Operating systems

Sheet1

(EED)

Name: Samuel Ayman Shawky

ID: 20010750

Q1.1) What are the two main categories of processor registers? What are they used for?

Answer)

The two main categories of processor registers are:

1-User-visible Registers

Minimize the main memory access by optimizing register use

And it's available for all programs ex:

- Data registers
- Address registers.

2-Control and status registers

They are specialized registers used by the CPU and peripheral devices to control operations and convey status information within a computer system.

ex

- Program counter (PC): contain address of next instruction.
- Instruction register (IR): hold the instruction being executed.
- Memory address register (MAR): holds the address of location to be accessed in the memory.

Q1.2) True or False?

A) GPUs are used for dealing with streaming signals such as audio or video **(False) -> DSPs**

B) SoC is a microchip with all the necessary system components on it **(TRUE)**

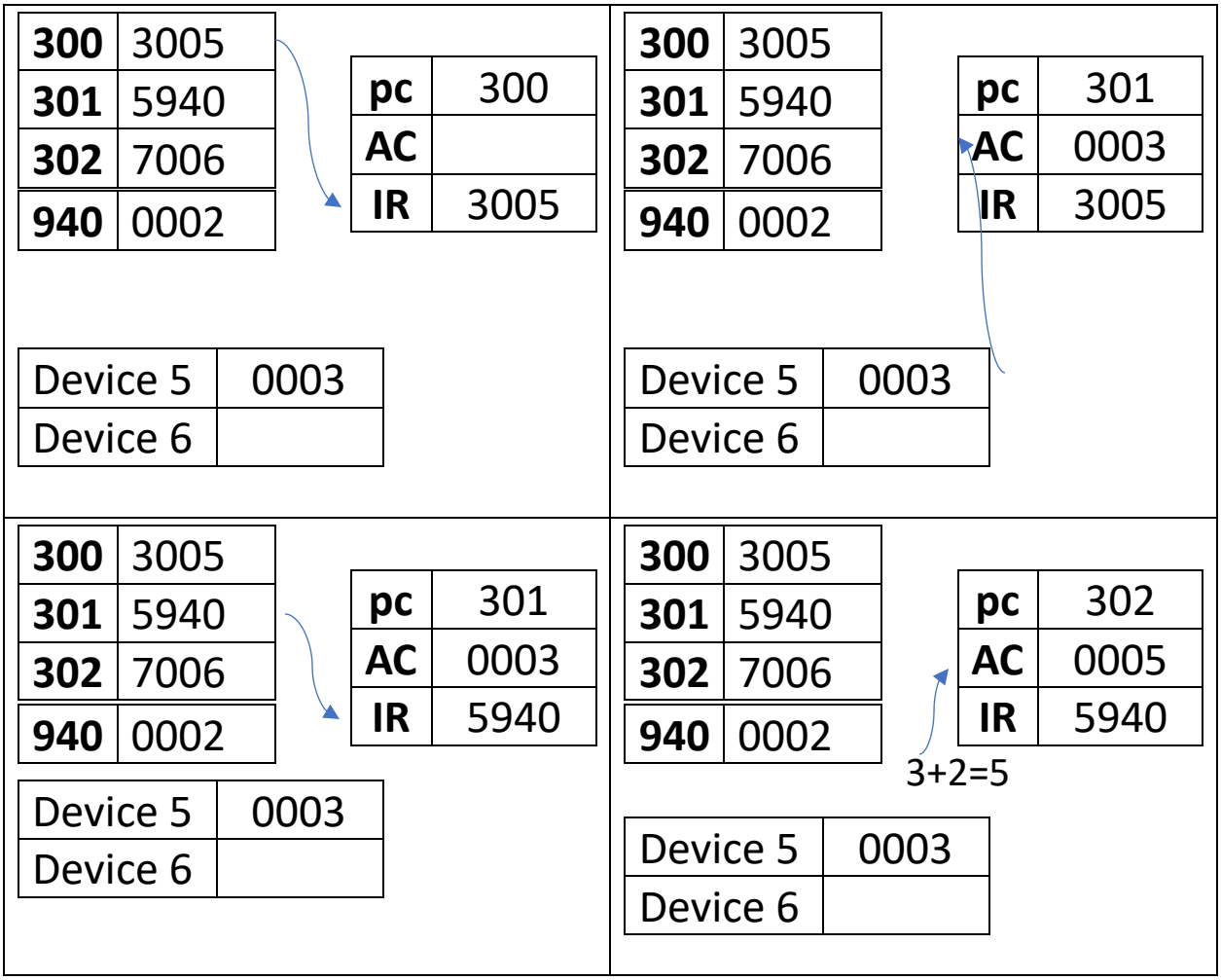
Q1.3)

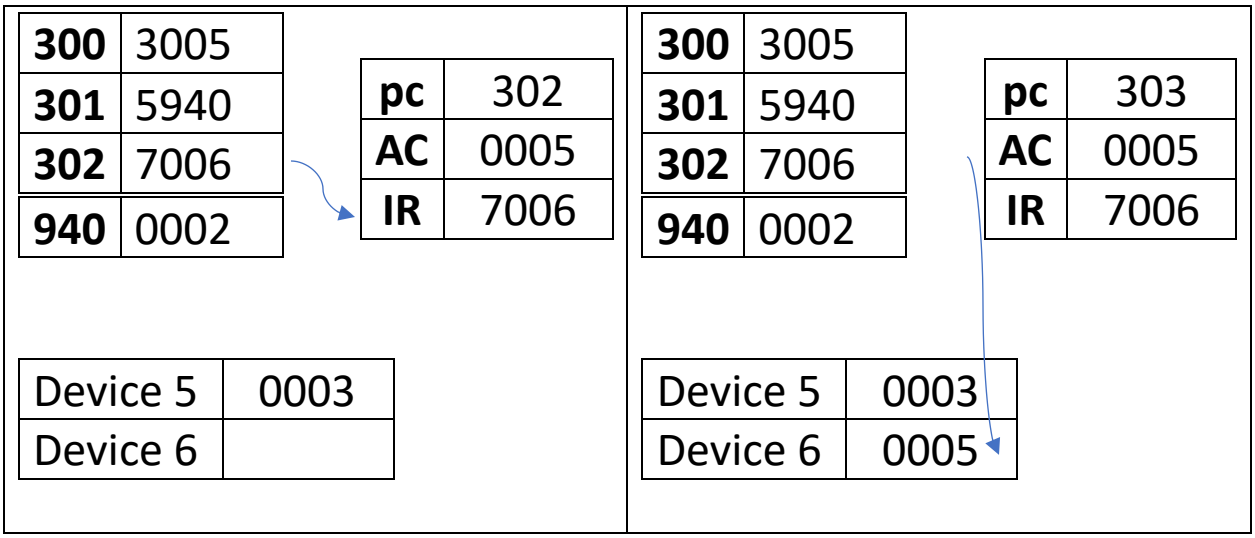
A) What are the four distinct actions that a machine instruction can specify?

Answer)

1. processor-memory: Transfer data between processor and memory
2. Processor-I/o: Transfer data between processor and I/O devices
3. Data-processing: Arithmetic or logic operations on data
4. Control : Alter sequence of execution

B)





Q1.4)

A)True or False?

Program interrupts are used to signal the normal completion of an operation or to signal a variety of error conditions **(False).**

-> (I/O interrupts)

B) Interrupt process involves overhead (in interrupt handling), Nevertheless, the processor can be employed much more efficiently with the use of interrupts. Explain?

Answer)

The interrupt process involves overhead, as it requires from processor to stop the current task and save its state and do another task as a response of the interrupt, despite this overhead interrupt improves the utilization of processor, as it enables multitasking and other services.

C) Explain, by example, handling the multiple interrupts via a priority scheme, and mention its advantages over the disabling-interrupts scheme?

Answer)

Printer interrupt with priority=5, communication interrupt with priority=8,disk interrupt with priority=3

if processor is executing an instruction and printer sent interrupt it finishes the current instruction and start printer operation while execution communication interrupt with higher priority=8 so the address of next instruction saved in stack and go to execute communication ,if disk interrupted with lower priority =3 ,communication interrupt has a higher priority can't be interrupted with a lower priority interrupt so communication continue execution then back to printer then execute disk operation then back to the first instruction that was executing

priority interrupts will be better than disabling-interrupts scheme in urgent operations take high priority and execute first but in disabling-interrupts scheme will wait in interrupt queue.

Q1.5)

A) Why are there different kinds of memory hierarchies? What are the characteristics distinguishing them?

Answer) There are different kinds of memory to balance between needs like speed and cost and size.

the characteristics distinguishing them are:

1-access speed, 2-capacity , 3-cost per Bit

B) Explain the term “locality of reference”?

Answer)

Locality of reference refers to the tendency of the computer program to access the same set of memory locations for a particular time period.

C)

Consider the following code:

```
for (i = 0; i < 20; i++)
    for (j = 0; j < 20; j++)
        a[i] = a[i] * j
```

1- Give one example of the spatial locality in the code?

Answer)

example of the spatial locality: the elements of array a ,a[0]&a[1]&… as it’s accessed many times in the outer loop.

2- Give one example of the temporal locality in the code?

Answer)

example of the temporal locality: a[i] used many times for each i to calculate. a[i] = a[i] * j.

Q1.6)

A) Consider a memory system with cache having the following parameters:

Sc = 32 KB Cc = 0.1 cents/bytes TC = 10 ns

Sm = 256 MB Cm = 0.0001 cents/bytes Tm = 100 ns

1-What was the total cost prior to the addition of the cache?

Answer)

cost=main memory.

cost=256*1024^2*0.0001=26843.5$

2. What is the total cost after the addition of the cache?

Answer)

cost= main-memory-cost + cache cost
=256*1024^2*0.0001+32*1024*0.1=30120.3$

3-What is the percentage decrease in time due to the inclusion of cache with respect to a system without cache memory considering a cache hit ratio of 0.85?

Answer)

T_{av} =0.85(10) +0.15(110) =25 ns,

access time decreased by 75%

B) Suppose that a large file is being accessed by a computer memory system comprising of a cache and a main memory. The cache access time is 60 ns. The time to access main memory (including cache access) is 300 ns. The file can be opened either in read or in write mode. A write operation involves accessing both the main memory and the cache (write-through cache). A read operation accesses either only the cache or both the cache and main memory depending upon whether the access word is found in the cache or not. It is estimated that read operations comprise 80% of all operations. If the cache hit ratio for read operations is 0.9, what is the average access time of this system?

Answer)

T_{cache}=60ns, T_{cache} + T_{memory}=300ns

T_{avr}=T_{read} +T_{write}

T_{avr}=0.8(0.9*60+0.1*300) +0.2*300=127.2ns

Q1.7) DIRECT MEMORY ACCESS

A computer consists of a CPU and an I/O device D connected to main memory M via a shared bus with a data bus width of one word. The CPU can execute a maximum of 10 instructions 6 per second. An average instruction requires five processor cycles, three of which use the memory bus. A memory read or write operation uses one processor cycle. Suppose that the CPU is continuously executing "background" programs that require 95% of its instruction execution rate but not any I/O instructions. Assume that one processor cycle equals one bus cycle. Now suppose that very large blocks of data are to be transferred between M and D.

1- If programmed I/O is used and each one-word I/O transfer requires the CPU to execute two instructions, estimate the maximum I/O data transfer rate, in words per second, possible through D?

5% only for I/O devices so no of instructions for I/O=10^6*0.05 = 50000 instruction/s 1 word require 2 instructions

so, transfer rate =25000 word/s

2- Estimate the same rate if DMA transfer is used?

Assuming i/o device uses all 3 memory buses

Assuming processor instructions after every block is neglected

Cycles=$(3/5)*10^6=6*10^5$ cycle/second

Max transfer rate =$6*10^5$ word /second

Q1.8) MULTIPROCESSOR AND MULTICORE ORGANIZATION

A) What are the advantages of An SMP organization over a uniprocessor organization?

Answer)

- Performance**:** SMP processors work simultaneously, tasks are distributed among all processors which improves the performance.
- Reliability: even one processor fails the machine doesn't halt
- Scaling: vendors can offer range of products and performances according to budget
- Incremental growth: can easily increase performance by adding processor.

B) What is a chip multiprocessor? What is the motivation for its development?

Answer)

Chip multiprocessor is two or more processors in a single piece of silicon, these processors called cores each core contain all components of an independent processor cores have caches L1&L2 in some cases have L3

The motivation for the development of multicore computers comes from reaching limits in improving single-core processor performance. Microprocessor systems experienced performance gains through increased clock frequency, closer cache memory placement. However, these approaches faced constraints due to hardware limitations. So,

designers turned to multicore architectures, incorporating multiple processors and cache memory onto a single chip.