M/M/1 Queue

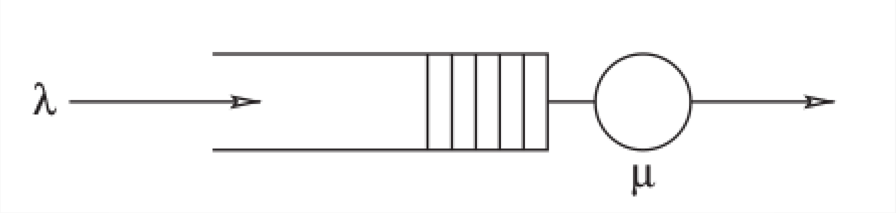
Lecture 5

Communication Theory III

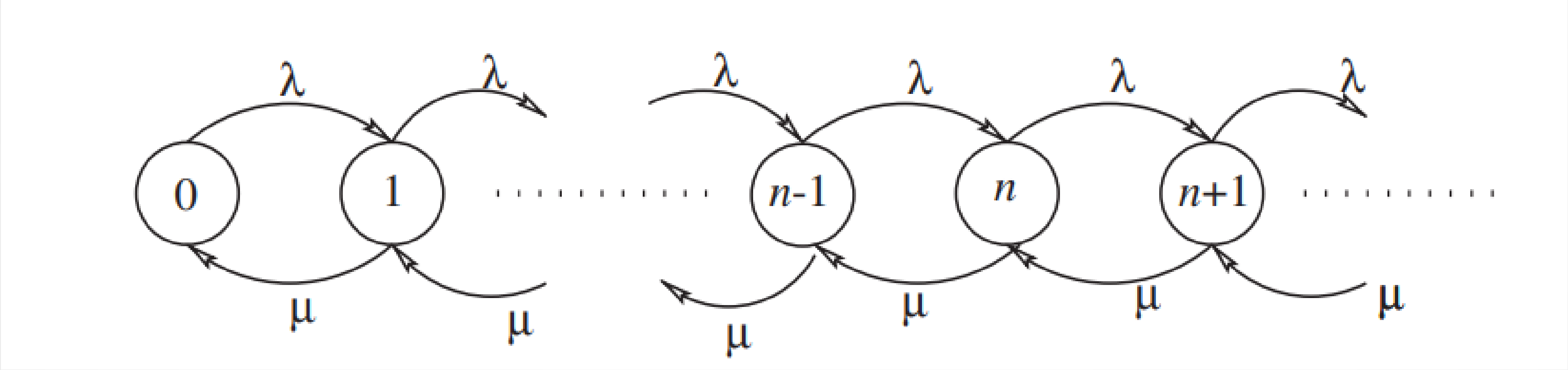
Eng. (Mrs.) PN Karunanayake

M/M/1 Queue

- The simplest of all queueing systems is the **M/M/1** queue. This is a **single server** queue with **FCFS** scheduling discipline, an arrival process that is **Poisson** and service time that is **exponentially distributed.**



State Diagram



Steady State Solution

- The probability that the system is in any given state n at any time,

$$P_n(t) = \text{Prob}\{n \text{ in system at time } t\}$$

In steady state

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0.$$

$$p_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n},$$

where $\rho \equiv \lambda/\mu$, is the utilization factor, a measure of the average use of the service facility.

Steady State Solution

$$\sum_{n=0}^{\infty} \rho^n$$

The geometric series converges if and only if $|\rho| < 1$. Therefore the existence of a steady state solution demands that $\rho = \lambda/\mu < 1$.

If $\lambda > \mu$ the, server will get further and further behind. Thus the system size will keep increasingly without limit.

$$\sum_{n=0}^{\infty} \rho^n = \frac{1}{1-\rho}$$

Hence $p_0 = 1 - \rho$

Steady State Solution M/M/1

- Steady-state solution for the M/M/1 queue is given by,

$$p_n = \rho^n (1 - \rho) \quad \text{for} \quad \rho = \lambda/\mu < 1,$$

- Example:
- The probability that the queue contains at least k customers,

$$\text{Prob}\{n \geq k\} = \sum_{i=k}^{\infty} p_i = (1 - \rho) \sum_{i=k}^{\infty} \rho^i = (1 - \rho) \left(\sum_{i=0}^{\infty} \rho^i - \sum_{i=0}^{k-1} \rho^i \right)$$

$$= (1 - \rho) \left(\frac{1}{1 - \rho} - \frac{1 - \rho^k}{1 - \rho} \right) = \rho^k.$$

Performance measures concerning the M/M/1 Queue

- **Mean Number in System.**

- Let N be the random variable that describes the number of customers in the system at steady state, and let L = E[N].

$$L = (1-\rho)\frac{\rho}{(1-\rho)^2} = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}.$$

- **Variance of Number in System**

$$\mathrm{Var}\,[N] = \rho\frac{1+\rho}{(1-\rho)^2} - \left(\frac{\rho}{1-\rho}\right)^2 = \frac{\rho}{(1-\rho)^2}.$$

Performance measures concerning the M/M/1 Queue

- **Mean Queue Length**
- Let N_q be the random variable that describes the number of customers waiting in the queue at steady state, and let $L_q = E[N_q]$

$$L_q = \rho L = L - \rho.$$

- **Average Response Time**
- Let R be the random variable that describes the response time of customers in the system. Little's law states that the mean number of customers in a queueing system in steady state is equal to the product of the arrival rate and the mean response time.

$$E[N] = \lambda E[R] \quad (L = \lambda W).$$

$$E[R] = \frac{1}{\lambda} E[N] = \frac{1}{\lambda} \frac{\rho}{1-\rho} = \frac{1/\mu}{1-\rho} = \frac{1}{\mu - \lambda}$$

Performance measures concerning the M/M/1 Queue

- **Average Waiting Time**

$$L_q = \lambda W_q$$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{\rho}{\mu - \lambda}.$$

- Example:
- Compute these performance measures using the parameters λ = 4, μ = 5.

Example 1

The arrival pattern of cars to the local oil change center follows a Poisson distribution at a rate of four per hour. If the time to perform an oil change is exponentially distributed and requires on average of 12 minutes to carry out, what is the probability of finding more than 3 cars waiting for the single available mechanic to service their car ?

Example 2

In a tool crib, workers come to take tools at 4/hours on the average. Waiting for them costs 10 $ per hour. The service time per worker is in the tool crib is 12mins. What will be total waiting cost of the workers per day if it is 8 hours a day?

Assume M/M/1 queueing system

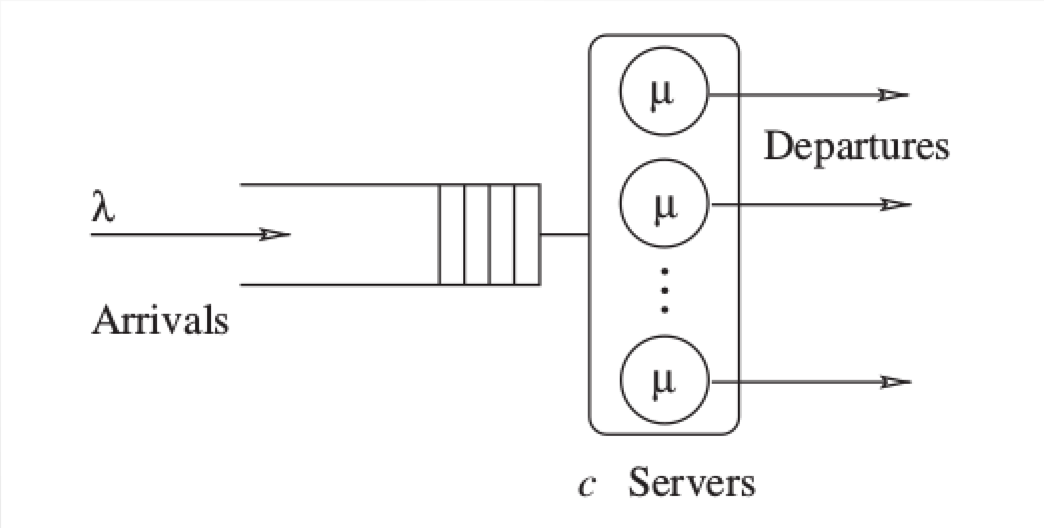
Example 3

An overhead crane of ABC Ltd. Moves jobs from one machine to another & must be used every time a machine requires loading or unloading. The demand for service is random. Data taken by recording the elapsed time between service calls followed an exponential distribution having a mean of a call every 24 minutes. In a similar manner the actual service time of loading or unloading took an average of 8 minutes. If the machine time is valued at 8.50$ per hour, how much does the down time cost every day.

Assume 8 working hours per day.

The M/M/c Queue – Multi Server Systems

- Consider c identical servers for the M/M/c queue (queue with parallel channels).
- Substitution of $\lambda_n = \lambda$ and $\mu_n = \mu$ for all n turns the general birth-death system into the M/M/1 Queue.



The M/M/c Queue

- Customers arrive according to a **Poisson process** with rate $\lambda_n = \lambda$ for all n.
- Served in **first-come first-served** order by any available server.
- Each of the **c servers provides independent and identically distributed** exponential service at rate μ.
- When the number of customers is greater than or equal to c, i.e., **n ≥ c**, all the servers are busy and the **effective service rate**, also called the **mean system output rate (MSOR)**, is equal to **cμ.**

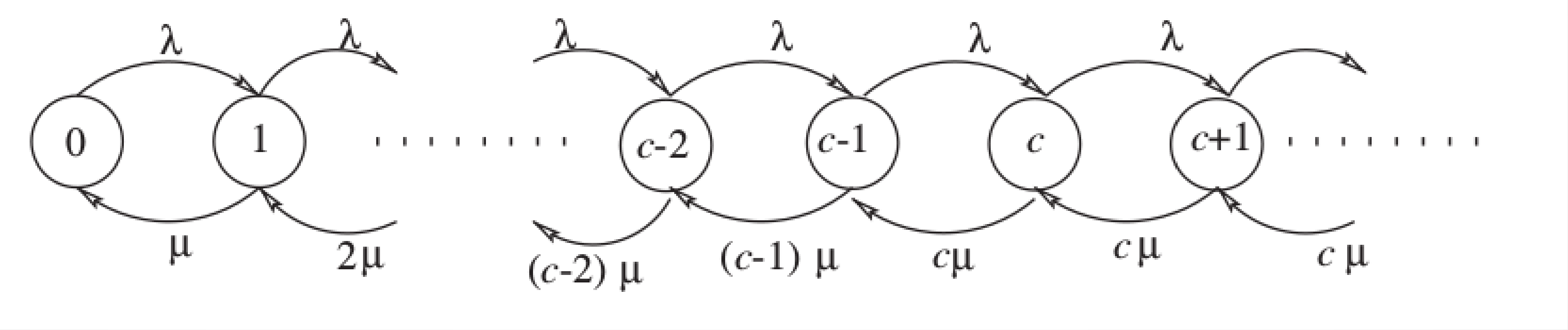
The M/M/c Queue

- If the number of customers is less than c, then only n out of the c servers are busy and the
  - **MSOR is equal to nµ**

Load dependent service center is used to designate a service center in which the **customers departure rate is a function of the number of customers present**.

$$\mu(n) = \min(n, c)\mu.$$

State Diagram of The M/M/c Queue.



Birth and death rates for M/M/c queue

$$\lambda_n = \lambda \quad \text{for all } n,$$

$$\mu_n = \begin{cases} n\mu, & 1 \le n \le c, \\ c\mu, & n \ge c. \end{cases}$$

$$p_n = p_0 \prod_{i=1}^{n} \frac{\lambda}{i\mu} = p_0 \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!} \quad \text{if } 1 \le n \le c,$$

$$p_n = p_0 \prod_{i=1}^{c} \frac{\lambda}{i\mu} \prod_{i=c+1}^{n} \frac{\lambda}{c\mu} = p_0 \left(\frac{\lambda}{\mu}\right)^n \frac{1}{c!} \left(\frac{1}{c}\right)^{n-c} \quad \text{if } n \ge c.$$

Birth and death rates for M/M/c queue

- Define ρ = λ/(cμ).
- In order for the system to be stable, ρ < 1.

• mean arrival rate must be less than the mean maximum potential rate with which customers can be served.

- substituting cρ for λ/μ,

$$p_n = p_0 \frac{(c\rho)^n}{n!} \quad \text{for } n \le c,$$

$$p_n = p_0 \frac{(c\rho)^n}{c^{n-c} c!} = p_0 \frac{\rho^n c^c}{c!} \quad \text{for } n \ge c.$$

$$p_0 = \left[\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} \frac{1}{1-\rho} \right]^{-1}$$

Performance Measures for the M/M/c Queue

- Expected queue length L_q
- *P_n 's involved are those for which $n \geq cn$'s*

$$L_q = \frac{(\rho c)^{c+1}/c}{c!(1-\rho)^2} p_0$$

$$L_q = \frac{(\lambda/\mu)^c \lambda\mu}{(c-1)!(c\mu-\lambda)^2} p_0.$$

- W_q , the mean time spent waiting prior to service;
- W , the mean time spent in the system,
- L, the mean number of customers in the system,

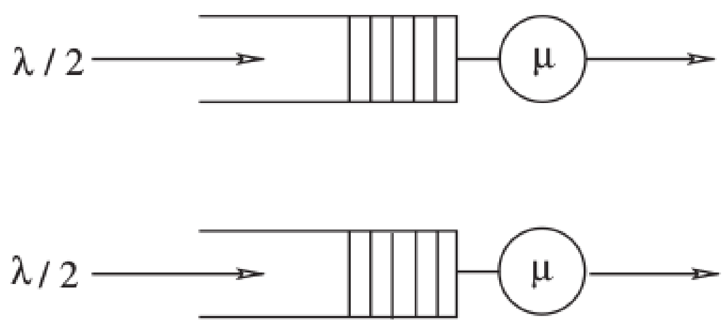
$$W_q = \left[\frac{(\lambda/\mu)^c \mu}{(c-1)!(c\mu-\lambda)^2}\right] p_0,$$

$$W = \left[\frac{(\lambda/\mu)^c \mu}{(c-1)!(c\mu-\lambda)^2}\right] p_0 + \frac{1}{\mu},$$

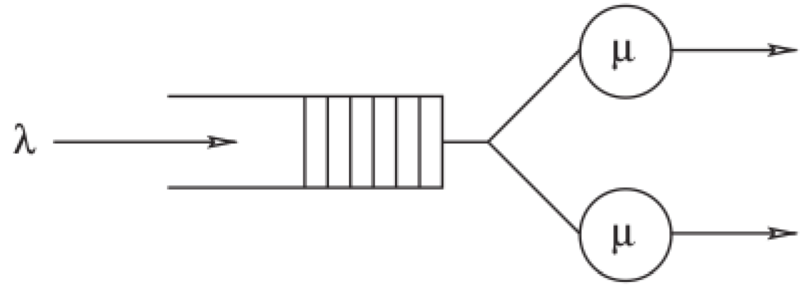
$$L = \left[\frac{(\lambda/\mu)^c \lambda\mu}{(c-1)!(c\mu-\lambda)^2}\right] p_0 + \frac{\lambda}{\mu}.$$

Comparing M/M/1 Queue with M/M/c Queue

2 of M/M/1 Queue and 1 of M/M/2 Queue



(a) Separate queue system



(b) Common queue

Performance of two M/M/1 Queues

- Arrival rate $\lambda/2$ and service rate μ.
- $\rho = (\lambda/2)/\mu = \lambda/(2\mu)$
- Mean Number for 1 M/M/1 Queue = $\rho/(1-\rho)$
- Mean Number for 2 M/M/1 Queues = $2\rho/(1-\rho)$
- The average response time can now be found using Little's law for 2 M/M/1 Queues:

$$E[R_1] = \frac{1}{\lambda}E[N_1] = \frac{1}{\lambda}\frac{2\rho}{(1-\rho)} = \frac{2}{2\mu-\lambda}.$$

Performance of M/M/2 Queue

- The mean number of customers in an M/M/c queue with arrival rate λ and service rate μ per server is given by,

$$E[N_2] = \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^c \lambda\mu}{(c-1)!(c\mu-\lambda)^2} p_0 \quad \text{with} \quad \frac{\lambda}{c\mu} = \rho \quad \text{or} \quad \lambda/\mu = c\rho$$

- With c = 2,

$$\begin{aligned} L_2 = E[N_2] = \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^2 \lambda\mu}{(2\mu-\lambda)^2} p_0 &= \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^2(\lambda/\mu)}{(1/\mu^2)(2\mu-\lambda)^2} p_0 \\ &= \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^3}{(2-\lambda/\mu)^2} p_0 \\ &= 2\rho + \frac{(2\rho)^3}{(2-2\rho)^2} p_0. \end{aligned}$$

Performance of M/M/2 Queue

- The probability of the system being empty, p_0 , is computed as:

$$p_0 = \left[1 + \sum_{n=1}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!}\left(\frac{1}{1-\rho}\right)\right]^{-1}$$
$$= \left[1 + 2\rho + \frac{(2\rho)^2}{2!}\frac{1}{1-\rho}\right]^{-1} = \frac{1-\rho}{1+\rho}.$$

- Thus,

$$L_2 = 2\rho + \frac{8\rho^3(1-\rho)}{4(1-\rho)^2(1+\rho)} = \frac{2\rho(1-\rho)(1+\rho) + 2\rho^3}{(1-\rho)(1+\rho)} = \frac{2\rho}{1-\rho^2}.$$

- Using Little's formula

$$E[R_2] = \frac{1}{\lambda}E[N_2] = \frac{2\rho/\lambda}{1-\rho^2} = \frac{1/\mu}{1-\rho^2} = \frac{4\mu}{4\mu^2 - \lambda^2}.$$

Performance of single superserver working twice as fast

- Use the M/M/1 queue with arrival rate λ and service rate 2μ.

$$L_3 = E[N_3] = \frac{\rho}{1-\rho} \quad \text{and} \quad E[R_3] = \frac{1/2\mu}{1-\lambda/2\mu} = \frac{1}{2\mu-\lambda}$$

Comparison of the three systems

$$E[N_1] = \frac{2\rho}{1-\rho} \quad \geq \quad E[N_2] = \frac{2\rho}{1-\rho} \cdot \frac{1}{1+\rho} \quad \geq \quad E[N_3] = \frac{\rho}{1-\rho},$$

$$E[R_1] = \frac{2}{2\mu - \lambda} \quad \geq \quad E[R_2] = \frac{1/\mu}{1-\rho^2} \quad \geq \quad E[R_3] = \frac{1}{2\mu - \lambda},$$

- Let $\alpha = 2\rho/(1 - \rho)$ and $\beta = 2/(2\mu - \lambda)$, then $\alpha = \lambda\beta$

$$E[N_1] = \alpha \quad \geq \quad E[N_2] = \alpha \cdot \frac{1}{1+\rho} \quad \geq \quad E[N_3] = \alpha/2,$$

$$E[R_1] = \beta \quad \geq \quad E[R_2] = \beta \cdot \frac{1}{1+\rho} \quad \geq \quad E[R_3] = \beta/2.$$

What can you say about the three Queues?

Erlang-C formula

- The probability that an arriving customer is forced to wait in the queue. (There is no server available)

$$\text{Prob}\{\text{queueing}\} = \sum_{n=c}^{\infty} p_n = p_0 \sum_{n=c}^{\infty} \frac{c^c}{c!}\rho^n = p_0 \frac{c^c}{c!}\left[\frac{\rho^c}{1-\rho}\right]$$

$$= \frac{(c\rho)^c}{c!(1-\rho)} p_0 = \frac{(\lambda/\mu)^c \mu}{(c-1)!(c\mu-\lambda)} p_0.$$

$$W_q = \frac{1}{\lambda}\frac{\rho C(c,\lambda/\mu)}{(1-\rho)} = \frac{C(c,\lambda/\mu)}{c\mu-\lambda},$$

$$W = \frac{1}{\lambda}\frac{\rho C(c,\lambda/\mu)}{(1-\rho)} + \frac{1}{\mu} = \frac{C(c,\lambda/\mu)}{c\mu-\lambda} + \frac{1}{\mu},$$

$$L = \frac{\rho C(c,\lambda/\mu)}{(1-\rho)} + c\rho = \frac{\lambda C(c,\lambda/\mu)}{c\mu-\lambda} + \frac{\lambda}{\mu}.$$