16 February 2004

M/M/c Queuing Theory for "The Challenge at Instron"

Assuming Poisson arrivals and exponential service times allows application of an M/M/c model.

The workload is 3483 jobs per year. A single server or mechanic represents 1880 hours per year of available labor. This capacity can also be interpreted as the hours during which the job shop is open for business and able to serve customers. The job arrival rate is then

$$\lambda = \frac{3483 \text{ jobs / year}}{1880 \text{ hours / year}} = 1.8527 \text{ jobs/hour.}$$

From the linear programming model with 6.78 heads, the 3483 jobs require 12,415 hours in assembly time and 334 hours in setup time, totaling 12,749 hours per year. The service rate is therefore

$$\mu = \frac{3483 \text{ jobs / year}}{12{,}749 \text{ hours / year}} = 0.2732 \text{ jobs/hour.}$$

The utilization for the M/M/c system is $\rho = \lambda / c\mu$, in this case equal to 100 percent. From queuing theory, we know that 100 percent utilization results in an infinite queue length. So as a first step to adding reserve capacity, let the staffing level grow to seven heads while retaining the same lot sizes. Now the utilization drops to 0.9688.

The expressions for the expected waiting time and queue lengths are fairly complicated and depend on the probability of there being no jobs in the system upon arrival, p_0 and on the probability of a job having to wait upon arrival, P_Q.

$$p_0 = \left(\sum_{k=0}^{c-1} \frac{(c\rho)^k}{k!} + \frac{(c\rho)^c}{c!\,(1-\rho)} \right)^{-1} \qquad P_Q = \frac{p_0 * (c\rho)^c}{c! * (1-\rho)}$$

In this case, $p_0 = 0.0002$ and $P_Q = 0.91$. The job shop is almost always busy and there's a ninety-one percent chance of finding a queue upon arrival. Expectations of waits and queues are shown below.

Expected waiting time, $E\{W_q\} = \rho * P_Q / (\lambda (1 - \rho)) = 15.38$ hours.

Expected queue length, $E\{L_q\} = \rho * P_Q / (1 - \rho) = 28.5$ jobs.

Expected system time, $E\{W\} = (1/\mu) + \rho * P_Q / (\lambda (1 - \rho)) = 19$ hours.

Expected number in system, $E\{L\} = (c * \rho) + (\rho * P_Q) / (1 - \rho) = 35$ jobs.

The expected values indicate that jobs are completed well within two weeks on average. But we must turn to waiting time distributions to discover how well the job shop operates for individual jobs.

The waiting time distribution for the M/M/c system is also complicated:

$$
Wq(t) = \begin{cases} 1 - \dfrac{c\,(\lambda/\mu)^c}{c!\,(c - \lambda/\mu)}\, p_0 & t = 0 \\[2ex] \dfrac{(\lambda/\mu)^c\,(1 - e^{-(\mu c - \lambda)\,t})}{(c-1)!\,(c - \lambda/\mu)}\, p_0 + W_q(0) & t > 0 \end{cases}
$$

This curve is plotted below with various values of labor heads. For the case with seven heads, 95% of the jobs wait about fifty hours for assembly and so are completed well within the two-week time frame. With eight heads, this value shrinks dramatically to less than ten hours waiting. As staffing drops below seven heads, however, the waiting time distribution is very sensitive to small changes. With a reduction of only one-tenth of a head, to 6.9 heads, the cycle time requirement is not met - only ninety percent of the jobs are completed on time.

[Note: fractional factorials are calculated using Stirling's formula: $n! = e^{-n}\, n^n \sqrt{2\pi n}$, approximately.]