M.Sc. IV Sem. (Mathematics)

Paper 2nd - Fundamentals of Computer Science - II

Unit V

Reference Book : **C. Ritchie, *Operating Systems incorporating UNIX and Windows,* BPB Publications, New Delhi.**

Topic : Process Management & Scheduling

Process Management :

Every action on a computer is run inside a process. A process can be viewed as the execution of a program and is used as a unit of work for a processor. As long as Von Neumann architecture is used to build computers, only one process can be run at a time. Modern operating systems are able to execute many processes at once (multi-tasking) even under a single C.P.U. **Process management is an operating system's way of dealing with running multiple processes.** Since most computers contain one processor with one core, multitasking is accomplished by simply switching processes quickly. Process management involves the computation and distribution of 'time shares'.

Another concept that must be considered within process management is that of a thread. A thread may be viewed as a sub-process, that is, it is a separate sequence of execution within the code of one process. Threads are becoming increasingly important in the design of distributed and client-server systems.

Scheduling :

In a multi-programming computer, several processes will be competing for use of the processor. At any instant, only one process will be running while the others will be ready, waiting for the processor. **The operating system has the task of determining the optimum sequence and timing of assigning processes to the processor. This activity is called scheduling.**

Objectives of Scheduling :

Objectives of scheduling are as follows :

1. The scheduling system should maximize the system throughput.
2. The scheduling system should be fair to all users. This does not mean all users must be treated equally, but consistently, relative to the importance of the work being done.
3. The scheduling system should provide tolerable response (for online users) or turn-around time (for batch users).
4. The scheduling system should degrade performance gracefully. If the system becomes overloaded, it should not collapse but avoid further loading or temporarily reduce the level of service (e.g. response time).
5. The scheduling system should be consistent and predictable.

Criteria for Scheduling :

In making decisions about the scheduling of processor work, a number of criteria can be taken into account by the operating system. These are :

a) Priority assigned to the job.
b) Class of job i.e. batch or online or real time.
c) Resource requirements e.g. expected run time, memory required etc.
d) Input output or C.P.U. bound i.e. whether job used I/O time or processor time.
e) Resources used to date, e.g. the amount of processor's time already consumed.
f) Waiting time to date e.g. the amount of time spent waiting for service so far.

Types of Scheduling :

Scheduling can be exercised at three distinct levels :

1) **High Level Scheduling (or Long Term or Job Scheduling)**

2) **Medium Level Scheduling (or Intermediate Scheduling)**

3) **Low Level Scheduling (or Short Term or Processor Scheduling)**

1. High Level Scheduling :

High level scheduling deals with the decision as to whether admits a new job to the system, i.e. it decides which newly submitted jobs are to be converted into processes and be put into the READY queue to compete for access to the processor. This activity is applicable to batch systems. The decision is based on the First-Come-First-Served (FCFS) and Shortest-Job-First (SJF) which selects the waiting job which has the shortest estimated run time.

2. Intermediate Scheduling :

Intermediate scheduling in concerned with the decision to temporarily remove a process from the system in order to reduce the system load or to reintroduce a process. Although a process sitting in the READY or BLOCKED queue is not actively using the processor, nor it competing for the processor use and will consume processor time in future. The medium level scheduler attempts to relieve temporary overloading by remove processes from the system for short periods. Meanwhile, the processes are held in the READY SUSPENDED or BLOCKED SUSPENDED state.

3. Low Level Scheduling :

Processor scheduling deals with the decision of which ready process is to be assigned to the processor. This level is often called 'dispatcher'. It is the most complex and significant of the scheduling levels whereas the high and medium level schedulers operate over seconds or minutes, the low level scheduling is making critical decisions many times every second. The low level scheduling is invoked whenever the current process relinquishes control, which will occur, when the process calls for an input-output transfer or some other interrupt arises.

Comparison among Scheduler

S.N.	Long-Term Scheduler	Short-Term Scheduler	Medium-Term Scheduler
1	It is a job scheduler	It is a CPU scheduler	It is a process swapping scheduler.
2	Speed is lesser than short term scheduler	Speed is fastest among other two	Speed is in between both short and long term scheduler.
3	It controls the degree of multiprogramming	It provides lesser control over degree of multiprogramming	It reduces the degree of multiprogramming.
4	It is almost absent or minimal in time sharing system	It is also minimal in time sharing system	It is a part of Time sharing systems.
5	It selects processes from pool and loads them into memory for execution	It selects those processes which are ready to execute	It can re-introduce the process into memory and execution can be continued.

Question :

What is the meant by the statements that a process is :

a) in the READY state ?
b) in the BLOCKED state ?