Operating System

Dr. Satyabrata Das

Associate Professor

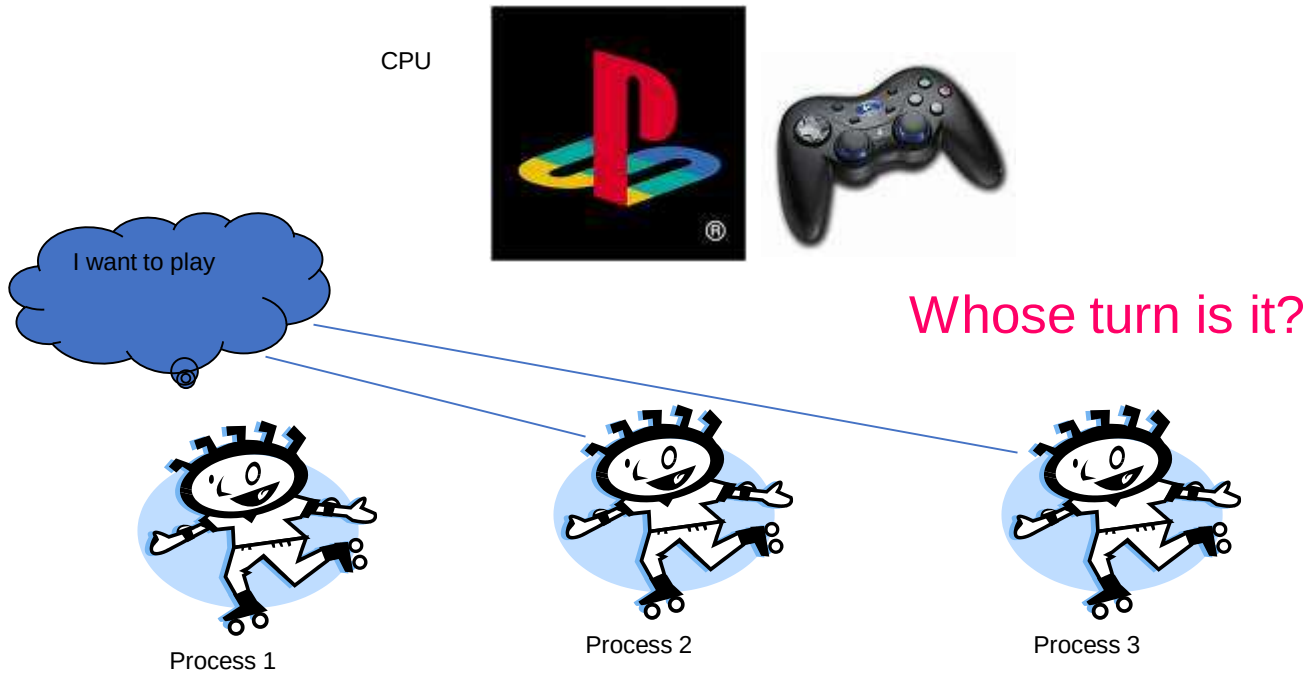
Dept. of IT, VSSUT, Burla

Outlines

- Introduction to CPU Scheduling
- Schedulers
  - (a) Long term Scheduler
  - (b) Short term scheduler
  - (c) Medium term scheduler
- Dispatcher
- Context switching
- Scheduling Criteria/Methodology

Scheduling

- Deciding which process/thread should occupy the resource (CPU, disk, etc)

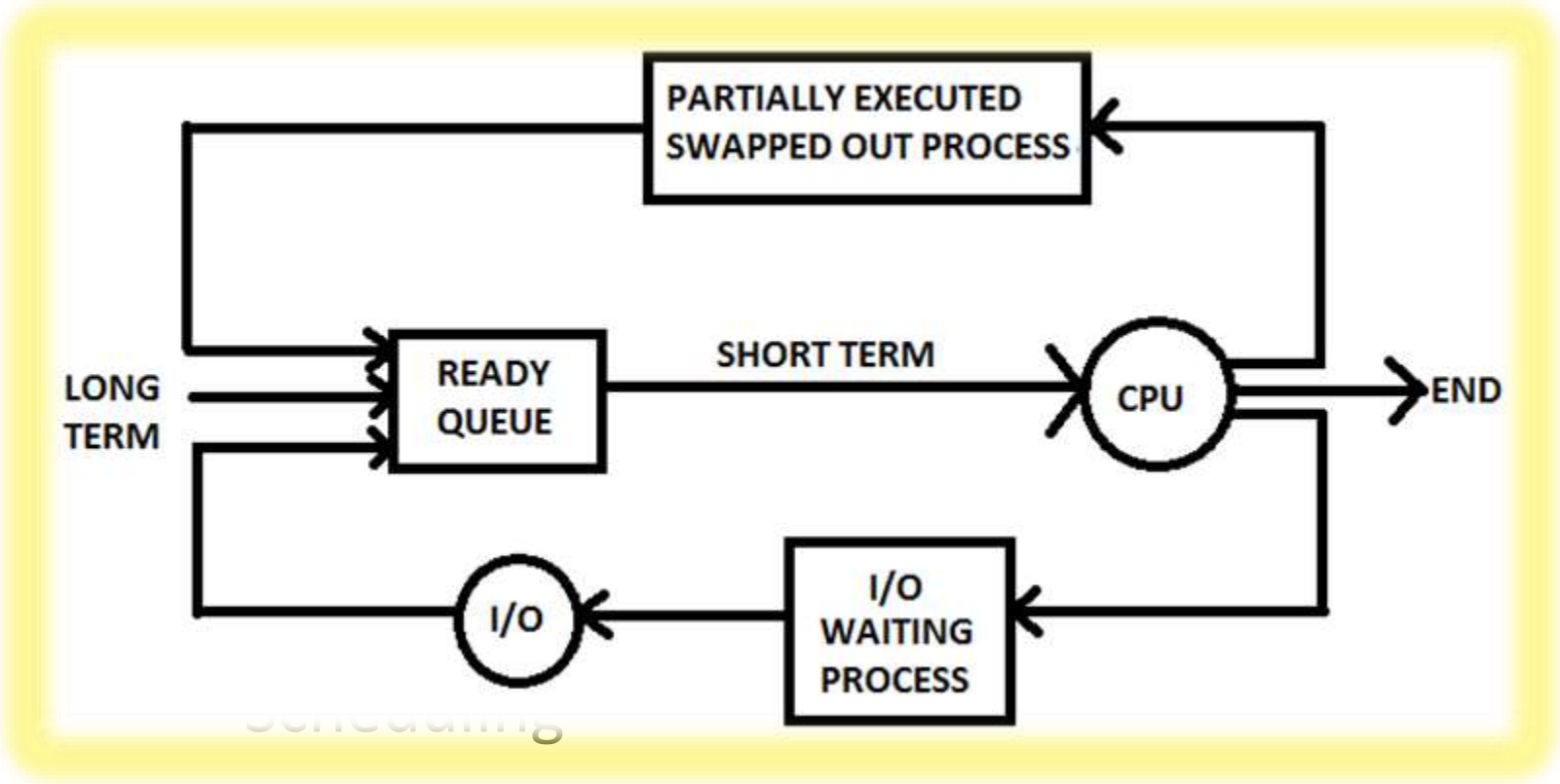


CPU Scheduling

Scheduling is a fundamental function of operating system. When a computer is multiprogrammed it has multiple processes competing for CPU at the same time. If only one CPU is available then a choice has to be made regarding which process has to execute next. This decision making process is known as scheduling and the part of the operating system that makes the choice is known as the scheduler. The algorithm it uses in making this choice is called scheduling algorithm.

CPU scheduling is the basis of multiprogrammed operating system. Scheduling is a fundamental operating system function, almost all computer resources are scheduled before use. The CPU is also one of the primary resources. So CPU is also scheduled before use. The CPU scheduling algorithm determines how the CPU will be allocated to the process.

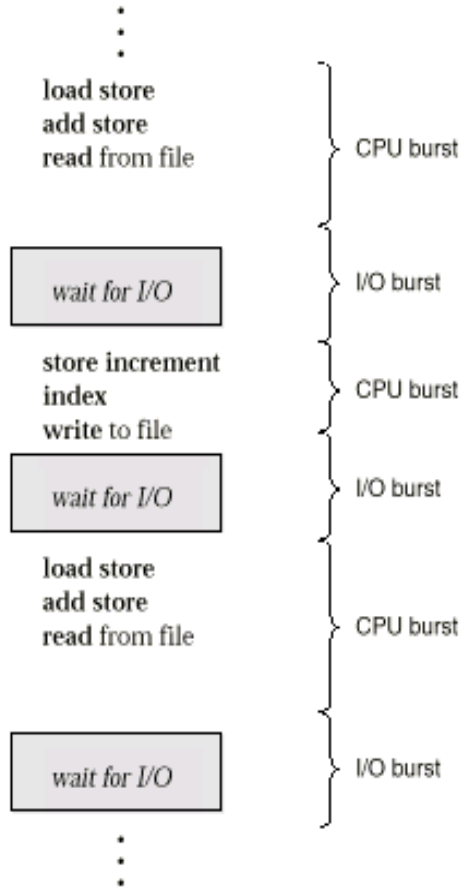
CPU scheduling contd..



CPU SCHEDULING

Scheduling Concepts

Multiprogramming	A number of programs can be in memory at the same time. Allows overlap of CPU and I/O.
Jobs	(batch) are programs that run without user interaction.
User	(time shared) are programs that may have user interaction.
Process	is the common name for both.
CPU - I/O burst cycle	Characterizes process execution, which alternates, between CPU and I/O activity. CPU times are generally much shorter than I/O times.
Preemptive Scheduling	An interrupt causes currently running process to give up the CPU and be replaced by another process.



CPU SCHEDULING

The Scheduler

Selects from among the processes in memory that are ready to execute, and allocates the CPU to one of them

- CPU scheduling decisions may take place when a process:
  1. Switches from running to waiting state
  2. Switches from running to ready state
  3. Switches from waiting to ready
  4. Terminates
- Scheduling under 1 and 4 is *nonpreemptive*
- All other scheduling is *preemptive*

CPU SCHEDULING

The Dispatcher

- Dispatcher module gives control of the CPU to the process selected by the short-term scheduler; this involves:
  - switching context
  - switching to user mode
  - jumping to the proper location in the user program to restart that program

- *Dispatch latency* – time it takes for the dispatcher to stop one process and start another running

Schedulers

A process migrates between the various scheduling queues throughout its lifetime process is called scheduler. The operating system must select for the scheduling process from these queues in some fashion. This selection process is carried out by the appropriate scheduler. In a batch system more processes are submitted then executed immediately. So these processes are spooled to a mass storage device like disk, where they are kept for later execution.

Types of Schedulers

There are 3 main schedulers

1. Long-term scheduler (jobs scheduler) – selects which programs/processes should be brought into the ready queue.
2. Medium-term scheduler (emergency scheduler) – selects which job/process should be swapped out if system is loaded.
3. Short-term scheduler (CPU scheduler) – selects which process should be executed next and allocates CPU.

Types of Schedulers Contd..

(1) Long Term Scheduler

It selects processes from disk and loads them into memory for execution. It controls the degree of the multiprogramming i.e., no of processes in memory. It executes less frequently than other scheduler. If the degree of multiprogramming is stable than the average rate of process creation is equal to the average departure rate of processes leaving the system. So the long term scheduler is required to be invoked when the process leaves the system. Due to longer interval between executions it can afford to take more time to decide which process should be selected for execution.

Types of Schedulers Contd..

(2)Short Term Scheduler

The short term scheduler selects among the process that are ready to execute and allocate the CPU to one of them. The primary distinction between these two schedulers is the frequency of their execution. The short term scheduler must select a new process for the CPU quite frequently. It must execute at least one in 100ms. Due to the short duration of time between executions, it must be very first.

Types of Schedulers Contd..

(3)Medium Term Scheduler

Some operating systems introduce an additional intermediate level of scheduling known as medium term scheduler. The main idea behind this scheduler is that sometimes it is advantageous to remove process from memory and thus reduce the degree of multiprogramming. At some later time, the process can be reintroduced into memory and its execution can be continued from where it had left off. This is called as swapping. The process is swapped out and swapped in later by medium term scheduler. Swapping is necessary to improve the process miss or due to some change in memory requirements the available memory limit is exceeded which required some memory to be freed up.