

Q-1 A: System calls for file management:

Interface between a process and operating system is provided by system calls.

System calls are usually made when a process in user mode requires access to a resource. Then it requests the kernel to provide a resource via system calls.

File management:

These system calls are responsible for file manipulation such as creating a file, reading a file, writing into device buffer etc.

There are four system calls → `open()`, `read()`, `write()`, `close()` for file management.

1) `open()`

↳ system call which is used to know the file descriptor of user-created files. Since read and write use the file descriptor as their first parameter, so to know the descriptor `open()` system call is used.

Syntax: `fd = open (file-name, mode, permission);`

here → `file-name` is the name of the file which we want to open.

→ `mode`: is used to define file opening mode: `(create, read, write)`.

→ `permission`: used to define permissions.

2) `read()`:

↳ which is used to read content from the file. It can also be used to read input from the keyboard (by specifying 0 in the file descriptor).

Syntax: `read (file-descriptor, buffer, max-length)`

`file-descriptor` or `name of the buffer where data is to be stored` ↓ ↓ ↓
maximum amount of data can be read.

2.) Difference between multiprogramming and multitasking

→ Multiprogramming: A computer running more than one programs at a time.

→ Multitasking: Tasks which are sharing a common resource (like One CPU)

Multiprogramming

→ Concept of context switching is used.

→ In this system, the OS simply switches and executes to another job when current job needs to wait.

→ here the idea is to reduce CPU idle time for as long as possible.

→ This increases CPU utilization by organising jobs.

→ Concept of context switching and time-sharing is used.

→ In this, processor is used in time sharing mode. Here switching happens when allowed time expires or there is other reason for current process to wait.

→ here, the idea is to further extend the CPU utilization by increasing responsiveness time sharing.

→ This also increases CPU utilization with increment in responsiveness.

Q-2 A.) Long Term Scheduler controls the degree of multiprogramming. (also called as Job scheduler).

Q-2. B.) Scheduling criteria.

1) CPU utilisation.

- The main objective of any CPU scheduling algo is to keep the CPU as busy as possible.
- CPU utilization can range from 0 to 100. but in a real time system, it varies from 40 to 90 percent which depends on the load upon system.

2.) Throughput

- A measure of the work done by system CPU is the number of processes being executed and completed per unit time. This is called throughput.
- The throughput may vary depending upon the length or duration of the process.

3.) Turnaround time:

- How long the particular process takes to execute is the important criteria for that process.
- The time elapsed from the time of submission of a process to the time of completion is known as turnaround time.
- This is the sum of time spent waiting to get into memory, waiting in ready queue, executing and waiting for I/O.

4.) Waiting time:

- A scheduling algo. doesn't affect the time required to complete the process once it starts execution.

- It only affects the waiting time of the process i.e. time spent by a process waiting in ready queue.

5.) Response time

- Turn-around time is not the best criteria in an interactive system.
- A process may produce some output fairly early and computing new results while previous results are being output to the user.
- Therefore, another criteria is the time taken from submission of the process of request until the first response is produced. This measure is called response time.

Q-2(c) ▷ Race condition:

- A race condition is an undesirable situation that occurs when a device or system attempts to perform two or more operations at the same time, but because of the nature of the device or system, the operation must be done in the proper sequence to be done correctly.

▷ Avoiding Race

- In order to prevent race conditions from occurring, we would typically put a lock around the shared data to ensure only one thread can share the data at a time.

- To avoid the race condition, we need mutual exclusion. Here Mutual exclusion is somehow of making sure that if one process is using a shared variable or file, the other processes will be excluded from doing the same things.

Q-2 c
continue.

The part of the program where the shared memory is accessed is called the critical region or critical section.

If we could arrange matters such that no two processes were ever in their critical regions at the same time, we could avoid race conditions.

► Rules for avoiding Race conditions:

1. No two processes may be simultaneously inside their critical regions. [Mutual exclusion]
2. No assumptions can be made about speeds or the number of CPUs.
3. No process running outside its critical region may block other processes.
4. No process should have to wait forever to enter its critical region.

1D). An I/O interface is required whenever the I/O device is driven by processor.

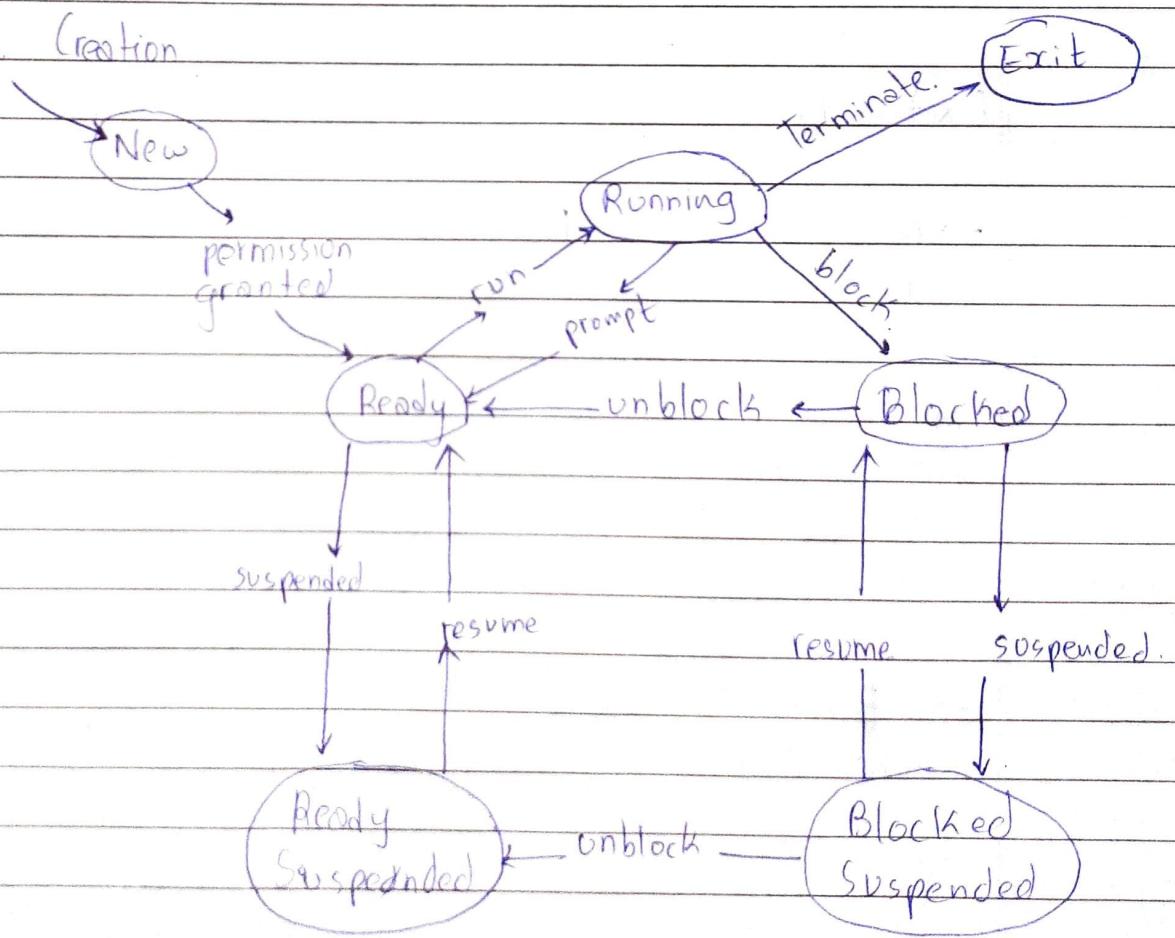
~~Memory~~ mapped memory I/O.

In this case a cartesian portion of the process address space is mapped to the device & communication occurs by reading and writing directly from I/O.

Q - 2 E Seven states of process model :

- 1 New
- 2 ready
- 3 running
- 4 Blocked
- 5 Blocked suspended
- 6 Ready suspended
- 7 exit

Process Creation



New: when the process is newly created.

Ready: when the process is created and ready to run.

Ready suspended : when the process leaves the ready state and suspended.

Blocked : when the process is blocked after running state.

Blocked suspended : when the process is blocked and also suspended.

Running : when the process is running on the CPU.

Exit : Process leaves the CPU and RAM.

→ Ready suspend : Process that was initially in the ready state but were swapped out of main memory and placed onto external storage by scheduler are said to be in suspend ready state.

The process will transition back to ready state whenever the process is again brought onto the main memory.

→ Block suspend :

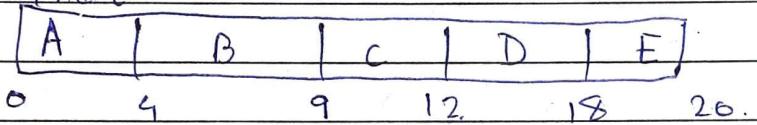
when work is finished it may go to suspend ready

Q - 3.

Process	Arrival time	Burst time
A	0	5
B	2	5
C	3	3
D	5	6
E	5	2

1) First come First serve:

Gant chart:



CT TAT WT

5 5 0
9 7 2

12 9 6

18 14 8

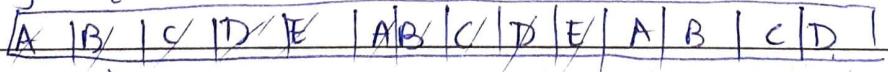
20 15 13

Avg TAT = 9.8 ms.

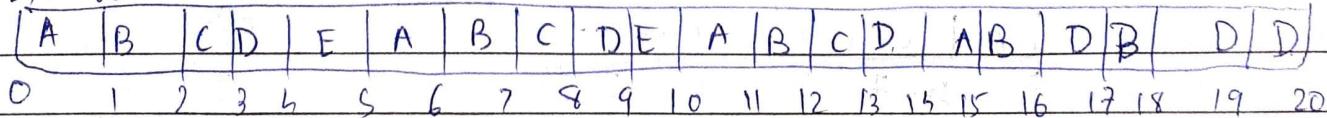
Avg WT = 5.8 ms.

2) Round Robin (q=1 ms.)

Ready queue



Gant chart.



CT TAT WT

15 15 11

Avg TAT = 12.4 ms.

18 16 17

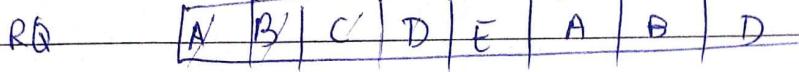
Avg WT = 8.4 ms.

13 10 7

20 15 10

10 5 3

3.) Round Robin ($q.t = 3 \text{ ms}$)



CT TAT WT

15 15 11

Avg TAT = 12.2 ms

17 15 10

Avg WT = 8.2 ms

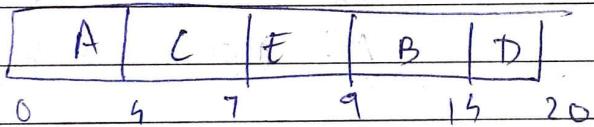
9 6 3.

20 16 10

15 9 7.

4.) shortest job first (SJF)

Gantt chart.



CT TAT WT

5 5 0.

Avg TAT = 8 ms

Avg WT = 5 ms

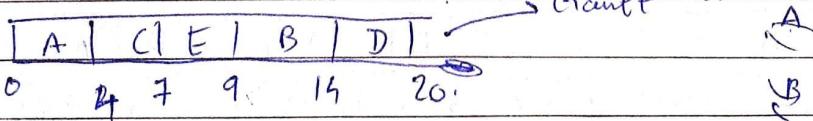
15 12 7

7 4 1

20 16 10

9 4 2

5.). SRTN (SRTN)



CT TAT WT

Avg TAT = 8 ms

Avg WT = 4 ms

Q)

Process ID

State

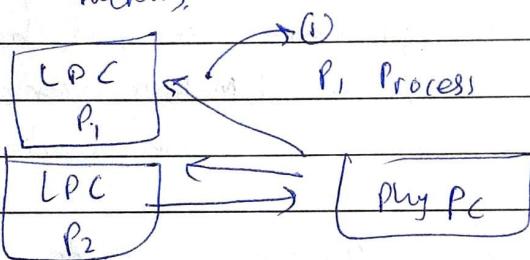
Priority

Program Counter

I/O information

etc.

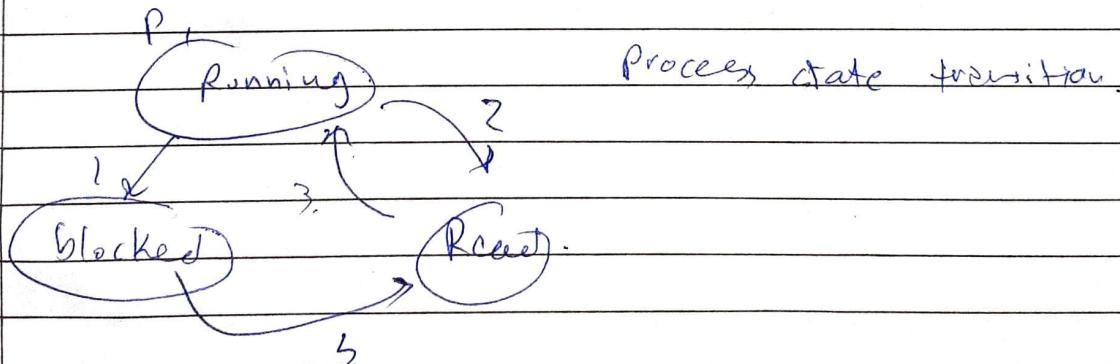
memory



order.

(P) CPU allocated to P_1

Q.D)



Process state transition.

1.) Running - blocked Process blocks wait for input or wait for an event or resource that is needed

2.) Running - Ready when process is either completed or terminated due to timeout i.e. end of time slice

3.) Ready \rightarrow Running \rightarrow Scheduler picks the process

4.) Blocked - ready Input becomes available as discussed in first point.