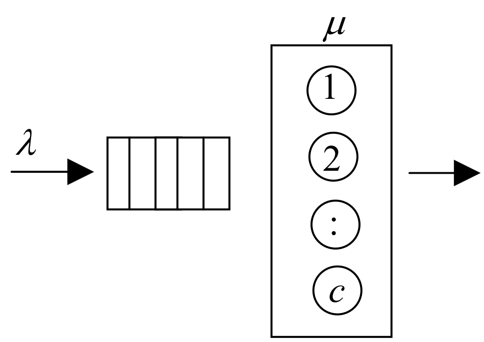
Queueing Theory (3)

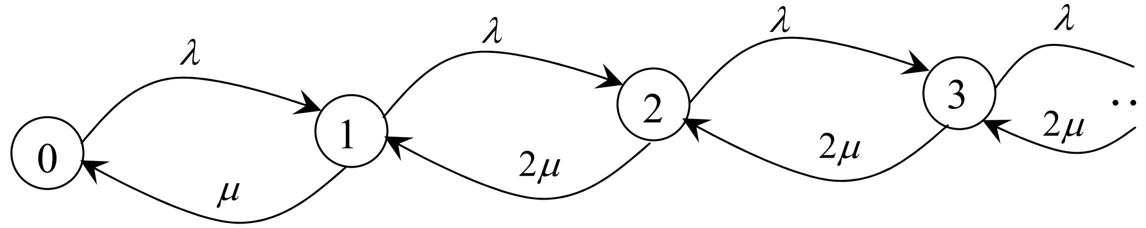
- **The *M*/*M*/*c* queue**
  - This is a queue with a Poisson arrival process with rate λ, exponential service times with rate μ and c servers.
  - It is a generalization of $M/M/1$ with multi-servers.



  - The number of customers in the $M/M/c$ system $L(t)$ is a birth death process with $\lambda_n = \lambda$, and

$$\mu_n = \begin{cases} n\mu, & \text{if } n < c \\ c\mu & \text{if } n \geq c \end{cases}$$

  - The expression for μ_n follows since the minimum of n exponential rvs with rate μ is exponential with rate $n\mu$.
  - The transition diagram for $c = 2$ is shown below.



- Recall that $\rho = \lambda/(c\mu)$ is the traffic intensity.
- Define $a = \lambda/\mu$. This is the mean number of busy servers.
- In the following we assume $\rho < 1$.
- Applying the general flow balance equation for a birth-death process, the limiting probabilities are given by

$$P_0 = \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_{n-1}\lambda_{n-2}\ldots\lambda_1\lambda_0}{\mu_n\mu_{n-1}\ldots\mu_2\mu_1}\right)^{-1}$$

$$= \left(1 + \sum_{n=1}^{c-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \sum_{n=c}^{\infty} \frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n\right)^{-1}$$

$$= \left(\sum_{n=0}^{c-1} \frac{a^n}{n!} + \sum_{n=c}^{\infty} \frac{a^n}{c!c^{n-c}}\right)^{-1}$$

$$= \left(\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!}\sum_{n=c}^{\infty} \frac{a^{n-c}}{c^{n-c}}\right)^{-1}$$

$$= \left(\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!}\sum_{m=0}^{\infty} \rho^m\right)^{-1}$$

$$= \left(\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!(1-\rho)}\right)^{-1}$$

- Then,

$$P_n = \begin{cases} \dfrac{a^n}{n!} P_0, & \text{if } n < c \\ \dfrac{a^n}{c!c^{n-c}} P_0 & \text{if } n \geq c \end{cases}$$

- Then, the mean number in queue is

$$\boxed{L_q} = \sum_{n=c}^{\infty}(n-c)P_n = \sum_{n=c}^{\infty}(n-c)\frac{a^n}{c!c^{n-c}}P_0 = \frac{a^c P_0}{c!}\sum_{n=c}^{\infty}(n-c)\frac{a^{n-c}}{c^{n-c}}$$

$$= \frac{a^c P_0}{c!}\sum_{m=0}^{\infty} m\rho^m = \boxed{\frac{a^c \rho}{c!(1-\rho)^2}P_0}$$

since $\sum_{m=0}^{\infty} m\rho^m = \frac{\rho}{(1-\rho)^2}$

- Then, Little's law implies that the mean waiting time is

$$W_q = \frac{L_q}{\lambda} = \frac{a^c \rho}{\lambda c!(1-\rho)^2}P_0 = \frac{a^c}{c!(c\mu)(1-\rho)^2}P_0$$

- In addition, the mean number in the system is

$$L = L_q + a = a + \frac{a^c \rho}{c!(1-\rho)^2}P_0$$

- And the mean time in the system, is

$$W = W_q + \frac{1}{\mu} = \frac{1}{\mu} + \frac{a^c}{c!(c\mu)(1-\rho)^2}P_0$$

- The probability that all servers are busy is

$$P_{c+} = \sum_{n=c}^{\infty} P_n = \sum_{n=c}^{\infty}\frac{a^c}{c!c^{n-c}}P_0 = \frac{a^c}{c!(1-\rho)}P_0 = \frac{P_c}{1-\rho}$$

- Let T_q be the waiting time in queue of a customer. Then,

$$P\{T_q > t\} = \frac{a^c P_0}{c!(1-\rho)}e^{-c\mu(1-\rho)t}$$

- This can be shown similar to the $M/M/1$ case.

- **Example 6 (enhanced bank service)**
  - In the $M/M/1$ bank model with $\lambda = 9$ customers/hour and $\mu = 10$ customers/hour. We find that the mean waiting time is $W_q = 54$ mins.
  - Management is considering adding more servers to bring the mean waiting time to less than 5 minutes. How many more servers should be added?
  - Try adding another server. For this $M/M/2$ system, $a = 0.9$ and $\rho = a/2 = 0.45$. Then,

$$P_0 = \left(\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!(1-\rho)} \right)^{-1} = \left(1 + 0.9 + \frac{0.9^2}{2!(1-0.45)} \right)^{-1} = 0.379$$

$$W_q = \frac{a^c}{c!(c\mu)(1-\rho)^2} P_0 = \frac{0.9^2}{2!(2\times 10)(1-0.45)^2} 0.379$$

$$= 0.025 \text{ hours} = 1.5 \text{ mins}$$

  - Adding one more server achieves the desired service level.

- **Example 7.**
  - An airline is planning a new telephone reservation center. Each agent will have a reservations terminal and can serve a typical caller in 5 minutes, the service time being exponentially distributed. Calls arrive randomly and the system has a large message buffering system to hold calls that arrive when no agent is free. An average of 36 calls per hour is expected during the peak period of the day. The design criterion for the new facility is that the probability a caller will find all agents busy must not exceed 0.1 (10%).
  - How many terminals should be provided?
  - This can be modeled as an $M/M/c$ with $\lambda = 36$ calls /hour, $\mu = 60/5 = 12$ calls /hour, and c is to be determined such that $P_{c+} < 0.1$.
  - The minimum number of terminals, c, needed is one that achieves stability. That is,

$$\rho = \lambda/(c\mu) < 1 \Rightarrow 36/(12c) < 1 \Rightarrow c > 3.$$

  - Try $c = 4$, then $\rho = 36/48 = 0.75$ and $a = 4\rho = 3$.

$$P_0 = \left(\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!(1-\rho)} \right)^{-1} = \left(1 + 3 + \frac{3^2}{2} + \frac{3^3}{6} + \frac{3^4}{4!(1-0.75)} \right)^{-1}$$
$$= 0.037736$$

$$P_{c+} = \frac{a^c}{c!(1-\rho)} P_0 = \frac{3^4}{4!(1-0.75)} 0.037736 = 0.509$$

  - So, four terminals won't do it.

- Try $c = 5$. Repeating the same computations yields

$$P_{c+} = 0.232 .$$

- Try $c = 6$. Repeating the same computations yields

$$P_{c+} = 0.0991 .$$

- So, six terminals are needed to achieve the desired service level.