



UNIT - 1

INTRODUCTION

Syllabus

Teaching and Examination Scheme:

Teaching Scheme			Credit	Examination Scheme					Total
Lect Hrs/ Week	Tut Hrs/ Week	Lab Hrs/ Week		External		Internal			
				T	P	T	CE	P	
3	0	2	4	60	30	20	20	20	150

Lect - Lecture, Tut - Tutorial, Lab - Lab, T - Theory, P - Practical, CE - CE, T - Theory, P - Practical

Contents:

Sr.	Topic	Weightage	Teaching Hrs.
1	INTRODUCTION: Concept of Operating Systems, Generations of Operating systems, Types of Operating Systems, OS Services, System Calls, Structure of an OS-Layered, Monolithic, Microkernel Operating Systems, Concept of Virtual Machine.	5%	4

Basic concepts

Software:-

- Program is a collection of code/instruction.
- Software is a collection of program.

Hardware:-

- Physical device is a collection of computer system which is called Hardware.

Example: Processor, RAM, Hard disk, I/O devices.

Types of Software:

Software is divide into 3 types:

- System software
- Utility software
- Application software

System software

The software which is used to perform all types of system level tasks of computer is called system software.

For example:

- Compiler
- Operating system
- Interpreter
- Linker
- Loader

Utility Software

The software, which provide an additional meaning to the computer system.

For Example:-

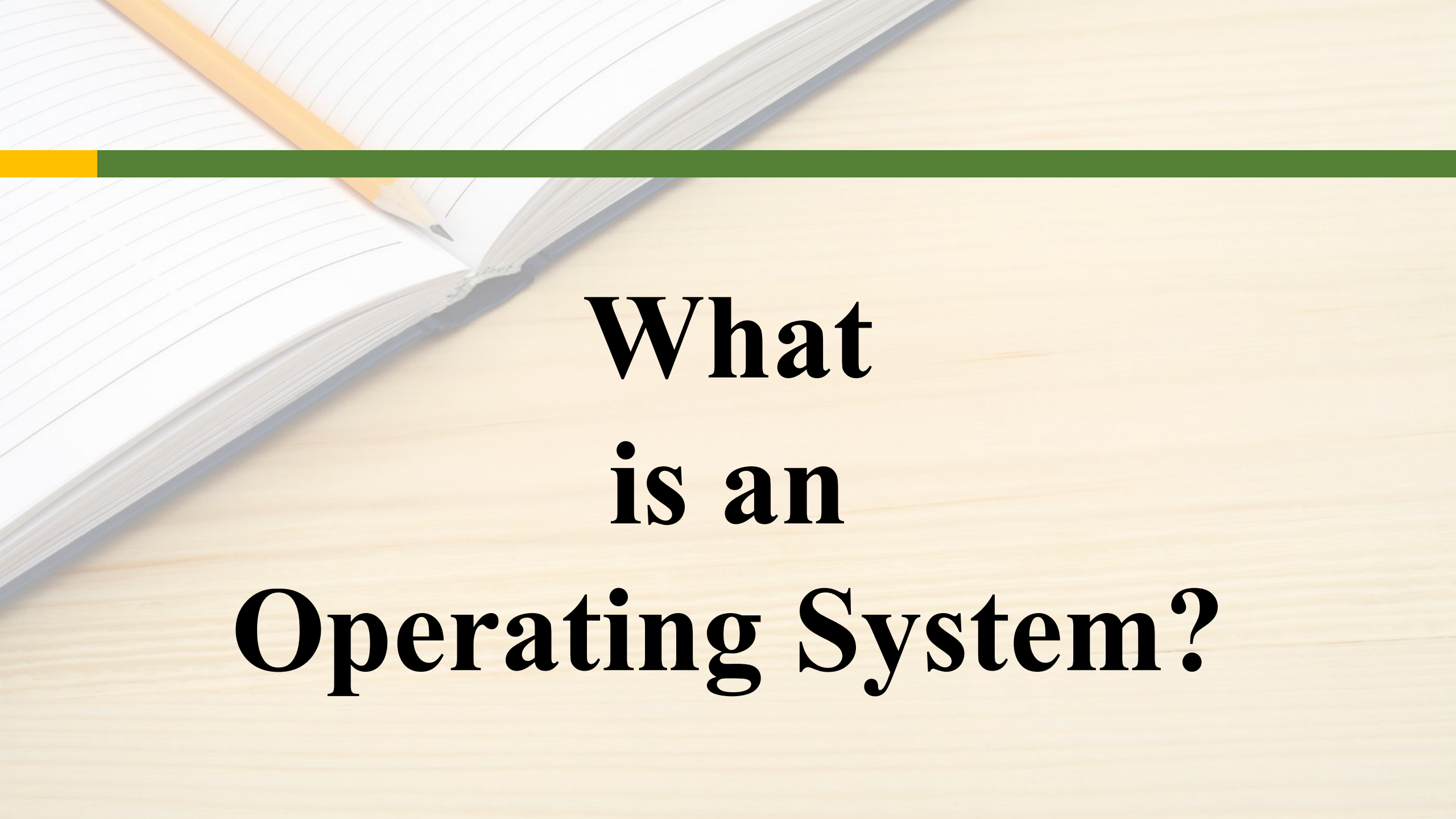
- Calculator
- MS-paint
- Browser
- Notepad
- Media Player

Application Software

The software which is created by users, using the different high level language and database system for any special purpose.

For Example:-

- Library Management system
- Banking Software
- Ticket Reservation system



**What
is an
Operating System?**

Definition

1. An operating system (OS) is a collection of system software that manages computer hardware resources and provides common services for computer programs.
2. A program that acts as an **intermediary/interface** between a user of a computer and the computer hardware.

Operating System

**Users and Processes
access the Computer's
resources through the
Operating System**

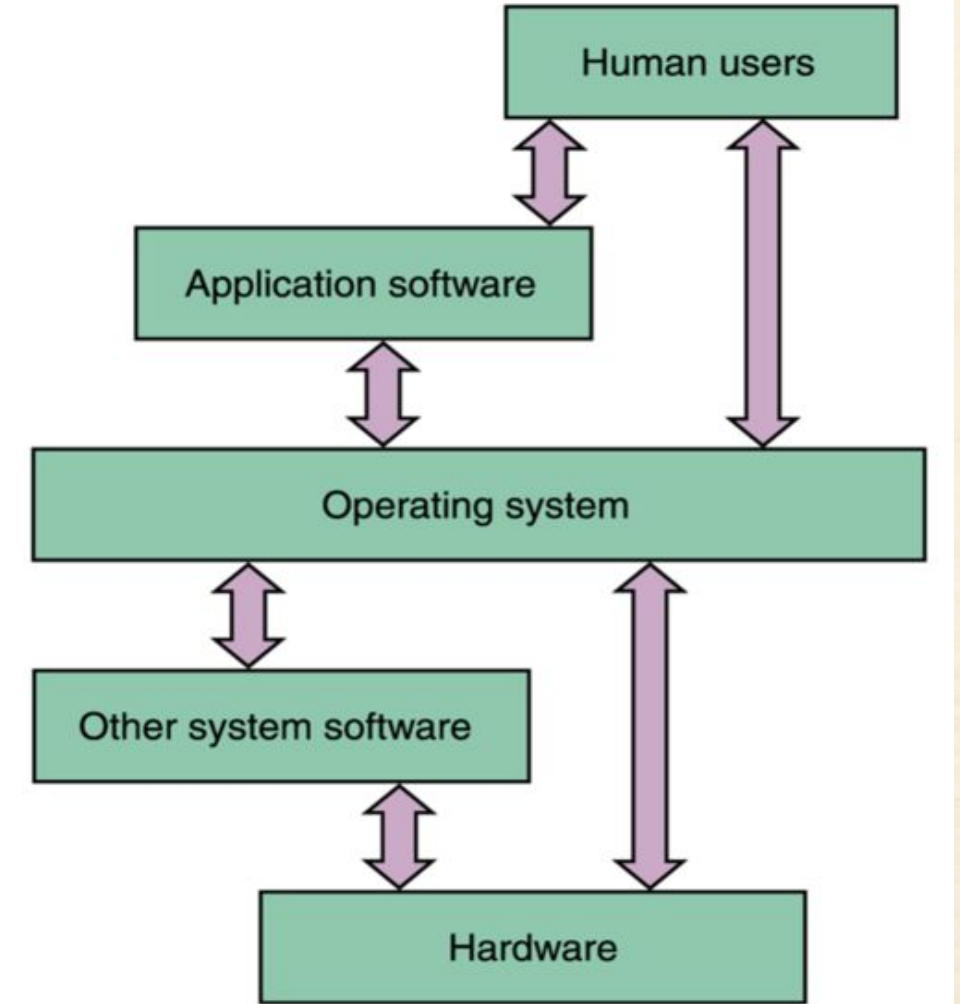


Figure – 1 Operating System Approach [6]

Goals of an Operating System

- Simplify the execution of user programs and make solving user problems easier
- Use computer hardware efficiently
 - Allow sharing of hardware and software resources.
- Make application software portable and flexible
- Provide isolation, security and protection among user programs
- Improve overall system reliability
 - Error confinement, Fault tolerance, Reconfiguration.

History/Generations of Operating Systems^[7]

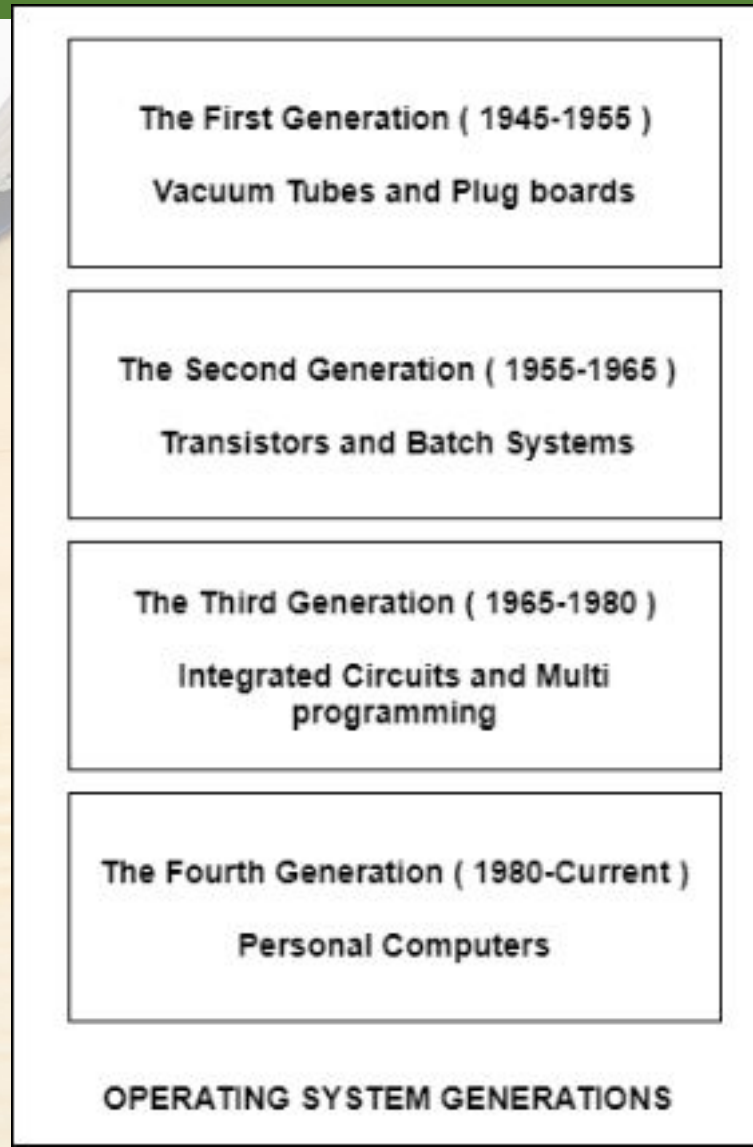


Figure – 2 Generations of OS [7]

The First Generation (1945 - 1955): Vacuum Tubes and Plugboards

- Digital computers were not constructed until the second world war. Calculating engines with mechanical relays were built at that time. However, the mechanical relays were very slow and were later replaced with vacuum tubes. These machines were enormous but were still very slow.



Figure – 3 Vacuum Tubes

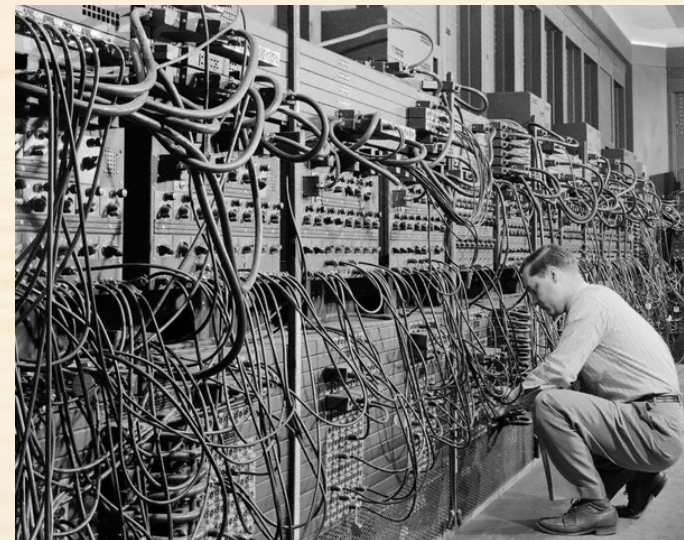


Figure – 4 1940s computers

The Second Generation(1955 - 1965): Transistors & Batch Systems

- Transistors led to the development of the computer systems that could be manufactured and sold to paying customers. These machines were known as mainframes and were locked in air-conditioned computer rooms with staff to operate them.
- The Batch System was introduced to reduce the wasted time in the computer. A tray full of jobs was collected in the input room and read into the magnetic tape.

The Third Generation (1965-1980): Integrated Circuits & Multiprogramming

- Until the 1960's, there were two types of computer systems i.e the scientific and the commercial computers. These were combined by IBM in the System/360. This used integrated circuits and provided a major price and performance advantage over the second generation systems.
- The third generation operating systems also introduced multiprogramming. This meant that the processor was not idle while a job was completing its I/O operation. Another job was scheduled on the processor so that its time would not be wasted.

The Fourth Generation (1980-Present): Personal Computers

- Personal Computers were easy to create with the development of large-scale integrated circuits. These were chips containing thousands of transistors on a square centimetre of silicon.
- The advent of personal computers also led to the growth of networks. This created network operating systems and distributed operating systems. The users were aware of a network while using a network operating system and could log in to remote machines and copy files from one machine to another.

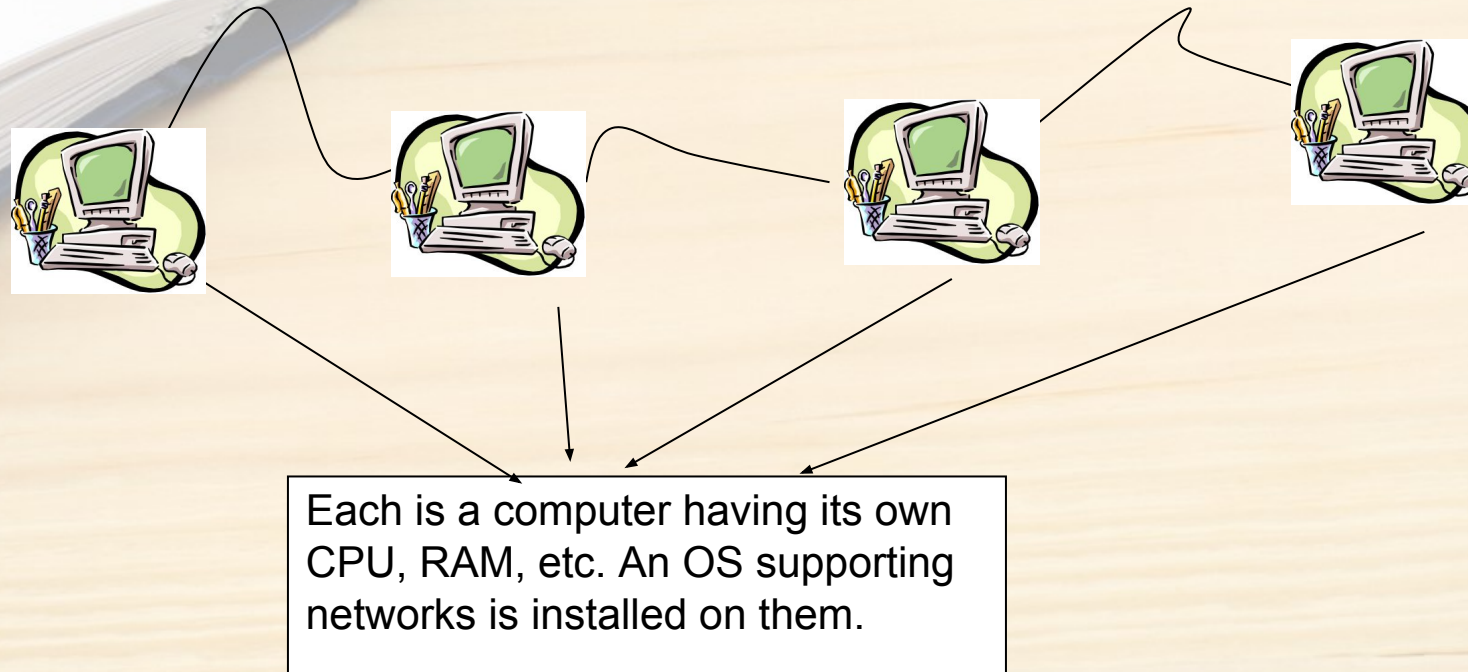


Figure – 6 Distributed OS

Microsoft Windows

- Microsoft created the Windows operating system in the mid-1980s.
- Most recent versions are Windows 10 (released in 2015), Windows 8 (2012), Windows 7 (2009), and Windows Vista (2007).
- Windows comes pre-loaded on most new PCs, which helps to make it the most popular operating system in the world.

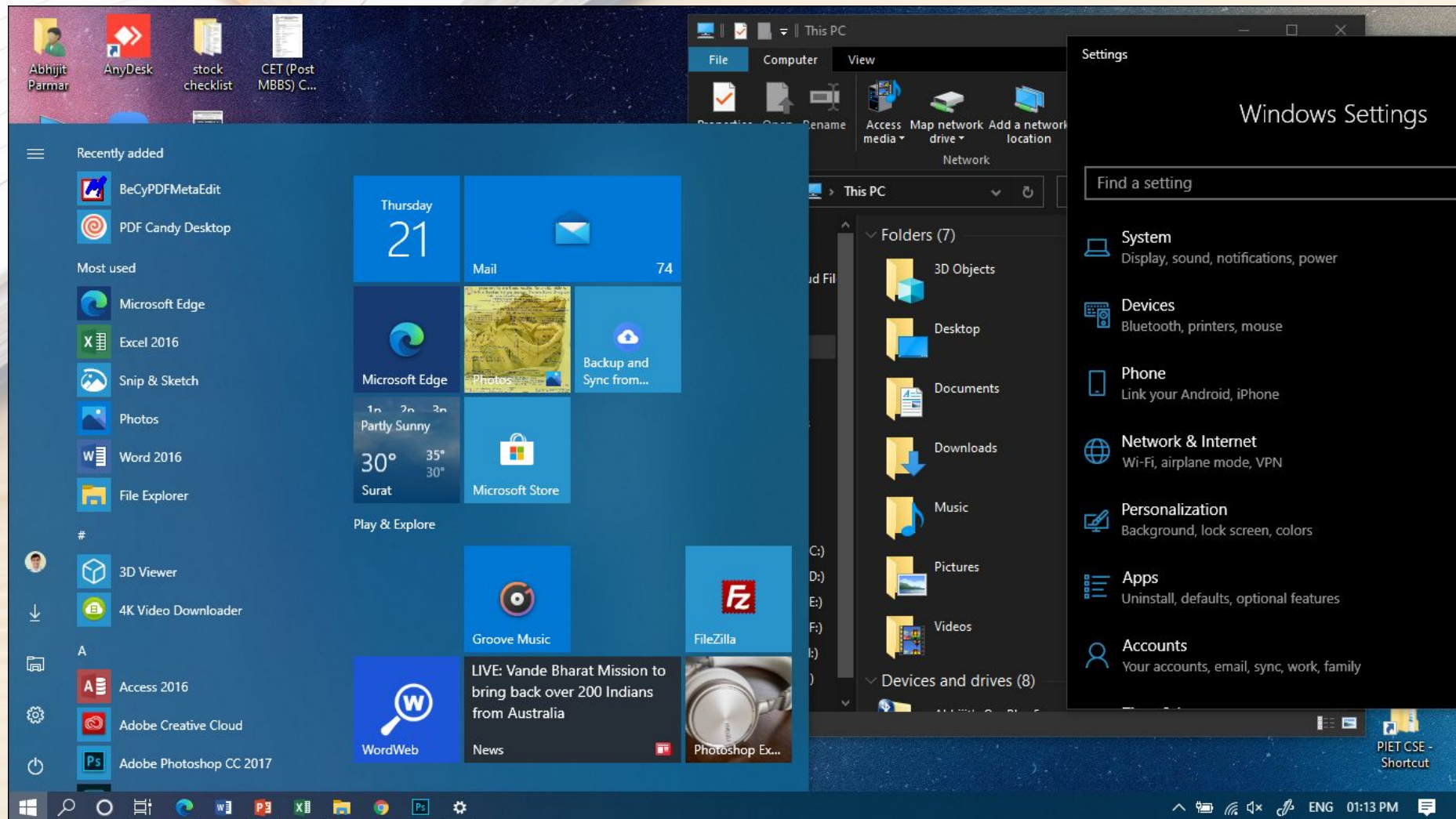


Figure – 7 Windows 10

Mac OS X

- Mac OS is a line of operating systems created by Apple.
- It comes preloaded on all new Macintosh computers, or Macs.
- Specific versions include El Capitan (released in 2015), Yosemite (2014), Mavericks (2013), Mountain Lion (2012), and Lion (2011).



Figure – 8 macOS

Linux

- Linux (pronounced LINN-ux) is a family of open-source operating systems, which means they can be modified and distributed by anyone around the world.
- The advantages of Linux are that it is free, and there are many different distributions or versions you can choose from.

linux



Figure – 9 Ubuntu OS

Operating systems for mobile devices

- Mobile devices such as smartphones, tablets, and MP3 players are different from desktop and laptop computers, so they run operating systems that are designed specifically for mobile devices.
- Examples of mobile OS - **Apple iOS and Google Android.**

Goals of OS

A background image showing a wooden desk with a spiral-bound notebook and a yellow pencil resting on it. The notebook is open, showing lined pages. A green horizontal bar is positioned below the title.

- Provide Convenience to user (users can access H/W easily)
- Throughput (No. of task performs per unit time)

Functionality of OS

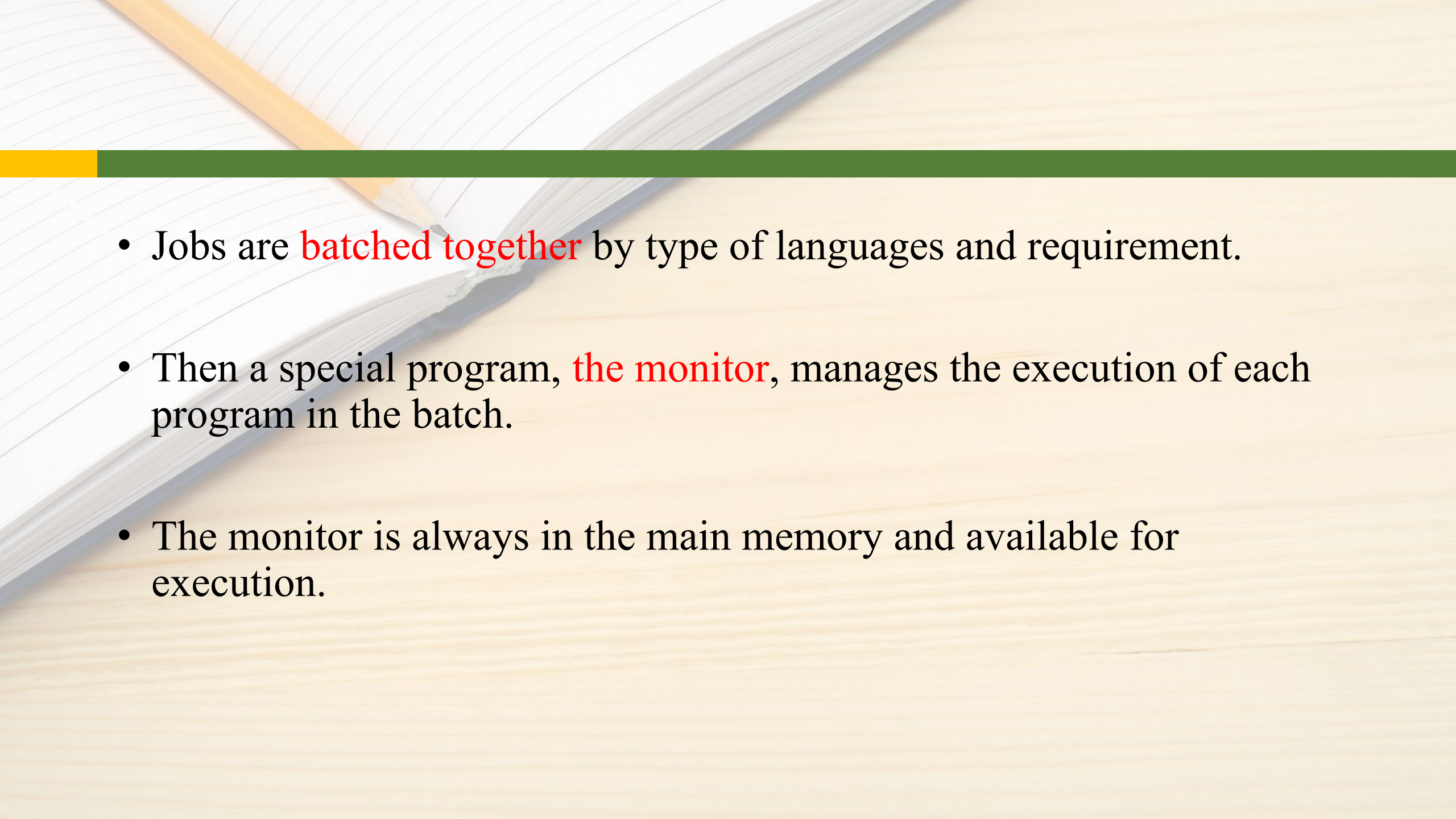
1. Resource Management
(Copy, store and process data)
2. Process Management
(No. of different process run by user using CPU scheduling algorithm)
3. Storage Management
(How to store permanent data on hard disk)
4. Memory Management
(Allocation and de-allocation of memory after execution of process)
5. Security
(Authentication: not only for user but also for process)

Types of Operating Systems on the basis of Functionality

1. Simple Batch System
2. Multiprogramming Batch System
3. Multitasking system
4. Multiprocessor System
5. Distributed Operating System
6. Real-time Operating System

SIMPLE BATCH SYSTEMS

- No direct interaction between user and computer.
- The user has to submit a job (written on cards or tape) to a computer operator.
- Then computer operator places a batch of several jobs on an input device.

- 
- Jobs are **batched together** by type of languages and requirement.
 - Then a special program, **the monitor**, manages the execution of each program in the batch.
 - The monitor is always in the main memory and available for execution.

Cont.

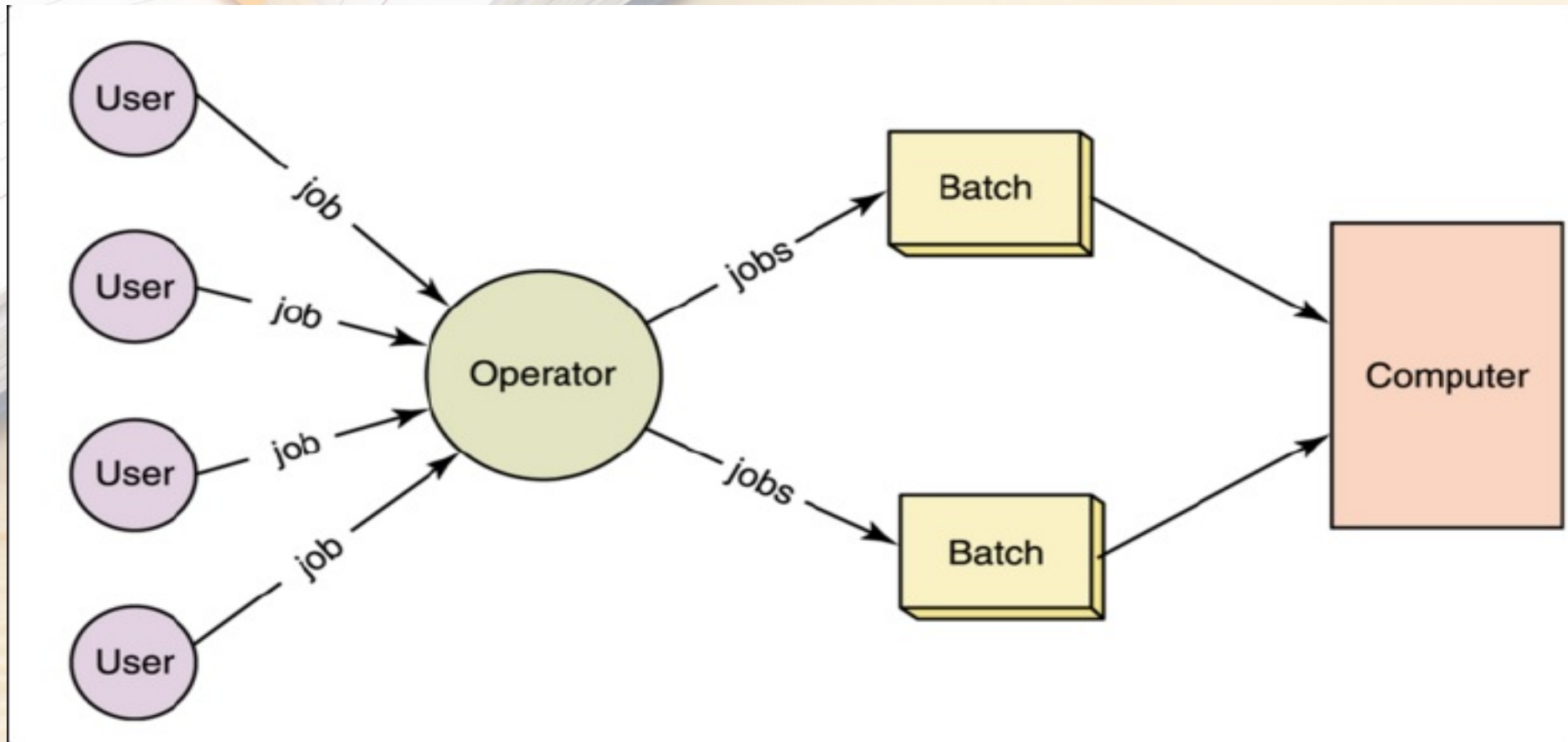


Figure – 10 Batch OS



□ *Advantages*

- Increased performance - next job start as the previous job finished.
- Suitable for executing large jobs that need little interaction

□ *Disadvantages*

- **Zero interaction** between user and computer.
- No mechanism to **prioritize** processes.

Multiprogramming Batch Operating System

- Several jobs are kept in main memory at the same time, and the CPU is multiplexed among them.

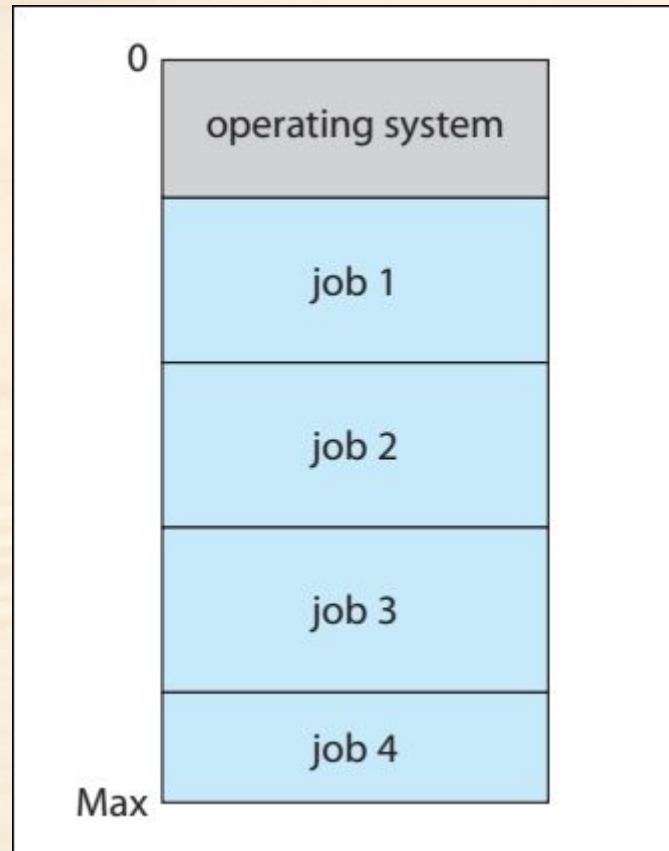
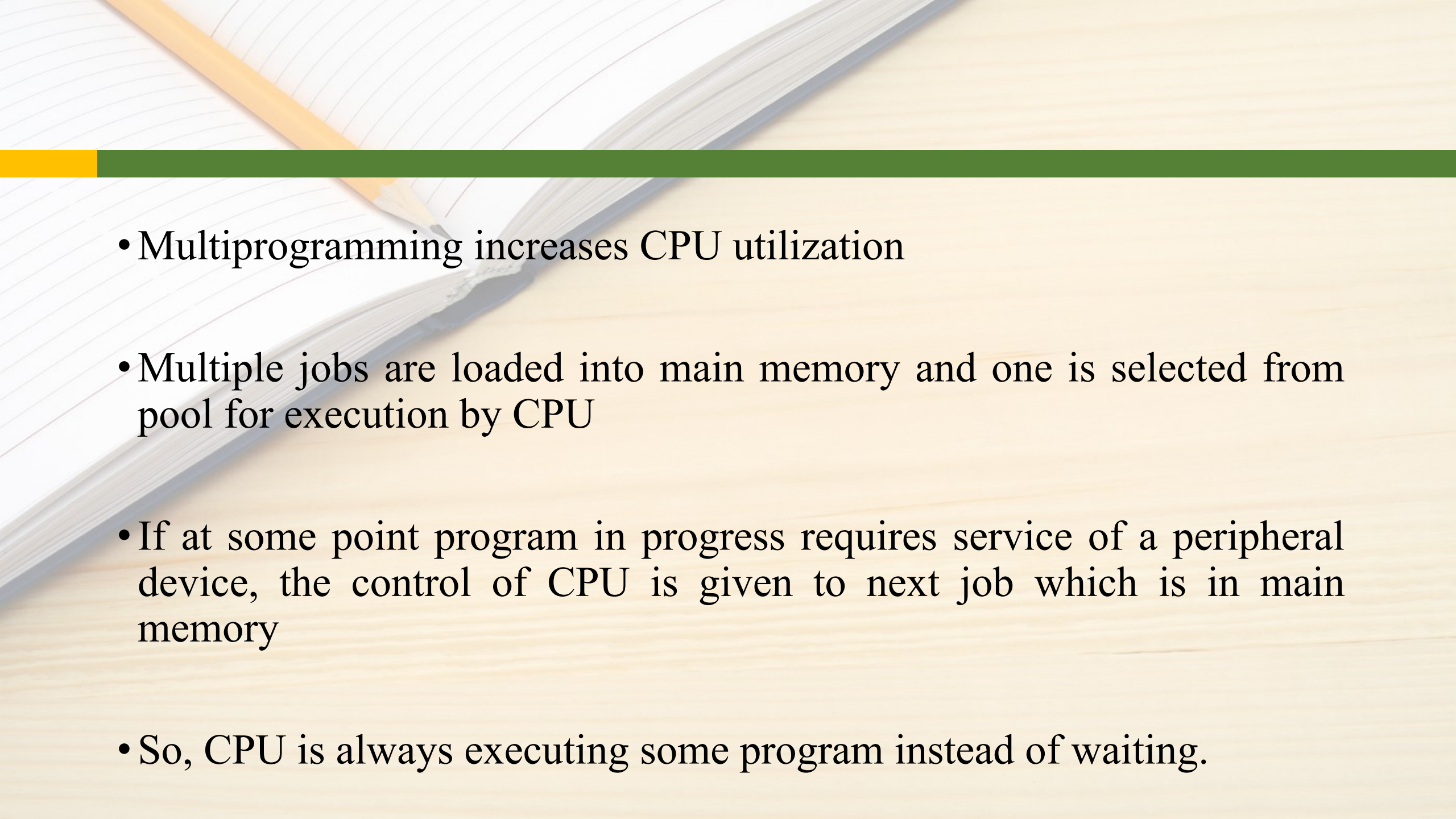
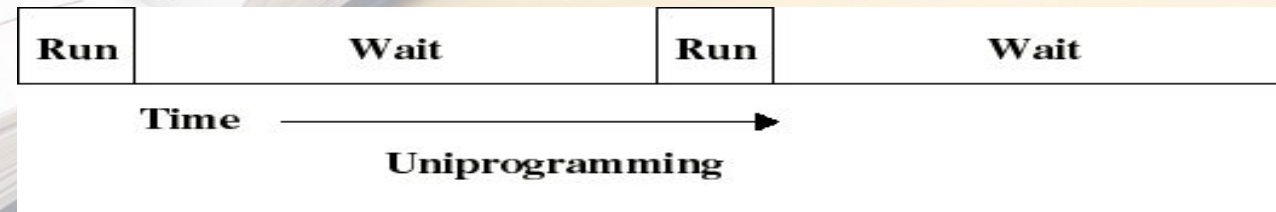


Figure – 11 Memory layout for a multiprogramming system [1]

- 
- Multiprogramming increases CPU utilization
 - Multiple jobs are loaded into main memory and one is selected from pool for execution by CPU
 - If at some point program in progress requires service of a peripheral device, the control of CPU is given to next job which is in main memory
 - So, CPU is always executing some program instead of waiting.

- CPU usage is poor when only one program is present in memory



- If memory can hold several programs, then CPU can switch to another one whenever a program is awaiting for an I/O to complete. This is multitasking (multiprogramming)

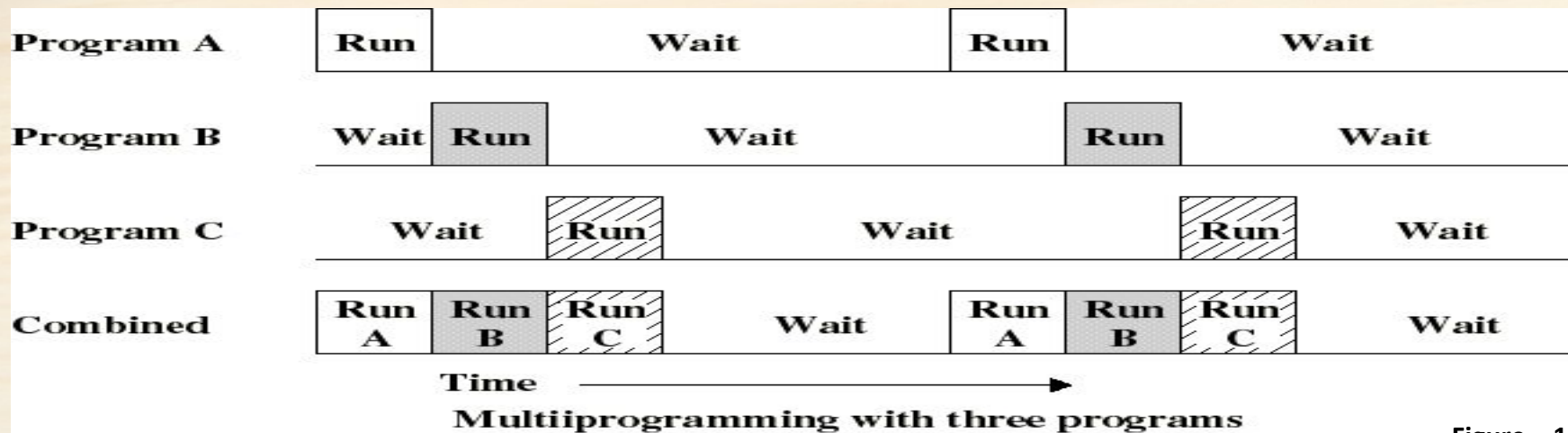


Figure – 12 Multiprogramming ([Source](#))

Effects of Multiprogramming

	Uniprogramming	Multiprogramming
Processor use	17%	33%
Memory use	30%	67%
Disk use	33%	67%
Printer use	33%	67%
Elapsed time	30 min.	15 min.
Throughput rate	6 jobs/hr	12 jobs/hr
Mean response time	18 min.	10 min.

Figure – 13 Effects of Multiprogramming ([Source](#))

The background of the slide features a close-up, slightly blurred image of an open notebook with lined pages. Two pencils, one yellow and one grey, are resting on the pages. A solid green horizontal bar is positioned across the upper portion of the image, partially obscuring the notebook. Below this bar, the text for the 'Advantages' section is displayed.

□ *Advantages*

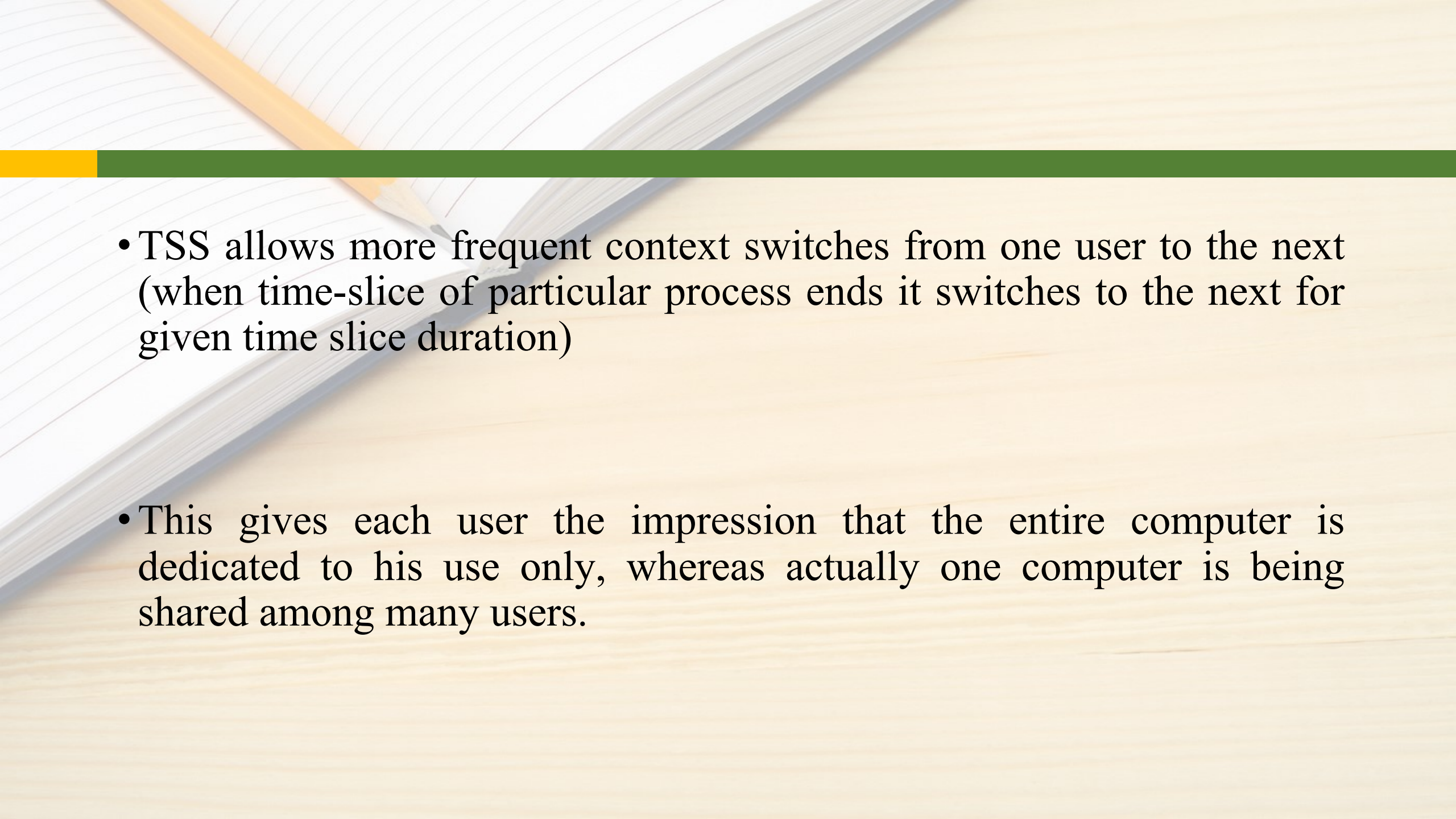
- High CPU utilization, so CPU never sits idle, if there are jobs available
- Many programs are allotted CPU almost simultaneously.
- Provides better resource utilization (Memory, I/O, CPU)
- More than one process can be executed simultaneously by user.

□ *Disadvantages*

- CPU scheduling is required.
- Memory management is required, to accommodate many jobs in memory
- Multiprogramming does not support interaction with users

Multitasking/Time Sharing System(TSS)

- Multiprogramming does not support interaction with users, TSS extends multiprogramming to handle multiple interactive jobs
- TSS uses CPU scheduling & multiprogramming to provide economical interactive systems of two or more users.
- Each user is given a time-slice for executing his job in Round-Robin Fashion (Every process will be given equal amount of CPU one by one in sequence). Job continues until the time slice ends.
- The CPU is multiplexed among several jobs that are kept in main memory.

- 
- TSS allows more frequent context switches from one user to the next (when time-slice of particular process ends it switches to the next for given time slice duration)
 - This gives each user the impression that the entire computer is dedicated to his use only, whereas actually one computer is being shared among many users.



□ *Advantages*

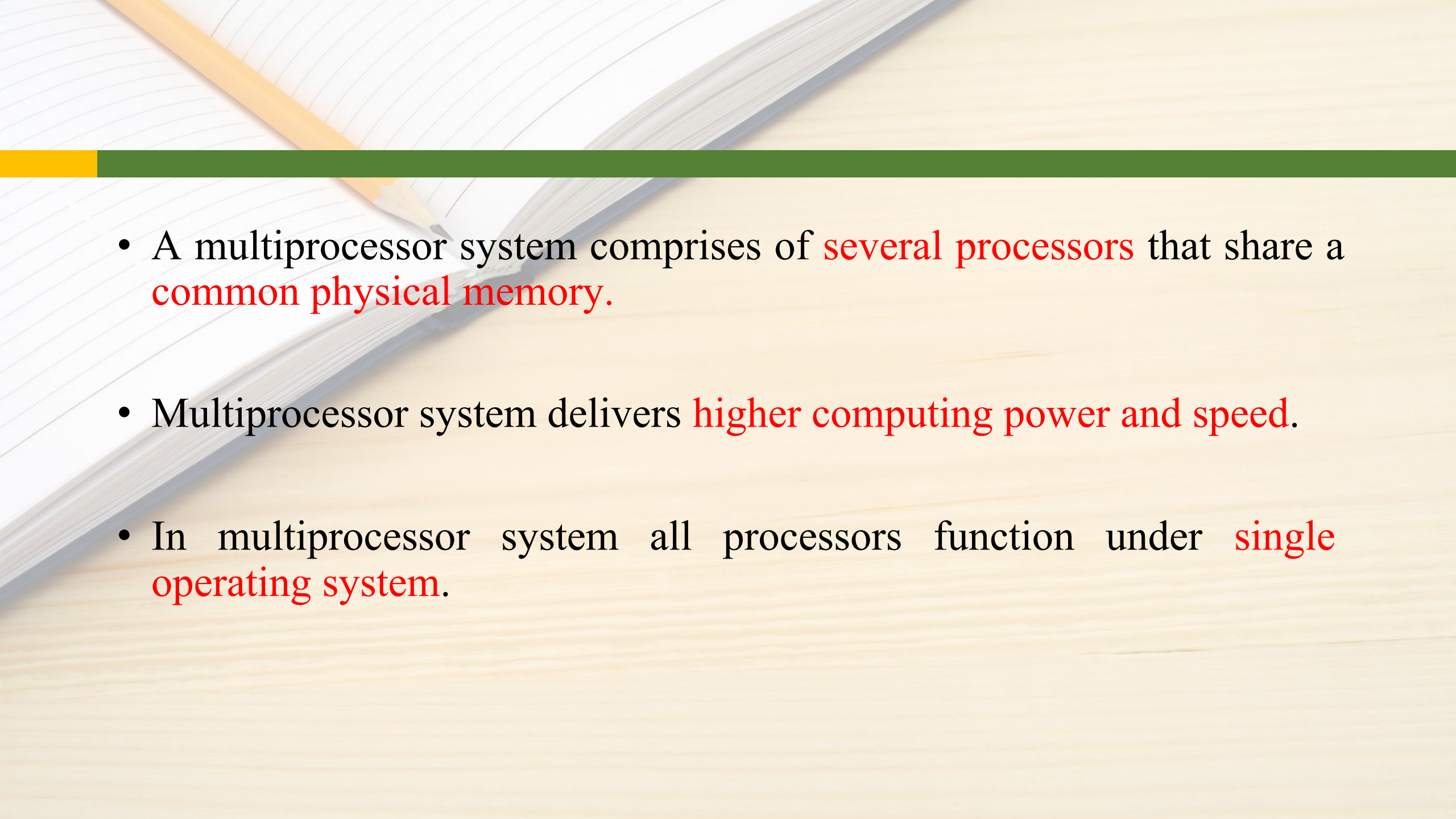
- Provides Quick Response
- Reduces CPU idle time

□ *Disadvantages*

- Security & Integrity of user's program & data is needed.
- If lots of users & applications are running then it may hang up the system. So, high configuration of hardware is required.

Multiprocessor/Parallel System

- **Multiprocessor systems** with more than one CPU works in close communication.
- **Tightly coupled system** – processors share memory and I/O devices, bus, system and communication usually takes place through the shared memory.

- 
- A multiprocessor system comprises of **several processors** that share a **common physical memory**.
 - Multiprocessor system delivers **higher computing power and speed**.
 - In multiprocessor system all processors function under **single operating system**.

A background image showing a yellow pencil resting on an open notebook with lined pages. A green horizontal bar is positioned across the middle of the image, partially obscuring the notebook and pencil.

□ *Advantages*

- **Increased *throughput*:** No. of jobs executed per unit time increased as there are more no. of processors.
- **Economical:** Buying one system with 3 CPU is cheaper than 3 systems with 3 different CPUs. The processors can share peripherals, cabinets and power supplies.
- **Increased reliability:** The failure of one processor will not stop the system, if functions with other available processors

Real Time System

- A real-time operating system (RTOS) promises a certain capability within a specified time constraint.
- It is defined as an operating system known to give maximum time for each of the critical operations that it performs, like OS calls and interrupt handling.

The background of the slide features a close-up, slightly blurred image of an open notebook with lined pages and a yellow pencil resting on it. A solid green horizontal bar is positioned across the upper portion of the image, partially obscuring the notebook. Below this bar, the text is presented in a clean, black serif font.

□ Hard real-time system

- The Real-Time Operating system which guarantees the **maximum time for critical operations** and complete them on time are referred to as **Hard Real-Time Operating Systems**.
- If the system fails to meet the deadline even once the system is considered to have Failed.

E.g. Defense applications, nuclear system etc. Missing deadlines creates hazards.

The background of the slide features a close-up, slightly blurred image of an open notebook with lined pages and a yellow pencil resting on it. A solid green horizontal bar is positioned across the upper portion of the image, partially obscuring the notebook. Below this bar, on the left side, is a small yellow rectangular element.

□ *Soft real-time system*

- The critical task will get priority over other tasks, but no assurity of completing it in a defined time. These systems are referred to as **Soft Real-Time Operating Systems**.
- It is less restrictive type of OS. even if the system fails to meet the deadline, the system is not considered to have failed. In this case the results of the requests are not worthless.

e.g. audio-video streaming etc.

Distributed System

- Distribute the computation among several physical processors.
- Distributed OS is an OS that runs on several machines and it controls the resources of several machines.
- ***Loosely coupled system*** – each processor has its own local memory; processors communicate with one another through various communications lines, such as high-speed buses or telephone lines.

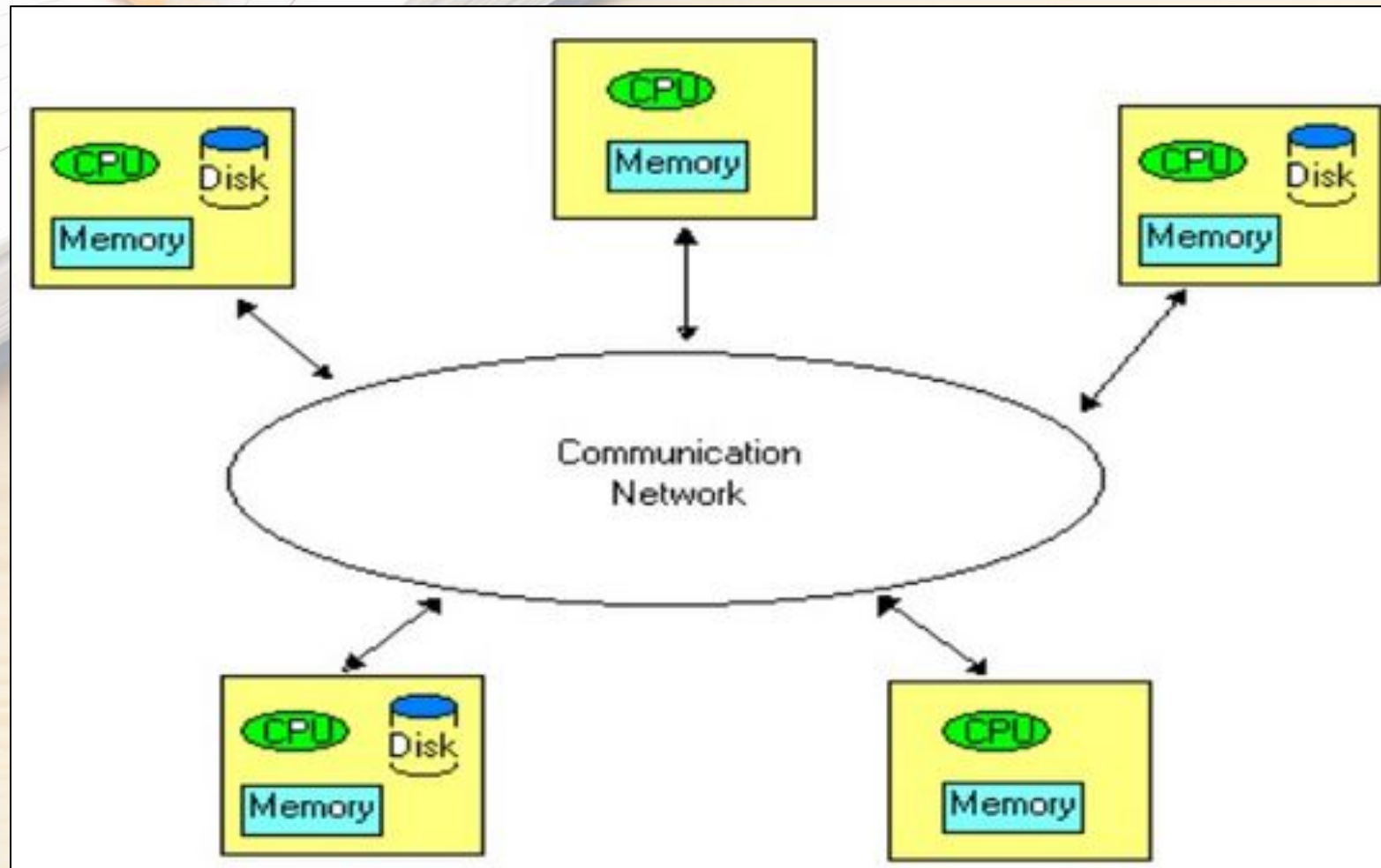


Figure – 14 Distributed OS ([Source](#))



□ *Advantages*

- Resources Sharing
- Computation speed up due to load sharing . So, Short response time and higher throughput.
- Higher Reliability: Degree of tolerance against failure
- Incremental Growth : to extend functionality of a system by simply adding additional resources to the system

Operating System Services^[1]

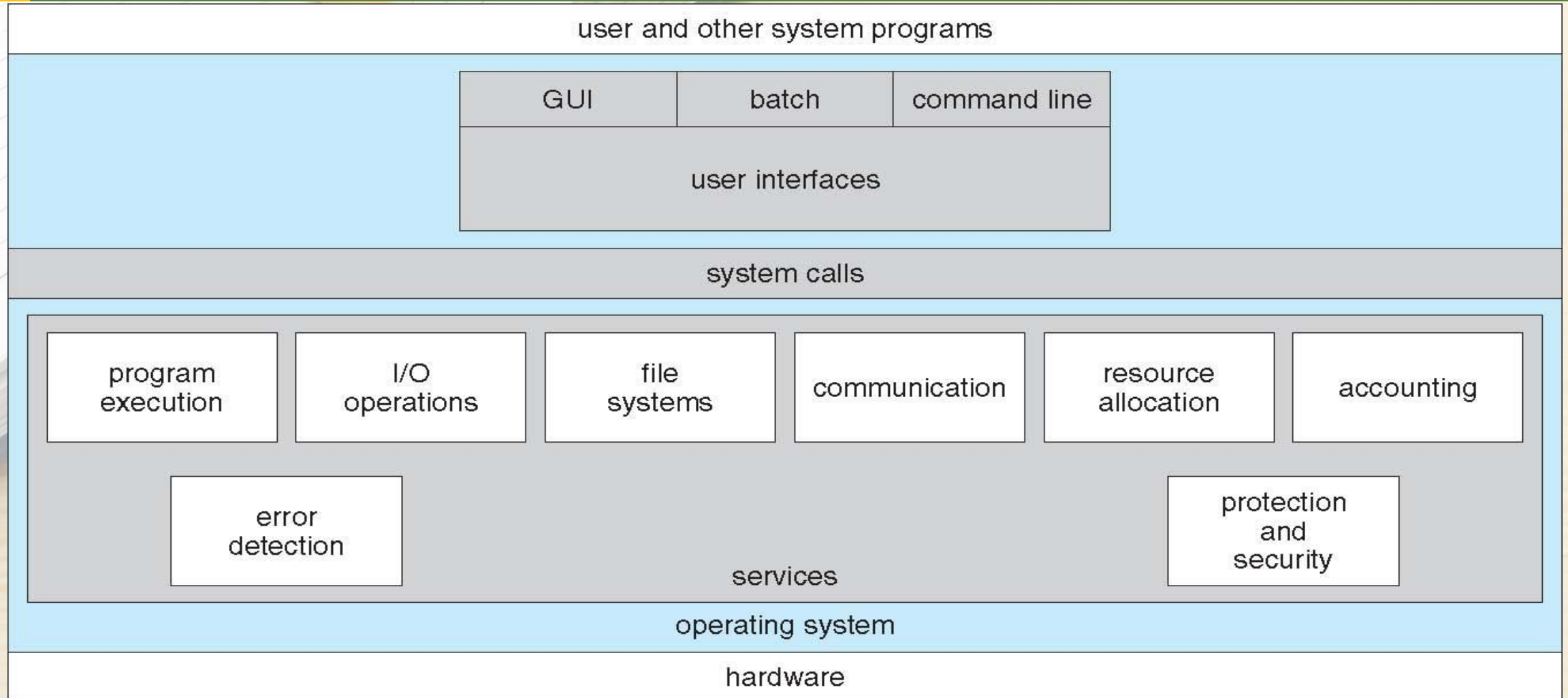


Figure – 15 A view of operating system services [1]

Operating System Services

- 1. User Interface:** Almost all operating systems have a user interface (UI). Varies between **Command-Line (CLI)**, **Graphics User Interface (GUI)**, **Batch Interfaces**
- 2. Program execution:** The system must be able to load a program into memory and to run that program, must be able to end execution, either normally or abnormally (indicating error)
- 3. I/O operations** – A running program may require I/O, which may involve a file or an I/O device, since user programs cannot execute I/O operations directly, the operating system must provide some means to do I/O.

Operating System Services

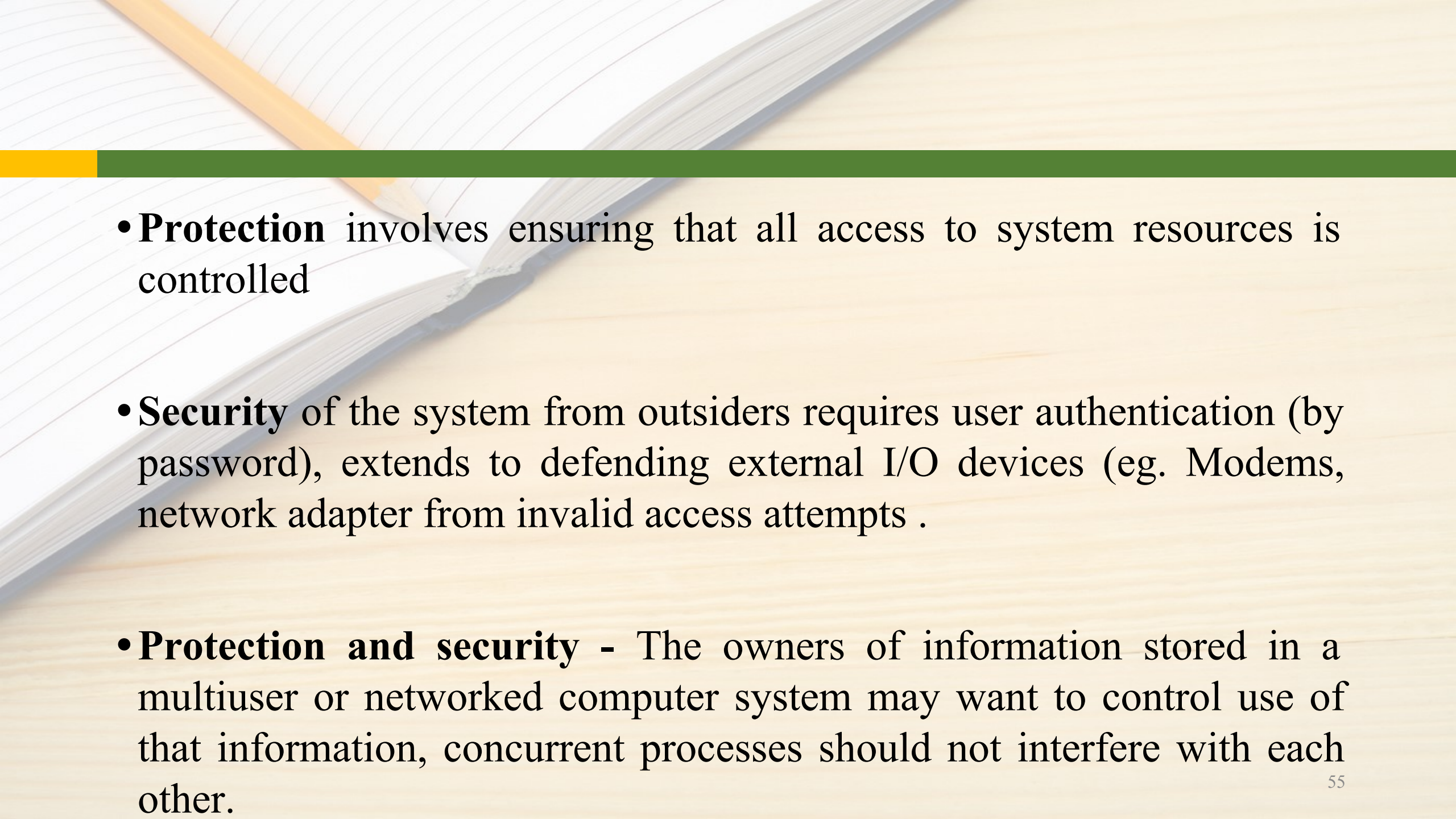
4. **File-system manipulation** – Programs need to read and write files and directories, create and delete them, search them, list file Information, permission management; allow or deny access to files/directories based on file ownership.
5. **Communications** – Exchange of information between processes executing either on the same computer or on different systems tied together by a network. Implemented via *shared memory* or *message passing*.

Operating System Services

- 6. Error detection** – OS needs to be constantly aware of possible errors
- May occur in the CPU and memory hardware, in I/O devices, in user program
 - For each type of error, OS should take the appropriate action to ensure correct and consistent computing
 - Debugging facilities can greatly enhance the user's and programmer's abilities to efficiently use the system

Some Additional OS Services

- **Resource allocation** - When multiple users or multiple jobs running concurrently, resources must be allocated to each of them. Many types of resources - Some (such as CPU cycles, main memory, and file storage) may have special allocation code, others (such as I/O devices) may have general request and release code.
- **Accounting** - To keep track of which users use how much and what kinds of computer resources. Used for accounting or usage statistics.

- 
- **Protection** involves ensuring that all access to system resources is controlled
 - **Security** of the system from outsiders requires user authentication (by password), extends to defending external I/O devices (eg. Modems, network adapter from invalid access attempts .
 - **Protection and security** - The owners of information stored in a multiuser or networked computer system may want to control use of that information, concurrent processes should not interfere with each other.

System Calls

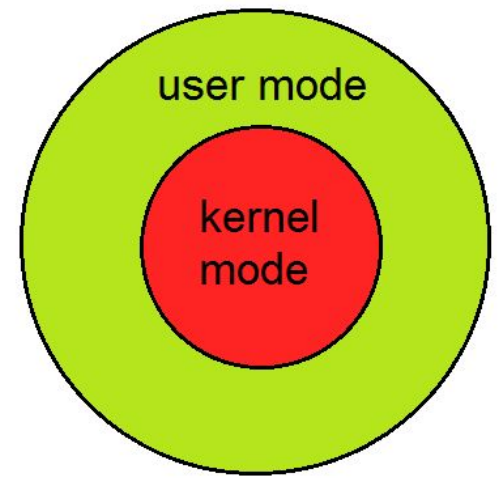


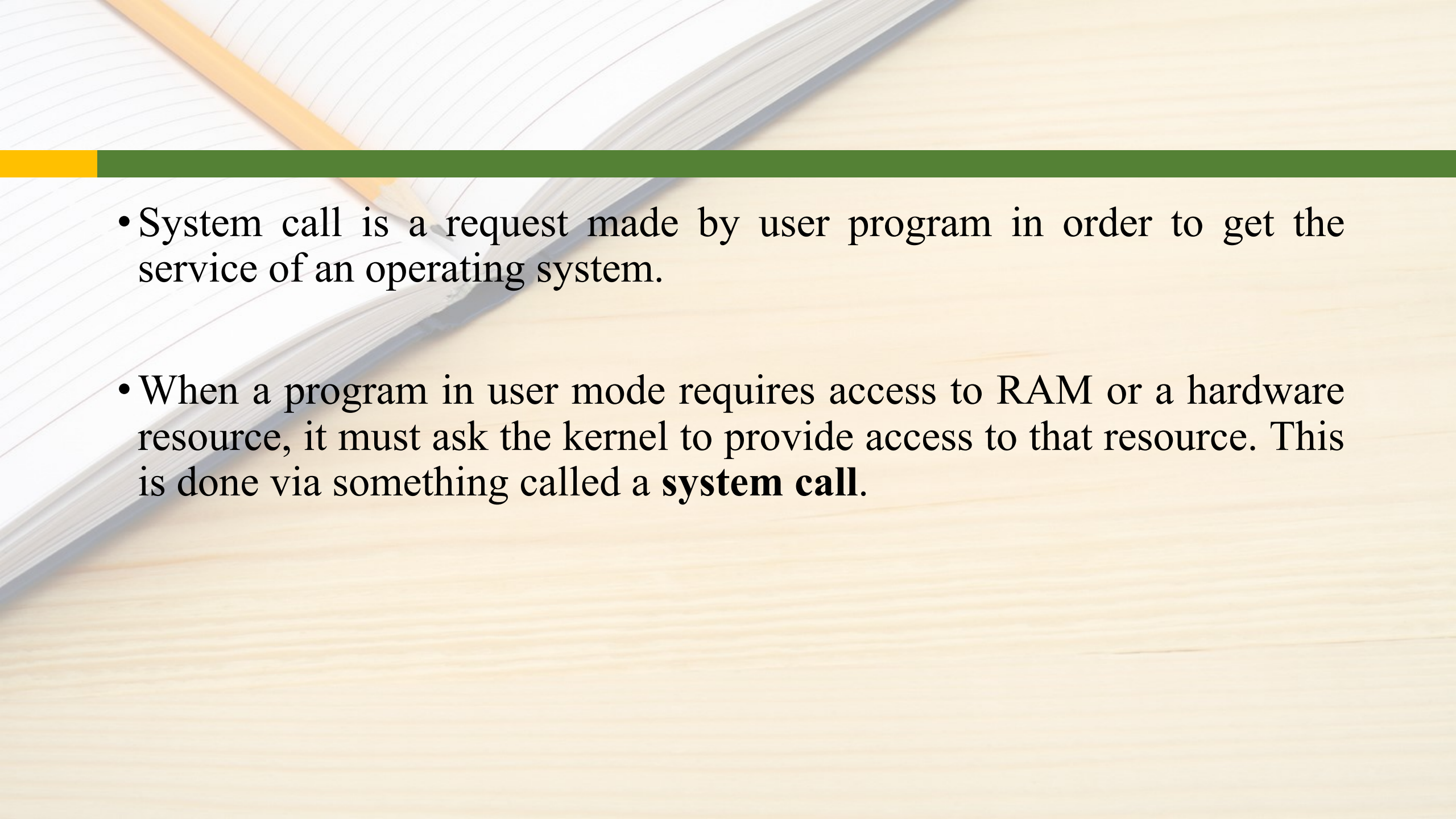
Figure – 16 OS modes ([Source](#))

□ Kernel Mode

- When CPU is in **kernel mode**, the code being executed can access any memory address and any hardware resource.
- Hence kernel mode is a very privileged and powerful mode.
- If a program crashes in kernel mode, the entire system will be halted.

□ User Mode

- When CPU is in **user mode**, the programs don't have direct access to memory and hardware resources.
- In user mode, if any program crashes, only that particular program is halted. That means the system will be in a safe state even if a program in user mode crashes.
- Hence, most programs in an OS run in user mode.

- 
- System call is a request made by user program in order to get the service of an operating system.
 - When a program in user mode requires access to RAM or a hardware resource, it must ask the kernel to provide access to that resource. This is done via something called a **system call**.

Some System Calls For Process Management

Process management

Call	Description
<code>pid = fork()</code>	Create a child process identical to the parent
<code>pid = waitpid(pid, &statloc, options)</code>	Wait for a child to terminate
<code>s = execve(name, argv, environp)</code>	Replace a process' core image
<code>exit(status)</code>	Terminate process execution and return status

Figure – 17 System Call [3]

Some System Calls For File Management

File management

Call	Description
<code>fd = open(file, how, ...)</code>	Open a file for reading, writing or both
<code>s = close(fd)</code>	Close an open file
<code>n = read(fd, buffer, nbytes)</code>	Read data from a file into a buffer
<code>n = write(fd, buffer, nbytes)</code>	Write data from a buffer into a file
<code>position = lseek(fd, offset, whence)</code>	Move the file pointer
<code>s = stat(name, &buf)</code>	Get a file's status information

Figure – 18 System Call [3]

Some System Calls For Directory Management

Directory and file system management

Call	Description
<code>s = mkdir(name, mode)</code>	Create a new directory
<code>s = rmdir(name)</code>	Remove an empty directory
<code>s = link(name1, name2)</code>	Create a new entry, name2, pointing to name1
<code>s = unlink(name)</code>	Remove a directory entry
<code>s = mount(special, name, flag)</code>	Mount a file system
<code>s = umount(special)</code>	Unmount a file system

Figure – 19 System Call [3]

Some System Calls For Miscellaneous Tasks

Miscellaneous

Call	Description
<code>s = chdir(dirname)</code>	Change the working directory
<code>s = chmod(name, mode)</code>	Change a file's protection bits
<code>s = kill(pid, signal)</code>	Send a signal to a process
<code>seconds = time(&seconds)</code>	Get the elapsed time since Jan. 1, 1970

Figure – 20 System Call [3]

Operating System layered structure

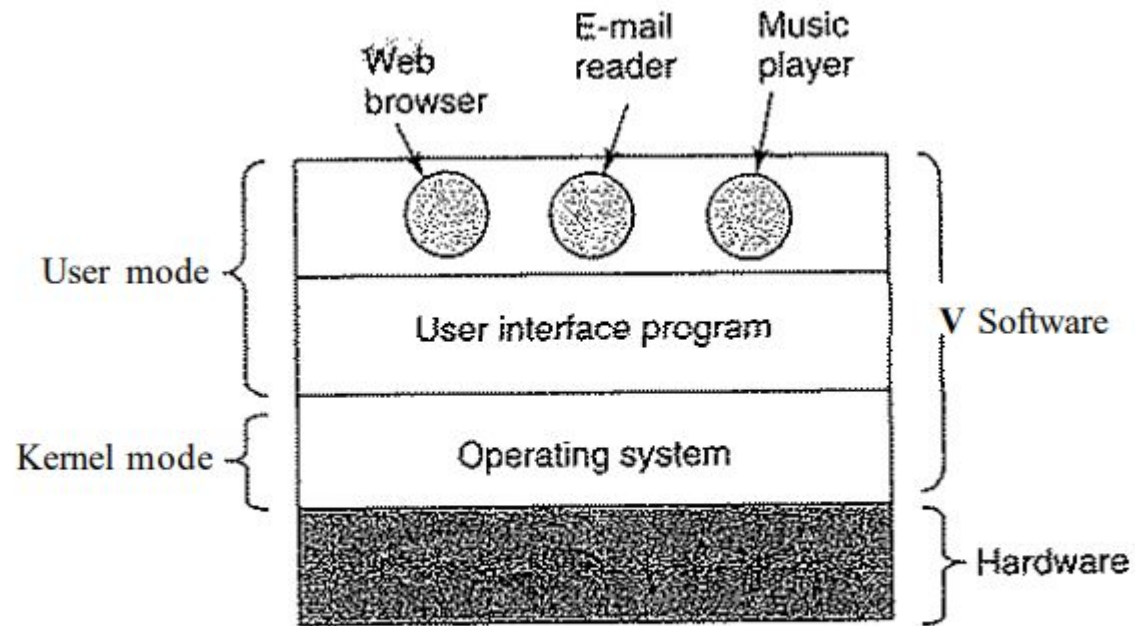
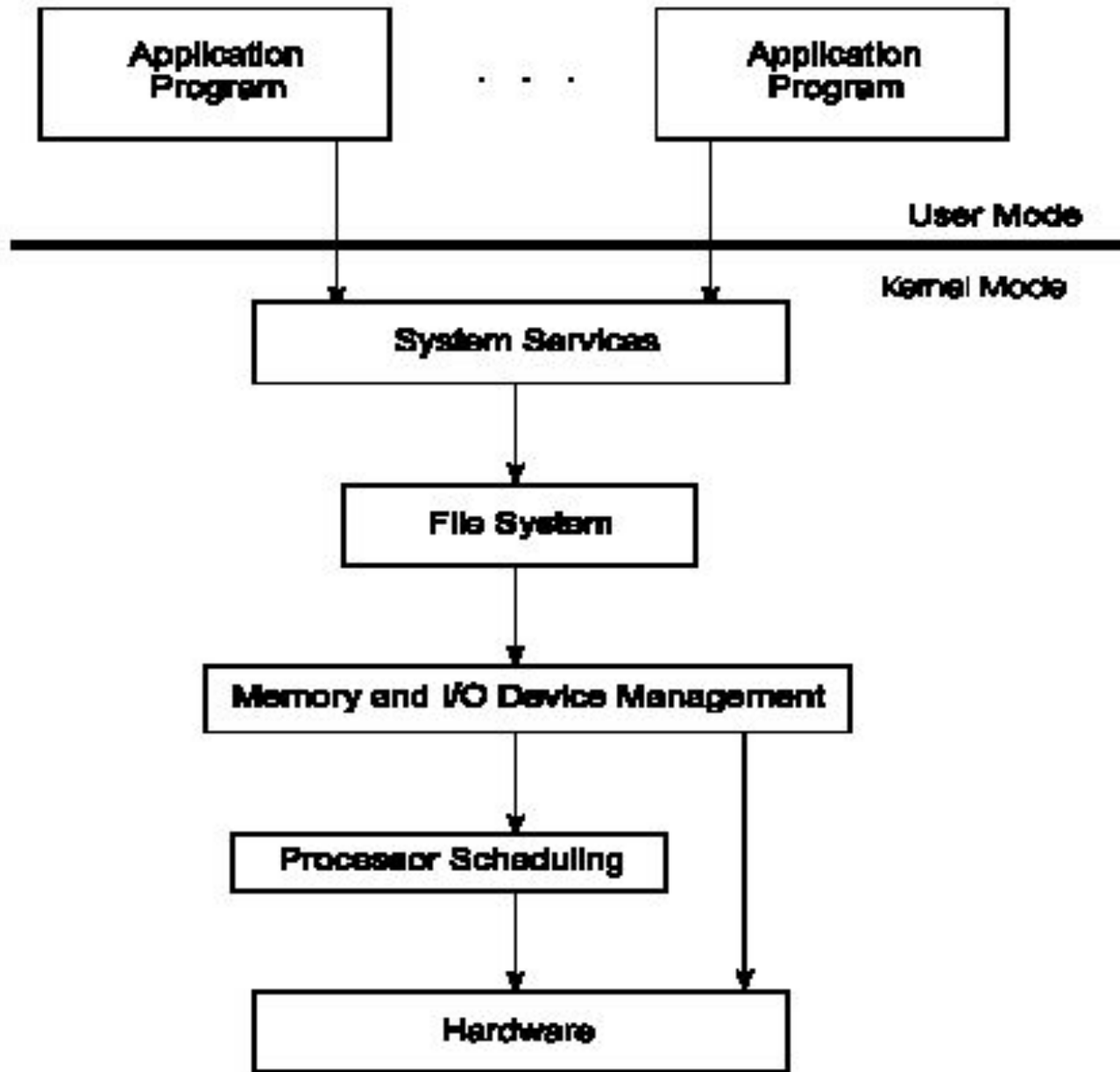


Figure – 21 Where the operating system fits in [3]



- With the layered approach, the bottom layer is the hardware, while the highest layer is the user interface.
- ✓ The main *advantage* is simplicity of construction and debugging.
- ✓ The main *difficulty* is defining the various layers.
- ✓ The main *disadvantage* is that the OS tends to be less efficient than other implementations

Figure – 22 Operating System Layers ([Source](#))

Operating System Structure - Components

- Process Management
- Main Memory Management
- File Management
- I/O System Management
- Secondary Management
- Networking
- Protection System
- Command-Interpreter System

Process Management

- A *process* is a program in execution. A process needs certain resources, including CPU time, memory, files, and I/O devices, to accomplish its task.
- The operating system is responsible for the following activities in connection with process management.
 - Process creation and deletion.
 - process suspension and resumption.
 - Deadlock handling
 - Provision of mechanisms for:
 - process synchronization
 - process communication

Memory Management

- Memory is a large array of words or bytes, each with its own address. It is a repository of quickly accessible data shared by the CPU and I/O devices.
- Main memory is a volatile storage device. It loses its contents in the case of system failure.
- The operating system is responsible for the following activities in connections with memory management:
 - Keep track of which parts of memory are currently being used and by whom.
 - Decide which processes to load when memory space becomes available.
 - Allocate and deallocate memory space as needed.

File Management

- A file is a collection of related information defined by its creator. Commonly, files represent programs (both source and object forms) and data.
- The operating system is responsible for the following activities in connections with file management:
 - File creation and deletion.
 - Directory creation and deletion.
 - Support of primitives for manipulating files and directories.
 - Mapping files onto secondary storage.
 - File backup on stable (nonvolatile) storage media.

I/O System Management

- The I/O system consists of:
 - A buffer-caching system
 - A general device-driver interface
 - Drivers for specific hardware devices

Secondary-Storage Management

- Since main memory (*primary storage*) is volatile and too small to accommodate all data and programs permanently, the computer system must provide *secondary storage* to back up main memory.
- Most modern computer systems use disks as the principle on-line storage medium, for both programs and data.
- The operating system is responsible for the following activities in connection with disk management:
 - Free space management
 - Storage allocation
 - Disk scheduling

Networking

- A *distributed* system is a collection processors that do not share memory or a clock. Each processor has its own local memory.
- The processors in the system are connected through a communication network. Communication takes place using a *protocol*.
- A distributed system provides user access to various system resources.
- Access to a shared resource allows:
 - Computation speed-up
 - Increased data availability
 - Enhanced reliability

Protection

- *Protection* refers to a mechanism for controlling access by programs, processes, or users to both system and user resources.
- The protection mechanism must:
 - Distinguish between authorized and unauthorized usage.
 - Specify the controls to be imposed.
 - Provide a means of enforcement.

Command-Interpreter System

- The program that reads and interprets control statements is called variously:
 - command-line interpreter
 - shell (in UNIX)
- Its function is to get and execute the next command statement.

Monolithic Approach

- Functionality of the OS is activated with simple function calls within the kernel, Monolithic kernel is one large program.
- Device drivers are loaded into the running kernel and become part of the kernel.

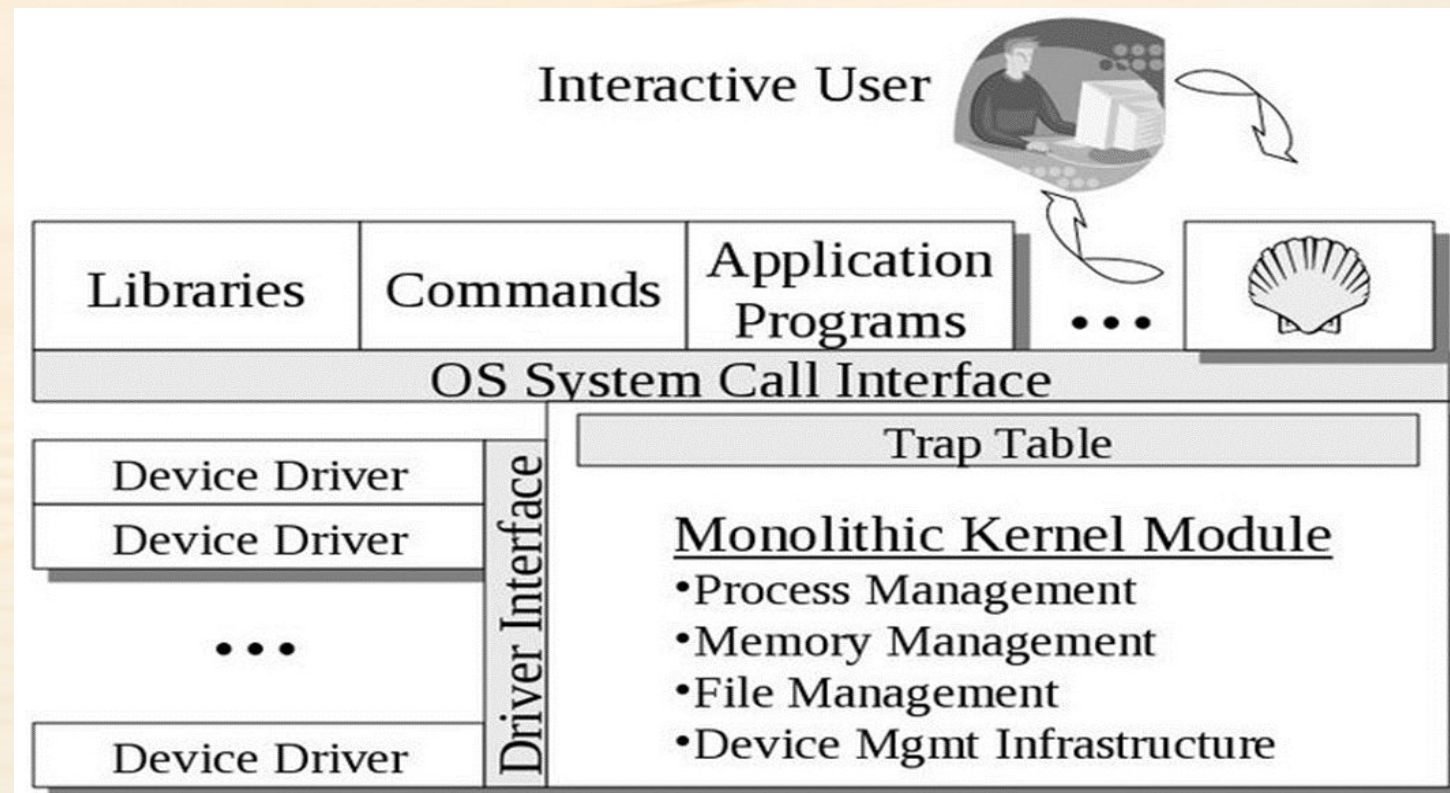


Figure –23 Monolithic Kernel [5]

Microkernel Approach

- Microkernel structures the OS by removing all unnecessary parts of the kernel and implement them as system and user level programs.
- They offers minimal process and memory management, and a communications facility.
- Communication between components is done by message passing.

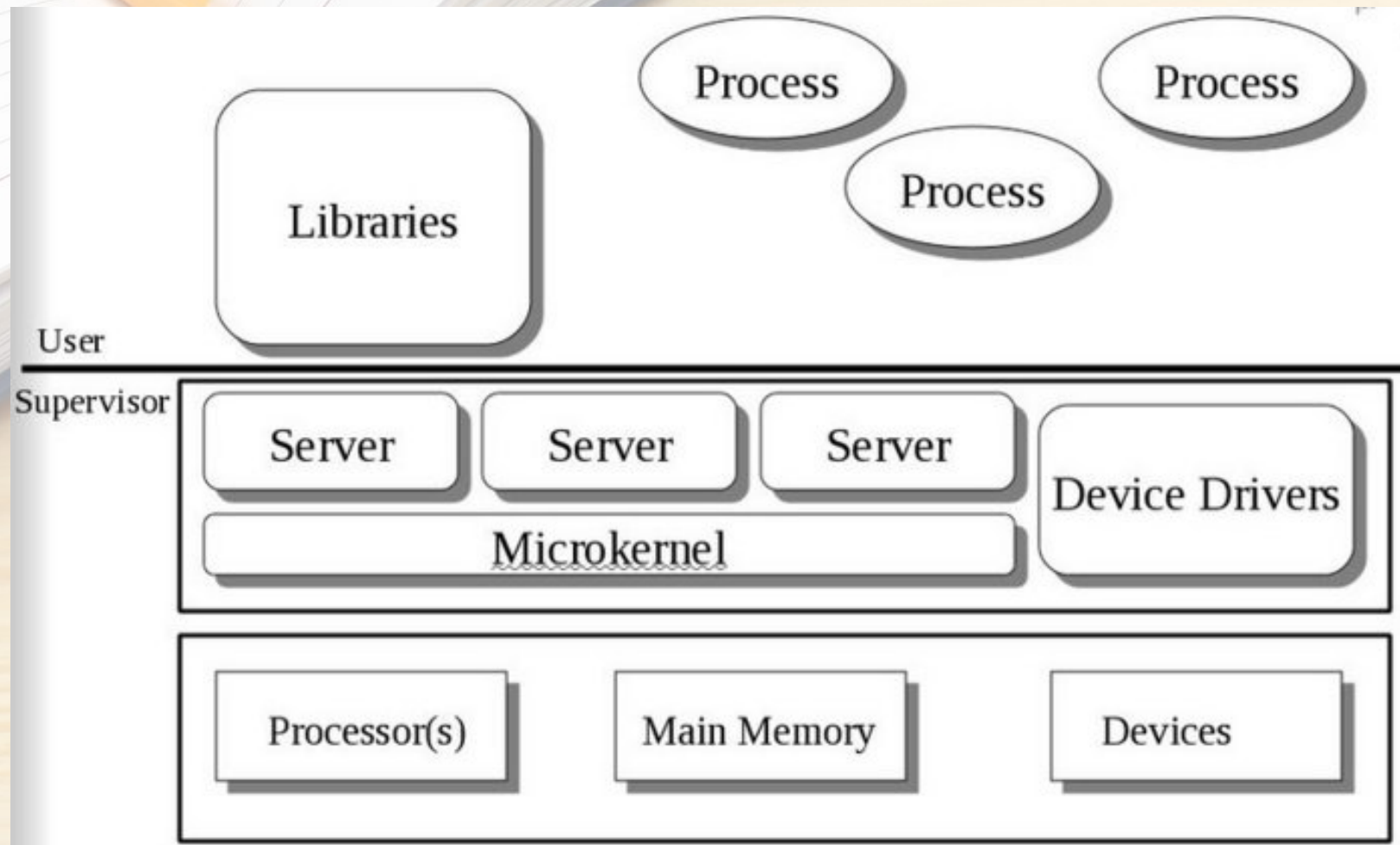


Figure – 24 Microkernel Approach [5]



□ *Advantage*

- ✓ Operating system can be easily extended
- ✓ Kernel is smaller, so very few changes are required in it.
- ✓ It offers more security and reliability.

□ *Disadvantage*

- It has poor performance due to increased system overhead of message passing.

Virtual Machine

- Virtual machine does abstract the hardware of a single computer (the CPU, Memory, Disk drives, Network Interface Cards, and so forth) **into several different execution environments** and thereby **creating the illusion** that each separate execution environment is running its own private computer/environment.

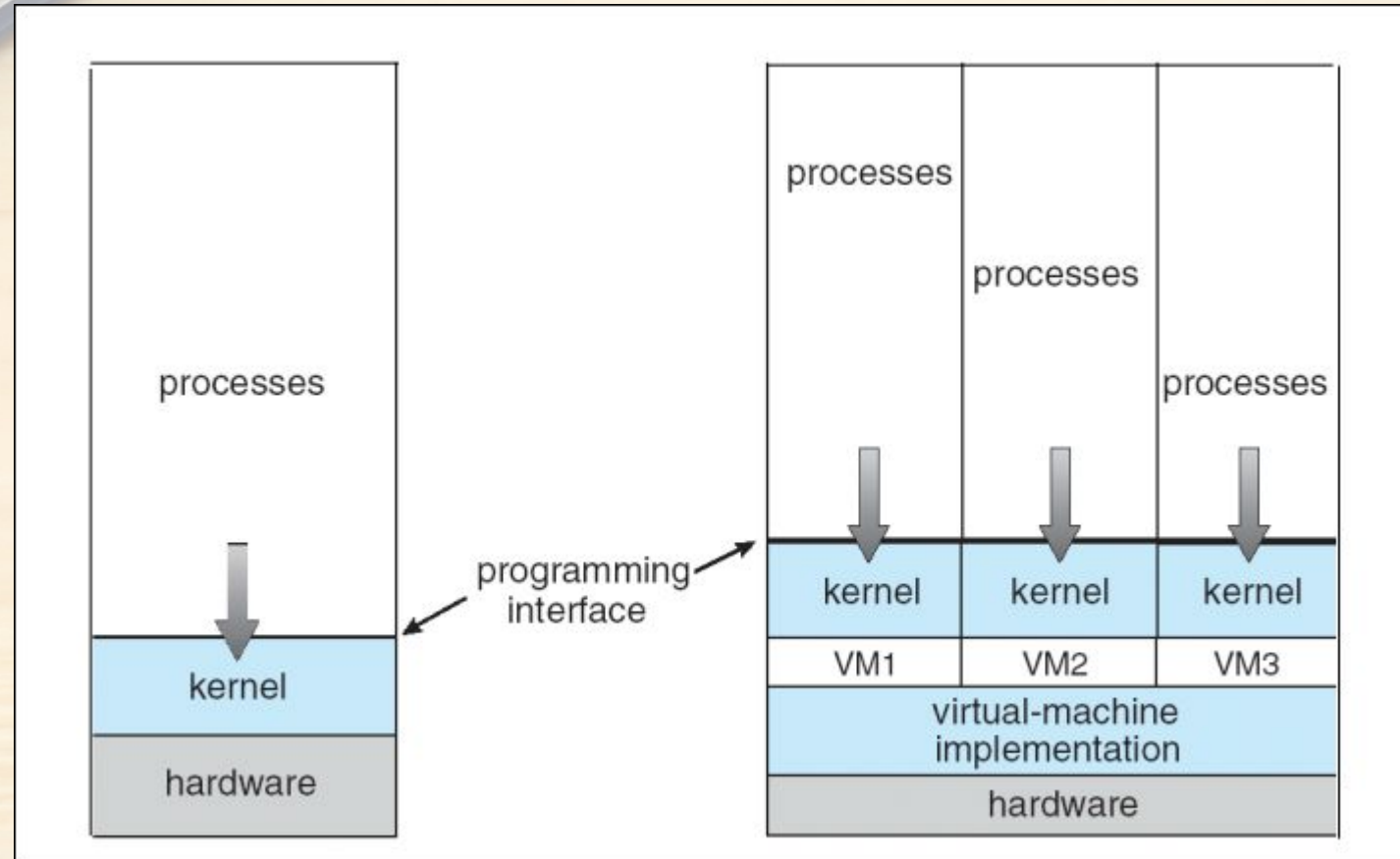
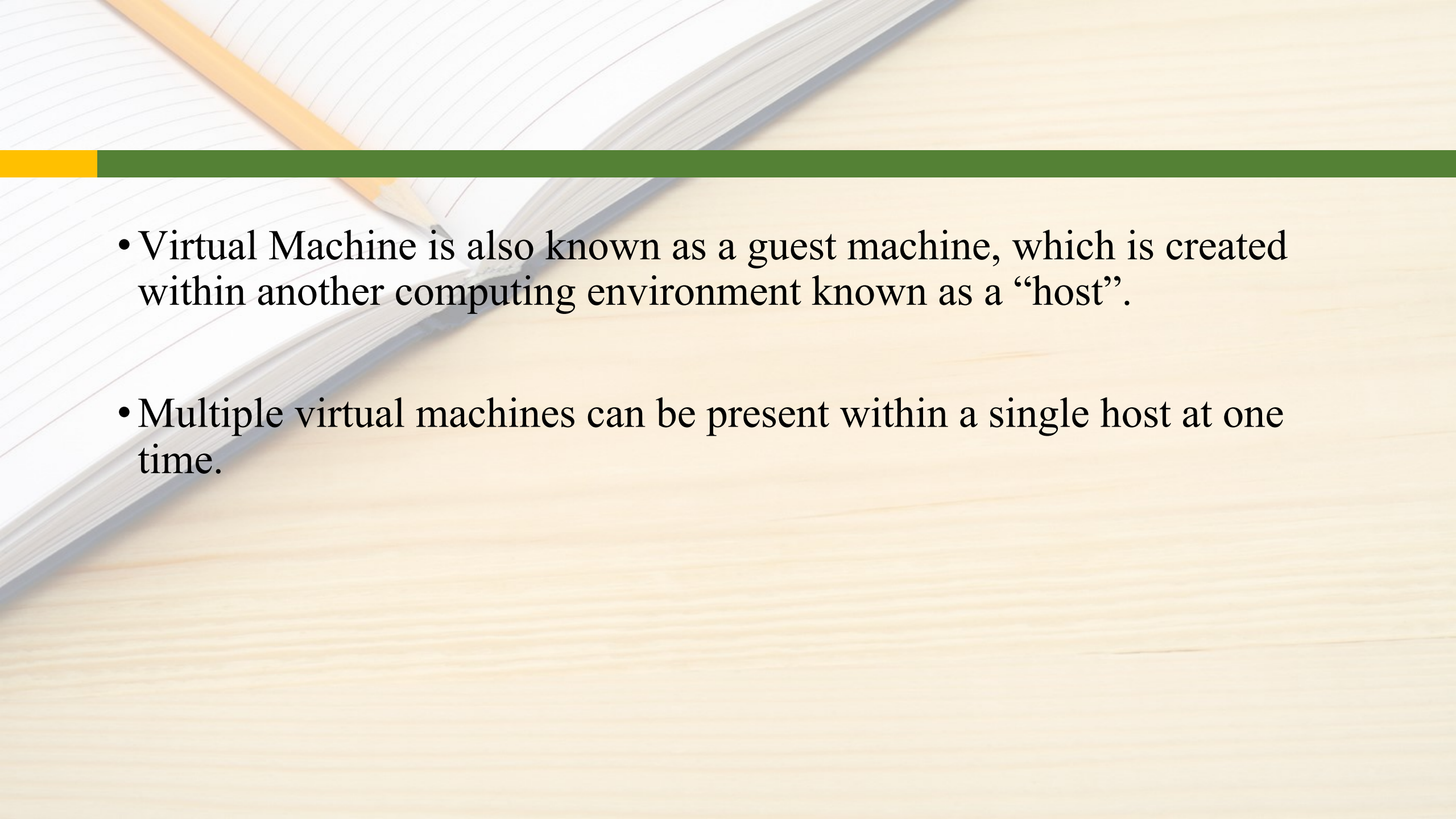


Figure – 25 Non-Virtual Machine & Virtual Machine [1]

- 
- Virtual Machine is also known as a guest machine, which is created within another computing environment known as a “host”.
 - Multiple virtual machines can be present within a single host at one time.

The background of the slide features a close-up, slightly blurred image of an open notebook with lined pages and a yellow pencil resting on it. A solid green horizontal bar is positioned across the upper portion of the image, partially obscuring the notebook. Below this bar, the text is presented in a clean, black font.

□ *Advantages*

- ✓ It allows multiple OS on a single physical computer without any interference.
- ✓ Virtual machines are widely available and easy to handle and maintain.
- ✓ It offers multiple application provisioning and support for critical solutions.

□ *Disadvantages*

- ✓ Little bit inefficient, unlike physical computers. As hardware resources are distributed in an indirect way.
- ✓ Multiple VMs running on a single incapable host may result in performance downgrade.

References

- [1]. Operating System Concepts - 9th Edition by Abraham Silberschatz, Peter Galvin, Greg Gagne, Wiley Asia Student Edition.
- [2]. Operating Systems: Internals and Design Principles - 5th Edition, William Stallings, Prentice Hall of India
- [3]. Modern Operating Systems- 4th Edition Andrew S. Tanenbaum, Pearson Education India
- [4]. Operating Systems: A Modern Perspective - 2nd Edition by Gary J. Nutt, Addison-Wesley
- [5]. [Monolithic Approach](#)
- [6]. [Basic Operating system concepts](#)
- [7]. [Operating system Generations](#)



Thank You

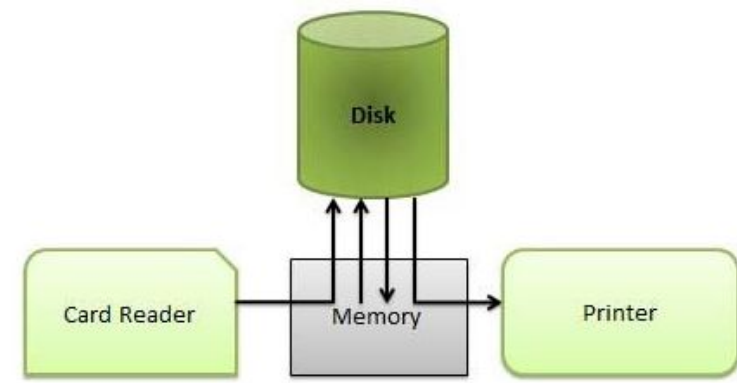
Recommended Extra Reading for Reference

- Spooling
- Buffering

Spooling

- SPOOL is an acronym for **simultaneous peripheral operations on-line**.
- A spool is a buffer that holds output for a device, such as a printer, that cannot accept interleaved data streams.
- Spooling overlaps input of one job with the computation of other jobs.
- The spooler may be reading the input of one job while printing the output of a different job.
- Spooling is useful because device access data with different rates.

Spooling



- The buffer provides a waiting station where data can rest while the slower device catches up.
- Buffer will store the job incoming from the Card reader and also it will contain the job after its completion of job, until the printer is not ready.
- Computer can perform I/O in parallel with the computation
- It becomes possible to have the computer read data from card reader to disk and to write out to printer while it was computing. This process is called spooling.

Spooling

□ Advantages:

- Performance of the system is increased
- CPU and I/O devices work more efficiently
- Leads naturally to multiprogramming
- Also used for processing data at remote sites

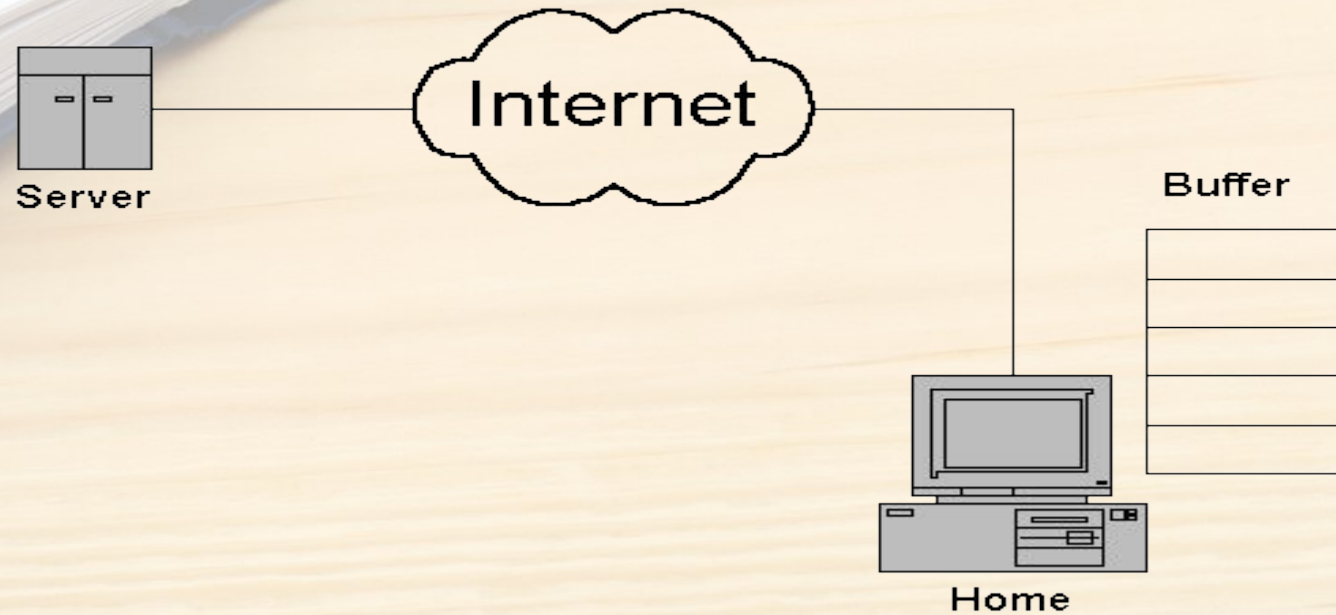
□ Example:

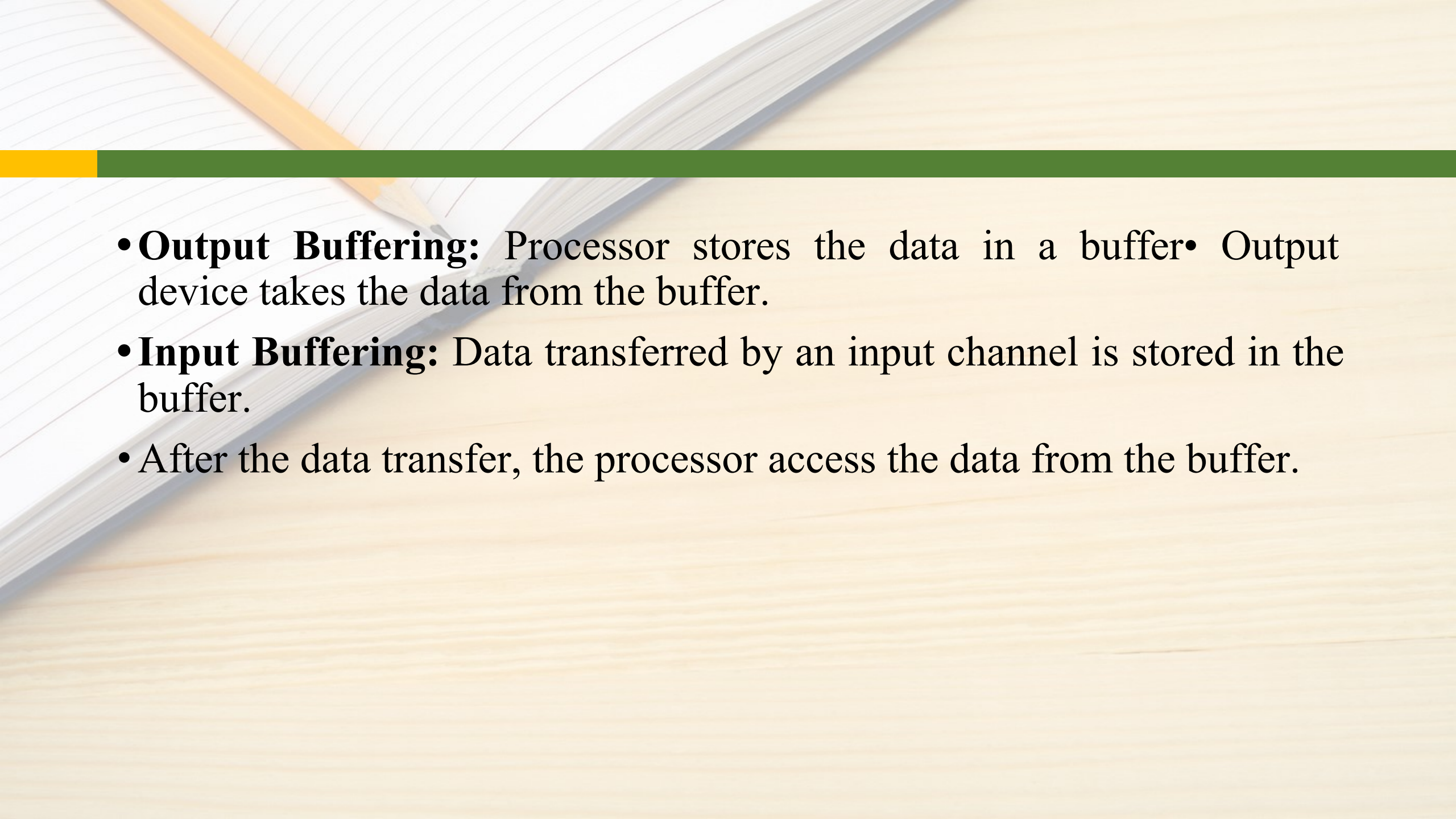
- Even experienced a situation when suddenly for some seconds your mouse or keyboard stops working? Meanwhile, we usually click again and again here and there on the screen to check if its working or not. When it actually starts working, what and wherever we pressed during its hang state gets executed very fast because all the instructions got stored in the respective device's spool.

Buffering

- A buffer is an area of main memory for holding data during input and output data transfers.
- Overlaps the I/O of job with that of its own computation.
- After the data have been read and the CPU is about to start the operation, the input device is instructed to begin the next input operation.
- Both the CPU and I/O device are busy.
- By the time CPU is ready for next operation, the input device would have finished reading it.
- CPU creates data and puts into a buffer until an output device can accept it.

Buffering



- 
- **Output Buffering:** Processor stores the data in a buffer• Output device takes the data from the buffer.
 - **Input Buffering:** Data transferred by an input channel is stored in the buffer.
 - After the data transfer, the processor access the data from the buffer.

Why Buffering?

- To cope with a speed mismatch between the producer and consumers of a data streams
- To adapt between devices that have different data- transfer sizes
- To support copy semantics for application I/O