# Delay Systems
# Lecture 5

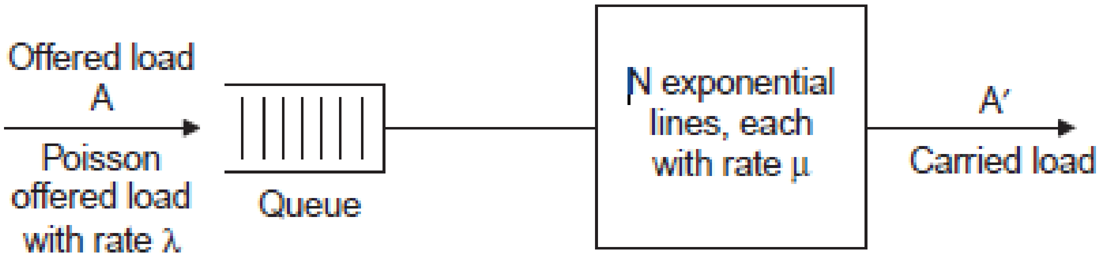
# Delay Systems

•The delay system places the call or message arrivals in a queue if it finds all *N* servers *(or* lines) occupied. This system delays non-serviceable requests until the necessary facilities become available.

•These systems are variously referred to as delay system, waiting-call systems and queueing systems. The delay systems are analysed using queueing theory which is sometimes known as waiting line theory.

•This delay system have wide applications outside the telecommunications. Some of the more common applications are data processing, supermarket check out counters, aircraft landings, inventory control and various forms of services.

Consider that there are k calls (in service and waiting) in the system and N lines to serve the calls. If $k \leq N$, k lines are occupied and no calls are waiting. If $k > N$, all N lines are occupied and $k - N$ calls waiting. Hence a delay operation allows for greater utilization of servers than does a loss system. Even though arrivals to the system are random, the servers see a somewhat regular arrival pattern. A queueing model for the Erlang delay system is shown in Fig.




Queueing model.

The basic purpose of the investigation of delay system is to determine the probability distribution of waiting times. From this, the average waiting time W as random variable can be easily determined. The waiting times are dependent on the following factors :

1. Number of sources
2. Number of servers
3. Intensity and probabilistic nature of the offered traffic
4. Distribution of service times
5. Service discipline of the queue.

In a delay system, there may be a finite number of sources in a physical sense but an infinite number of sources in an operational sense because each source may have an arbitrary number of requests outstanding. If the offered traffic intensity is less than the servers, no statistical limit exists on the arrival of calls in a short period of time. In practice, only finite queue can be realised. There are two service time distributions. They are constant service times and exponential service times. With constant service times, the service time is deterministic and with exponential, it is random.

- The service discipline of the que involves two important factors.

1. Waiting calls are selected on of first-come, first served (FCFS) or first-in-first-out (FIFO) service.
2. The second aspect of the service discipline is the length of the queue. Under heavy loads, blocking occurs. The blocking probability or delay probability in the system is based on the queue size in comparison with number of effective sources. We can model the Erlang delay system by the birth and death process with the following birth and death rates respectively.

Under equilibrium conditions, the state probability distribution $P(k)$ can be obtained by substituting these birth rates into the following equation

$$P(k) = \frac{\lambda_0 \lambda_1 \ldots\ldots \lambda_{k-1}}{\mu_1 \mu_2 \ldots\ldots \mu_k} P(0) \; k = 1, 2, \ldots\ldots$$

we set

$$P(k) = \begin{cases} \dfrac{1}{k!}\left(\dfrac{\lambda}{\mu}\right)^k P(0) & 0 \le k \le S \\[2ex] \dfrac{\left(\dfrac{\lambda}{\mu}\right)^k}{N!\,N^{k-N}} P(0) & k \ge N. \end{cases}$$

Under normalised condition,

$$\sum_{k=0}^{\infty} P(k) = 1 \qquad \text{or} \qquad \sum_{k=0}^{N-1} \frac{A^k}{k!} P(0) + \sum_{k=N}^{\infty} \frac{A^k}{N!\,N^{k-N}} P(0) = 1$$

$$\frac{1}{P(0)} = \sum_{k=0}^{N-1} \frac{A^k}{k!} + \frac{N^N}{N!} \sum_{k=N}^{\infty} \left(\frac{A}{N}\right)^k$$

$$= \sum_{k=0}^{N-1} \frac{A^k}{k!} + \frac{N^N}{N!} \left[\left(\frac{A}{N}\right)^N + \left(\frac{A}{N}\right)^{N+1} + \left(\frac{A}{N}\right)^{N+2} + \ldots\ldots\right]$$

$$= \sum_{k=0}^{N-1} \frac{A^k}{k!} + \frac{N^N}{N!} \left(\frac{A}{N}\right)^N \left[1 + \frac{A}{N} + \left(\frac{A}{N}\right)^2 + \ldots\ldots\right]$$

$$= \sum_{k=0}^{N-1} \frac{A^k}{k!} + \frac{A^N}{N!} \left[\frac{1}{1 - A/N}\right] = \sum_{k=0}^{N-1} \frac{A^k}{k!} + \left[\frac{A^N}{N!} + \frac{A^N}{N!}\left(\frac{A}{N-A}\right)\right]$$

$$\frac{1}{P(0)} = \sum_{k=0}^{N} \frac{A^k}{k!} + \frac{A^N}{N!}\left(\frac{A}{N-A}\right)$$

$$P(0) = \frac{1}{\displaystyle\sum_{k=0}^{N} \frac{A^k}{k!} + \frac{A^N}{N!}\left(\frac{A}{N-A}\right)}$$

$$P(k) = \frac{A^k}{k!} P(0), k = 1, 2, \ldots\ldots, N$$

Now, the probability of waiting (the probability of finding all lines occupied) is equal to

$$P(W > 0) = C(N, A)$$

From 8.63 and 8.64

$$C(N, A) = \frac{A^N / N!}{\sum_{k=0}^{N} \frac{A^k}{k!} + \frac{A^N}{N!}\left(\frac{A}{N-A}\right)}$$

$$\frac{1}{C(N, A)} = \frac{\sum_{k=0}^{N} \frac{A^k}{k!}}{\frac{A^N}{N!}} + \frac{\frac{A^N}{N!}\left(\frac{A}{N-A}\right)}{\frac{A^N}{N!}!}$$

$$\frac{1}{C(N, A)} = \frac{1}{B} + \frac{A}{N-A}.$$

$$\text{Prob. (delay)} = P(> 0)\, C(N, A) = \frac{BN}{N - A(1 - B)}$$

where B = Blocking probability for a LCC system

N = Number of servers

A = Offered load (Erlangs)

## Erland Second Formula /Erlang C Formula / Erlang Delay Formula

For single server systems (N = 1), the probability of delay reduces to ρ, which is simply the output utilization or traffic carried by the server. Thus the probability of delay for a single server system is also.

The distribution of waiting times for random arrivals, random service times, and a FIFO service discipline is

$$(P > t) = P(> 0)\, e^{-(N-A)t/h}$$

where P(> 0) = probability of delay (equation 8.67)

$h$ = average service time of negative exponential service time distribution.

By integrating equation 8.68, the average waiting time for all arrivals can be determined as

$$W(t)_{avg.} = \frac{C(N, A)\, h}{N - A}$$

$W(t)_{avg.}$ is the expected delay for all arrivals. The average delay of only those arrivals that get delayed is commonly denoted as

$$T_W = \frac{h}{N - A}$$

# Question Sample

. *A message switching network is to be designed for 90% utilization of its transmission link. Assuming exponentially distributed message lengths and an arrivals rate of 10 messages per min. What is the average waiting time and what is the probability that the waiting time exceeds 3 minutes ?*

**Given data :** $\rho = 90\% = 0.9$

$\lambda = 10$ *messages / minute.*

**Sol.** Assuming single channel, N = 1

For N = 1, prob (delay) = $P(> 0) = \rho = 0.9$. Also $A = \rho = 0.9$.

The average service time $$h = \frac{\text{Prob. (delay)}}{\lambda} = \frac{0.9}{10} = 0.09$$

Average waiting time $$W(t)_{avg} = \frac{P(>0)\,h}{N - A} = \frac{0.9 \times 0.09}{1 - 0.9} = 0.81 \text{ min.}$$

Prob. of the waiting time exceeding 3 minutes

$$= P(> 5) = P(> 0)\, e^{-(N-A)t/h} = 0.9 \times e^{-(1-0.09)3/0.09} = 0.032.$$