PROCESS MANAGEMENT

Process Concept

Def^n:- The program under execution is called as the process.

- → It should reside in the main m/m.
- → It is occupied the CPU to execute the instr^n

Process will have various Attributes:-

1) process id
2) process state
3) program counter
4) priority
5) General Purpose Register
6) List of open files
7) List of open devices.
8) protection info.

1) Process id:

Process id is the unique identification no which is assinged by the OS at the time of process creation.

No two processes have same process id

2) Process State:

→ It contains the current state of the process where it is residing.

→ process have various state, in which particular state it is residing - this info provided by Process state. ↓ at a particular time

3) Program Counter:

→ program counter contains the address of next ins to be executed.

4) Priority:

→ priority is the parameter which is assigned by the OS at the time of process creation.

5) General Purpose Register:

→ what are the registers (all registers) used by the process, that information will be maintained in general purpose register.

6) List of open files:

→ what are all the files used by or opened by the process will be maintained in the list of open files attribute of process.

7) List of open devices:

→ what are all the devices used or opened by the process, that info will be maintained in the list of open devices attribute of process.

All the attributes of the process is called as the context of the process.

The context of the process will be stored in PCB (process control Block).

P.id	P.S
P.C	priority
L.O.F	L.O.D
G.P.R	Prot^{n}

Process Control Block (PCB)

→ Every process will have its own PCB. The PCB's of the process will be stored in the main memory.

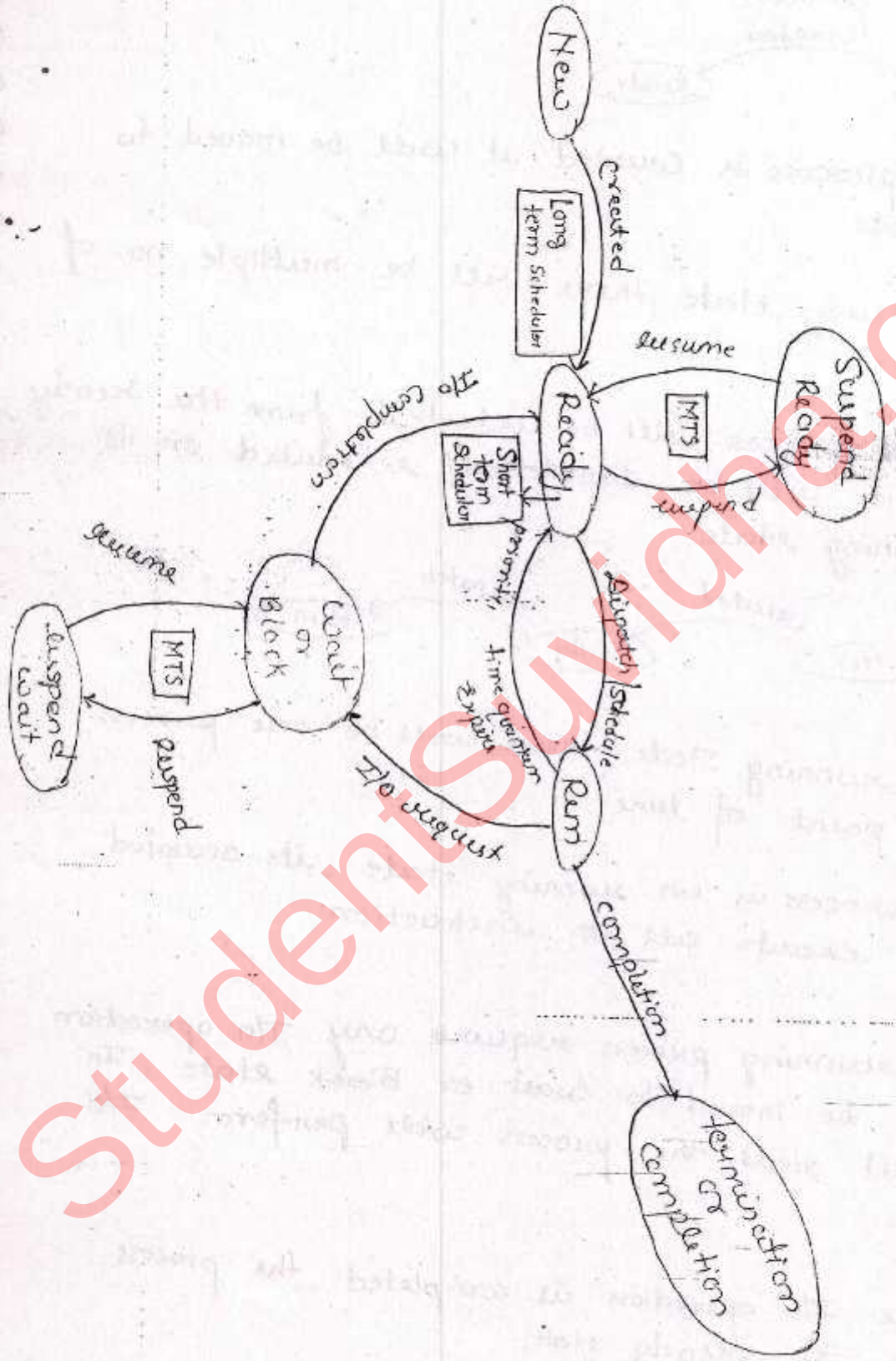
→ The process will have various state:

1. New
2. Ready
3. Run
4. wait (or) block
5. termination (or) completion
6. Suspend Ready
7. Suspend wait (or) Suspend Block

Various operations performed on the process :-

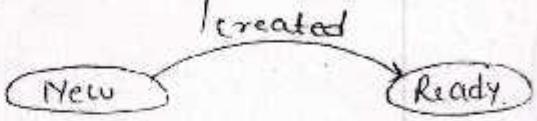
1. Creation
2. Scheduling
3. Execution
4. Termination (or) killing
5. Suspending

process state diagram



New

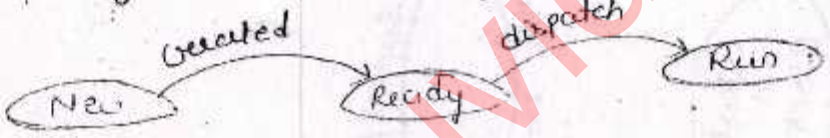
→ Initially the process will be in the new state, it means process is under creation or being created.



→ Once the process is created, it will be moved to Ready state.

→ In the ready state, there will be multiple no. of process.

→ One of the process will be selected from the ready state & it will be dispatched/scheduled on to the running state

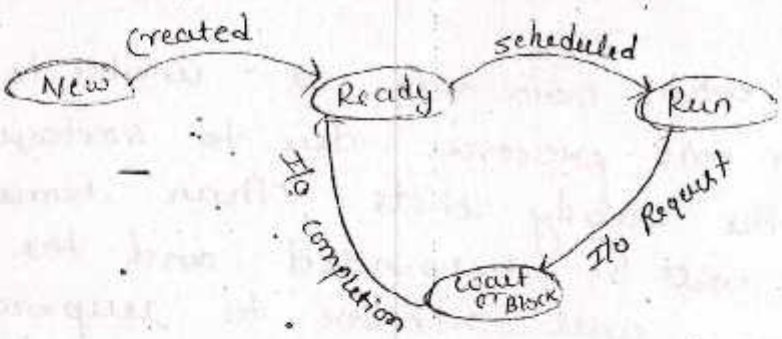


→ In the running state, there will be one process at any point of time

→ when process is in running state, its occupied CPU & execute all its instruction.

→ If the running process requires any I/o operation it will be moved to wait or Block state. In the wait state, the process will perform I/O operation.

Once the I/o operation is completed, the process will go to ready state.



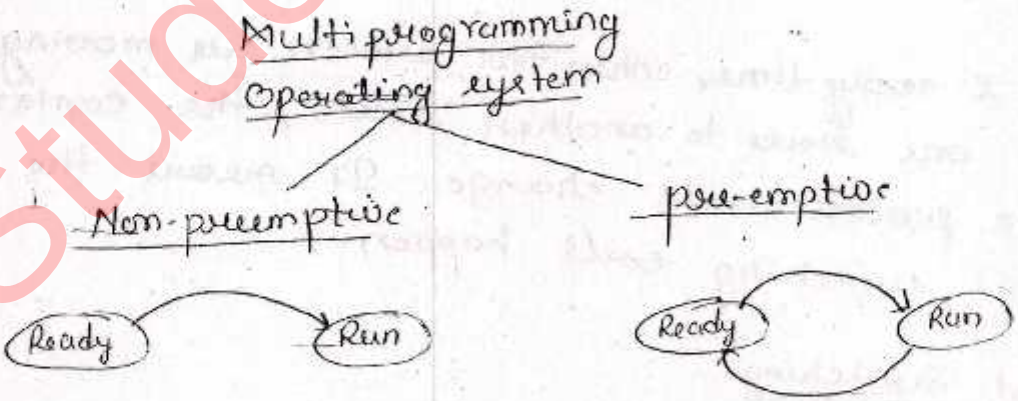
→ There will be multiple no. of processes in the wait state, it means the multiple processes will perform the I/o operation simultaneously.

→ When the process is in the ready, running and block states, it means process is residing in the main memory.

→ Once the running process completed its execution, it will go to the termination state.

Multiprogramming OS is categorized into two type:-

1. Non-preemptive
2. Preemptive (also called as multi-tasking or time sharing)



→ There is a case when main m/m is unable to handle more than one processes due to sortage of resources in the ready state. Than some of the processes will be suspended and the suspended processes will remove to suspend ready state.

→ Whenever the resources are sufficient, the processes will be resumed back to ready state

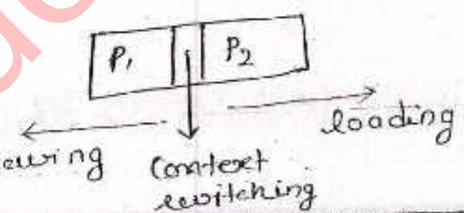
→ When the process is in the suspend ready it is residing in the backing store (secondary m/m)

→ Whenever the resources are not sufficient to manage more than one processes in the wait state, then some of the processes will be suspended & the suspended processes will remove to suspend wait state.

→ Whenever the resources are sufficient, the processes will be resumed back to wait state.

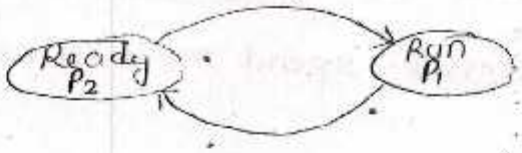
→ Each & every time, when the process is moving from one state to another state, the context of the process will change. It means the context switching will happen.

Context Switching

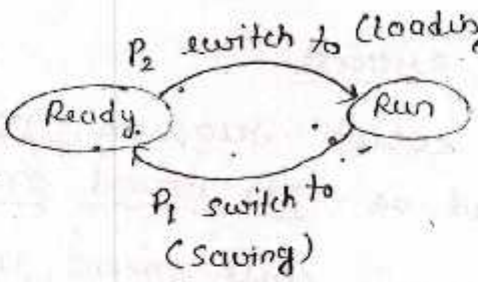


Saving the context of one process & loading the context of another process is called as context switching.

Initialy



Context Switching



Note:- In some special cases only if one process is there, it is also called as context switching.

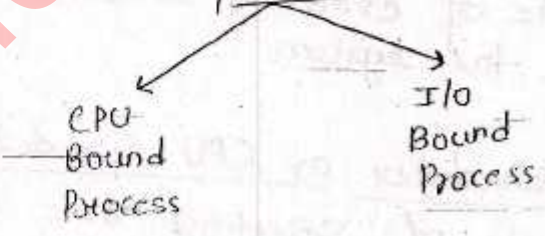
Eg:- Round Robin Scheduling with only one process.

→ If the context of the process is more, than context switching time will also increased, which is undesirable.

→ The context switching time is considered as overhead (burden) for the system.

↔ The processes based their execution time are of two types

Process



1. CPU Bound process-

The processes which required the more CPU time are called as CPU Bound processes.

→ CPU Bound processes - spend more time in the running state

2. I/O Bound process-

The processes which require more amount of I/O time are called as I/O Bound processes.

→ I/O Bound processes will spend more time only in the waiting state.

↔ Degree of Multiprogramming

The no of processes present in the main mem. at any point of time is called as degree of multiprogramming.

↔ In the Operating System, there are three different Schedulers.

1. Long term Scheduler or Job Scheduler.

It is responsible of creating and bringing the new process into the system

2. Short term Scheduler or CPU Scheduler:

It is responsible of selecting one of the process in the ready state for scheduling on to the running state.

3. Mid term Scheduler or Medium Term Scheduler

→ It is responsible of suspending & resuming the processes.

→ The job done by the mid term scheduler is called as "Swaping".

⇔ Dispatcher

→ The Dispatcher is responsible of saving the context of one process and loading the context of another process.

→ The context switching will be done by the dispatcher.

Schedulers are example of operating system processes.

→ The longterm scheduler should select the good combination of CPU bound & I/O bound process in order to get good throughput of the system.

→ If LTS select only I/O bound process — all are in wait state - ⊛ CPU become idle & ready queue will ⊛ throughput decrease. be always empty.

→ If LTS select only CPU bound process ⊛ CPU utilization more waiting queue will be always empty. ⊛ throughput decrease.

→ The long term scheduler controls the degree of multiprogramming.

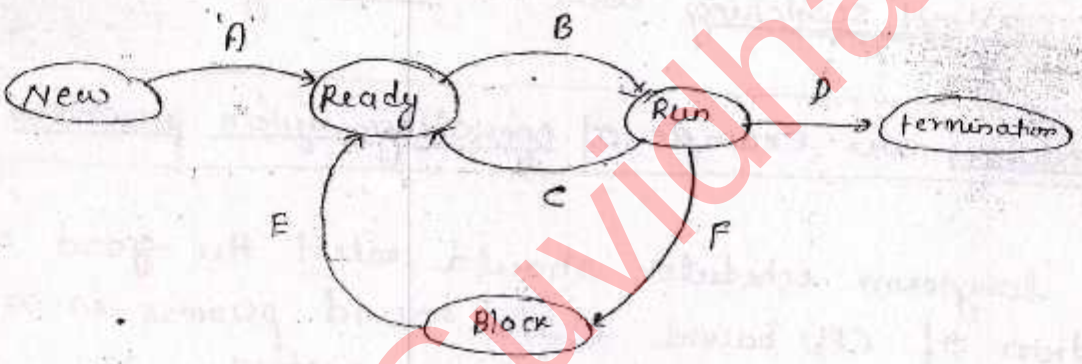
Q. Consider a system with the 'n' CPU processors then what is the min^{n} & max^{m} no. of processes that may present in the ready, running & wait state

Sol:-

State	min	max
Ready	0	all or any no.
Run	0	n
wait	0	all or any no.

'n' processor having 'n' running state

Q:-



Consider the below statement:

S1: If a process makes a trans^{n} 'D', it would result another process making a transition 'A' immediately.

S2: The process P_2 in the Block state can make a transition E while another process P_1 is in the running state.

S3: The OS uses pre-emptive scheduling.

S4: The OS uses non-preemptive scheduling

Which of the above are true?

a) 1, 2
b) 1, 3
c) 2, 3
d) 2, 4

When the process enter into lifecycle, it have various time.

Type of times w.r.t process state

Arrival Time:- The time when the process is arrived into the ready state is called arrival time of the process.

Burst Time :- The time required by the process for its execution is called as Burst time of the process.

Completion Time: The time when the process is completed its execution is called as completion time of process.

Turn Around Time: The time difference b/w the completion time and arrival time is called as T.A.T.

$$T.A.T = C.T - A.T$$

Waiting Time: The time difference b/w the turn around time and burst time is called as waiting time.

$$W.T = T.A.T - B.T$$

Response Time: The time difference b/w the first response and arrival time is called as response time.