schedulers

A scheduler is an os module that selects the next process to be admitted to the system for running purpose.

→ There are 3 different types of schedulers present
1) Long term scheduler
2) short " "
3) Medium " "

Long term scheduler :- / Job Scheduler

→ More than one processes are stored in a mass storage device (disk) in a batch queue.
→ It selects those processes from the disk & loads them into the ready queue of main memory.
→ Thus it controls the degree of multiprogramming i.e the no. of processes in the memory.
→ It is invoked when a process leaves the system
→ Execution frequency of this scheduler is less in comparision to others
→ Its main aim is to provide a balance mix of jobs cpu bound & I/O bound jobs so that neither cpu nor I/O devices becomes idle. Becoz if all the processes are cpu bound, then I/O devices will be idle & if all are I/O bound, cpu will be idle

Short term scheduler / cpu scheduler

→ It selects a process among the processes in the ready queue & allocates cpu to it by the help of a dispatcher
→ It is invoked more frequently than other schedulers
→ Its main aim is to maximize the system performance

→It is invoked each time the cpu requires a new process for execution, which happens in the following situations

a) When the running process makes an I/O request

b) running process is terminated

c) after every time slice

d) new process is created & enters the ready queue

e) interrupt & fatal exceptions.

Note

Dispatcher

It is a module that connects the cpu to the process that is selected by the short term scheduler.

→It switches from one process to another, jumps to the proper location & starts executing new process.

Medium term scheduler

→It introduces an intermediate level of scheduling.

→ Since the main memory is of limited capacity, so more no. of processes can't be kept here at a time. So when the process becomes suspended in the mid of execution due to some I/O request, or some other reason, then ~~process~~ that partially executed process is removed from main memory & placed in the suspend queue of the secondary memory. This is called swap-out. It reduces the ~~degree~~ CPU contention & decreases the degree of multiprogramming & also reduces system load.

- These swapped out processes are reintroduced later into the ready queue for their execution. This is called swap-in

Q the entire operation is called swapping. This task is done by the ~~sho~~ medium term scheduler.

→ It is invoked when the memory space is freed or when a process departs or when the no. of ready processes are less than the ready queue limit.