LECTURE #7

EQUILIBRIUM BEHAVIOR OF BIRTH-DEATH QUEUEING SYSTEM

204528

Queueing Theory and Applications in Networks

Assoc. Prof. Anan Phonphoem, Ph.D. (รศ.ดร. อนันต์ ผลเพิ่ม)

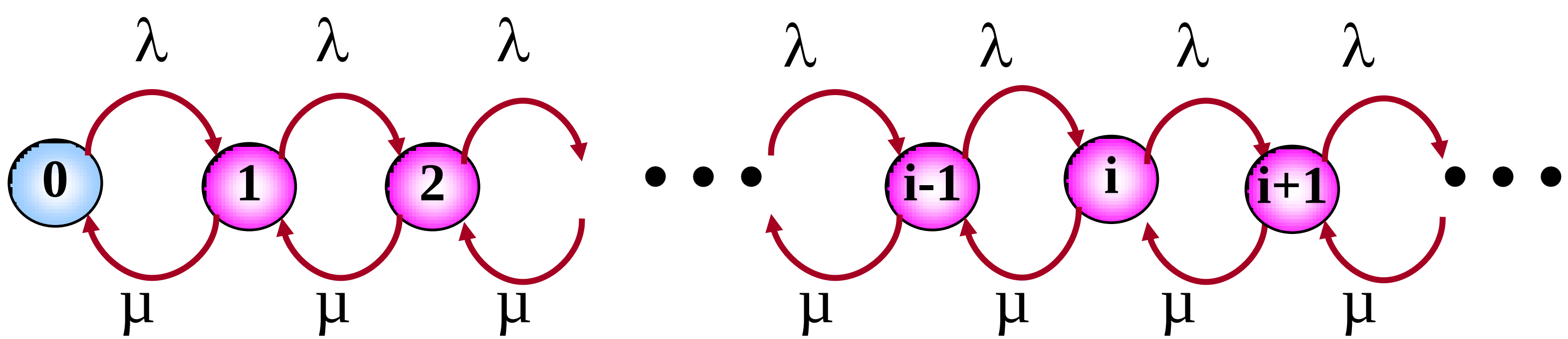
Computer Engineering Department, Kasetsart University

Outline

- M/M/1
- Discouraged Arrivals
- Responsive Servers (M/M/∞)
- Finite Storage (M/M/1/K)

A Birth-Death Process : M/M/1

- Assumption
  - $\lambda_i = \lambda$ for $i \geq 0$
  - $\mu_i = \mu$ for $i \geq 1$
  - The system begins at time t_0 with 0 member



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

A Birth-Death Process : M/M/1

- A Birth-Death Process
  - Constant coefficients λ and μ
- Interarrival Time / Service Time / #Servers
- Memoryless / Memoryless / 1 Server
- M/M/1 = A single-server queue with a Poisson arrival and an exponential distribution for service time

Equilibrium Solution

$$p_0 = \left(1 + \sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{\lambda_k}{\mu_{k+1}}\right)^{-1}$$

$$p_i = p_0 \left(\prod_{k=0}^{i-1} \frac{\lambda_k}{\mu_{k+1}} \right) \qquad \forall i \geq 1$$

M/M/1 Equilibrium Solution

$$p_i = p_0 \left(\prod_{k=0}^{i-1} \frac{\lambda_k}{\mu_{k+1}} \right) \qquad \forall i \geq 1$$

- From $\lambda_k = \lambda$ for $k \geq 0$ and $\mu_k = \mu$ for $k \geq 1$

$$p_i = p_0 \left(\prod_{k=0}^{i-1} \frac{\lambda}{\mu} \right) \qquad \forall i \geq 1$$

$$p_i = p_0 \left[\frac{\lambda}{\mu} \right]^i \qquad \forall i \geq 0$$

M/M/1 Equilibrium Solution

$$p_0 = \left(1 + \sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{\lambda_k}{\mu_{k+1}}\right)^{-1}$$

- From $\lambda_k = \lambda$ for $k \geq 0$ and $\mu_k = \mu$ for $k \geq 1$

$$p_0 = \left(1 + \sum_{i=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^{i}\right)^{-1}$$

Birth-Death Process Classification

Define:

$$S_1 \triangleq \sum_{i=0}^{\infty} \prod_{k=0}^{i-1} \frac{\lambda_k}{\mu_{k+1}}$$

.....From p_0

$$S_2 \triangleq \sum_{i=0}^{\infty} \left(\lambda_i \prod_{k=0}^{i-1} \frac{\lambda_k}{\mu_{k+1}} \right)^{-1}$$

Ergodic: $S_1 < \infty$ and $S_2 = \infty$

Recurrent Null: $S_1 = \infty$ and $S_2 = \infty$

Transient: $S_1 = \infty$ and $S_2 < \infty$

Ergodicity

- E_j = **Ergodic** if
  - E_j = ***Aperiodic*** and ***Recurrent Nonnull***
- $f_j = 1$, $M_j < \infty$, and $\beta = 1$
- A Markov Chain is **ergodic**
  - If **all** states of a Markov Chain are **ergodic**
  - If number of states is **finite** and **all** states of a Markov Chain are **aperiodic**, and **irreducible**

Stability Condition

- $p_0 > 0$
- The sufficient condition for ergodicity in M/M/1 is $\lambda < \mu$

$$p_0 = \left(1 + \sum_{i=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^{i}\right)^{-1}$$

$$= \frac{1}{1 + \dfrac{\lambda / \mu}{1 - \lambda / \mu}} = 1 - \frac{\lambda}{\mu}$$

M/M/1

- $\rho = \lambda / \mu$
- For stability conditions $0 \leq \rho < 1$

$$p_0 = 1 - \rho$$

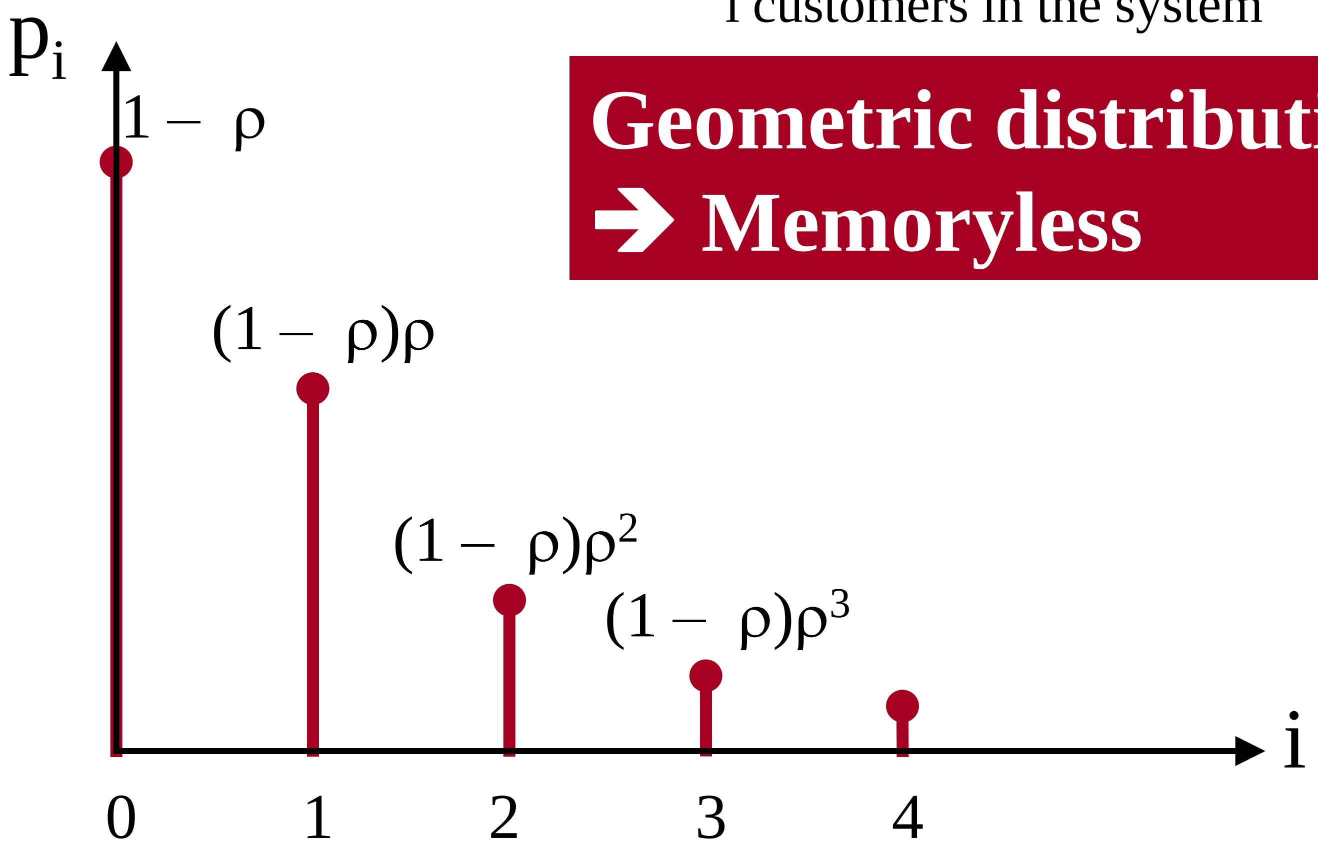
$$p_i = p_0 \left[\frac{\lambda}{\mu}\right]^i$$

$$p_i = (1 - \rho)\rho^i$$

The Solution of p_i

i customers in the system

Geometric distribution
➔ Memoryless



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

The average # of customers

$$\overline{N} = \sum_{i=0}^{\infty} i\, p_i = (1-\rho)\sum_{i=0}^{\infty} i\,\rho^i$$

$$= (1-\rho)\rho\frac{\partial}{\partial\rho}\sum_{i=0}^{\infty}\rho^{\,i}$$

$$= (1-\rho)\rho\frac{\partial}{\partial\rho}\frac{1}{(1-\rho)}$$

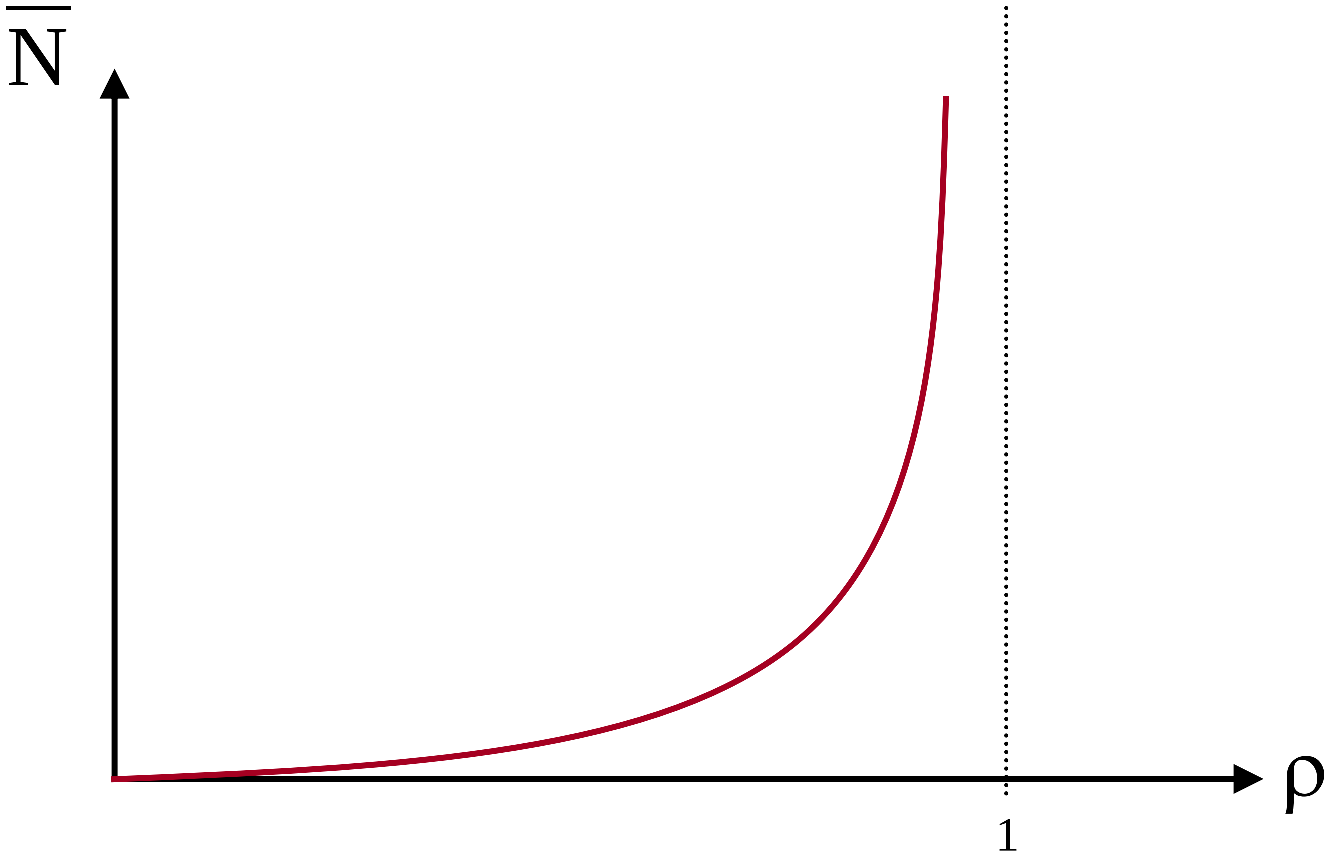
$$= (1-\rho)\rho\,\frac{1}{(1-\rho)^2}$$

$$= \frac{\rho}{(1-\rho)}$$

The average # of customers

$$\overline{N} = \frac{\rho}{(1-\rho)} = \frac{\lambda/\mu}{(1-\lambda/\mu)} = \frac{\lambda}{(\mu-\lambda)}$$

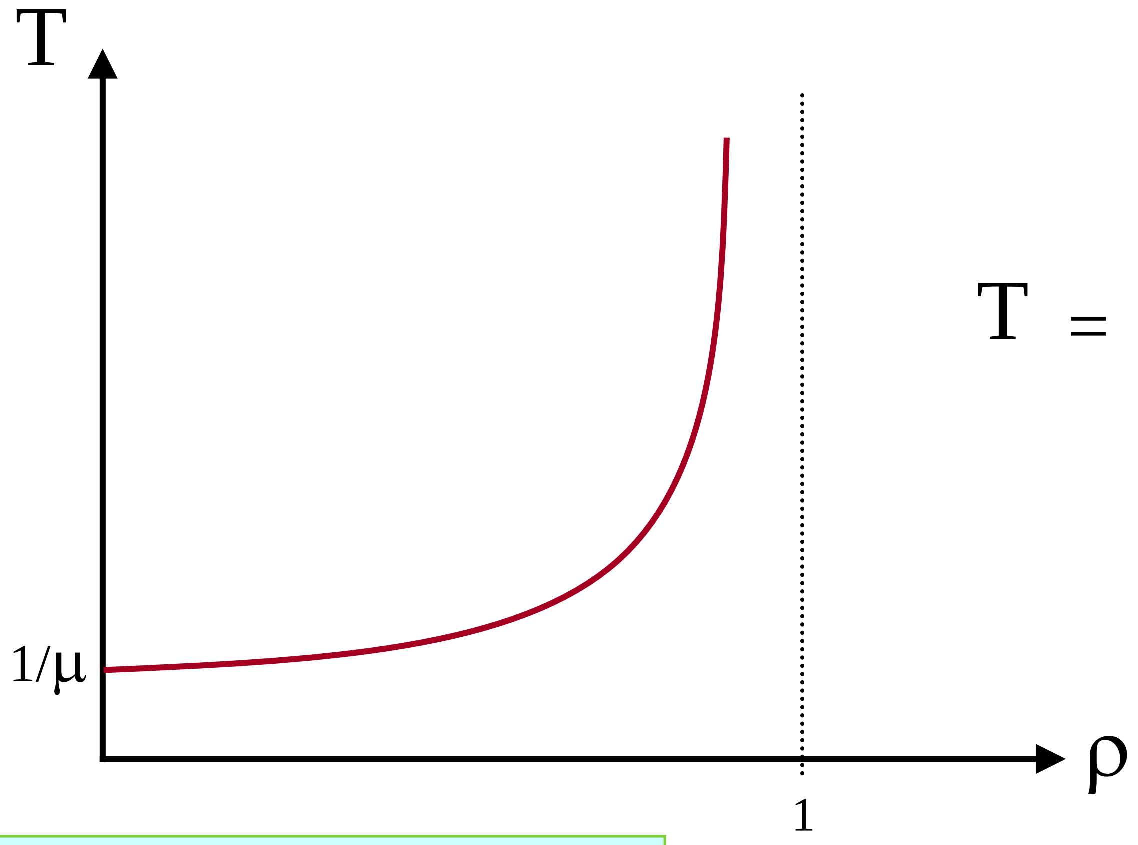
The average # of customers



The average delay

$$
\begin{aligned}
T &= \frac{\overline{N}}{\lambda} \\
&= \frac{1}{\lambda}\left(\frac{\rho}{(1-\rho)}\right) \\
&= \frac{1/\mu}{(1-\rho)} \\
&= \frac{1}{(\mu-\lambda)}
\end{aligned}
$$

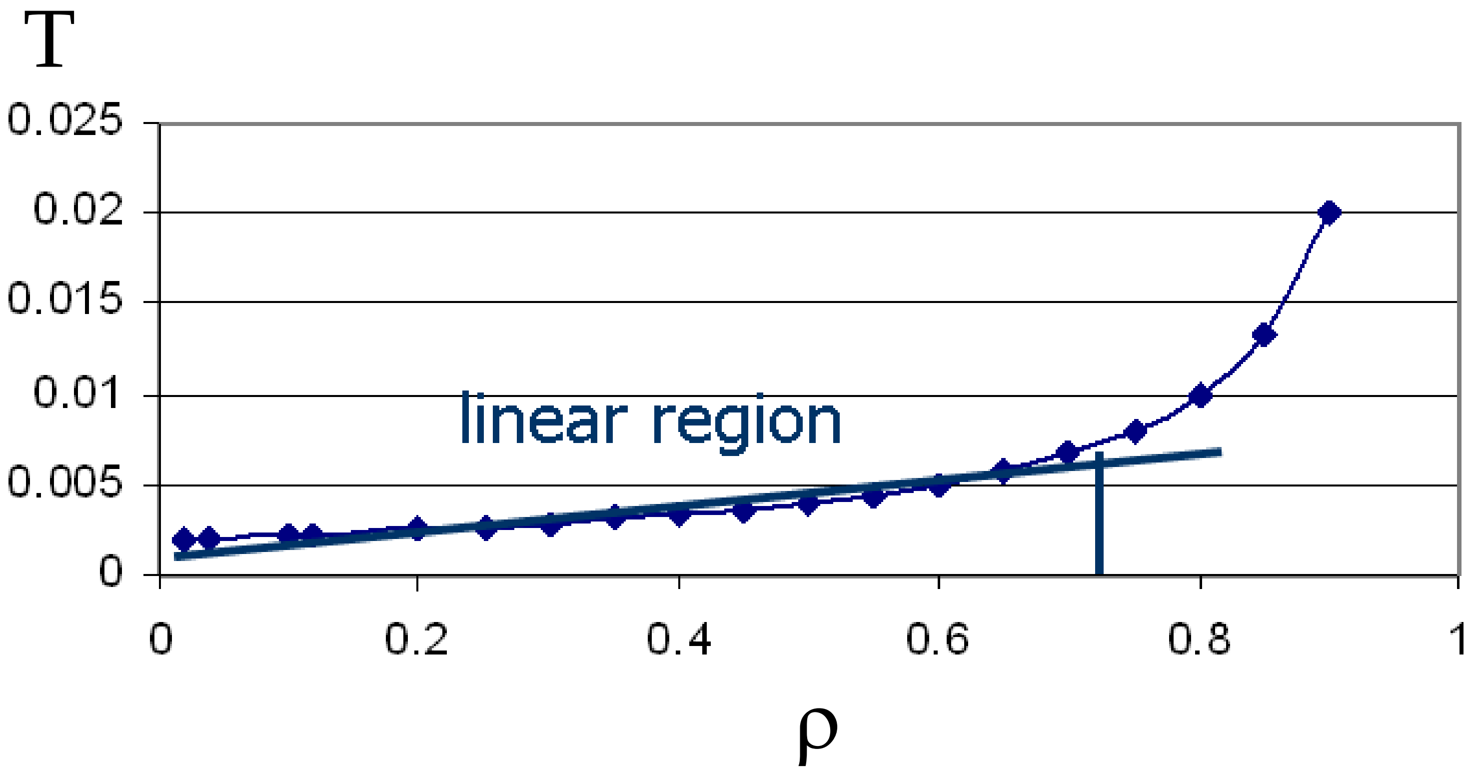
The average delay



$$T = \frac{1/\mu}{(1-\rho)}$$

$\rho = 0$ ➔ only service time

Stable Region



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

The average # of customers in Q

- W = Total Delay – Service time = T – S

$$\overline{W} = \frac{1/\mu}{(1-\rho)} - (1/\mu)$$

$$= \frac{\rho/\mu}{(1-\rho)}$$

$$\overline{N_Q} = \lambda\,\overline{W} = \frac{\rho^2}{(1-\rho)}$$

Example 1

- A wireless channel of IEEE 802.11b = 11Mbps
- For a video file needed to be sent at 2 Mbps
- λ = 2 Mbps
- μ = 11 Mbps
- $\rho = \lambda/\mu = 2/11 = 0.182$

Example 1

- The avg.Tx time $= \dfrac{1/\mu}{(1-\rho)}$

  $= (1/11)/(1 - 0.182)$

  $= 111.14$ ms

- The avg.# of waiting bits $= 0.182^2/(1 - 0.182)$

  $= 0.04049$

Example 2

Example from CS352, Rutgers University

- On a network gateway, measurements show that the packets arrive at a mean rate of 125 packets per second (pps)
- Gateway takes about 2 milliseconds to forward them
- Assuming an M/M/1 model
  - What is the probability of buffer overflow if the gateway had only 13 buffers ?
  - How many buffers are needed to keep packet loss below one packet per million?

Example 2

- Measurement of a network gateway:
  - mean arrival rate : 125 Packets/s
  - mean response time : 2 ms
- Assuming exponential arrivals:
  - What is the gateway's utilization?
  - What is the probability of n packets in the gateway?
  - mean number of packets in the gateway?
  - The number of buffers so P(overflow) is $<10^{-6}$?

Example 2

- Arrival rate λ =
- Service rate μ =
- Gateway utilization $\rho = \lambda/\mu$ =
- Prob. of n packets in gateway =
- Mean number of packets in gateway =

Example 2

- Arrival rate $\lambda = 125$ pps
- Service rate $\mu = 1/0.002 = 500$ pps
- Gateway utilization $\rho = \lambda/\mu = 0.25$
- Prob. of n packets in gateway

$$(1-\rho)\rho^{n} = 0.75(0.25)^{n}$$

- Mean number of packets in gateway

$$\frac{\rho}{1-\rho} = \frac{0.25}{0.57} = 0.33$$

Example 2

- Probability of buffer overflow:

  = P(more than 13 packets in gateway)

  $= \rho^{13} = 0.25^{13} = 1.49\text{x}10^{-8}$

  = 15 packets per billion packets

- To limit the probability of loss to less than 10^{-6}:

  $\rho^n \leq 10^{-6}$

  $n > \log(10^{-6})/\log(0.25)$

  $= 9.96$

HW: M/M/1

- A Bangkok gas station named “Smile Pump”
- Normally, they provide all kinds of Diesel and Gasohal (91, 95).
- Due to the Fuel crisis, the station installs a NGV gas pump which is the only one NGV in the area.

HW: M/M/1

- Car (mostly Taxi) arrives at Poisson with rate 6 per hour. Each car take 5 min. exponentially distributed time for filling up.
  - What is the system utilization ?
  - How long does each car spend at the gas station ?
  - How big is the waiting area that Gas station needs to prepare ?
  - How long does each car have to wait in the Queue?
  - How many car in the queue ?

HW: M/M/1

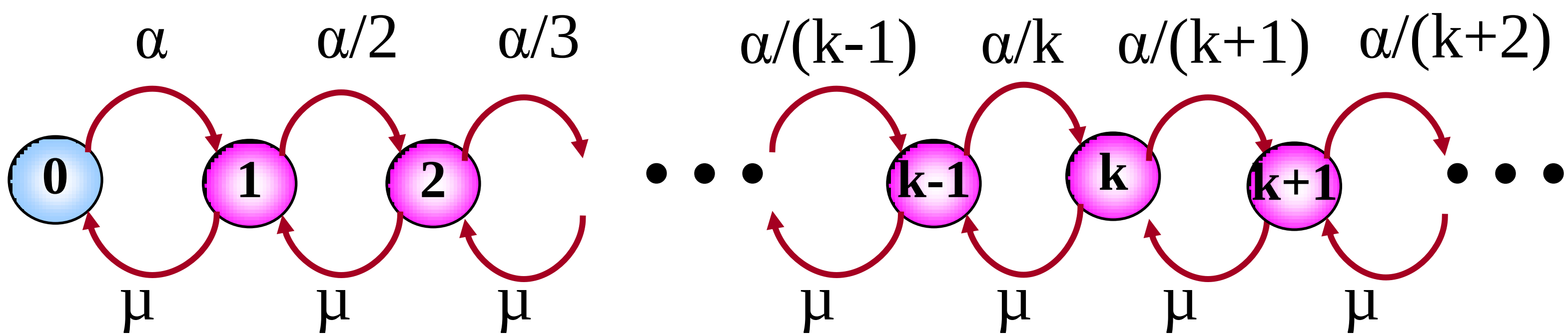
- During a Bangkok Music Festival, a lot of people coming to Bangkok. The arrival rate of cars increase to 10 per hour.
  - What is the system utilization ?
  - How long does each car spend at the gas station ?
  - How long does each car have to wait in the queue?
  - How many car in the queue ?

HW: M/M/1

- During a Songkran Festival, people leaving Bangkok. The arrival rate of cars drop to 3 per hour.
  - What is the system utilization ?
  - How long does each car spend at the gas station ?
  - How long does each car have to wait in the queue?
  - How many car in the queue ?

DISCOURAGED ARRIVALS

Discouraged Arrivals



- Assumption
  - $\lambda_k = \alpha / (k+1)$ for $k \geq 0$
  - $\mu_k = \mu$ for $k \geq 1$

Discouraged Arrivals

$$p_k = p_0 \left(\prod_{i=0}^{k-1} \frac{\alpha/(i+1)}{\mu} \right)$$

$$= p_0 \left(\frac{\alpha}{\mu} \right)^k \left(\frac{1}{k!} \right)$$

$$p_0 = \left(1 + \sum_{k=1}^{\infty} \left(\frac{\alpha}{\mu} \right)^k \left(\frac{1}{k!} \right) \right)^{-1}$$

$$= e^{-\alpha/\mu}$$

Discouraged Arrivals

$$\rho = 1 - e^{-\alpha/\mu}$$

- The ergodic condition is $\alpha/\mu < \infty$

$$p_k = p_0 \frac{(\alpha/\mu)^k}{k!}$$

$$= \frac{(\alpha/\mu)^k}{k!} e^{-\alpha/\mu}$$

Poisson Distribution

Discouraged Arrivals

$$\overline{N} = \frac{\alpha}{\mu}$$

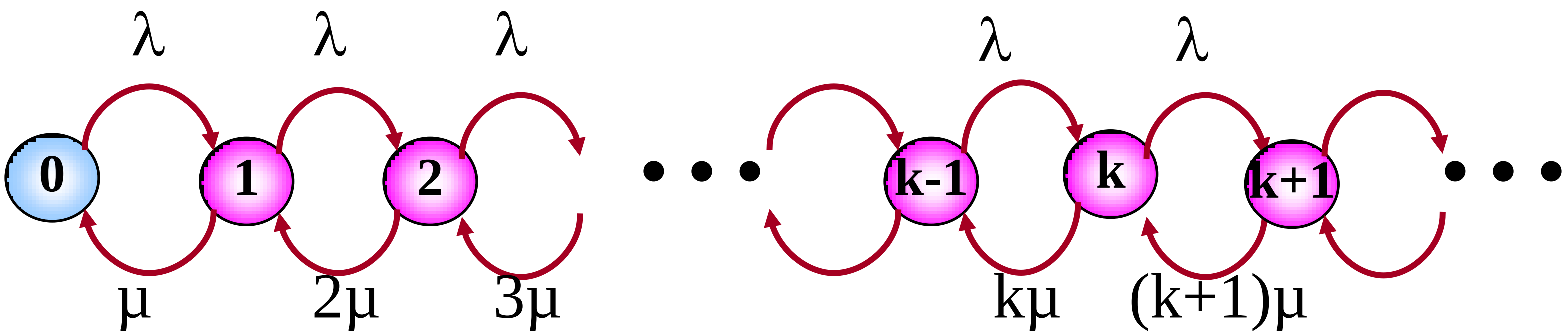
$$\lambda = \mu\rho = \mu(1 - e^{-\alpha/\mu})$$

- From Little's result

$$T = \frac{\alpha}{\mu^2 (1 - e^{-\alpha/\mu})}$$

RESPONSIVE SERVERS (M/M/∞)

Responsive Servers (M/M/∞)



- Assumption
  - $\lambda_k = \lambda$ for $k \geq 0$
  - $\mu_k = k\mu$ for $k \geq 1$
- Infinite number of servers

Responsive Servers (M/M/∞)

$$p_k = p_0 \left(\prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} \right)$$

$$= p_0 \left(\frac{\lambda}{\mu} \right)^k \left(\frac{1}{k!} \right)$$

$$p_0 = \left(1 + \sum_{k=1}^{\infty} \left(\frac{\lambda}{\mu} \right)^k \left(\frac{1}{k!} \right) \right)^{-1}$$

$$= e^{-\lambda/\mu}$$

Responsive Servers (M/M/∞)

$$\rho = 1 - e^{-\lambda/\mu}$$

- The ergodic condition is $\lambda/\mu < \infty$

$$p_k = p_0 \frac{(\lambda/\mu)^k}{k!}$$

$$= \frac{(\lambda/\mu)^k}{k!} e^{-\lambda/\mu}$$

Poisson Distribution

Responsive Servers (M/M/∞)

$$\overline{N} = \frac{\lambda}{\mu}$$

- From Little's result

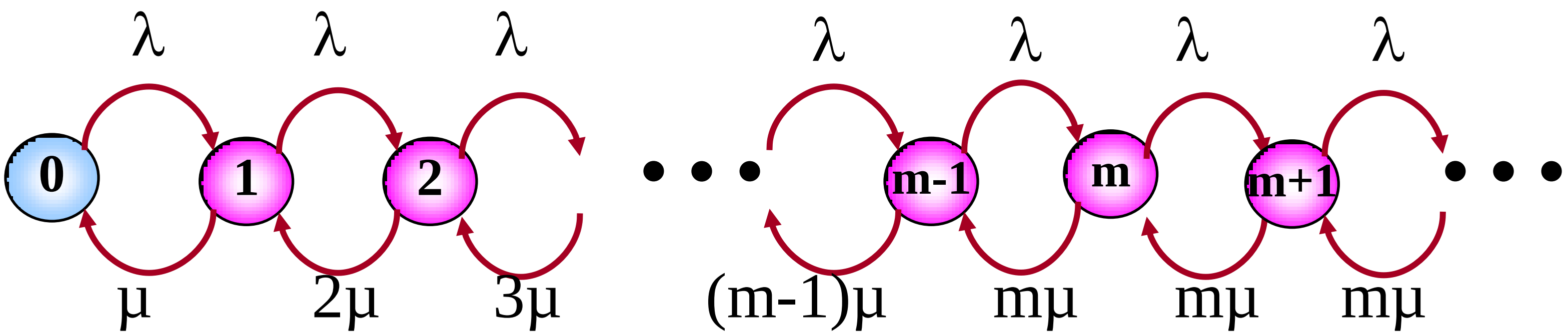
$$T = \frac{\overline{N}}{\lambda} = 1 / \mu$$

← Always get serve

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

M-SERVER (M/M/M)

m-Server (M/M/m)



- Assumption
  - $\lambda_k = \lambda$ for $k \geq 0$
  - $\mu_k = \min [\, k\mu, m\mu \,]$

$$= \begin{cases} k\mu & 0 \leq k \leq m \\ m\mu & m \leq k \end{cases}$$

m-Server (M/M/m)

- For $k \leq m$

$$p_k = p_0 \left(\prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} \right)$$

$$= p_0 \left(\frac{\lambda}{\mu} \right)^k \left(\frac{1}{k!} \right)$$

m-Server (M/M/m)

- For $k \geq m$

$$p_k = p_0 \prod_{i=0}^{m-1} \frac{\lambda}{(i+1)\mu} \prod_{j=m}^{k-1} \frac{\lambda}{m\mu}$$

$$= p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{m!\, m^{k-m}}$$

m-Server (M/M/m)

$$p_k = \begin{cases} p_0 \dfrac{(m\rho)^k}{k!} & k \le m \\ p_0 \dfrac{(\rho)^k m^m}{m!} & k \ge m \end{cases}$$

$$\rho = \frac{\lambda}{m\mu} < 1$$

m-Server (M/M/m)

$$p_0 = \left(1 + \sum_{k=1}^{m-1} \frac{(m\rho)^k}{k!} + \sum_{k=m}^{\infty} \frac{(m\rho)^k}{m!} \frac{1}{m^{k-m}}\right)^{-1}$$

$$p_0 = \left(\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \left(\frac{(m\rho)^m}{m!}\right)\left(\frac{1}{1-\rho}\right)\right)^{-1}$$

m-Server (M/M/m)

$$P[\text{queueing}] = \sum_{k=m}^{\infty} p_k$$

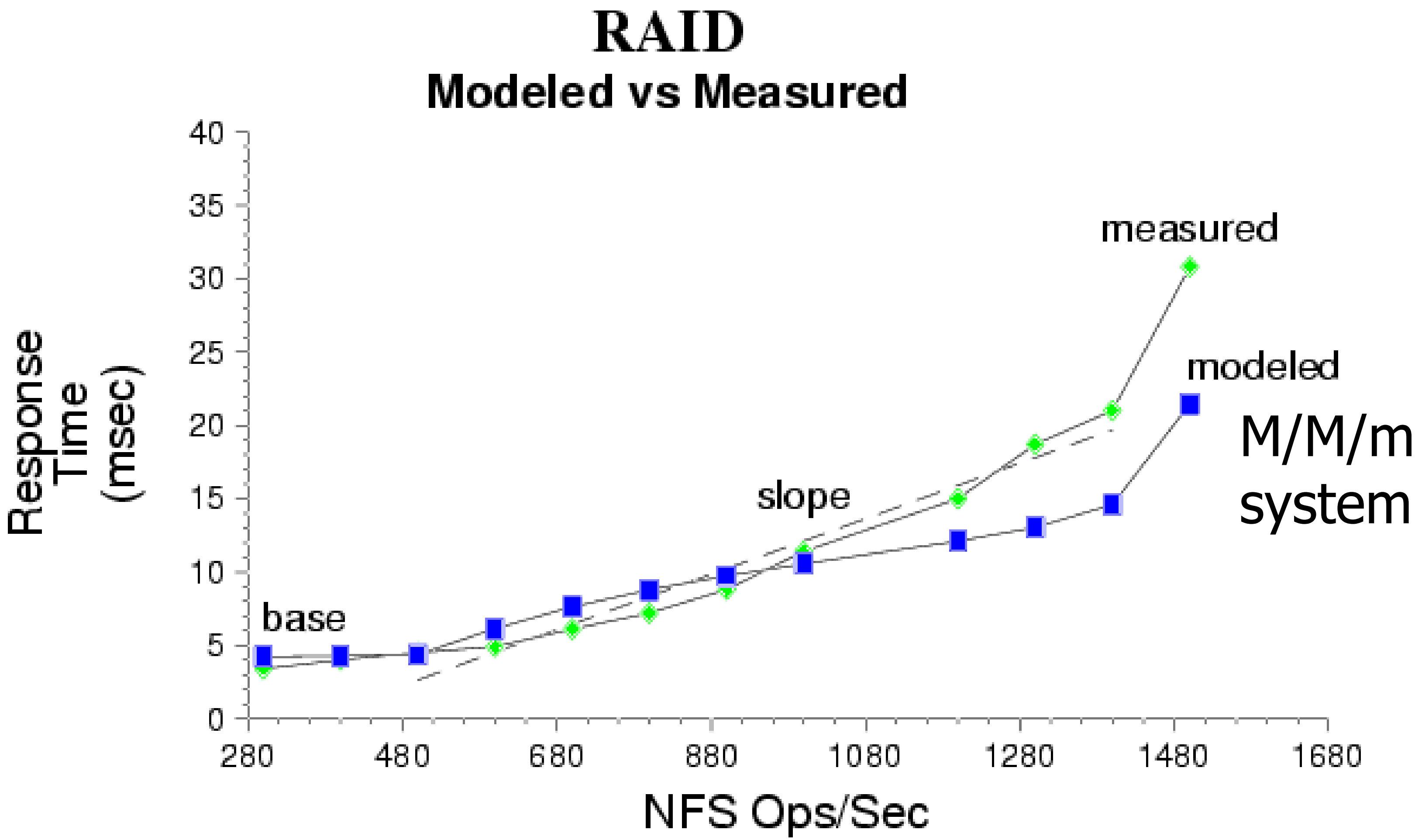
$$= \sum_{k=m}^{\infty} p_0 \frac{(m\rho)^k}{m!} \frac{1}{m^{k-m}}$$

$$P[\text{queueing}] = \frac{\left(\frac{(m\rho)^m}{m!}\right)\left(\frac{1}{1-\rho}\right)}{\left(\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \left(\frac{(m\rho)^m}{m!}\right)\left(\frac{1}{1-\rho}\right)\right)}$$

Erlang's C formula
$C(m, \lambda/\mu)$

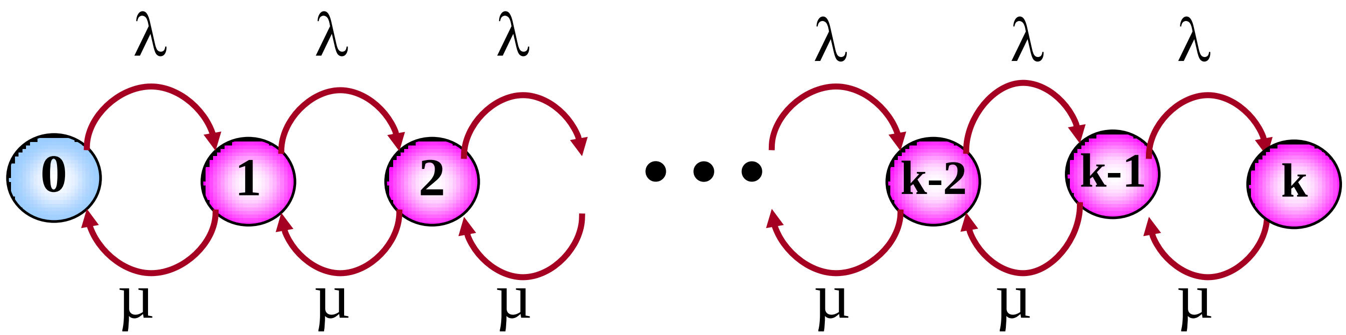
Empirical Example

Example from CS352, Rutgers University



FINITE STORAGE (M/M/1/K)

Finite Storage (M/M/1/K)



- Assumption
  - $\mu_k = \mu \qquad 1 \le k \le K$
  - $\lambda_k = \begin{cases} \lambda & k < K \\ 0 & k \ge K \end{cases}$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Finite Storage (M/M/1/K)

- For $k \leq K$

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{\mu}$$

$$= p_0 \left(\frac{\lambda}{\mu} \right)^k$$

- For $k > K$

$$p_k = 0$$

Finite Storage (M/M/1/K)

$$p_0 = \left(1 + \sum_{k=1}^{K}\left(\frac{\lambda}{\mu}\right)^{k}\right)^{-1}$$

$$= \left(1 + \frac{(\lambda/\mu)(1-(\lambda/\mu)^{K})}{1-(\lambda/\mu)}\right)^{-1}$$

$$= \frac{1-(\lambda/\mu)}{1-(\lambda/\mu)^{K+1}}$$

Finite Storage (M/M/1/K)

$$p_k = \begin{cases} \dfrac{1-(\lambda/\mu)}{1-(\lambda/\mu)^{K+1}}\left(\dfrac{\lambda}{\mu}\right)^k & 0 \le k \le K \\ 0 & \text{Otherwise} \end{cases}$$

Finite Storage (M/M/1/K)

- For K = 1 (Blocked call cleared)

$$p_k = \begin{cases} \dfrac{1}{1 + (\lambda/\mu)} & k = 0 \\[2ex] \dfrac{\lambda/\mu}{1 + (\lambda/\mu)} & k = 1 = K \\[2ex] 0 & \text{Otherwise} \end{cases}$$