Priority Scheduling with General Services

Table of Contents

Setting

Analyzing a priority scheduling system with general services. Let's chose n different number of priorities:

```
prio = 3
```

For every priority there's a different average arrival time ($\frac{1}{\lambda_j}$) and a different average service time ($\frac{1}{\mu_j}$) that can be selected.

In our example $\frac{1}{\lambda_j} = \frac{0.5\, n^2}{j}$ $\frac{1}{\mu_j} = 0.3\, j$

```
avgInterArrivalTime = 1×3
    4.5000    2.2500    1.5000

avgServiceTime = 1×3
    0.3000    0.6000    0.9000
```

Theoretical results

Every j class is characterized by a **server utilization factor** $\rho_j = \frac{\lambda_j}{\mu_j}$ $j = 1, \ldots, n$

The system is stable if the general ρ is $\rho = \sum_{j=1}^{n} \rho_j < 1$

```
rho = 1×3
    0.0667    0.2667    0.6000

The system is stable, the general rho of the system is:
general_rho = 0.9333
```

Let's calculate the theoretical per-class average queueing time and average response time.

Average queueing time for non-preemption system

$$W_j^q = \frac{\sum_{i=1}^{n} \rho_i E[Z_{B,i}]}{\left(1 - \sum_{i=1}^{j} \rho_i\right)\left(1 - \sum_{i=1}^{j-1} \rho_i\right)} \qquad j = 1, \ldots, n$$

Average queueing time for preemption-resume system

$$W_j^q = \frac{\sum_{i=1}^{n} \rho_i E[Z_{B,i}]}{\left(1 - \sum_{i=1}^{j} \rho_i\right)\left(1 - \sum_{i=1}^{j-1} \rho_i\right)} + \frac{\sum_{i=1}^{j-1} \rho_j}{1 - \sum_{i=1}^{j-1} \rho_i} \frac{1}{\mu_j} \qquad j = 1, \ldots, n$$

Average response time $W_j = W_j^q + \dfrac{1}{\mu_j}$

For a non-preemptive system

The per-class average response time is:

```
avgResponseTime = 1×3
    0.6857    1.1786    9.0000
```

For a preemptive-resume system

The per-class average response time is:

```
avgResponseTime_pree = 1×3
    0.3107    0.7875    9.4500
```

The extended service time is the time from first entrance in service to exit.

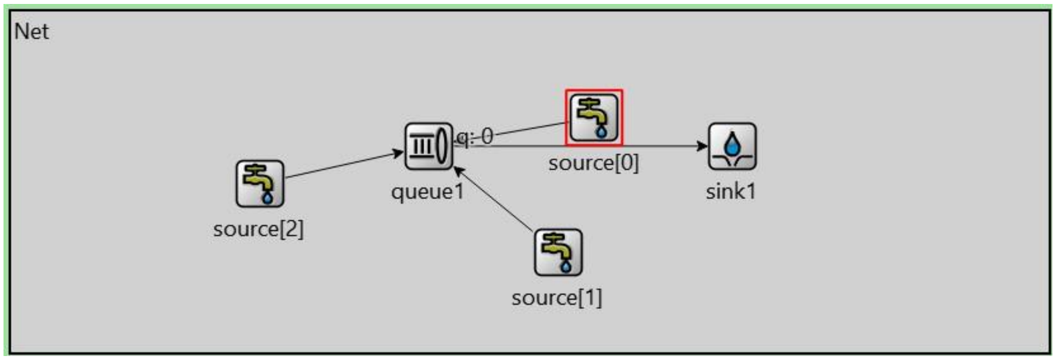
Extended service time $T_j^{\text{serv}} = \dfrac{\sum_{i=1}^{j-1} \rho_j}{1 - \sum_{i=1}^{j-1} \rho_i} \dfrac{1}{\mu_j} + \dfrac{1}{\mu_j} \qquad j = 1, \ldots, n$

The per-class extended service time is:

```
extendedServiceTime = 1×3
    0.3000    0.6429    1.3500
```

Simulation results

Statistic acquired through Omnet++ for a 3 priority class system.



Per-class server utilization factor:

```
simul_rho = 1×3
    0.0663    0.2662    0.6015
```

For a non-preemptive system

The per-class average response time is:

```
simul_avgResponseTime = 1×3
    0.6851    1.1787    8.9452
```

For a preemptive-resume system

The per-class average response time is:

```
simul_avgResponseTime_pree = 1×3
    0.3107    0.7872    9.3947
```

The per-class extended service time is:

```
simul_extendedServiceTime = 1×3
    0.3000    0.6428    1.3494
```

Comparison

Let's compare the result obtained with the simulation and the theoretical one, then we can see the difference.

Per-class server utilization factor difference for the system:

```
ans = 1×3
    0.0004    0.0004    0.0015
```

Per-class average response time difference for the non-preemptive system:

```
ans = 1×3
    0.0007    0.0001    0.0548
```

Per-class average response time difference for the preemptive-resume system:

```
ans = 1×3
```

```
    0.0000    0.0003    0.0553
```

Per-class extended service time difference for the preemptive-resume system:

```
ans = 1×3
10^-3 ×
    0.0000    0.0626    0.5526
```

Per-class extended service time for the non-preemptive system has no meaning because it's equal to the per-class service time.