Terminology

Time-sharing / multiprogramming
- Typical proc. alternates between CPU bursts and I/O bursts
- To maximize CPU utilization:
  - multiple processes are kept in memory simultaneously
  - when one process is waiting, another process executes

CPU bound process: spends more time doing computations, generates I/O requests infrequently
I/O bound process: spends more time doing I/O than computing

Job queue: contains all procs. in the system
Ready queue: contains all procs. that reside in main memory and are ready to run
Device queue: contains all procs. waiting for a particular device
Preemptive vs. non-preemptive scheduling

- Process switch can occur when a process
  1. needs to wait for some resource (sleeps)
  2. exits
  3. returns from kernel mode to user mode but is not the most eligible process to run

- **Non-preemptive scheduling**: scheduling takes place only in 1 and 2
  - when a process gets the CPU, it keeps it until it sleeps/exits
  - used in MS Windows (?)

- **Preemptive scheduling**: case 3 is also permissible
First-come first-served

Method:
1. Maintain a FIFO queue.
2. When a process enters the ready queue, it is placed at the end of the queue.
3. When the CPU is free, it is allocated to the process at the head of the queue.

Properties:
- Non-preemptive
- Unsuitable for time-sharing systems (\(\because\) each user should get a share of the CPU at regular intervals)
- Average waiting time is not minimal
- Convoy effect: many processes may have to wait for one long process to finish
  Example: 1 CPU-bound proc. + many I/O bound procs.
First-come first-served

Example:

<table>
<thead>
<tr>
<th>Ready processes</th>
<th>Burst time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>24</td>
</tr>
<tr>
<td>$P_2$</td>
<td>3</td>
</tr>
<tr>
<td>$P_3$</td>
<td>3</td>
</tr>
</tbody>
</table>

Processes arrive in the order $P_1, P_2, P_3$

Gantt chart:

<table>
<thead>
<tr>
<th></th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
</tr>
</thead>
</table>

Average waiting time: $(0 + 24 + 27)/3 = 17ms$
**Shortest job first**

**Method:**

1. When the CPU is available, assign it to the process with the shortest next CPU burst.

2. Break ties on a FCFS basis.

**Properties:**

- Optimal in terms of average waiting time
- Suitable for job scheduling in a batch system (use time limit specified by user at time of submission)
- Length of the next CPU request is generally not known

**Pre-emptive SJF:** (shortest remaining time first)

1. When a new process arrives at the ready queue, compare its CPU burst with remaining time for current process.

2. If new process has shorter burst, preempt current process.
Shortest job first

Example:

<table>
<thead>
<tr>
<th>Ready processes</th>
<th>Arrival time</th>
<th>Burst time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>$P_2$</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>$P_3$</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>$P_4$</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Gantt chart:

<table>
<thead>
<tr>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_4$</th>
<th>$P_3$</th>
</tr>
</thead>
</table>

Average waiting time: \( (0 + 7 + 15 + 9)/4 = 7.75\, ms \)

(compare pre-emptive version, FCFS)
CPU burst prediction

- Some function of the measured lengths of previous CPU bursts may be used

- Exponential average:

\[
\tau_{n+1} = \alpha t_n + (1 - \alpha) \tau_n
\]

\[
= \alpha t_n + \alpha(1 - \alpha) t_{n-1} + \ldots + (1 - \alpha)^j \alpha t_{n-j} + \ldots
\]

\[+(1 - \alpha)^{n+1} \tau_0\]

\[t_n\] - length of \(n\)-th CPU burst

\[\tau_{n+1}\] - predicted value for the next CPU burst

\[\alpha\] - relative weight given to recent history
Priority scheduling

Method:

1. Compute a priority for each process.
   - Internal priorities: computed using time limits, memory requirements, ratio of avg. I/O burst to avg. CPU burst, etc.
   - External priorities: computed on the basis of external political / administrative factors
2. Allocate CPU to process with highest priority.
3. Break ties on a FCFS basis.

Properties:

- Can be preemptive or non-preemptive (cf. SJF)
- Starvation (indefinite blocking) may occur
  (if high priority processes keep arriving, low priority process may have to wait indefinitely for CPU)
Priority scheduling with **aging**: priority may be increased in proportion to waiting time
- prevents starvation

**Example:**

<table>
<thead>
<tr>
<th>Ready processes</th>
<th>Burst time</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>$P_2$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$P_3$</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$P_4$</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>$P_5$</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

(low numbers $\Rightarrow$ high priority)

**Average waiting time:** 8.2 ms
**Round robin**

**Time quantum** (or time slice): maximum interval of time between two invocations of the scheduler

- a process can be allocated the CPU for one quantum at one time
- usually between 10–100ms

**Method:**

1. Maintain a FIFO queue of ready processes.
2. Allocate CPU to first process from queue; set timer for 1 time quantum.
3. If running process releases CPU, or timer expires: preempt current process and switch context to the next process in the ready queue; add previously running process to tail of ready queue.
Round robin

Example:

<table>
<thead>
<tr>
<th>Ready processes</th>
<th>Burst time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>24</td>
</tr>
<tr>
<td>$P_2$</td>
<td>3</td>
</tr>
<tr>
<td>$P_3$</td>
<td>3</td>
</tr>
</tbody>
</table>

Time quantum = 4ms

Average waiting time: 5.66ms

Properties:

- Suitable for time-sharing systems (∵ every process gets the CPU for $q$ time units after waiting for $(n - 1)q$ time units)
- Duration of time quantum:
  - large time quantum $\Rightarrow$ RR $\rightarrow$ FCFS
  - small time quantum $\Rightarrow$ context-switching overhead ↑
Multilevel queue

Method:

1. Partition the queue into several separate queues; assign a fixed priority value to each.

2. Assign each process to some fixed queue, based on its properties. Example: system procs. / interactive procs. / interactive editing procs. / batch procs. / student procs.

3. Select a queue based on:
   - fixed priority, OR
   - priority-based proportional time slicing.

4. Select a job from the queue using a suitable scheduling algorithm (e.g. FCFS, RR).

Properties:

- Preemptive
Multilevel feedback queue

- Processes may be moved between scheduling queues

- Parameters:
  - # of queues
  - scheduling algorithm / time slice for each queue
  - initial queue selection policy
  - promotion/demotion policies

Example:

- 3 queues, \( Q_0, Q_1, Q_2 \)

- scheduling policies:
  \( Q_0 = RR \) (quantum = 8ms)  \( Q_1 = RR \) (quantum = 16ms)  \( Q_2 = FCFS \)

- on entry to ready queue, processes assigned to \( Q_0 \)

- on exit from \( Q_0 \), process is placed at tail of \( Q_1 \)
- on exit from \( Q_1 \), process is placed at tail of \( Q_2 \)

  OPTIONAL: if process waits too long in \( Q_2 \), promote it to \( Q_1 \)
Scheduling criteria

- CPU utilization

- Throughput: number of processes that are completed per unit time
  - long processes ⇒ throughput ↓
  - short processes ⇒ throughput ↑

- Turnaround time: interval from the time of submission of a process to the time of completion

- Waiting time: total amount of time spent by a process in the ready queue

- Response time: time from the submission of a request until the first response is produced
  (amount of time taken to start responding, not including the time taken to complete the output)

NOTE: maximum (minimum)/average/variance may be suitable for evaluation
Real-time scheduling

Hard real-time systems:

- Critical tasks must be completed within a guaranteed amount of time
- Resource reservation:
  - processes are submitted with deadlines
  - scheduler may admit the process and guarantee completion, or reject
- Duration of operating system functions must be predictable and bounded
- Consists of special-purpose software running on dedicated hardware

Soft real-time systems:

- Critical processes receive priority over “ordinary” processes
- May be implemented as a general-purpose system
Real-time scheduling

Preemptible vs. non-preemptible kernels:

- Non-preemptible kernels
  - context switch can happen only at restricted points
    - completion of system call/interrupt
    - `sleep()`
    - specially inserted preemption points
  - delays may be unpredictable
  - easier to implement

- Preemptible kernels
  - suitable for soft real-time systems
  - harder to implement
Real-time scheduling

Priority inversion

- High priority process may have to wait for resource held by a low priority process

- **Priority inheritance**: processes that are accessing resources required by high priority process inherit the high priority until they release the resource