Lecture 2

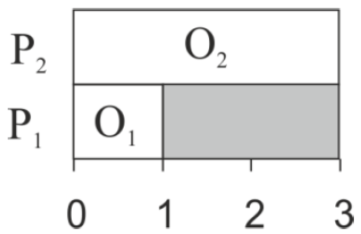
Homogeneous and heterogeneous processors

Homogeneous processors
– All processors are identical in speed and capability.
– Any task can be executed on any processor with the same performance.

Heterogeneous processors
– processors are different (they may have different speeds, resources, or specializations).
– The time to finish a task depends on which processor it is assigned to.

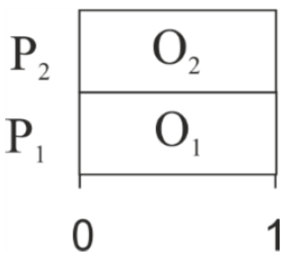
To illustrate the difference between homogeneous and heterogeneous processors, assume we have two operations to be executed: $O_1 = 1$ (u.t), $O_2 = 3$ (u.t).

Case 1 *(Homogeneous processors):* two processors P_1, P_2 with equal speed.



In this case the total time to execute the operations (makespan) will be 3 units of time.

Case 2 *(Heterogeneous processors):* two processors, where P_2 is three times faster than P_1.



In this case, the total time to execute the operations (makespan) will be 1 unit of time.

Remark

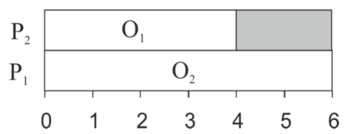
To achieve optimal scheduling with heterogeneous processors, tasks with longer processing times should be assigned to the faster processors, while tasks with shorter processing times should be assigned to the slower processors.

Example (1)

Given two operations $O_1 = 2$ (u.t), $\quad O_2 = 6$ (u.t). Assume we have two processors, where P_2 is slower than P_1, with a speed that is half that of P_1. Find the optimal scheduling.

Solution:

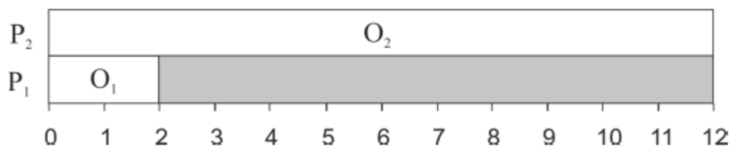
Execution time of O_1 on P_2 =processing time of O_1 / speed of P_2.
Execution time(O_1 on P_2) $= \frac{2}{1/2} = 4$ units of time .



Thus, the total time to complete both operations in the above scheduling is 6 units of time, and this is the optimal scheduling.

Remark

If the same operations are scheduled as below



Execution time(O_2 on P_2) $= \frac{6}{1/2} = 12$ units of time .

Thus, the total time to complete both operations in the above scheduling is 12 units of time. This is a feasible scheduling, but not the optimal.

Example (2)

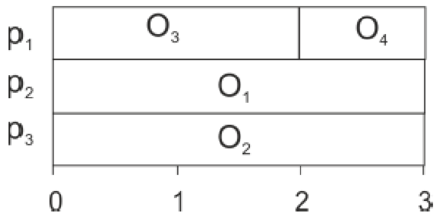
Given $O_1 = 6,\ O_2 = 9,\ O_3 = 2,\ O_4 = 1$ (all in u.t). Let us have three processors, where the speed of P_2 is twice that of P_1, and the speed of P_3 is three times that of P_1. Find the optimal scheduling.

Solution

The speed of $P_1 = 1$
The speed of $P_2 = 2$
The speed of $P_3 = 3$

$$\text{Execution time}(O_2 \text{ on } P_3) = \frac{9}{3} = 3 \,\text{units of time} .$$

$$\text{Execution time}(O_1 \text{ on } P_2) = \frac{6}{2} = 3 \,\text{units of time} .$$



Thus, the total time to complete all the operations in the above scheduling is 3 units of time, and this is the optimal scheduling.

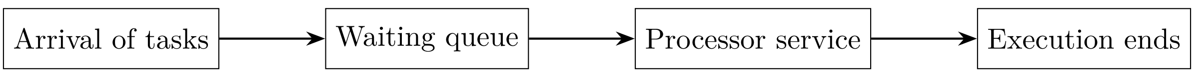
Homework

Given $O_1 = 10,\ O_2 = 8,\ O_3 = 7,\ O_4 = 2$ (u.t).
Let us have three processors, where the speed of P_2 is twice that of P_1, and the speed of P_3 is a quarter that of P_1. Find the optimal scheduling.

Types of scheduling

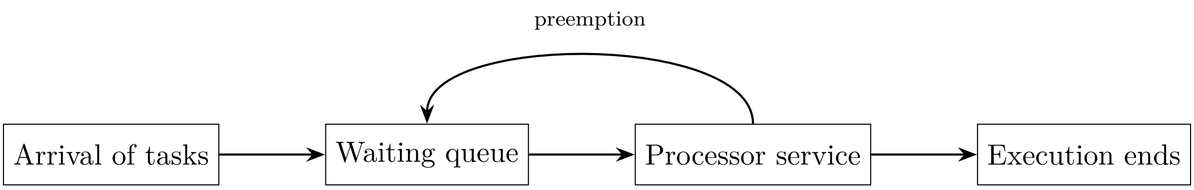
1) Non-preemptive scheduling

In this type of scheduling, tasks can not switch from the executing state to the waiting state until they are fully executed.



2) Preemptive scheduling

In this type of scheduling, tasks can switch from the executing state to the waiting state before they are fully executed.



Scheduling criteria

Processor scheduling algorithms vary in their properties, and some algorithms may outperform others. To assess the effectiveness of these algorithms in executing a given set of tasks, it is essential to understand the key criteria for comparing their efficiency. This understanding will help us achieve optimal scheduling.

1. **Fairness**: This means that each task will have a chance to execute by allocating the processor to it.
2. **Processor utility**: it is the ratio of tasks processed by the processor to the total number of tasks.
3. **Processor throughput**: it is the ratio of the number of tasks executed in the system per unit of time.
4. **Return time**: it is the time period for executing the entire task, or the time period between the arrival of the task and its completion.
5. **Waiting time**: it is the time period that the task spends in the queue waiting for the processor to be allocated to it.
6. **Response time**: it is the time period between the arrival of the task and the start of execution.
7. **Scheduling Length(Makespan)**: is the period of time it takes to execute all the tasks.

Most scheduling algorithms work to increase the utility and Processor throughput and reduce return time, waiting time, response time, and scheduling length, as well as achieve fairness between processes. Optimum is achieving balance in each of these criteria.