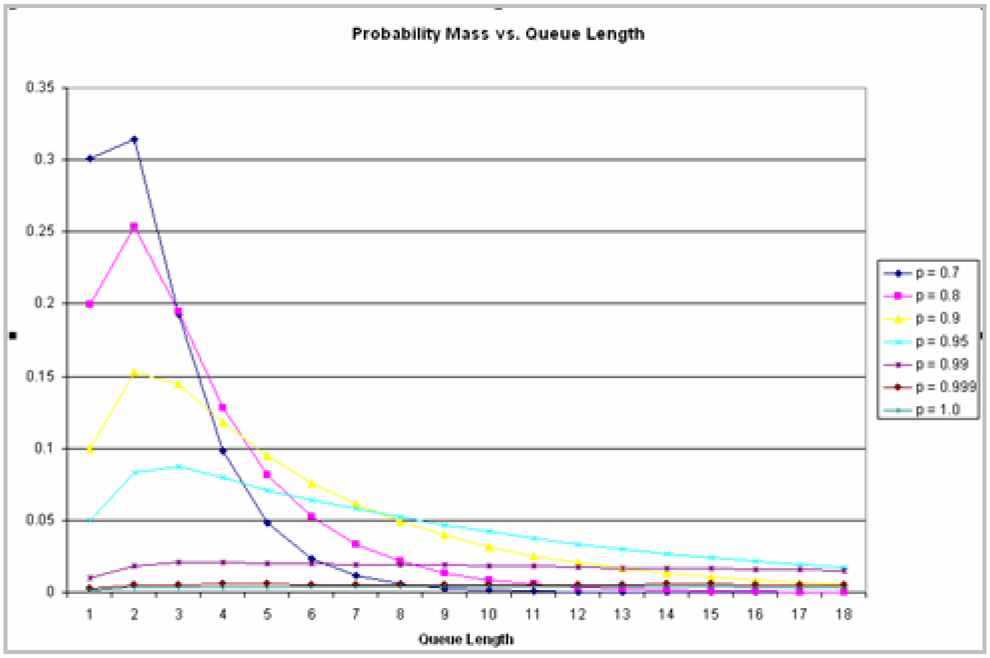
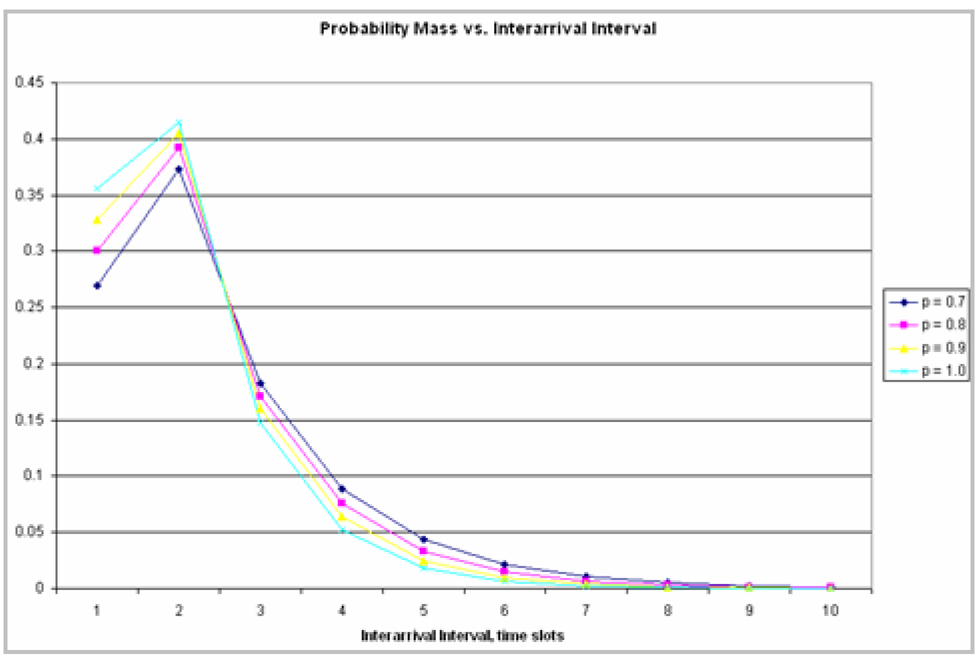
David Bozarth
CES 500
23 May 2006

Simulation of an ATM switch to study queue buffering and queue discipline

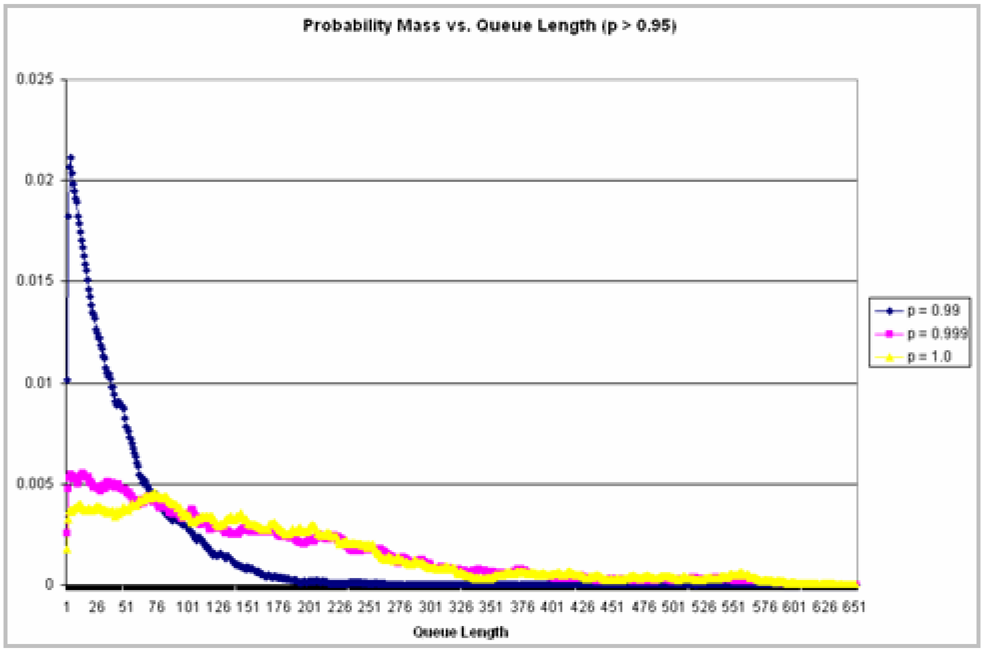
A simulation program (SimSwitch) was built according to the project specifications. The Appendix describes implementation details.

The probability mass functions of queue length and interarrival interval are shown averaged over 16 queues.

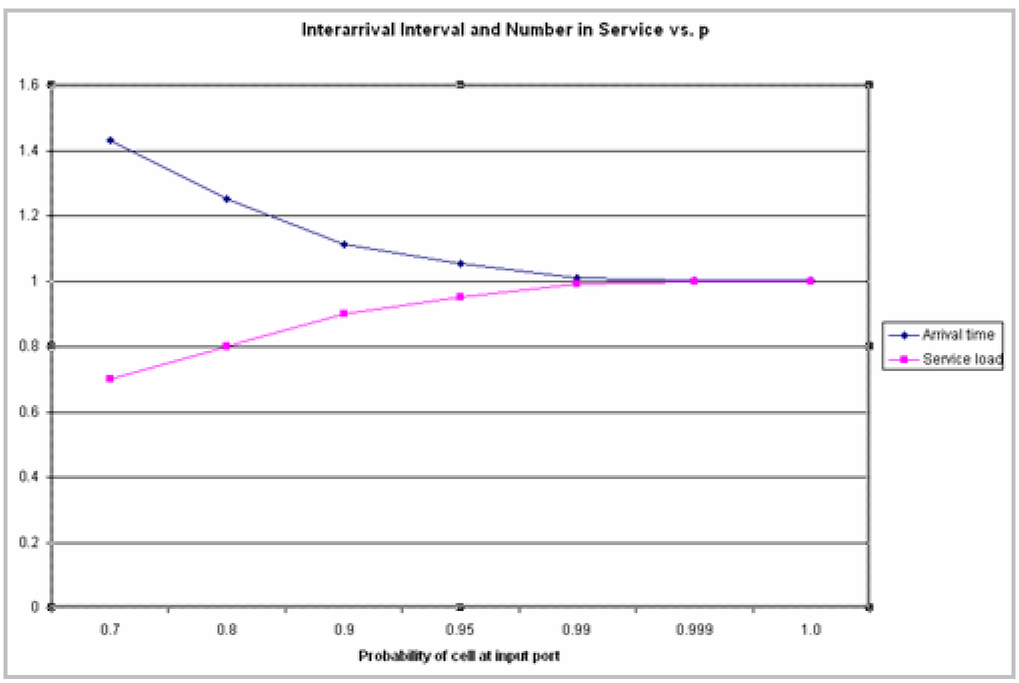




A closer look at queue length for higher values of p:



Additionally, the simulator was programmed to report mean queue length, mean system load, and mean interarrival time for each queue. Mean service load was derived from queue length and system load. These quantities are plotted against values of p.



The graph suggests a reciprocal relation between mean interarrival interval and mean service load:
$\lambda = L_s$.

The table below shows close agreement among **p** and the observed values of $\mathbf{L_s}$ and $\boldsymbol{\lambda}$.

p	0.7	0.8	0.9	0.95	0.99	0.999	1.0
Lq (obs)	1.452382	2.314236	4.699301	9.575826	45.70341	144.3284	170.1619
L (obs)	2.151699	3.11434	5.599597	10.52574	46.69351	145.3259	171.1601
Ta (obs)	1.429981	1.249839	1.110755	1.05272	1.010029	1.00101	1.000012
Ls (obs)	0.699318	0.800104	0.900295	0.949911	0.990104	0.997533	0.998277
λ (obs)	0.69931	0.800103	0.900289	0.94992	0.990071	0.998991	0.999988
Lq (calc)	1.626373	3.202478	8.128649	18.01803	98.72114	988.9308	84909.53
L (calc)	2.325682	4.002581	9.028938	18.96795	99.71121	989.9298	84910.53

Arrivals occur according to a Poisson process with parameter $\boldsymbol{\lambda}$. Each queue is attended by a single server. Since for a single queue in this system exactly one cell is serviced for each time slot, then $\boldsymbol{\mu} = \mathbf{1}$. These considerations suggest that the queueing discipline is M/D/1.

During lecture we found that the following two relationships hold for the M/D/1 queue:

$$L_Q = \rho^2 / (2\,(1 - \rho)) \quad [\text{M/D/1}]$$

$$L = L_Q + \rho \quad [\text{M/D/1}]$$

Since $\boldsymbol{\mu = 1}$ implies $\boldsymbol{\lambda = \rho}$, then according to the above relationships we should have

$$L_Q = 0.81 \text{ and } L = 1.51 \quad (p = 0.7)$$

$$L_Q = 4.05 \text{ and } L = 4.95 \quad (p = 0.9)$$

Comparing the table data with the above relationships, the claim of M/D/1 discipline is *not supported*.

Consider the possibility that the queueing discipline is M/M/1. The table above shows relationships between observed values of L_Q and L, and values of L_Q and L that were calculated (assuming M/M/1) using the observed mean arrival rate $\boldsymbol{\lambda}$.

Consider the following justification for the claim that $\boldsymbol{\lambda} = \mathbf{L_s}$ for an M/M/1 queue with $\boldsymbol{\mu} = \mathbf{1}$.

Since $\mu = 1$, then $\rho = \lambda \Rightarrow L_q = \dfrac{\lambda}{1-\lambda}$ for M/M/1.

Then $L_s = L - L_q = \dfrac{\lambda}{\mu - \lambda} - \dfrac{\lambda^2}{1-\lambda}$ for M/M/1.

$$L_s = \frac{\lambda(1-\lambda) - \lambda^2(\mu-\lambda)}{(\mu-\lambda)(1-\lambda)} = \frac{\lambda^3 - 2\lambda^2 + \lambda}{(\mu-\lambda)(1-\lambda)} = \frac{\lambda(\lambda-1)^2}{(\lambda-\mu)(\lambda-1)}$$

$L_s = \dfrac{\lambda(1-\lambda)}{(\mu-\lambda)} = \lambda$ for M/M/1, and $\mu = 1$

The preceding observations and reasoning support the claim that the queue discipline is **M/M/1**.

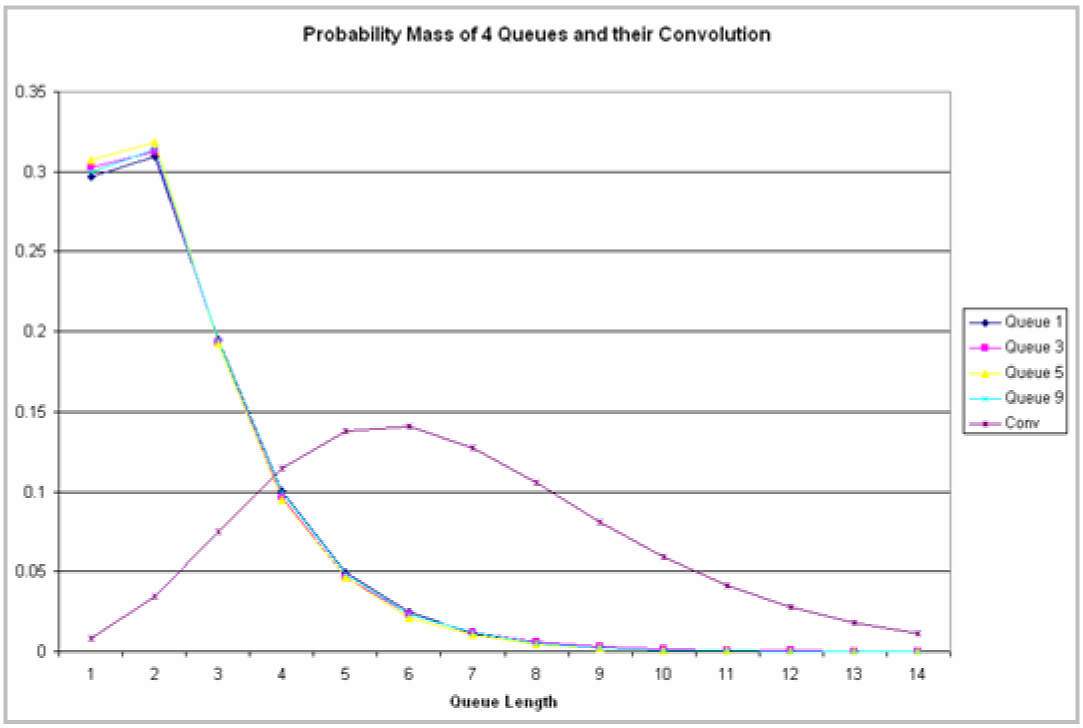
Agreement between observed and calculated (assuming M/M/1) values of L_Q and L is discernable at light arrival loading, but diverges for heavier arrival loading. This is because of the identity **λ = ρ** which approaches unity as **p** approaches unity:

$L_Q = \rho^2 / (1 - \rho)$ [M/M/1]

$L = \rho / (1 - \rho)$ [M/M/1]

Four output ports sharing a single queue

I applied the z-transform $\mathbf{p_k \Leftrightarrow p_k{*}z^k}$ to the individual queue length vectors for four queues, which yielded four polynomials. I used the Mathematica **Expand** function to find the product of the four polynomials, then applied the Mathematica **InverseZTransform** function to get the convolution of the four vectors.



With this configuration, the range of queue lengths is extended and the mean queue length is tripled. Because the probability of queue length zero is greatly diminished, the probability of "dead time" for the server attending the composite queue, is reduced. Therefore this discipline is more efficient purely from a cost viewpoint. (If other considerations like reneging are in play, the advantage could be reduced.)

With the composite queue, the service rate for an individual queue is effectively quadrupled, so it appears from the perspective of a customer in the queue that there are four servers. So the queueing discipline is **M/G/4**.

APPENDIX: Simulation Tools and Techniques

The simulation program was built on the basic Java 1.4 development platform. It uses numerous one-dimensional arrays and a special Java data structure ArrayList, which combines features of an array and a list. An array of ArrayList objects (one ArrayList for each queue) was used to store the number of occurrences of each queue length. A second array of ArrayList objects stored the number of occurrences of each interarrival interval.

The timing logic of the program was designed as follows:

```
During each time slot, the following events occur in the indicated order:
    0. For each output queue:
         If queue length is nonzero, then queue length is decremented by one & number in
service set to one. Otherwise, number in service is set to zero.
    1. A received cell appears at each of zero or more inputs with probability P.
    2. For each such cell:
        a. One destination port is assigned as a uniformly distributed random variable.
        b. The queue length at the destination port is incremented by one.
        c. The list of occurrences of interarrival intervals for the destination queue is
updated.
        d. The most recent arrival time for the destination queue is updated.
    3. For each output queue:
        a. The list of occurrences of queue lengths for that queue is updated.
        b. The total 'number in system' for that queue is updated.

After all time slots have been processed, three disk files are written in comma-separated
text format:
    1. Probability mass function of queue length for each queue.
    2. Probability mass function of interarrival interval for each queue.
    3. Aggregate queue parameters (Lq, L, Ta [mean interarrival interval]) for each queue.
```

Separate runs of the simulator produced comma-separated text files for each value of **p** under test. The text files were imported into a spreadsheet program. For each queue length and interarrival interval, the pmf values were averaged over all queues. The resulting pmf values and queue parameters were used to generate the plots and table shown above.

Generation and processing of the z-transform polynomials was a labor-intensive process involving many manual cut-and-paste operations between software applications. Better knowledge about data transfer and programming facilities in Mathematica or MATLAB could make future processes like this much easier.