

Tutorial

Chapter 1

Introduction

Give two main objectives of operating system design

- Resource allocation, for example, CPU time, memory space, I/O devices.
- Control program, which is controlling user programs to prevent errors and improper use of computer.

Why hardware is abstracted is abstracted by OS and application? Give an example of such abstraction.

- OS and application programs abstract the hardware of a computer. This abstraction helps in controlling and coordinating the use of hardware among various application programs for various users.
 - OS Example: Windows, Linux
 - Applications: Compilers, Editors.

What is the difference between: mono-programming and multi-programming.

- In a mono-programming, a single job occupies the system from start until end. In such a case CPU sits idle when there is a need of user action.
- A multiprogramming system increases CPU utilization by organizing several jobs into memory so that CPU always has something to execute.

Direct Memory Access (DMA) 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?**
- b. How does the CPU know when the memory operations are complete?**
- 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 user programs? If so, describe what forms of interference are caused.**

The CPU can initiate a DMA operation by writing values into special registers that can be independently accessed by the device. The device initiates the corresponding operation once it receives a command from the CPU. When the device is finished with its operation, it interrupts the CPU to indicate the completion of the operation.

Both the device and the CPU can be accessing memory simultaneously. The memory controller provides access to the memory bus in a fair manner to these two entities. A CPU might therefore be unable to issue memory operations at peak speeds since it has to compete with the device in order to obtain access to the memory bus

What is the difference between: mono-processor and multi-processor.

- In a mono-processor machine there is only one processor which is responsible for the execution.
- In a multi-processor machine there are more than one processor that may share the execution.

What is a program?

A sequence of instructions that a computer can interpret and execute.

What is a process?

A process is a running instance of a program, including all variables and other state.

What is a Time sharing?

A way of sharing out computer facilities between number of people who want to use the computer at the same time.

What are the three main purposes of an operating system?

- To provide an environment for a computer user to execute programs on computer hardware in a convenient and efficient manner.
- Resource allocator. To allocate the resources of the computer as needed to solve the problem given. The allocation process should be as fair and efficient as possible.
- As a control program it serves two major functions: (1) supervision of the execution of user programs to prevent errors and improper use of the computer, and (2) management of the operation and control of I/O devices.

What is the main advantage of multiprogramming?

Multiprogramming makes efficient use of the CPU by overlapping the demands for the CPU and its I/O devices from various users. It attempts to increase CPU utilization by always having something for the CPU to execute

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?

An interrupt is a hardware-generated change of flow within the system.

An interrupt handler is summoned to deal with the cause of the interrupt; control is then returned to the interrupted context and instruction.

A trap is a software-generated interrupt.

An interrupt can be used to signal the completion of an I/O to obviate the need for device polling.

A trap can be used to call operating system routines or to catch arithmetic errors.

Define the essential properties of the following types of operating systems:

- a) Batch**
- b) Time sharing**
- c) Real time**
- d) Network**
- e) Distributed**
- f) Clustered**
- g) Handheld**

Batch:

Jobs with similar needs are batched together and run through the computer as a group by an operator or automatic job sequencer. Performance is increased by attempting to keep CPU and I/O devices busy at all times through buffering, off-line operation, spooling, and multiprogramming. Batch is good for executing large jobs that need little interaction; it can be submitted and picked up later.

Time sharing:

Uses CPU scheduling and multiprogramming to provide economical interactive use of a system. The CPU switches rapidly from one user to another. Instead of having a job defined by spooled card images, each program reads its next control card from the terminal, and output is normally printed immediately to the screen.

Real time:

Often used in a dedicated application. The system reads information from sensors and must respond within a fixed amount of time to ensure correct performance.

Network:

A network operating system is an operating system that provides features such as file sharing across the network, and that includes a communication scheme that allows different processes on different computers to exchange messages.

Distributed:

Distributes computation among several physical processors. The processors do not share memory or a clock. Instead, each processor has its own local memory. They communicate with each other through various communication lines, such as a high-speed bus or telephone line.

Clustered:

Like parallel systems, clustered systems gather together multiple CPUs to accomplish computational work. Clustered systems differ from parallel systems, however, in that they are composed of two or more individual systems coupled together.

Handheld:

operating systems are designed to provide an environment in which a user can easily interface with the computer to execute programs. many varieties of handheld computers have come into fashion. These devices are mostly standalone, used singly by individual users. Some are connected to networks, either directly by wire or (more often) through wireless modems.

What network configuration would best suit the following environments?

a) office:

Local Area Network (LAN) connects network devices over a relatively short distance

What network configuration would best suit the following environments?

- **office**
- **university campus**

Local Area Network (LAN) connects network devices over a relatively short distance

What network configuration would best suit the following environments?

- **state/region**

Metropolitan Area Network (MAN) - a network spanning a physical area larger than a LAN but smaller than a WAN, such as a city. A MAN is typically owned and operated by a single entity such as a government body or large corporation

What network configuration would best suit the following environments?

- **nation/country**

WAN - Wide Area Network (WAN) covers a large physical distance. The Internet is the largest WAN

What are the tradeoffs inherent in handheld computers?

- Limitation of size in memory. Usually in megabytes compared to gigabytes in computers.
- Processor speed is limited because of less power ability. the operating system has to be designed for low power consumption.
- Lack of physical space limits the I/O capability. The input method are small screen based keyboard or hand-writing recognition while small display screen for output.

What is the main difficulty that a programmer must overcome in writing an operating system for a real-time environment

The main difficulty is keeping the operating system within the fixed time constraints of a real-time system. If the system does not complete a task in a certain time, it could cause a breakdown of the entire system it is running. Therefore when writing an operating system for a real-time system, the writer must be sure that his scheduling schemes don't allow response time to exceed the time constraint

Which of the following instructions should be privileged?

- a) Set value of timer.**
- b) Read the clock.**
- c) Clear memory.**
- d) Issue a trap instruction.**
- e) Turn off interrupts.**
- f) Modify entries in device-status table.**
- g) Switch from user to kernel mode.**
- h) Access I/O device.**

Set value of timer, clear memory, turn off interrupts, modify entries in device-status table, access I/O device. The rest can be performed in user mode