1. What purpose does the datapath in a CPU serve?

Datapath of a CPU is a network of registers and ALUs connected by buses. It provides temporary storage of data and functional units for transforming data.

2. What does the control unit in a CPU do?

The control unit of a CPU is in charge of sequencing operations performed by the datapath, and makes sure that correct data are where they need to be at the correct time.

3. How does the ALU know which function to perform?

the ALU knows which operation to perform as it is controlled by the signals coming from the control unit.

4. Explain the di↵erences between data buses, address buses, and control buses.

a data bus is the collection of wires that specifically carries data, which is the actual information that needs to be moved from one location in a CPU to another. An address bus is a collection of wires that carries address information for accessing memory or register IDs for accessing specific registers. A control bus is the collection of wires carries the information specifying the particular operations that need to be performed by the ALU, or read/write operations that need to be performed by the memory, or IO requests for IO communications, etc.

5. (Explain the relation between clock cycle time and clock frequency.

Clock cycle time means the amount time each clock cycle takes, while the clock frequency is the reciprocal of the cycle time. For example, an 1 GHz clock signal has a cycle time of 1 ns.

6. What are the main roles of operating Systems (OS)?

There are four main roles that an OS plays:
• Provides functions/SW packages to control HW, and serves as an interface of HW to SW.
• Manage HW resources to maximize the overall system efficiency.
• Make the system user-friendly and easy to use.
• Protect HW of the system from harmful actions of user applications or external

7. Consider a hypothetical microprocessor generating a 16-bit address (for example, assume that the program counter and the address registers are 16 bits wide) and having a 16-bit data bus.
    a. What is the maximum memory address space that the processor can access directly if it is connected to a “16-bit memory”?

b. What is the maximum memory address space that the processor can access directly if it is connected to an "8-bit memory"?

c. What architectural features will allow this microprocessor to access a separate "I/O space"?

d. If an input and an output instruction can specify an 8-bit I/O port number, how many 8-bit I/O ports can the microprocessor support? How many 16-bit I/O ports? Explain.

Answer :

a. The Maximum memory address space = 2^16 = 64 Kbytes.

b. The Maximum memory address space = 2^16 = 64 Kbytes.

Therefore, in (a) and (b), the microprocessor is to access 64K bytes, but the difference thing between them is that the access of 8-bit memory will transfer a 8 bits and the access of 16-bit memory may transfer 8 bits or 16 bits word.

c. Separate I/O instructions are needed because during its execution will generate separate its own signals I/O signals. That signals will be different from the memory signals which is generated during the execution for memory instructions. Therefore, one more output pin will be needed to carry I/O signals.

d.With an 8-bit I/O port number the microprocessor can support 2^8 = 256 8-bit input ports, and 2^8 = 256 8-bit output ports. With an 8-bit I/O port number the microprocessor can support 2^8 = 256 16-bit input ports, and 2^8 = 256 16-bit output ports.

Thus, the size of the I/O port will not change the number of I/O ports since the number of I/O ports depends on the number of bits which is used to represent the I/O port number (equals to 8 bits in both cases).

8. What is The differences among sequential access, direct access, and random access?

- **Sequential access**

  Memory is organized into units of data, called records. Access must be made in a specific linear sequence.

- **Direct access**

Individual blocks or records have a unique address based on physical location. Access is accomplished by direct access to reach a general vicinity plus sequential searching, counting, or waiting to reach the final location.

- **Random access**

  Each addressable location in memory has a unique, physically wired-in addressing mechanism. The time to access a given location is independent of the sequence of prior accesses and is constant.

9. The general relationship among access time, memory cost, and capacity?
   - Faster access time, greater cost per bit
   - Greater capacity, smaller cost per bit
   - Greater capacity, slower access time

10. The general relationship among access time, memory cost, and capacity?
    1. List I/O Module Function
    - Control & Timing
    - CPU Communication
    - Device Communication
    - Data Buffering
    - Error Detection
    2. Is it true that All peripherals are slower than CPU and RAM? Elaborate on what is needed to deal with that? (2 points)
    - Yes, need to add DMA or IO controller

11. Consider a hypothetical 32-bit microprocessor having 32-bit instructions composed of two fields: the first byte contains the opcode and the remainder the immediate operand or an operand address.

    a. What is the maximum directly addressable memory capacity (in bytes)?
    b. Discuss the impact on the system speed if the microprocessor bus has
       1. a 32-bit local address bus and a 16-bit local data bus, or
       2. a 16-bit local address bus and a 16-bit local data bus.
    c. How many bits are needed for the program counter and the instruction register?

a. $2^{24} = 16$ MBytes

b. **(1)** If the local address bus is 32 bits, the whole address can be transferred at once and decoded in memory. However, because the data bus is only 16 bits, it will require 2 cycles to fetch a 32-bit instruction or operand.
(2) The 16 bits of the address placed on the address bus can't access the whole memory. Thus a more complex memory interface control is needed to latch the first part of the address and then the second part (because the microprocessor will end in two steps). For a 32-bit address, one may assume the first half will decode to access a "row" in memory, while the second half is sent later to access a "column" in memory. In addition to the two-step address operation, the microprocessor will need 2 cycles to fetch the 32 bit instruction/operand.

c. The program counter must be at least 24 bits. Typically, a 32-bit microprocessor will have a 32-bit external address bus and a 32-bit program counter, unless on-chip segment registers are used that may work with a smaller program counter. If the instruction register is to contain the whole instruction, it will have to be 32-bits long; if it will contain only the op code (called the op code register) then it will have to be 8 bits long.

12. Consider a computer system that contains an I/O module controlling a simple keyboard/printer teletype. The following registers are contained in the processor and connected directly to the system bus:

INPR: Input Register, 8 bits
OUTR: Output Register, 8 bits
FGI: Input Flag, 1 bit
FGO: Output Flag, 1 bit
IEN: Interrupt Enable, 1 bit

Keystroke input from the teletype and printer output to the teletype are controlled by the I/O module. The teletype is able to encode an alphanumeric symbol to an 8-bit word and decode an 8-bit word into an alphanumeric symbol.

a. Describe how the processor, using the first four registers listed in this problem, can achieve I/O with the teletype.

b. Describe how the function can be performed more efficiently by also employing IEN.

a. Input from the Teletype is stored in INPR. The INPR will only accept data from the Teletype when FGI=0. When data arrives, it is stored in INPR, and FGI is set to 1. The CPU periodically checks FGI. If FGI =1, the CPU transfers the contents of INPR to the AC and sets FGI to 0.

When the CPU has data to send to the Teletype, it checks FGO. If FGO = 0, the CPU must wait. If FGO = 1, the CPU transfers the contents of the AC to OUTR and sets FGO to 0. The Teletype sets FGI to 1 after the word is printed.

b. The process described in **(a)** is very wasteful. The CPU, which is much faster than the Teletype, must repeatedly check FGI and FGO. If interrupts are used, the Teletype can issue an interrupt to the CPU whenever it is ready to accept or send data. The IEN register can be set by the CPU (under programmer control)

13. For a synchronous read operation (Figure 3.19), the memory module must place the data on the bus sufficiently ahead of the falling edge of the Read signal to allow for signal settling. Assume a microprocessor bus is clocked at 10 MHz and that the Read signal begins to fall in the middle of the second half of T3.

a. Determine the length of the memory read instruction cycle.
b. When, at the latest, should memory data be placed on the bus? Allow 20 ns for the settling of data lines.

a. With a clocking frequency of 10 MHz, the clock period is 10^{-9} s = 100 ns. The length of the memory read cycle is 300 ns.

b. The Read signal begins to fall at 75 ns from the beginning of the third clock cycle (middle of the second half of T_3). Thus, memory must place the data on the bus no later than 55 ns from the beginning of T_3.