Queueing Theory Part 2

LA Math Circle
High School II
Dillon Zhi

November 8, 2015

Last week we studied the "M/M/1/K" queue: The system has one server and has a capacity of K customers. (That is, at any time, at most one customer is being served and at most $K-1$ customers are waiting.) The customers arrive at the rate λ and the server serves customers at the rate μ.

In the "steady state", the probability P_j that the system has exactly j customers in it satisfied these equations:

$$\begin{aligned}
\lambda P_0 &= \mu P_1 \\
\lambda P_1 + \mu P_1 &= \mu P_2 + \lambda P_0 \\
\lambda P_2 + \mu P_2 &= \mu P_3 + \lambda P_1 \\
&\cdots
\end{aligned}$$

Solving these equations with the additional equation $P_0 + P_1 + \cdots + P_K = 1$ resulted in the formula

$$P_j = a^j \frac{1-a}{1-a^{K+1}}$$

where $a = \lambda/\mu$.

1. Now consider a system with s servers and a capacity of K customers. (That is, at any time, at most s customesr are being served and at most $K - s$ customers are waiting.) The servers share a queue. The customers arrive at the rate λ and *each* server serves its customers at the rate μ.

   Find the probabilities $P_0, P_1, ..., P_K$, where P_j is the probability that the system has j customers in the steady state.

2. Which of these systems is "best"?

   - s servers with each server serving at the rate μ, sharing a queue
   - s servers with each server serving at the rate μ, each with its own separate queue
   - one fast server serving at the rate $s\mu$

3. a. Show that the mean number of customers in the "M/M/1/K" queue is

$$\frac{a}{1-a} + \frac{K+1}{1-a^{K+1}} a^{K+1}$$

if $a \neq 1$, and the mean number of customers in the system is $K/2$ if $a = 1$.

b. Plot a graph of the mean number of customers versus offered load for $K = 3$.

c. Suppose the system has infinite capacity. What is the formula for the mean number of customers in terms of a? Plot a graph of this relationship.

4. Consider two servers working in parallel, serving customers from a single queue with infinite capacity. Customers arrive in the same manner as in the previous problems, at the rate λ. The faster server serves at the rate μ_1, and the slower server serves at the rate μ_2. A job arriving when both servers are idle is assigned to the faster server. A job being served in one server cannot be transferred to the other.

   a. Determine the distribution of the number of customers in the system.

   b. Use this distribution to find the average number of customers in the system.

   c. Is it ever better to not use the slower machine at all?