Name: Amber Khurshid
Section: BAI-5A
Roll no: 22P-9295

OS Assignment #01

Q1. Direct memory access is used for high speed I/O devices in order to avoid increasing the CPU's execution load.

a) How does the CPU interface with the device to coordinate the transfer?

The CPU sets up the DMA by giving the device's address, the memory location, and the amount of data to transfer. After that the DMA controller takes over, reducing the need for the CPU to stay involved in transfer process.

b) How does the CPU know when the memory operations are complete?

Once the data transfer is done, the DMA controller sends a signal (interrupt) to the CPU to let it know if can continue processing.

c) The CPU is allowed to execute other programs while the DMA controller is transferring data. Does this process interfere with the execution of the used program? If so, describe what forms of interference are caused.

Although DMA does not interfere directly with user programs, there can be minor delays if both the CPU and DMA try to access memory at the same time. However these delays are typically minimal.

2. Describe the differences between symmetric and asymmetric multiprocessing. What are the advantages and disadvantages of multiprocessor systems?

Symmetric Multiprocessing:

In symmetric multiprocessing, all processors share the same memory and work equally on tasks, leading to better work load distribution.

Asymmetric Multiprocessing:

In asymmetric multiprocessing, one main processor controls the system, while the others handle specific tasks. This set up is simpler to design but does not use resources efficiently.

Advantages:

Increased performance, reliability, and fault tolerance. They can handle more tasks simultaneously and are more robust.

Disadvantages:

They can be complex and costly, and resource conflicts or synchronization issues between processors may arise.

3. What is the purpose of interrupts? What are the differences between a trap and an interrupt? Can traps be generated intentionally by a user program? If so, for what purpose?

Purpose of Interrupts:
Interrupts are signals sent to the CPU, telling it to pause its current task to deal with something more urgent like completing an I/O call. This enables real-time responses to external events.

Differences btw trap and Interrupts
A trap is a software-generated interrupt, typically triggered by errors or system calls, whereas interrupts are hardware based.

Intentional traps:
Yes, user programs can deliberately trigger traps, often to request services from the operating system through system calls, allowing controlled communication with OS.

4. Rank the following storage systems from slowest to fastest.
a) Hard disk drives
b) Registers
c) Optical Disk
d) Main memory
e) Non volatile memory
f) Magnetic tapes
g) Cache.

Magnetic Tapes, Optical Disks, Hard Disk Drives, Non volatile memory, Main Memory, Cache, Registers.

5. Differentiate between multiprogramming systems and multiprocessing systems.

Multiprogramming Systems:

They allow multiple programs to reside in memory at the same time, enabling better CPU utilization. The CPU switches between programs to maximize usage, although only one process runs at a time.

Multiprocessing Systems:

This involves multiple CPUs working simultaneously to execute different parallel processes concurrently. It provides parallelism and increases processing power.

6. Give two reasons why caches are useful. What problems do they solve? What problems do they cause? If a cache can be made as large as the device for which it is caching, why not make it that large and eliminate the device?

Reasons why they are useful:

- Caches provide faster access to frequently used data compared to main memory.
- They reduce average time to access memory, improving overall system performance.

Problems they solve:

They solve the problems of the speed disparity between the CPU and main memory.

Problems they create:

Caches introduce complexity in maintaining data consistency in multiprocessor systems.

Why not make caches Huge?

Making caches as huge as storage devices would be extremely expensive and still wouldn't be as fast as smaller, quicker caches. Large caches also consume more power and are less efficient.

7. Distinguish between the client-server and peer to peer models of distributed systems.

In client server models, clients request services from a central server. It offers centralized control and better security management. Whereas in peer to peer models all nodes in the system act as both clients and servers, sharing resources equally. This model provides better scalability and redundancy but it is harder to secure.

8. What is the purpose of system calls, and how do system calls relate to the OS and to the concept of dual mode operation?

System calls provide a way for user programs to interact with OS. This allows controlled access to hardware resources.

Dual mode protects the system by separating user-mode from kernel-mode. It prevents user programs from directly accessing critical resources.