) because of great speeds and low printing cost. Lack of descenders on letters (many models rectify inability to reproduce well on a photocopier are major this) and the problems. Dot Matrix Printers have the ability to produce multiple copies of documents in a single print
Ink-jet	High-quality print character sets: choice of type set is controlled by program; print quality quite good than dot matrix; cheap; useful for low volume applications; high printing cost;
Laser	Very high quality, prints at speeds unapproachable by other technologies. Price is affordable. Excellent for high volume high quality works.

Figure 9: Some Printer Technologies

Check Your Progress 1

1. State true or false.

(a) In case where graphical user interfaces are common mouse should not be used.

True ☐ False ☐

(b) Keyboard is one of the most common input device.

True ☐ False ☐

(c) Scanners are devices used for outputting pictures.

True ☐ False ☐

(d) Projection displays can be used for classroom teaching.

True ☐ False ☐

(e) Keyboard, VDU and printers are essential for computers.

True ☐ False ☐

2. Differentiate between the CRT display and LCD display.

………

………

3. Compare and contrast the Laser and Dot Matrix Printers.

………

………

4. Define the term "Source Data Automation". Give two examples.

………

………

2.3 INPUT/OUTPUT MODULE INTERFACE

The input/output module interfaces (I/O module) is normally connected to the computer system on one end and one or more Input/Output devices on the other. Normally, an I/O module can control one or more peripheral devices. An I/O module is needed because of:

(a) Diversity and variety of I/O devices makes it difficult to incorporate all the peripheral device logic (i.e. its control commands, data format etc.) into CPU. This in turn will also reduce flexibility of using any new development.

(b) The I/O devices are normally slower than that of memory and CPU, therefore, it is suggested not to use them on high-speed bus directly for communication purpose.

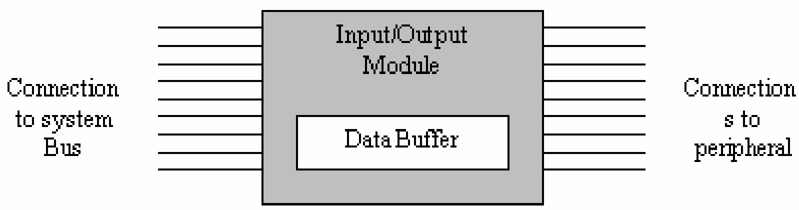
(c) The data format and word length used by the peripheral may be quite different than that of a CPU.

An I/O module is a mediator between the processor and an I/O device/devices. It controls the data exchange between the external devices and main memory; or external devices and CPU registers. Therefore, an I/O module provide an interface internal to the computer which connect it to CPU and main memory and an interface external to the computer connecting it to external device or peripheral. The I/O module should not only communicate the information from CPU to I/O device, but it should also co-ordinate these two. In addition since there are speed differences between CPU and I/O devices, the I/O module should have facilities like buffer (storage area) and error detection mechanism.

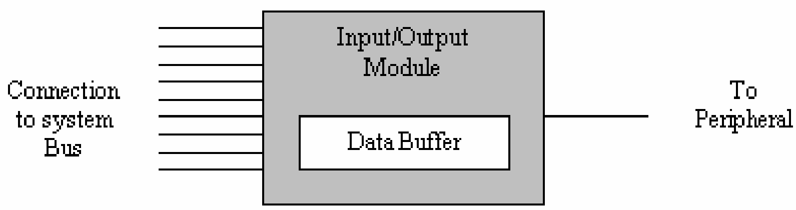
If an I/O. module takes more processing burden then it is termed as I/O channel or processor. The primitive I/O modules, which require detailed control by processor, are called I/O controller or device controller. These I/O controllers are normally used in micro-computers, while I/O processors are mainly used for Mainframe, because the I/O work for microcomputer is normally limited to single user's job. Therefore we do not expect a very huge amount of I/O to justify the investment in I/O processor, which are expensive.

2.4 EXTERNAL INTERFACES

Our discussion on I/O system will not be complete if we do not discuss about external interfaces. External interface is the interface between the I/O module and the-peripheral devices. This interface can be characterised into two main categories: (a) parallel interface (b) serial interface. Figure 4 gives a block diagram from these two interfaces.



(a) Parallel interface



(b) Serial interface

Figure 4: Parallel and Serial Interface

In parallel interface multiple bits can be transferred simultaneously. The parallel interfaces are normally used for high-speed peripherals such as tapes and disks. The dialogue that takes place across the interface includes the exchange of control information and data. A common parallel interface is centronics.

In serial interface only one line is used to transmit data, therefore, only one bit is transferred at a time. Serial interface is used for serial printers and terminals. The most common serial interface is RS-232-C.

Check Your Progress 2

1. State true or false.

(a) Devices are normally connected directly to system bus.

True ☐ False ☐

(b) Input/output Module is needed only for slower I/O devices.

True ☐ False ☐

(c) Data buffering is helpful for smoothing out the speed difference between CPU and input/ output devices.

True ☐ False ☐

(d) Parallel interfaces are commonly used for connecting printers to a computer.

True ☐ False ☐

Q.2 What is a device controller?

....

....

2.5 WHAT IS PARALLEL PROCESSING?

In the previous section, we have defined the Input/Output interfaces of a computer. Let us now, explore about the newer technologies and trends in computers.

The computers can be made faster by increasing their clock speed or by using improved VLSI technology. But there is a practical limitation to both of them. The only way, that seems possible, theoretically, that can increase the speed in an unlimited fashion is to have multiple modules, that is break the original problems into independent modules and execute them simultaneously. More the number of processors doing the job simultaneously more the speed up will be. Parallel processing is a term used to denote a large class of techniques that are used to provide simultaneous data processing tasks for the purpose of increasing the computational speed of a computer.

One of the most popular classifications of parallel computer is by M.J. Flynn, called Flynns Classification, and is based on the multiplicity of instruction streams and data streams in a computer system. The sequence of instruction read from the memory constitute the instruction steam, and the data they operate of in the processor constitute the data stream. Flynn's classification divides the computers into four categories:

- single instruction single data (SISD)
- single instruction multiple data (SIMD)
- multiple instruction single data (MISD)
- multiple instruction multiple data (MIMD)

SISD organisation is available in most serial computers today. Instructions are executed sequentially, through may be overlapped in their execution stage (pipelining). All the functional units are under the supervision of one control unit. The Von Neumann architecture falls under this category.

SIMD organisation consists of multiple processing elements supervised by the same control unit. All processing units receive the same instruction broadcast but works or different data streams. One of the very common example is the execution of a 'for' loop, in which the same set of instructions are executed for, may be a different set of data.

MISD organisation consists of N processor units, each working on a different set of instruction working on the same set of data. The output of one unit becomes the input to the other unit. There is no commercial computer of this kind.

MIMD organisation implies interactions between N processors because all memory streams are derived from the same data stream shared by all processors. If the interaction between the processors is high, it is called a tightly coupled system, or else it is called a loosely coupled system. Most multiprocessors fit into this category.

Parallel processors can be categorised into several categories. These include:

* Array Processors

* Distributed Architecture

* Multiprocessors

* Data flow Architectures

Array processors are parallel architectures, which deal with repetitive operations. These are an example of SIMD architecture, with applications varying from mathematical array operations and in structures in which data objects are known well in advance.

Distributed Architecture are composed of relatively autonomous subsystems which are capable of handling complete system and execution functions, and which cooperate together to run a large application.

Multiprocessors are architectures composed of multiple computing units, which operate in a synchronous mode. Generally, both shared and local memory can be available for these processors. The communication between the processors takes place either through the shared memory area or through the interprocessor messages.

Dataflow Architectures are functionally distributed architectures, in which the operations are triggered with the arrival of data at these processors. They may be viewed as very general MIMD parallel architectures.

Most of these parallel processor computers follow an important concept called pipelining. In the subsequent sections, we will discuss some of the newer technological concepts.

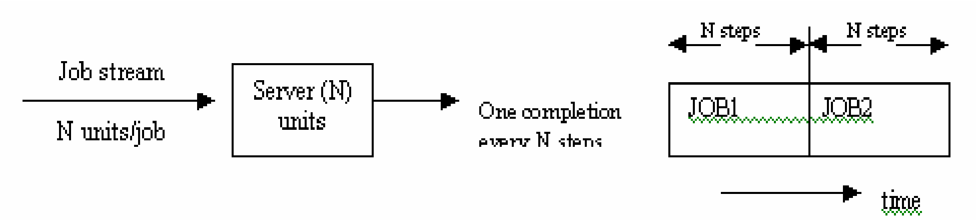
2.6 PIPELINING

The basic idea behind pipeline design is quite natural; it is not specific to computer technology. In fact the name pipeline stems from the analogy with petroleum pipelines in which a sequence of hydrocarbon products is pumped through a pipeline. The last product might well be entering the pipeline before the first product has been removed from the terminus. The key contribution of pipelining is that it provides a way to start a new task before an old one has been completed. The concept can be better understood by taking an example of an industrial plant. To achieve pipelining, one must subdivide the input task into a sequence of subtasks, each of which can be executed by a specialised hardware stage that operates concurrently with other stages in the hardware. Successive tasks are streamed into the pipe and get executed in an overlapped fashion at the subtask level. Hence the completion rate is not a function of total processing time, but rather of how soon a new process can be introduced.

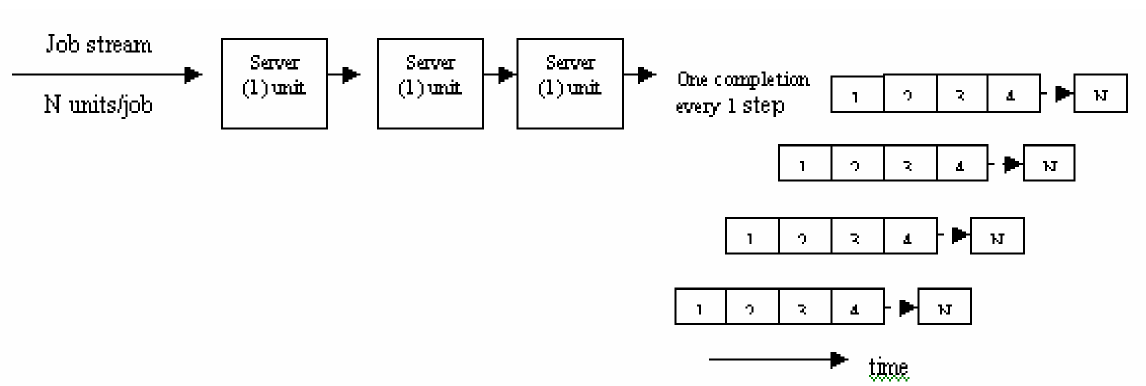
The subdivision of labour in assembly lines has contributed to the success of mass production in modern industry. By the same token, pipeline processing has led to the tremendous improvement of system throughput in modern digital computers.

Now consider for execution of an instruction a computers performs N processing steps. A server unit as represented in figure 5(a) can perform any one of the N steps in one unit time. If we process M instructions by this server, then the rate of completion is one instruction of every N steps, and the time behaviour of the job stream is as described in Figure 5(a). Compare Figure 5(a) with Figure 5(b) which shows N servers concatenated in a sequence each performing only a single step, the job flows through the collection of servers by visiting server 1, then server 2, and so on, and finally emerging from server N after N steps. The time behaviour of this system is described by figure 5(b).

Figure 5(b) is an ideal model of a constant speed assembly line, such as an automobile assembly plant.



(a) Sequential execution with N-unit server



(b) Pipelined execution with 1-unit servers

Figure 5: Two ways of executing N-unit jobs in a stream

Thus, any arithmetic task is divided into steps. Thus, a sort of concurrent operation can be obtained in the pipeline. This will help in increasing the throughput of the system.

Though, superficially it seems that it should be possible to execute any application in various pipeline stages, it is not actually so. Only those applications, which can be broken into independent subtasks, can take the advantage of pipelining. It is most efficient for those applications that need to repeat the same task many times with different sets of data.

2.7 VECTOR PROCESSING

Some of the applications such as:

* Weather forecasting
* Petroleum explorations
* Seismic data analysis
* Flight simulations
* Artificial intelligence and expert systems
* image processing and its applications

are so complex that without the use of sophisticated computers it would not be possible to complete these in reasonable amount of time.

These computational problems are beyond the capabilities of a conventional computer, as these problems require a vast number of computations that may take a conventional computer days or even weeks to be completed. The vector computers have emerged as the most important high performance architecture for numerical problems. It has two key qualities of efficiency and wide applicability.

These scientific problems may involve manipulations of large arrays, where the same computational are carried out with different sets of data. These numbers can be easily formulated as vectors and matrices of floating point number (real numbers). A vector can be defined as an ordered set of a one-dimensional array of data items. Each element in a vector is a scalar quantity, which may be floating point number, an integer, a logical value, or a character (byte).

Supercomputers

A commercial computer with vector instructions and pipelined floating-point arithmetic operations is referred to as a supercomputer. Supercomputers are very powerful, high-performance machines used mostly for scientific computations. To speed-up the operation, the components are packed tightly together to minimise the distance that the electronic signals have to travel.

A supercomputer is a computer best known for its high computational speed, fast and larger memory systems, and the extensive used of parallel processing. It is equipped with multiple functional units and each unit has its own pipeline configuration. Although the supercomputer is capable of general purpose applications found in all other computers, it is specifically optimised for the type of numerical calculations involving vectors and matrices of floating point numbers.

A measure used to evaluate computers in their ability to perform a given number of floating point operations per second is referred to as flops. The term megaflop is used to denote million flops and gigaflops to denote billion flops. A typical supercomputer has a basic cycle time of 4 to 20 nsecs. If the processor can calculate a floating-point operation through a pipeline each cycle time, it will have the ability to perform 50 to 250 mflops. This would be sustained from the time the first answer is produced and does not include the initial setup time of the pipeline.

2.8 INTRODUCTION TO RISC

Aim of the computer architect's -is to design computers, which are more cost effective than their predecessors. Cost effectiveness includes:

1. Cost of hardware to manufacture the machine

2. Cost of programming

3 Cost incurred in correcting both the initial hardware and subsequent programs on account of the new/old/compatible architecture.

2.8.1 Reasons For Increased Complexity

Let us see what are the reasons for increased complexity that leads to complex instruction set computers (CISC) and what exactly we mean by this.

1. Speed of memory Vs. the speed of CPU

The CPU, in general, was about 10 times as fast as main memory. Thus, more and more instructions were put into hardware. Since then many higher level instructions have been added to machines in an attempt to improve performance. This trend began towards the imbalance in speed.

The advances in semiconductor memory have made several changes to the assumptions about the relative difference in speed between CPU and the main memory. Semiconductor memories are both fast and relatively inexpensive. Cache memories have further reduced the difference between the CPU and the memory Speed.

2. Code density

With early computers memory was very expensive. It was therefore cost effective to have very compact programs. It was assumed that more instructions would result in small programs. While code compaction is important, however, the cost of 10 percent more memory is often far cheaper than the cost of squeezing code by 10 percent out of the CPU architecture innovations.

3. Support for high level language

With the increasing use of more and more high level languages, manufacturers have become eager to, provide more powerful instructions to support them. It is argued that richer instruction set would alleviate the software crisis and would simplify the compilers. However, RISC designers did not agree to it. They conducted research to find out what constructs are most frequently used in average programs.

Their research shows that integer constants appeared almost as frequently as arrays or structures; more than 80 percent of the scalars were local variables to subroutines; and more than 90 percent of the arrays or structure were global variables.

In addition it was found that procedure calls and returns are an important aspect of High-level language programs. These are also the most time consuming operations in complied high-level language programs. Thus, it will be profitable to consider ways of implementing these operations efficiently. Two aspects are significant in this respect: the number of parameters and variables that procedure dealt with; and the depth of nesting.

Further, it has been statistically found that 98% of dynamically called procedures pass lower than six arguments, and 92% of them use fewer than six scalar variables.

Thus, in the 1980s, a new philosophy evolved called RISC. These machines would have fewer instructions - (reduced set) and the instructions would be simple and would generally execute in one cycle, hence, the reduced instruction set computers (RISCs). Optimising compilers are used in these machines to compile the programs into simple machine instructions.

2.8.2 Principles of RISC

The following are the characteristics of RISC architecture:

1. Hardware Instructions should be kept simple unless there is a very good reason to do otherwise.

2. Microcode, a programming way of implementing control unit, is not magic. Moving software into microcode does not make it better, it just makes it hard to change.

3. Simple decoding and pipelined execution are more important than program size. At the peak rate a new instruction is started every cycle. Pipelined execution gives a peak perforinance, of one instruction every step, so the longest piece determines the performance rate of the pipelined machine. So ideally each piece should take the same amount of time.

4. The large number of registers should be utilised optimally.

Check Your Progress 3

State true or false:

1. Vector processors are most suited for high performance computations.

   True ☐ False ☐

2. Pipelined processing improves the system throughput.

   True ☐ False ☐

3. Applications, which can be broken into subtask, can be implemented as pipelined system.

   True ☐ False ☐

4. 80486 is a RISC processor.

   True ☐ False ☐

2.9 SUMMARY

In this unit, we have discussed about the Input/Output devices, input/Output module and interfaces. The discussion is only at the introductory level, however, you can refer to further readings for more details. The input/output technologies are developing all the time with the advancement of technology. Some newer technologies such as data gloves, virtual reality are already available and more and more sophisticated technologies will be available in near future, however all the new things can be categorised under one of the techniques we have discussed in the unit.

In addition in this block, we have introduced the concepts of pipelining; vector processing and Reduced Instruction Set Computers.

2.10 MODEL ANSWERS

Check Your Progress 1

1. (a) False
   (b) True
   (c) False
   (d) True

(e) True

2.

	CRT	LCD
(i)	CRT is bulky	LCD is compact
(ii)	High energy requirement	Low energy requirement
(iii)	Normally high resolution graphics	Low resolution

3.

	Laser printer	Dot Matrix Printer (DMP)
(i)	High Cost	Low Cost
(ii)	Form Characters using laser beam	Use Dot Matrix to print characters
(iii)	Non-impact printer	Impact printer
(iv)	High quality output	Output is of low quality
(v)	Noiseless operation	Noisy
(vi)		Produces multiple copies in a single print

4. In source data automation data is entered essentially where it is being produced through automatic devices. Examples are OMR & OBR.

Check Your Progress 2

1 (a) False (b) False (c) True (d) True

2. A device controller is an I/O module, which interacts with the I/O devices as per the instructions provided to it by the CPU.

Check Your Progress 3

1. True 2. True 3. True 4. False

2.11 FURTHER READINGS

1. Stallings, William, Computer Organisation and Architecture, Third edition, Maxwell Macmillan International Editions.
2. Mano, M. Morris, Computer System Architecture and Organisation, Second Edition, McGraw-Hill International Editions, 1988.