- **Process Scheduling**

The process scheduling is the activity of the process manager that handles the removal of the running process from the CPU and the selection of another process on the basis of a particular strategy.

Process scheduling is an essential part of a Multiprogramming operating systems. Such operating systems allow more than one process to be loaded into the executable memory at a time and the loaded process shares the CPU using time multiplexing.

- **Categories of Scheduling**

There are two categories of scheduling:

1. **Non-preemptive:** Here the resource cannot be taken from a process until the process completes execution. The switching of resources occurs when the running process terminates and moves to a waiting state.

2. **Preemptive:** Here the OS allocates the resources to a process for a fixed amount of time. During resource allocation, the process switches from running state to ready state or from waiting state to ready state. This switching occurs as the CPU may give priority to other processes and replace the process with higher priority with the running process.

- **Process Scheduling Queues**

The Operating System maintains the following important process scheduling queues:

- o **Job queue** − This queue keeps all the processes in the system.
- o **Ready queue** − This queue keeps a set of all processes residing in main memory, ready and waiting to execute. A new process is always put in this queue.
- o **Device queues** − The processes which are blocked due to unavailability of an I/O device constitute this queue.
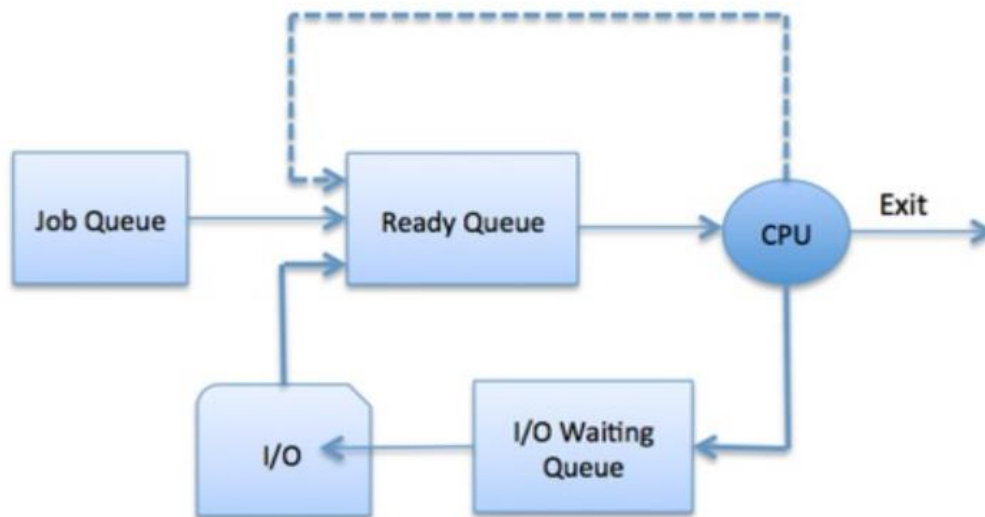
Process Scheduling Queues

- **Two-State Process Model:**

  Two-state process model refers to running and non-running states which are described below

  - o **Running**

    When a new process is created, it enters into the system as in the running state.

- o **Not Running**

    Processes that are not running are kept in queue, waiting for their turn to execute.

- **Schedulers**

Schedulers are special system software which handle process scheduling in various ways. Their main task is to select the jobs to be submitted into the system and to decide which process to run. Schedulers are of three types:

- o **Long-Term Scheduler**

    It is also called a job scheduler. A long-term scheduler determines which programs are admitted to the system for processing. It selects processes from the queue and loads them into memory for execution. Process loads into the memory for CPU scheduling.

- o **Short-Term Scheduler**

    It is also called as CPU scheduler. Its main objective is to increase system performance in accordance with the chosen set of criteria. It is the change of ready state to running state of the process. CPU scheduler selects a process among the processes that are ready to execute and allocates CPU to one of them.

- o **Medium-Term Scheduler**

    It is a part of swapping. It removes the processes from the memory, A suspended processes cannot make any progress towards completion. In this condition, to remove the process from memory and make space for other processes.

- **Context Switching**

A context switching is the mechanism to store and restore the state or context of a CPU in Process Control block.

When the scheduler switches the CPU from executing one process to execute another, the state from the current running process is stored into the process control block. After this, the state for the process to run next is loaded from its own PCB and used to set the PC, registers, etc. At that point, the second process can start executing.

When the process is switched, the following information is stored for later use.

- Program Counter
- Scheduling information
- Base and limit register value
- Currently used register
- Changed State
- I/O State information
- Accounting information