

Multiprocessing: scheduling

COSC 208, Introduction to Computer Systems, 2021-11-12

Outline

- Scheduling processes
- First In First Out (FIFO) scheduling
- Shortest Job First (SJF) scheduling
- Preemption
- Shortest Time-to-Completion First (STCF) scheduling
- Round Robin (RR) scheduling

Scheduling processes

- OS decides which process to run and for how long
- *What factors should the OS consider when making these decisions?*
 - Time for process to complete
 - Overhead of context switching
 - Fairness
 - User interaction
- For now, consider one of these metrics: turnaround time
 - $T_{\text{turnaround}} = T_{\text{complete}} - T_{\text{arrive}}$
- For now, assume a process starts and runs to completion—i.e., no I/O and no preemption

First In First Out (FIFO) scheduling

| Process | Arrival time | Duration |
|---------|--------------|----------|
| A | 0 | 15 |
| B | 5 | 15 |
| C | 10 | 15 |

- *What is the average turnaround time for the processes using FIFO?*

- *What happens if A's duration is 60?*

- *What happens if C's duration is 60?*

- *How can we change the schedule so the average turnaround time when A's duration is 60 is more like the average turnaround time when C's duration is 60?*

Shortest Job First (SJF) scheduling

| Process | Arrival time | Duration |
|---------|--------------|----------|
| A | 0 | 60 |
| B | 0 | 15 |
| C | 0 | 15 |

- *What is the average turnaround time for the above processes using SJF?*

| Process | Arrival time | Duration |
|---------|--------------|----------|
| A | 0 | 60 |
| B | 5 | 15 |
| C | 10 | 15 |

- *What is the average turnaround time for the above processes using SJF?*

- We're back to FIFO—*What happened!? How can we fix this?*

Preemption

- OS only regains control when a system call occurs—e.g., read/write file, yield
 - Syscalls may occur infrequently, or never, due to program design, bugs, or malicious behavior
- *How does an OS forcibly regain control?*—set a timer that raises an interrupt
 - Interrupt causes a trap instruction to be executed
 - Interrupts can also be raised by devices—e.g., Network Interface Card (NIC)
- *What must the OS do if it decides to run another process?*—perform a context switch
 - Save the machine state associated with the process that was running—in particular, the contents of all registers are saved in the process's control structure
 - Restore the machine state associated with the process that should run—again, the contents of all registers are loaded from the process's control structure

Shortest Time-to-Completion First (STCF) scheduling

- Allow preemption
- If a process arrives that has less computation remaining than the currently running process, then preempt the current process and run the new process
- Also known as Preemptive Shortest Job First (PSJF)

| Process | Arrival time | Duration |
|---------|--------------|----------|
| A | 0 | 60 |
| B | 5 | 15 |
| C | 10 | 15 |

- *What is the average turnaround time for the above processes using STCF scheduling?*

- Now consider response time: $T_{\text{response}} = T_{\text{first_run}} - T_{\text{arrive}}$
- *What is the average response time for the same processes using STCF scheduling?*
- Now consider wait time: $T_{\text{wait}} = \sum (T_{\text{start_run}} - T_{\text{become_ready}})$

- *What major assumption have we made thus far that is impractical in a real system?*

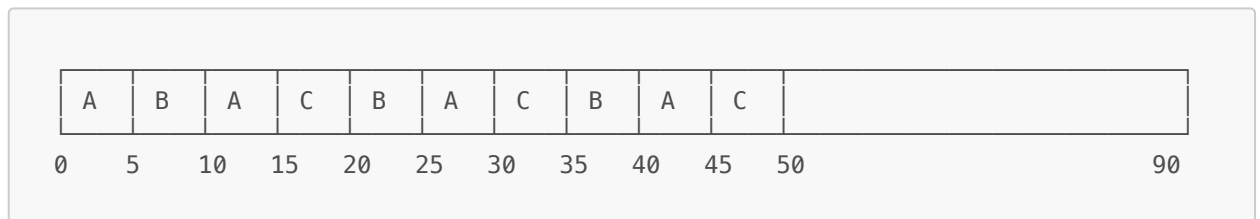
Round Robin (RR)

- Let each process run for a small amount of time, then switch to the next process; when you get to the last process, then start again with the first process and repeat

| Process | Arrival time | Duration |
|---------|----------------|----------|
| A | Just before 0 | 60 |
| B | Just before 5 | 15 |
| C | Just before 10 | 15 |

- What is the average turnaround time and response time for the above processes assuming we let a process run for 5 seconds before switching processes?

- Average Turnaround =
 - Average Response =
 - Average Wait =
- In practice, there is a queue of processes that are in the ready state, resulting in the following schedule:



- Determine the schedule for the above process with a time quantum of 10.

- Average Turnaround =
 - Average Response =
 - Average Wait =
- What happens to average response time as we increase the time quantum?