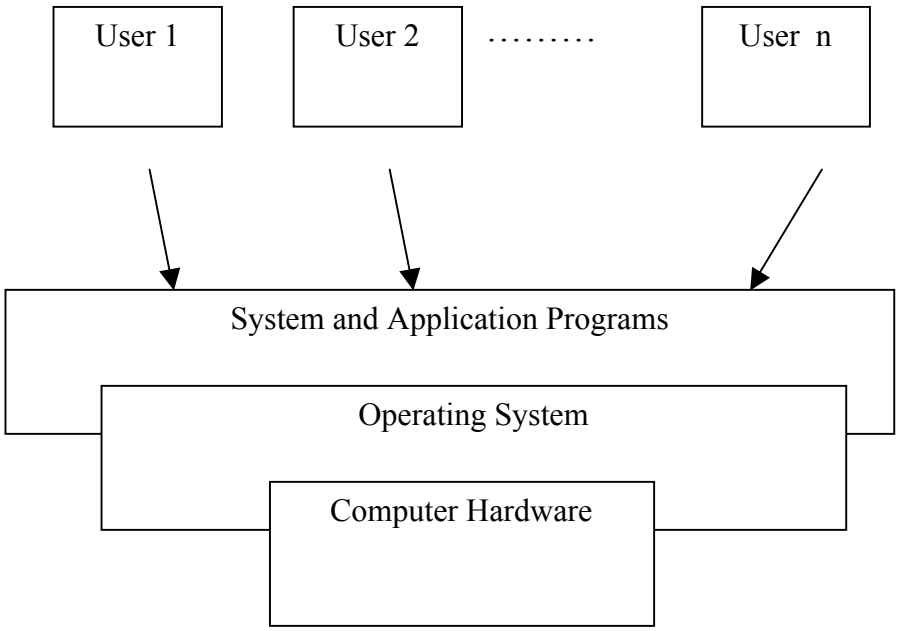
The Operating System

- Important part of every computer system.
- It is a program that manages the computer hardware.
- Provides basis for application programs.
- Acts as an intermediary between the user of the computer and the computer hardware.

Operating systems are varied in task accomplishing :

- Main frame OS designed to optimize utilization of the hardware.
- PC OS supports complex games, business applications, … etc

A Computer System can be divided into 4 components :-



Components of a Computer System

- **Hardware** (Memory, CPU , I/O Devices) >>>> provides basic computing resources.

- **Application Programs** (Word, Excel, Compilers, Web Browsers)defines the way in which these resources are used to solve a computing problem of the user.

- **Operating System** : controls and coordinates the use of the hardware among the various application programs for the various users.

- **Users**. Users vary by the interface use :-

  Most users sit in front of the PC consisting of monitor, keyboard, mouse. These systems are designed for single user >>> monopolize its resources to max the work. In

  This case, the OS is designed for

  - Ease of use
  - Some attention is paid to performance.
  - None attention is paid to resource utilization.

  Some users sit in front of a terminal that is connected to a mainframe or minicomputer. Other users accessing the same computer from different terminals. >>> share resources and exchange information.

  This case, the OS is designed for

  - maximize resource utilization. This to assure that
    - All available CPU time, memory, and I/O are used efficiently.
    - No individual user takes more than her fair time.

  Some computers have no little or no user view such as embedded computers in home devices and cars. They might have numeric keypad or might turn an indictor or a light on or off to show status.

System View (Purpose Of The OS)

We can **view** the OS as a ***resource allocator***. A computer system has many resources (hardware, software that may require a CPU time, memory space, file storage space, I/O devices in order to solve a problem). In this case, The OS acts a manager of these resources and must decide how to allocate these resources without allowing a conflict to occur.

We can view the OS as a ***control program***. Control the various I/O devices and user programs to prevent an error or improper use of the computer.

System Goals

1) **Convenience for the user (The small PCs)** : They exist because it suppose to make it easier to compute with them than without them.

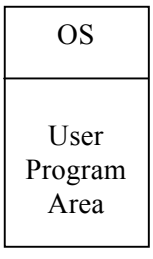
2) **Efficiency** : Efficient operation of the computer system. This is true when it come to large, shared, multiuser systems. These systems are expensive to design and implement. They are very complex and some times, it might require change in the hardware design in order to simplify the function of the OS.

Historical Overview

Mainframe Systems

First computer used to tackle many commercial and scientific applications.

- ***Batch System*** : Computer run one and only one application at a time.



  - They are big, run from a consol.
  - Card readers, tape drives were common input devices.
  - Line printers and tape derives were common output devices.
  - No direct interaction with the computer , usage of JCL to run the programs.
  - The Jobs were in form of punch cards.

- After hours or days, the job was completed and the output was produced.

To Speed up the process, the operators **batched** together jobs with similar needs and run them as group. Most of the time, the CPU was idle because the speed of the I/O devices (mechanical devices such as card reader) is far slower from electronic devices (CPU). Technology improved >>> disk drives (faster than I/O devices). Keep all the jobs on a disk, perform ***job scheduling***

- ***Multiprogramming Systems*** :- A single user cannot keep either the CPU or the I/O devices busy at all times. Multiprogramming increases the CPU utilization by organizing the jobs so that CPU always has a job to execute. It never stays ***idle***. (Example in our daily life, a layer does not work in one case only, while one case is getting the paper typed or waiting for a trail, he works on another case .)

  Must have a job pool. All jobs in the pool are waiting for memory. ***Job scheduling*** is used to run each job. Job is loaded into main memory for execution. Having few jobs in memory require some type of ***memory management***.

OS
Job 1
Job 2
Job 3

- ***Time shared Systems*** (or ***Multitasking Systems***): Allow the user to interact with the computer system while it is running. Interactive provides a direct communication between the user and the system via keyboard or a mouse and wait for an immediate result. The ***response time*** should be short.
  - ***A time sharing OS*** allows many users to share the computer simultaneously. Since the response time is very short and the system switch rapidly between jobs, every user think that the system is fully dedicated to him/her.
  - ***A time sharing OS*** uses job scheduling and multiprogramming to provide each user with a small portion of time shared computer.
  - ***A time sharing OS*** uses memory management and protection. Jobs might have to be swapped in and out of memory to a disk that serves as a backing store for main

memory. This achieved via ***virtual memory***. (when the program is larger than the ***physical memory***).

- ***A time sharing System*** provides a file system. File system resides on collection of disks and hence disk management must be provided.

- ***A time sharing system*** provides a mechanism for concurrent execution which require a sophisticated CPU-Scheduling, job synchronization and communication.

Desktop Systems

In early years :-

- CPU in PCs lacked the feature needed to protect an OS from user programs.
- File protection was not necessary.
- PC OS were neither multiuser nor multitasking.

Now, It changed to IBM's OS/2 multitasking, Apple's new virtual memory and multitasking. Linux, a Unix-like OS available for PCs. Microcomputers were able to adopt some of the technology developed for larger OS >>>> Hardware cost of microcomputers was lowered. >>> more sophisticated file protection mechanism.

Multiprocessor Systems

Known as ***parallel systems*** or ***tightly coupled*** systems. They have more than one processor in close communication, sharing the computer bus, the clock, and sometimes memory and peripherals.

Advantages :-

Increase Throughout : More processors means more work done in less time. It is not always the case because of overhead in keeping all the parts working correctly.

Economy of Scale : Save more money than multiple single processor machines because they can share peripherals , storage, power supply.

Increase Reliability : with proper distribution of functions, failure of one processor will not halt the system, it only slows it down. If we have 10 processors, and if one processor fails, then the result in only 10% slow down. (***Fault Tolerant Machines***) , (***2 identical systems***).

1.5 Distributed Systems

A ***network*** is communication path between two or more systems. Able to share computational tasks and provide set of features.

Networks Vary By

- ***Protocols used*** (TCP/IP or ATM). TCP/IP is the most common network protocol. (Transmission Control Protocol / Internet Protocol).

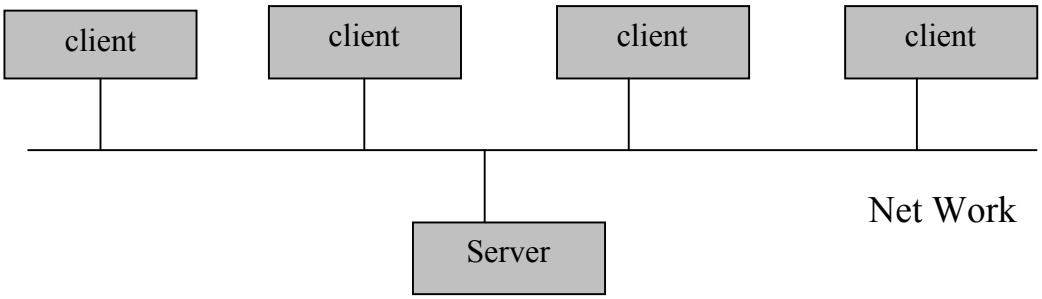
  To an OS, a network protocol needs :-
  - An interface device - network adaptor with a device driver to manage it
  - A software to pack the data when sending and unpacking the data when receiving it,

- ***Distance Between Nodes***
  - ***LAN*** (Local Area Network) – exist within a room, a floor, or a building.
  - ***WAN*** (Wide Area Network) – cities, countries.
  - With new technology, new form of network was created such as ***MAN*** (Metropolitan-Area Network) that connects buildings within a city, and ***Small-Area Network*** such as the ***BlueTooth*** that communicate over a short distance of several feet.

- ***Transport of Media***
  - Copper wires, fiber stands, and wireless transmission between satellites, microwave dishes, and radios

1.5.1 Client-Server Systems

As PCs become faster, cheaper, more powerful, designers have shifted from centralized system architecture. Terminals are replaced by PCs. User-interface functionality that used to be handled by the centralized systems are now handled by the PCs. Thus, Centralized Systems acts as ***Servers*** to satisfy requested granted by the ***client*** systems.



General Structure of a client-server system.

Server System can be categorized as :-

Compute-Server Systems : Provide an interface to which clients can send requests to perform an action, in response to which they execute the action and send back the results to the client.

File-Server Systems : provide a file-system interface where the clients can create , update, and delete files.

1.5.2 Peer-to-Peer Systems

Growth of computer networks (Internet, WWW) had a big influence on the recent development of the OSs. When PC was introduced in the 70s, they were designed as standalone computers. With the beginning of wide spread usage of the internet for mail, ftp .. etc, PCs became connected to computer networks and the network connectivity became an essential component of any computer system.

Virtually, all modern PCs and workstations are capable of running web browsers for accessing ***hypertext*** documents on the web. OSs such as (windows XP, Mac OS .. etc) include a system software (such as TCP/IP, PPP) that enable computers to access the internet via local-area network or telephone connection (for browsing, email, file transfer, remote login).

The computer networks used in these applications consist of collection of processors that do not share memory or a clock , instead, each processor has its own memory, the processors communicate with each other via various types of communication tools (such as high speed bus or telephone lines). These systems are called ***loosely coupled systems*** or ***distributed systems***).

Network Operating System : A 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.

1.6 Clustered Systems

They gather together multiple CPUs to accomplish computational work. They differ from ***parallel systems*** (more than one processor in close communication sharing the computer bus, clock, and some memory), in that they compose of two or more individual systems coupled together.

The generally accepted definition is that clustered computers share storage and are closely linked via LAN network.

Clustering is used to create ***high availability*** .

- Limited to two or four hosts due to complexity of connecting
- A layer of cluster software runs on the cluster nodes.
- Each node can monitor one or more of the other (over the LAN)
- If a monitored machine fails, the monitoring machine can take the ownership of its storage and restart the application that were running on the failed machine.
- The fail machine can remain down until it is fixed, the user will only see a brief interruption of the service

Asymmetric Clustering :- One machine is hot standby mode (does nothing, just wait) while other is running the application.

Symmetric Clustering ： Two or more hosts are running applications and they monitor each other.

Parallel Clustering : Over the ***WAN***. Multiple hosts access the same data on shared storage. (Oracle Parallel Server).

- Oracle DB designed to run on parallel clusters
- Each machine run Oracle
- Layer of software tracks the access to shared disk
- Each machine has full access to all data in the database .

Distributed Locking Manager (***DML***) : most clusters do not allow shared access to data on disk, thus ***Distributed File System*** must provide access control and locking to files to ensure no conflicting operations occur.

1.7 Real-Time Systems

- Another form of special-purpose operating system.
- It is used when rigid time requirement has been placed on the operation of a processor or flow of data.
- It is often used as a control of device in a dedicated application (such as medical imaging systems, industrial control system , weapon systems.. etc.).
- It has well-defined, fixed time constraints. Processing must be done within fixed amount of time or the system will fail.

A ***hard real-time*** system :

- Guarantees that critical tasks to be completed on time.
- This goal requires that all delays in the system must be bounded from the time of stored data to the time that it take the OS to finish any requested made of it.
- Secondary storage is missing or limited, data instead being stored in short term memory or read-only memory (ROM).
- Most advanced OS features are missing since they tend to separate the user from the hardware.
- Real-time and time sharing operating systems can not be mixed.

A ***soft real-time*** system :

- Critical real-time task gets priority over other tasks, and retain priority until task is completed .
- A real-time task can not waiting for a long time for the kernel to run it.
- They can mixed with other types of operating systems.
- Has more limited utility than hard real-time systems.
- Given their lack of deadline support, they are ***more risky*** to be used in industrial control and robotics.
- They are useful in areas such as multimedia, virtual reality .. etc.
- Because of expanded uses of soft real-time functionality, it appears in most OSs such as major version of UNIX.

1.8 Handheld Systems

Personal Digital Assistants (PDAs), such as ***Palm*** or cellular telephones with connectivity to the network such as the internet.

Due to the limited size of such devices , developers of handheld applications and operating systems face many challenges (such as small amount of memory, slow processors, small display screens) .

Memory : many of such machines carry 68 mb, or 128 mb, 256 mb, 512 mb, or 1 gb of memory. Thus, the OS and applications must manage memory efficiently. PCs have more memory than handheld devices.

Speed of the Processor : PCs have faster processors than handheld devices. Faster processor require more power which require larger battery that would have to be replaced or recharged more frequently. To minimize the size of these devices, smaller , slower processors which consumes less power are typically used.

Small display screens : PCs have larger screens up to 21 inch. Handheld devices have screens that are no more than 3 inches square. Familiar tasks such as reading e-mail or browsing the web must be conducted on smaller screens. ***Web-clipping*** is used (small subset of a web page is deliverable and displayed on the device).

They may use wireless technology such as BlueTooth allowing remote access to e-mail and web-browsing. (cell phones fall in this category but not PDAs). To download data to these devices, one must first download data to a PC or workstation, and then download the data to the PDA. Some PDAs allow direct transformation of data via infrared link. Their use begins to expand as network connections become more available and other options such as cameras and MP3 players expand their utilities.

1.10 Computing Environments

Brief overview of how each type of previous systems are used in verity of computing environment settings

1.10.1 Traditional Computing :-

Consider "typical office environment ", few years ago, this environment consisted of :-

- PC connected to network
- Servers providing file and print services
- Remote access was awkward
- Portability was achieved by laptop computers carrying some of the user's workspace.

Current trend :-

- more ways to access these environment
- companies implement ***portals*** which provide web accessibility to their internal servers.
- Handheld computers can synchronize with PCs to allow very portable use of company information.
- Can connect to ***wireless networks*** to use company's web portal.

At Home

- Few years ago, most users have single computer, slow modem that connects to the internet of the office or both. Cost of connection was very high.
- Now, low cost of internet connection, faster connection to the internet and allow access to company's data from the web. Some homes have ***firewalls*** to protect these home environment from security breaches. Firewalls are costly now and did not exist decades ago.

1.10.2 Web-Based Computing

- PC is the most prevalent access device with workstations, handheld, PDAs and even cell phones also provide access.

- Web computing has increased on networking. Devices that were not previously networked now have wired or wireless access. Devices have faster network connectivity .

1.10.3 Embedded Computing.

- Embedded real-time system such as (car engines, manufacturing robot to VCR and microwave ovens).
- They have very specific tasks
- Lacking OS advance features such as virtual memory and even disks,
- Little or no user interface.

The use of embedded systems continues to expand. The power of these devices, both as standalone units and as a member of networks and the web are increasing.

- Entire house hold can be computerized so that central computer – either embedded or general purpose computer – can control heating and lighting, alarm systems, and even coffee makers.
- Web access can let a home-owner to tell the house to heat up before he arrives
- Someday, the refrigerator may call the store when it notice that the milk is gone.