Module-2, Unit-1
Component of Central Processing Unit (CPU) and External Interface

Question 1: What are the main components of a Central Processing Unit (CPU) and their function? Explain roles that are performed by processor registers.

Solution 1:

The main components of a central processing unit are: processing elements, storage elements, control unit and internal CPU interconnection system.

The responsibilities of the processing elements are to process the input data and provide the result. One such processing element is ALU (Arithmetic Logic Unit). Through ALU, we can perform Arithmetic operations, like, Addition, Subtraction, Multiplication, etc. and Logic operations, like, NOT, AND, OR, EXOR, etc.

While working with the data in the processor (like adding two numbers), we have to keep the data inside the processor and to keep the data we use storage elements. These storage elements of the processor are known as Registers.

While processing the data in the processor, we need to transfer the data from one place to other. For example, while adding two numbers, first we have to transfer the data from the registers to the input of the ALU. After performing the ALU operation, the result generated by ALU has to be transferred from the output of the ALU to the register for storage. Therefore, we should have the connection between all the components and for that we need an interconnection system inside the processor.

The function of the control unit is to generate the appropriate signals at the right moment. For example, while adding two numbers in the ALU, after placing the data to the input of the ALU, the control unit needs to generate the appropriate signals for the ALU to indicate that the operation performed to the input data is an addition operation. In the next moment, the control unit needs to generate the signals such that the result of the ALU operation is transferred to the destination register.

The registers of a processor are the storage elements to store data. Processor registers are generally divided into two categories: General Purpose Registers (GPR) and Special purpose Registers. General purpose registers are used to keep the data needed for an operation and can be used to store the user data. As we know that the processor works on Von Neumann Stored Programmed principle, that is, we are storing the information in a storage device, which is known as main memory. During processing, data are brought to the processor and kept in a processor register. To store these data general purpose registers are used.

During program execution, we need to fetch the instructions and data from the main memory and we have to specify the memory location of the data and instruction. To provide the memory address, a special register is used which is known as MAR (Memory Address Register). Also a special register called MBR/MDR (Memory Buffer/Data Register) is used for data transfer. Also, execution of a program is carried out by executing the instruction by instruction in sequence. The program is also stored in main memory and instructions are fetched one after another. After fetching one instruction, we have to keep the address of the next instruction so that after completion of the current instruction, next instruction can be fetched from the memory. This information is kept in a special purpose register known as Program Counter (PC). After fetching the instruction, we keep the instruction in a special purpose register called Instruction Register (IR).

Question 2.

Explain the basic construction of processing elements (e.g., ALU) of a processor.

Solution 2:

Every processor has one or more processing elements to process the data. The most important processing element is ALU (Arithmetic and Logic Unit) and ALU is responsible to carry out any arithmetic and logic operations. In general, it is a hardware unit and implemented by digital integrated circuits. If we carry out any operation in hardware, then it performs better. But how many operations that will be pushed into the ALU is a design decision issue. If we include more operations in the ALU, then the circuitry is a complex one and to implement, it requires more silicon space. Also another issue is the data handling capabilities. In general, processor cannot take care of the complete range of any number systems; it depends on the size of input and output. For example, if the size of input is 8 bits, then it can handle numbers up to $255 = 2^8-1$ (for unsigned integer). For signed integer, the range is half. Due to these limitations, very few and primitives operations are included in the ALU or any other processing elements. If we have limited number of operations in the ALU, then the capability of the processor is also limited. To enhance the power of the processor, some of the instructions are implemented with the help of programs that are known as micro routines. Consider a simplified ALU which has addition and subtraction operations. But multiplication is not included in the ALU. But still we may have a multiplication instruction in the processor. For that, we incorporate a simple micro routine which performs the multiplication by repeated addition method. For that the addition operator of the ALU is used. So, the processor is equipped with the multiplication operation, but the process is a slower one. If the multiplication is implemented in hardware (using Booth's multiplication algorithm), then it gives a faster response.

Question 3: What are the issues related to internal CPU interconnections.

Solution 3:

The processor contains some registers and they are categorized into two groups - one is general purpose registers and the other is special purpose registers. Generally, number of general purpose registers is limited and typically in the range of eight or sixteen. To perform any operation, sometimes we have to move information from one register to another register. Also to perform ALU operations, data need to be brought to the input of the ALU. The results generated by the ALU finally need to be stored in the destination registers. So, if we look for the possibilities of data movement, it appears that we need connections between all components. If we provide direct connection from each component to every other component, then the interconnection system will be a complicated one. Therefore, we need a systematic approach for interconnection network.

In general, CPU buses are used for interconnection network. Consider a single internal CPU bus organization. In this case, all the components are connected through this internal CPU bus. For example, consider two general purpose register R1 and R2. We need to transfer the information from R1 to R2. In that case, first the data from the register R1 is placed into the internal CPU bus, and then the data is transferred from internal data bus to register R2. To enable register R1 and register R2, the control unit generates the appropriate signals. With single internal data bus, at any point of time only one data transfer is possible. So, for better performance, processors are designed with multiple internal buses also.

Question 4: What is a bus and how the other devices are connected to the processor through bus.

A bus is a communication pathway connecting two or more devices. Typically, a bus consists of multiple communication pathways, or lines. Each line is capable of transmitting signals representing binary 1 and binary 0. Generally, electrical signals are transmitted through these lines. A high voltage (typically around 5 volts) represents binary 1 and low voltage (typically near 0 volt) represents binary 0. In case of multiple lines, we can transmit signals in parallel. For example, an 8-bit unit of data can be transmitted over a bus consisting of 8 lines.

As we have seen that Central Processor Unit (CPU) consists of some storage elements (registers), processing elements (ALU) and control unit. CPU works on Von Neumann Stored Program Principle. That is, to work with the processor, we need to transfer information from storage unit to the processor. For this purpose, storage or memory unit needs to be connected with the processor. The memory unit is connected to the processor with a bus. We have to consider the following three issues:

- In the memory unit, we are storing the information and processor fetches this information from memory. The information stored in the memory is transferred to the processor through bus. This bus is known as data bus.

- In a memory unit, there are several memory locations. Processor needs to indicate the address of the memory location from which it is going to fetch the data. This memory address information is provided through another bus, known as address bus.

- Generally processor fetches information from memory. After processing the data, the result needs to be again stored in memory location. So, we have different operations in the memory. This information, whether it is a memory read operation (fetching) or memory write operation (storing) should be provided by the processor. This type of controlling information is provided through another bus which is known as control bus.

These three buses (data, address, control) combined together is termed as system bus. Through this system bus we connect memory module to the processor. The same system bus can also be used to connect Input/Output devices to the processor.

At any given point of time, single information (depending on the size of the bus; for 8 lines bus, 8 bits of information can be transferred together) can be transferred through the bus. To enhance the performance of the system, concept of multi bus is used. In the multi bus organization, (different) information can be transferred simultaneously through different buses, provided there are no conflicts between source and destination devices.