Analysis of M/M/n/K Queue

with

Multiple Priorities

• For a *P*-priority system, class *P* of highest priority

• Independent, Poisson arrival processes for each class with l_i as average arrival rate for class *i*

• Service times for each class are independent of each other and of the arrival processes and are exponentially distributed with mean $1/m_i$ for class *i*

• Both Non-preemptive and Preemptive Priority Service disciplines are considered

Since the service times are exponentially distributed (i.e. memory less), the results for preemptive resume and preemptive non-resume will be identical

Solution Approach

- Define System State appropriately
- Draw the corresponding State Transition Diagram with the appropriate flows between the states
- Write and solve the balance equations to obtain the system state probabilities

Note that we have given here the solution approach that may be taken to solve a queueing problem of this kind. This has been illustrated with simple examples. More complex cases may be similarly formulated and solved with a corresponding increase in the solution complexity

M/M/-/- Queue with Preemptive Priority

For a *P*-priority queue of this type, define the system state as the following *P*-tuple

$$(n_1, n_2, \ldots\ldots, n_P)$$

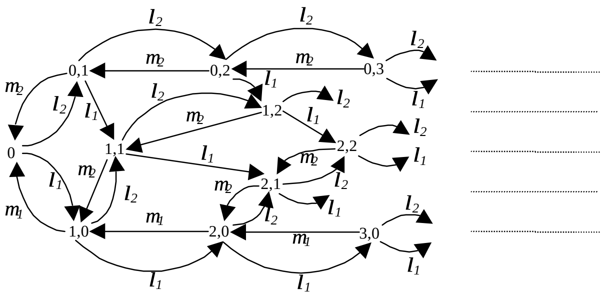
where

n_i = Number of jobs of priority class *i* in the queue

$i = 1, \ldots.., P$

Note that the server will always be engaged by a job of the highest priority class present in the system, i.e. by a job of class *j* with service rate m_j if n_j ³ 1 and $n_{j+1} = \ldots.. = n_P = 0$.

We illustrate the approach first for a *2*-priority M/M/1/∞ queue



Copyright 2002, Sanjay K. Bose 5

The corresponding balance equations for the *2*-priority M/M/1/∞ queue will be given by

$$p_0(l_1 + l_2) = p_{0,1}m_2 + p_{1,0}m_1$$
$$p_{0,1}(l_1 + l_2 + m_2) = p_{0,2}m_2 + p_0 l_2$$
$$p_{1,0}(l_1 + l_2 + m_1) = p_{2,0}m_1 + p_{1,1}m_2 + p_0 l_1$$
$$p_{1,1}(l_1 + l_2 + m_2) = p_{1,0} l_2 + p_{1,2}m_2 + p_{0,1} l_1$$

....

....

....

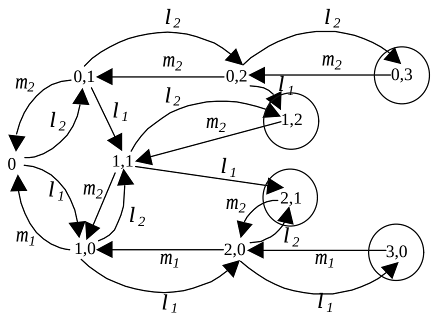
These may be solved to obtain the desired state probabilities

We illustrate next how this approach may be generalized to apply to queues with finite capacity and/or multiple servers

Copyright 2002, Sanjay K. Bose 6

2-Priority M/M/1/3 Queue (Preemptive Priority)

(An example of a queue with finite capacity)



New arrivals will be lost if they come when the system is in any of the circled states

State Transition Diagram for the 2-Priority M/M/1/3 Queue with Preemptive Priority

The corresponding balance equations are

$$
\begin{cases}
p_0(l_1 + l_2) = p_{0,1}m_2 + p_{1,0}m_1 \\
p_{0,1}(l_1 + l_2 + m_2) = p_{0,2}m_2 + p_0 l_2 \\
p_{1,0}(l_1 + l_2 + m_1) = p_{2,0}m_1 + p_{1,1}m_2 + p_0 l_1 \\
p_{1,1}(l_1 + l_2 + m_2) = p_{1,0} l_2 + p_{1,2}m_2 + p_{0,1} l_1 \\
p_{1,2}m_2 = p_{1,1} l_2 + p_{0,2} l_1 \\
p_{2,1}m_2 = p_{1,1} l_1 + p_{2,0} l_2 \\
p_{0,2}(l_1 + l_2 + m_2) = p_{0,1} l_2 + p_{0,3}m_2 \\
p_{2,0}(l_1 + l_2 + m_1) = p_{1,0} l_1 + p_{2,1}m_2 + p_{3,0}m_1 \\
p_{0,3}m_2 = p_{0,2} l_2 \\
p_{3,0}m_1 = p_{2,0} l_1
\end{cases}
$$

Normalization Condition

$$
p_0 + p_{0,1} + p_{1,0} + p_{1,1} + p_{0,2} + p_{2,0} + p_{1,2} + p_{2,1} \\
p_{0,3} + p_{3,0} = 1
$$

• We can solve for the state probability distribution by solving any nine of the ten balance equation along with the equation for the normalization condition

• Job loss probability (or the blocking probability)

$= p_{1,2}+p_{2,1}+p_{3,0}+p_{0,3}$

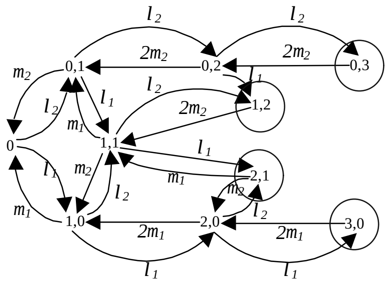
• Other desired probabilities may also be found from these state probabilities. Some examples are -

P{server busy serving low priority job} = $p_{1,0}+p_{2,0}+p_{3,0}$

P{one high priority job in the system} = $p_{0,1}+p_{1,1}+p_{2,1}$

2-Priority M/M/2/3 Queue (Preemptive Priority)

(An example of a queue with finite capacity and multiple servers)



New arrivals will be lost if they come when the system is in any of the circled states

Solve in the usual manner for the system state probabilities

M/M/-/- Queue with Non-preemptive Priority

We can propose two different methods of representing the system state for a M/M/c/K queue of this type with P priority classes.

Approach I : If $P < c$, then this approach gives a more compact representation using a *2P-tuple* than the more general Approach II given next.

State Representation $(n_1,\ldots\ldots,n_P, s_1,\ldots\ldots,s_P)$

where

n_j = number of jobs of class j in system $j=1,\ldots,P$

s_k = number of servers currently busy serving jobs of priority class k $k=1,\ldots,P$

Approach II : *This requires a (P+c)-tuple of the following form*

State Representation $n_1,\ldots\ldots,n_P, s_1,\ldots\ldots,s_c)$

where

n_j = number of jobs of class j in system $j=1,\ldots,P$

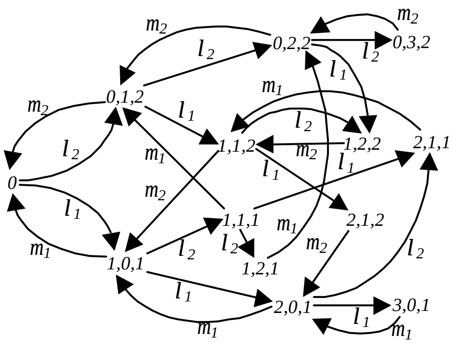
s_k = priority class of the service currently on-going at server k $k=1,\ldots,c$

Note that $n_1+\ldots\ldots+n_P \le K$ for a finite capacity system

We have used the representation of Approach II in the example described subsequently

2-Priority M/M/1/3 Queue (Non-preemptive Priority)

(An example of a single server queue with finite capacity)



State Transition Diagram

The balance equations for this queue are

$$p_0(l_1 + l_2) = p_{0,1,2}m_2 + p_{1,0,1}m_1$$
$$p_{0,1,2}(l_1 + l_2 + m_2) = p_0 l_2 + p_{0,2,2}m_2$$
$$p_{1,0,1}(l_1 + l_2 + m_1) = p_0 l_1 + p_{2,0,1}m_1 + p_{1,1,2}m_2$$
$$p_{1,1,1}(l_1 + l_2 + m_1) = p_{1,0,1} l_2$$
$$p_{1,1,2}(l_1 + l_2 + m_2) = p_{0,1,2} l_1 + p_{1,2,2}m_2 + p_{2,1,1}m_1$$
$$p_{0,2,2}(l_1 + l_2 + m_2) = p_{0,1,2} l_2 + p_{1,2,1}m_1 + p_{0,3,2}m_2$$
$$p_{2,0,1}(l_1 + l_2 + m_1) = p_{1,0,1} l_1 + p_{2,1,2}m_2 + p_{3,0,1}m_1$$

….and ….

$$p_{1,2,2} m_2 = p_{1,1,2} l_2 + p_{0,2,2} l_1$$
$$p_{2,1,2} m_2 = p_{1,1,2} l_1$$
$$p_{2,1,1} m_1 = p_{1,1,1} l_1 + p_{2,0,1} l_2$$
$$p_{1,2,1} m_1 = p_{1,1,1} l_2$$
$$p_{0,3,2} m_2 = p_{0,2,2} l_2$$
$$p_{3,0,1} m_1 = p_{2,0,1} l_1$$

with the following normalization condition

$$p_0 + p_{1,0,1} + p_{0,1,2} + p_{0,2,2} + p_{0,3,2} + p_{2,0,1} + p_{3,0,1} + p_{1,1,1} + p_{1,1,2} + p_{1,2,1} + p_{2,1,1} + p_{1,2,2} + p_{2,1,1} = 1$$

• These equations may be solved on the usual way to obtain the individual state probabilities as per the definition of the system state

• These state probabilities may then be used to compute other performance parameters and probabilities that may be of interest.

• For example, the blocking probability of this system will be given by ($p_{0,3,2} + p_{3,0,1} + p_{1,2,2} + p_{2,1,2} + p_{1,2,1} + p_{2,1,1}$)

• Other, similar probabilities and performance measures may also be calculated

Approach may be extended in the usual fashion to analyze other similar systems as follows -

• More than two priority classes

• Other buffer capacity values or even queues with infinite buffer capacities

• Different capacity limits for the different priority classes

• Queues with more than one server