Introduction to Queuing Networks

Solutions to Problem Sheet 6

---

1. (a) Node 1 refers to the bookshop and Node 2 to the check-out. Node 1 is an $M/G/\infty$ queue while Node 2 is an $M/M/1$-FIFO queue. The external arrival rates are $\gamma_1 = 1$ and $\gamma_2 = 0$ per minute, while the service rates are $\mu_1 = 1/15$ and $\mu_2 = 1/2$ per minute (reciprocals of the corresponding mean service times). The routing probabilities are $p_{10} = 0.7$, $p_{12} = 0.3$, $p_{20} = 0.9$ and $p_{21} = 0.1$. The traffic equations are:

$$\lambda_1 = 1 + 0.1\lambda_2, \quad \lambda_2 = 0.3\lambda_1,$$

which have the unique solution $\lambda_1 = 100/97$, $\lambda_2 = 30/97$.

(b) We first compute the mean number in each of the two queues (i.e., the mean number of browsing customers, and the mean number waiting to pay at the check-out). Node 1 is an $M/G/\infty$ queue with load $\rho_1 = \lambda_1/\mu_1 = 1500/97$. Hence, the invariant distribution of N_1, the number of customers at this queue, is Poisson(ρ_1), with mean $E[N_1] = \rho_1 = 1500/97$. Node 2 is an $M/M/1$ queue with load $\rho_2 = \lambda_2/\mu_2 = 60/97$. So the mean number at this queue (using known results for the $M/M/1$ queue) is $E[N_2] = \rho_2/(1-\rho_2) = 60/37$. Hence, the expected total number of customers in the bookshop is

$$E[N] = E[N_1] + E[N_2] \approx 15.46 + 1.62 = 17.1.$$

The total external arrival rate is $\gamma = \gamma_1 + \gamma_2 = 1$ customer per minute. Hence, by Little's law, the mean sojourn time is $E[W] = E[N]/\gamma = 17.1$ minutes.

(c) The number of customers who are browsing is the number in the first queue, which is N_1. As noted above, the invariant distribution of N_1 is a Poisson distribution with mean $\rho_1 = 15.46$. Since new customers enter the shop according to a Poisson process, it follows by the PASTA property (Poisson arrivals see time averages) that the distribution of customers seen by a new customer upon arrival is the same as the invariant distribution. Hence, the new customer will not find a free chair if $N_1 \geq 20$. Using the hint, we approximate the probability of this event as

$$\begin{aligned} P(N_1 \geq 20) &\approx \frac{21}{21-15.46}\frac{1}{\sqrt{40\pi}}\exp\Big(20 - 15.46 - 20\log_e \frac{20}{15.46}\Big) \\ &\approx 0.338\exp(4.54 - 20\log_e 1.294) \approx 0.184. \end{aligned}$$

2. (a) Node 1 denotes the prison system, and the external arrival rate γ_1 to this node is 4000 per month. Node 1 is an $M/G/\infty$ queue, with mean service time $1/\mu_1$ equal to 9 months. Individuals leaving prison depart the criminal justice system with probability 0.4 and have probability 0.6 of becoming recidivists. This is denoted by their joining Node 2, which has no external arrivals ($\gamma_2 = 0$), and is an $M/M/\infty$ queue with mean service time $1/\mu_2$ equal to 24 months. All recidivists eventually re-enter prison. (Equivalently, the picture could be drawn with all departures from Node 1 going to Node 2, and departures from Node 2 going to Node 1 with probability 0.6 and leaving the system with probability 0.4.)

(b) For the system as described above, the traffic equations are

$$\lambda_1 = \gamma_1 + \lambda_2, \quad \lambda_2 = 0.6\lambda,$$

which have the solution $\lambda_1 = 10,000$ per month and $\lambda_2 = 6000$ per month. (The equivalent picture described above would have $\lambda_1 = \lambda_2 = 10,000$ per month.)

(c) Prisons constitute an $M/G/\infty$ queue with arrival rate $\lambda_1 = 10,000$ per month, and service rate $\mu_1 = 1/9$ per month. The $M/G/\infty$ queue has invariant distribution which is Poisson with parameter $\rho = \lambda/\mu$.

The mean of this Poisson distribution is ρ, which in this case is 90,000.

(d) The variance of a Poisson(ρ) distribution is also ρ. Hence, the standard deviation of the number in prison is $\sqrt{\rho} = 300$. Using the suggested approximation, there should be 90,900 prison places provided in the UK.

3. (a) Node 1 refers to the disk and Node 2 to the cache. The external arrival rate to the disk is $\gamma_1 = 4$ (80% of total arrival rate), while that to the cache is $\gamma_2 = 1$ (20% of all requests, arriving at rate 5, can be served directly from cache). Node 1 has a single server operating a FIFO (first-in-first-out) / FCFS (first-come-first-served) policy while node 2 has a single server operating a PS (processor sharing) service discipline. The routing probabilities are given by $p_{12} = 1$, $p_{20} = 1$ and all other routing probabilities are zero.

The mean service time at the first queue is 200 ms = 0.2s, so the service rate is $\mu_1 = 1/0.2 = 5$. The mean service time at the second queue is given by

$$\begin{aligned} E[S_2] &= \int_0^\infty t f_T(t)dt = 200 \int_0^\infty \frac{t/100}{(1+\frac{t}{100})^3} \frac{dt}{100} \\ &= 200 \int_0^\infty \Big(\frac{1}{(1+x)^2} - \frac{1}{(1+x)^3} \Big) dx = 100, \end{aligned}$$

in units of milliseconds. Hence, the mean service time at the second queue is 0.1s, so the service rate is $\mu_2 = 10$.

(b) We first compute the mean number of jobs (requests) at each queue. Node 1 is an $M/M/1$-FIFO queue with arrival rate $\lambda_1 = 4$ and service rate $\mu_1 = 5$. Hence, its load $\rho_1 = \lambda_1/\mu_1 = 0.8$. By the known formula for the mean queue length at an $M/M/1$ queue,

$$E[N_1] = \frac{\rho_1}{1-\rho_1} = \frac{0.8}{0.2} = 4.$$

Node 2 is an $M/G/1$-PS queue with total arrival rate $\lambda_2 = 4 + 1 = 5$ (including both external arrivals at rate 1, and re-routed customers from Node 1 at rate 4) and service rate $\mu_2 = 10$, so that its load is $\rho_2 = 0.5$. Moreover, by the insensitivity property of $M/G/1$-PS queues, it has the same invariant distribution as an $M/M/1$ queue with the same load. Consequently, it has the same mean number of jobs,

$$E[N_2] = \frac{\rho_2}{1-\rho_2} = \frac{0.5}{0.5} = 1.$$

Hence, the mean number of jobs in the system, in equilibrium, is $E[N] = E[N_1] + E[N_2] = 5$. The total external arrival rate of jobs is $\gamma = 5$ per second. Consequently, by Little's law, the mean sojourn time of a job is $E[W] = E[N]/\gamma$ = 1 second.

(c) Suppose a FIFO discipline is used at the second queue, which has total arrival rate $\lambda_2 = 5$ and service rate $\mu_2 = 10$. In order to compute the mean sojourn time at the second queue, we use the Pollaczek-Khinchin formula, which says that

$$E[W_2] = \frac{\lambda_2 E[S_2^2]}{2(1-\rho_2)} + E[S_2] = 5E[S_2^2] + 0.1.$$

It remains to compute $E[S_2^2]$, the second moment of the service time distribution. We have

$$E[S_2^2] = \int_0^\infty \frac{0.02t^2}{(1+\frac{t}{100})^3} dt = \infty,$$

since the integrand is asymptotic to a constant times $1/t$, whose integral diverges. Therefore, the mean sojourn time at the second queue is infinite, which implies that the mean sojourn time in the system as a whole is infinite. We conclude that in this case FIFO is not preferable to PS!

4. We will model the system as a queueing network with 3 nodes. Node 1 corresponds to the pre-processing stage and is a $\cdot/G/1 - LIFO$ (last-in-first-out) queue with service rate $\mu_1 = 1/5$ per minute. (The service rate is the reciprocal of the mean service time.) Node 2 corresponds to the manual server (Treatment 1) and is a $\cdot/M/1$ queue with service rate $\mu_2 = 1/20$ per minute. Node 3 corresponds to Treatment 2 and is a $\cdot/G/\infty$ queue with $\mu_3 = 1/60$ per minute.

The external arrival rate is γ and the arrival process is Poisson. Routing decisions are Markovian, independent of current and past states and other routing decisions. The routing probabilities are $r_{12} = \alpha$, $r_{13} = 1 - \alpha$, $r_{20} = 1$, $r_{31} = 0.1$ and $r_{30} = 0.9$, with all other routing probabilities being zero. You might want to draw a diagram to represent the queueing network, marking the service disciplines and service rates at the nodes, and marking the routing probabilities on the edges between nodes.

Recall that $\cdot/M/1-FIFO$, $\cdot/G/\infty$, $\cdot/G/1-PS$ and $\cdot/G/1-LIFO$ queues are quasi-reversible: if the arrival process is Poisson, then so is the departure process in equilibrium. Hence, the queues at all 3 nodes in our system are quasi-reversible.

Since we have **Poisson arrivals, Markov routing and quasi-reversible queues**, the queueing network is a **generalised Jackson network** (or BCMP network). This means that its **invariant distribution is product-form:**

$$\pi(n_1, n_2, n_3) = \prod_{i=1}^{3} \pi_i(n_i) = (1-\rho_1)\rho_1^{n_1} \cdot (1-\rho_2)\rho_2^{n_2} \cdot \frac{\rho_3^{n_3}}{n_3!} e^{-\rho_3}.$$

Here $\rho_i = \lambda_i/\mu_i$ is the load at the i^{th} server, and λ_i is the total arrival rate at the i^{th} server, including both external and re-routed arrivals.

The first two queues have an invariant distribution which is geometric because they are single-server queues, while the third queue, being an infinite-server queue, has an invariant distribution which is Poisson.

The next step is to compute the arrival rates λ_i at the different queues by solving the network traffic equations. Assuming that the system is stable (all the work entering each queue eventually leaves it), the traffic equations are:

$$\lambda_1 = \gamma + 0.1\lambda_3, \quad \lambda_2 = \alpha\lambda_1, \quad \lambda_3 = (1-\alpha)\lambda_1,$$

where γ is the external arrival rate. Solving these equations, we get

$$\lambda_1 = \frac{\gamma}{0.9 + 0.1\alpha}, \quad \lambda_2 = \frac{\alpha\gamma}{0.9 + 0.1\alpha}, \quad \lambda_3 = \frac{(1-\alpha)\gamma}{0.9 + 0.1\alpha}.$$

(a) For stability, we require $\rho_1 < 1$ and $\rho_2 < 1$. There is no restriction on ρ_3 because node 3 is an infinite-server queue. Substituting $\mu_1 = 1/5$ and $\mu_2 = 1/20$, we obtain the stability condition

$$\gamma < \min\left\{0.18 + 0.02\alpha, \frac{0.045}{\alpha} + 0.005\right\}.$$

The first term is smaller if $\alpha < 1/4$ while the second term is smaller if $\alpha > 1/4$. Hence, we can write the stability condition as

$$\gamma < \gamma(\alpha) = \begin{cases} 0.18 + 0.02\alpha, & 0 \le \alpha \le 1/4, \\ (0.045/\alpha) + 0.005, & 1/4 \le \alpha \le 1. \end{cases}$$

Clearly, $\gamma(\alpha)$ is maximised at $\alpha = 1/4$, and the maximum value is $\gamma(1/4) = 0.185$.

(b) We now take $\alpha = 0.25$, and $\gamma = \gamma(\alpha)/2 = 0.0925$. Substituting these values in the traffic equations, we get

$$\lambda_1 = 0.1, \lambda_2 = 0.025, \lambda_3 = 0.075, \quad \Rightarrow \quad \rho_1 = 0.5, \rho_2 = 0.5, \rho_3 = 4.5.$$

Now, the mean number in system is the sum of the mean numbers in each queue, which is $\rho/(1-\rho)$ if the queue size distribution is geometric, and ρ if it is Poisson. Hence,

$$E[N] = \frac{\rho_1}{1-\rho_1} + \frac{\rho_2}{1-\rho_2} + \rho_3 = 6.5.$$

Finally, the mean length of time needed to process a single item is the mean sojourn time of an item, which is given by **Little's law**:

$$E[W] = \frac{1}{\gamma} E[N] = \frac{6.5}{0.0925} = 70.27 \text{ minutes.}$$

Note that Little's law is valid, not just for a single queue, but also for networks of queues. You can think of the network as a black box. External arrivals enter the black box, are processed for some time, and then leave the black box. The processing times may have a complicated distribution, and may not even be independent of each other or of the inter-arrival times. Nevertheless, Little's law still applies. It is extremely general and only requires stability of the queueing system.

5. (a) There are N nodes in series. All nodes are $./G/\infty$ queues. The service time is general as its distribution has not been specified. There are (effectively) infinitely many servers at each node. As each job gets its own server, the question of service discipline is moot.

For departures from each intermediate node (i.e., nodes between 1 and $N-1$ inclusive, there is probability 0.5 of going to the next queue, probability 0.25 of returning to the first queue (due to an error having been found in the application), and probability 0.25 of leaving the system (giving up). Departures from queue N leave the system (with a driving licence). Hence the routing probabilities are

$$r_{j,j+1} = \frac{1}{2},\ r_{j,1} = \frac{1}{4},\ r_{j,0} = \frac{1}{4}\ 1 \le j \le N-1, \quad r_{N0} = 1.$$

(b) Let λ_j denote the total arrival rate into the j^{th} node. The traffic equations are:

$$\begin{aligned} \lambda_1 &= \eta + \sum_{j=1}^{N-1} \frac{1}{4}\lambda_j, \\ \lambda_j &= \frac{1}{2}\lambda_{j-1},\ j = 2, \ldots, N. \end{aligned}$$

Thus, $\lambda_j = 2^{-(j-1)}\lambda_1$ for $j = 1, \ldots, N$ and

$$\lambda_1 = \eta + \frac{1}{4}\sum_{j=1}^{N-1} 2^{-(j-1)}\lambda_1 = \eta + \frac{1}{2}(1 - 2^{-(N-1)})\lambda_1,$$

from which it follows that

$$\lambda_1 = \frac{2^N}{2^{N-1}+1}\eta, \quad \lambda_j = \frac{2^{N+1-j}}{2^{N-1}+1}\eta,\ 2 \le j \le N.$$

(c) As each node is an $M/G/\infty$ queue, the number of customers in node j in steady state is Poisson with mean λ_j/μ_j. Here $\mu_j = 1$ for all j. Therefore, the total number of customers in system in steady state is given by

$$\sum_{j=1}^{N} \lambda_j = \frac{2^N\eta}{2^{N-1}+1}\sum_{j=1}^{N} 2^{-j+1} = \frac{2(2^N-1)}{2^{N-1}+1}\eta.$$

As the external arrival rate is η, it follows by Little's law that the mean sojourn time is

$$\frac{2^N}{2^{N-1}+1}$$

weeks. As all queues are infinity server queues, this quantity is finite for all η.

(d) (not seen) The rate at which applicants leave the system with a licence is the departure rate from node N, which is the arrival rate into this node, namely

$$\lambda_N = \frac{2\eta}{2^{N-1}+1}.$$

As the rate at which applicants enter the system is η, it follows that the fraction of applicants who are successful is

$$\frac{2}{2^{N-1}+1}.$$