Loren Howard, Yujian Zhang, Ken Wakaba

Professor Ahmed Ezzat

COEN 383

11 October 2017

Process Scheduling Algorithms Report

First-Come-First-Serve is a non-preemptive policy that selects and runs a process based on FIFO. Although it is the easiest to implement, it does not perform well when it comes to the metrics we have discussed. With FCFS, longer processes are going to fare better than shorter ones. FCFS serves no problem if all processes run the same length of time and arrive at the same moment, but this is never the case. If a long process were to arrive before, short processes, the short processes would have to wait until the long one is done in its entirety. This is what is known as the Convoy Effect. Due to this, the wait time as well as the turnaround time of the processes within FCFS could be poor. Response time too will be high if we have a large discrepancy in process execution time.

To solve the problem of FCFS, another policy is used known as Shorted Job First (SJF). As the name suggest, SJF runs the next shortest job within the queue. It is still a non-preemptive job so it will allow all processes to complete in its entirety before choosing the next shortest job. Response time for SJF is excellent but only for short processes. Meaning that even with this priority in place, processes could still end up waiting a while to complete, as well as have potentially high turnaround time, although it is an improvement from FCFS.

Shortest Remaining Time (SRT) is much like SJF, except it is a preemptive policy. A job is chosen based on the shortest remaining run time of a process within the queue. When a new process comes int, it is compared with the remaining time of the current process and if it is less, the current job is stopped and the new job is run. Due to the preemptive policy, SRT has very similar response time to that of SPN, but it has more consistent shorter response times. Although it is a slight improvement, SRT as well as SJF still do not have particularly good response time. If multiple processes arrive at the same time, the processes on the tail end of the queue still must wait for all the others before to run completely before being scheduled. Although great for turnaround time, the wait time and response time are poor.

Round Robin, the solution to the poor response time problem with the above policies, is also a one that uses preemption. Each process is run for a certain time, or quanta, and once the quanta is complete, the scheduler runs the next process in the queue. The shorter the quanta, the better the response time is for each process. While this policy is great for response time, it has awful wait time and turnaround time. This issue comes from the fact that RR only runs each process for a short amount of time before giving it to the next process. Having a small quantum also presents the problem of constant context switches, affecting the cost and performance. On the other hand, setting the quantum to be too long may cause poor response time, almost close to that of FCFS.

For Highest Priority First, the scheduler assigns a priority to each process, ranking them from highest priority to lowest. Depending on whether it is a preemptive or non-preemptive policy, the lower priority processes will get interrupted if a higher priority algorithm comes in. When looking at our metrics, higher priority processes have smaller wait times and response times than lower priority.