L3: Operating system operations, Computing environments

Dr. Rajashree Dash

Associate Professor,
Department of CSE,
ITER, Siksha O Anusandhan Deemed to be University

Outline

Operating system operations

- Modern operating systems are interrupt driven. If there are no processes to execute, no I/O devices to service, and no users to whom to respond, an operating system will sit quietly, waiting for something to happen.
- Events are almost always signaled by the occurrence of an interrupt or a trap.
- An interrupt service routine is provided to deal with the interrupt.
- A properly designed operating system must ensure that an incorrect (or malicious) program cannot cause other programs to execute incorrectly.

Dual mode operation

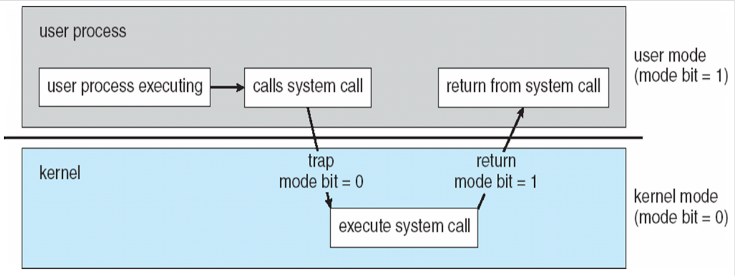

Figure: Dual mode operation

Timer

- Timer is used to prevent a user program to get stuck in an infinite loop or to fail to call system services and never return control to operating system.
- Timer is set to interrupt the computer after some time period. Period may be fixed or variable. A variable timer is implemented by a fixed rate clock and a counter.
- OS sets the counter (Privileged instruction). Every time the clock ticks, counter is decremented. When counter reaches zero, an interrupt occurs.
- Before turning control to user, OS ensures that timer is set to interrupt. As long as counter is positive, control is returned to user program. When counter becomes negative, OS terminates the program that exceeds allotted time.

Functions of Operating System

- Process management
- Memory management
- File management
- Mass storage management
- I/O subsystem
- Protection and security

Process management

Process management activities include:

- Creating and deleting both user and system processes
- Suspending and resuming processes
- Providing mechanisms for process synchronization
- Providing mechanisms for process communication
- Providing mechanisms for deadlock handling

Memory management

Memory management activities include:

- Keeping track of which parts of memory are currently being used and by whom
- Deciding which processes (or parts thereof) and data to move into and out of memory
- Allocating and deallocating memory space as needed

File management

File management activities include:

- Creating and deleting files and directories
- Primitives to manipulate files and directories
- Mapping files onto secondary storage
- Backup files onto stable (non-volatile) storage media

Mass storage management

Mass storage management activities include:

- Free-space management
- Storage allocation
- Disk scheduling

I/O subsystem

I/O subsystem is responsible for

- Memory management of I/O including buffering (storing data temporarily while it is being transferred), caching (storing parts of data in faster storage for performance), spooling (the overlapping of output of one job with input of other jobs)
- General device-driver interface
- Drivers for specific hardware devices

Protection and security

- Protection – any mechanism for controlling access of processes or users to resources defined by the OS
- Security – defense of the system against internal and external attacks
- Systems generally first distinguish among users, to determine who can do what
  - User identities (user IDs, security IDs) include name and associated number, one per user
  - User ID then associated with all files, processes of that user to determine access control
  - Group identifier (group ID) allows set of users to be defined and controls managed, then also associated with each process, file
  - Privilege escalation allows user to change to effective ID with more rights

Computing Environments

- Traditional computing involves desktop and laptop PCs, usually connected to a computer network.
- Mobile computing refers to computing on handheld smartphones and tablet computers, which offer several unique features.
- Distributed systems allow users to share resources on geographically dispersed hosts connected via a computer network.
- Virtualization involves abstracting a computer's hardware into several different execution environments.
- Cloud computing uses a distributed system to abstract services into a "cloud," where users may access the services from remote locations.
- Real-time operating systems are designed for embedded environments.

Distributed system

- A distributed system (Loosely coupled system) is a collection of physically separate, possibly heterogeneous, computer systems that are networked to provide users with access to the various resources that the system maintains.
- It is a collection of processors that do not share memory or clock. Each processor has its own local memory and they communicate through various communication lines.
- Distributed systems depend on networking for their functionality.
- Services may be provided through either the client-server model or the peer-to-peer model.

Client-Server Model

- Many systems now servers, responding to requests generated by clients.
- Server systems can be broadly categorized as compute servers and file servers:
- The compute-server system provides an interface to which a client can send a request to perform an action. In response, the server executes the action and sends the results to the client. (Example: Database server)
- The file-server system provides a file-system interface where clients can create, update, read, and delete files. (Example: Web server)

Client-Server Model

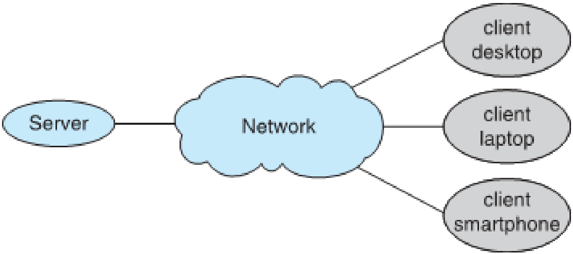

Figure: Client Server Model

Peer to Peer Model

- It does not distinguish clients and servers.Instead all nodes are considered peers.
- Each may act as client, server or both depending on whether it is requesting or providing a service.
- In such system, nodes must join network of peers.Once a node has joined the network, it can begin providing services to and requesting services from other nodes in the network.

Peer to Peer Model

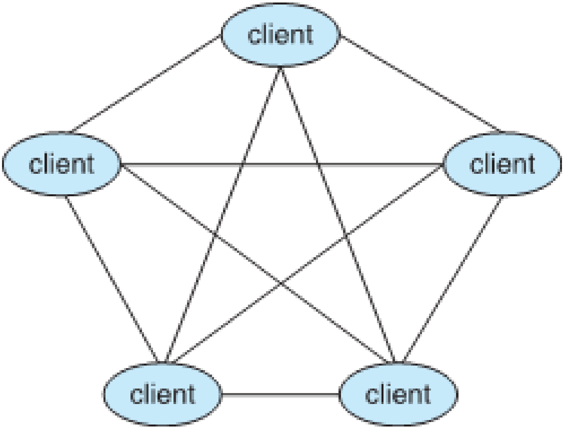

Figure: Peer to Peer Model with no centralized service

Peer to Peer Model

Determining what services are available is accomplished in one of two general ways:

- Using a centralized lookup service on the network (Example: Napster system)
- Using broadcasting technique (Example: Gnutella system)

Skype is an example of hybrid peer-to-peer computing, that includes a centralized login server, but it also incorporates decentralized peers and allows two peers to communicate.

Virtualization

- Virtualization involves abstracting a computer's hardware into several different execution environments.
- It is a technology that allows operating systems to run as applications within other operating systems.
- It is one member of a class of software that also includes emulation. Emulation is used when the source CPU type is different from the target CPU type.
- That same concept can be extended to allow an entire operating system written for one platform to run on another.

Cloud computing

- Cloud computing is a type of computing that delivers computing, storage, and even applications as a service across a network.
- It is a logical extension of virtualization, because it uses virtualization as a base for its functionality. Different types of cloud computing:
  - Public cloud – available via Internet to anyone willing to pay
  - Private cloud – run by a company for the company's own use
  - Hybrid cloud – includes both public and private cloud components
  - Software as a Service (SaaS) – one or more applications available via the Internet (i.e., word processor)
  - Platform as a Service (PaaS) – software stack ready for application use via the Internet (i.e., a database server)
  - Infrastructure as a Service (IaaS) – servers or storage available over Internet (i.e., storage available for backup use)

Real-Time Systems

- Real-Time Systems are designed for embedded environments such as automobiles, robotics and soon.The embedded systems tend to have very specific tasks.
- Real time system is used when there are rigid time requirements on the operation of a processor or the flow of data and thus is often used as a control device in a dedicated application.
- A real-time system has well-defined, fixed time constraints. Processing must be done within the defined constraints, or the system will fail.
- Sensors bring data to the computer.The computer must analyze the data and possibly adjust controls to modify the sensor inputs.
- Systems that control scientific experiments, medical imaging systems, industrial control systems, and certain display systems are real time systems.