CS-417
COMPUTER SYSTEMS MODELING
Spring Semester 2020
Batch: 2016-17
(LECTURE # 21)

FAKHRA AFTAB

LECTURER

DEPARTMENT OF COMPUTER & INFORMATION SYSTEMS ENGINEERING

NED UNIVERSITY OF ENGINEERING & TECHNOLOGY

Recap of Lecture # 20

Transient State and Steady State

M/M/1 Queue Analysis (Birth-Death Process)

Balance/Flow Equations for BD-Process

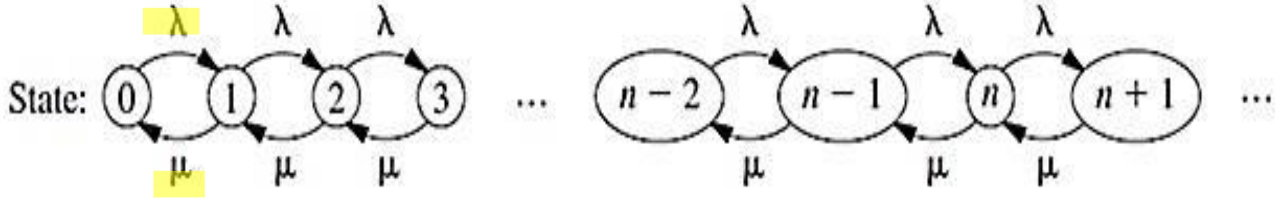
Chapter # 6 (Cont'd)

FUNDAMENTALS OF QUEUING MODELS

M/M/1 Analysis (Cont'd)

- Assumptions
  - all *inter-arrival times* are independently and identically distributed according to an *exponential distribution*, all *service times* are independent and identically distributed according to *another exponential distribution and* the *number of servers* is *1*.
- Just the special case of the birth-and-death process where
  - the queuing system's *mean arrival rate* and *mean service rate per busy server* are constant (λ and μ, respectively) regardless of the state of the system.
- When the system has just a *single server* (s = 1), the implication is that the parameters for the birth-and-death process are $\lambda_n = \lambda$ (n = 0, 1, 2, …) and $\mu_n = \mu$ (n = 1, 2, …).

(*a*) *Single-server case* (s = 1) $\quad \lambda_n = \lambda, \quad$ for n = 0, 1, 2, ...
$\quad \mu_n = \mu, \quad$ for n = 1, 2, ...



M/M/1 Analysis (Cont'd) - BD Process

Q2: *What is the probability of at least k customers in the system?*

$\rho^n P_0$

$$P[N \geq k] = \sum_{n=k}^{\infty} P_n$$

$$= \rho^k P_0 + \rho^{k+1} P_0 + \cdots \infty$$

$$= \boldsymbol{P_0} \left(\frac{\rho^k}{1-\boldsymbol{\rho}}\right) = (1 - \rho) \left(\frac{\rho^k}{1-\boldsymbol{\rho}}\right)$$

$$P[N \geq k] = \rho^k$$

M/M/1 Analysis (Cont'd) - BD Process

Q3: *Determine the expected number of customers in the system.*

$$L = E[N] = \sum_{n=0}^{\infty} n\, P_n$$

$$= \sum_{n=1}^{\infty} n\rho^n (1-\rho)$$

$$= (1-\rho)\left(\sum_{n=1}^{\infty} n\rho^n\right)$$

$$= (1-\rho)\,\rho\,\frac{d}{d\rho}\left(\sum_{n=1}^{\infty} \rho^n\right)$$

$$= \rho(1-\rho)\,\frac{d}{d\rho}\left(\frac{\rho}{1-\rho}\right)$$

$$= \rho(1-\rho)\left(\frac{1-\rho+\rho}{(1-\rho)^2}\right)$$

$$L = E[N] = \frac{\rho}{1-\rho}$$

$$\frac{d}{dx} x^n = n x^{n+1}$$

$$\frac{a}{1-r}$$

$$\frac{d}{dx} \frac{u}{v}$$

M/M/1 Analysis (Cont'd) - BD Process

Q4: *Determine the expected time spent in the system.*

Starting with Little's Law, $L = \lambda W$

$$\Longrightarrow W = \frac{L}{\lambda} = \frac{\rho}{1-\rho} \cdot \frac{1}{\lambda}$$

$$\Longrightarrow W = \frac{\lambda/\mu}{1-\lambda/\mu} \cdot \frac{1}{\lambda}$$

$$\Longrightarrow W = \frac{1}{\mu-\lambda}$$

$$\Longrightarrow W = \frac{1}{\mu\,(1-\rho)} = \frac{W_s}{(1-\rho)}$$

M/M/1 Analysis (Cont'd) - BD Process

Q5: *Determine the mean waiting time.*

$$\Rightarrow W_q = W - Ws$$

$$\Rightarrow W_q = \frac{1}{\mu-\lambda} - \frac{1}{\mu}$$

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

$$\Rightarrow W_q = \rho\, W = \frac{\rho}{1-\rho} . \frac{1}{\mu}$$

$$\Rightarrow W_q = \frac{L}{\mu}$$

M/M/1 Analysis (Cont'd) - BD Process

Q6: *Determine the expected number of customers in the queue.*

$$\Longrightarrow L_q = \lambda W_q$$

$$\Longrightarrow L_q = \frac{\lambda}{\mu} . L$$

$$\Longrightarrow L_q = \frac{\rho^2}{1-\rho}$$

M/M/1 Analysis (Cont'd) - BD Process

Q7: $Determine\ the\ mean\ number\ of\ customers\ in\ the\ service\ facility.$

$$\Longrightarrow L_s = L - L_q$$

$$\Longrightarrow L_s = \frac{\rho}{1-\rho} - \frac{\rho^2}{1-\rho}$$

$$\Longrightarrow L_s = \rho$$

M/M/1 Analysis (Cont'd) - BD Process

Q8: $Give\ the\ formulae\ for\ the\ distribution\ of\ total\ time, service\ time\ \&\ waiting\ time.$

By hypothesis;

$W_s(t) = P\{s \leq t\} = 1 - e^{-\mu t} = 1 - e^{-t/W_s}$

Similarly,

$W(t) = P\{W \leq t\} = 1 - e^{-t/W}$

Also,

$W_q(t) = P\{q \leq t\} = 1 - \rho e^{-t/W}$ (When queue discipline is FCFS)

M/M/1 Analysis (Cont'd) - BD Process

Q8: $Give\ the\ formulae\ for\ the\ distribution\ of\ total\ time, service\ time$ $\&\ waiting\ time.$

By hypothesis;

$W_s(t) = P\{s \leq t\} = 1 - e^{-\mu t} = 1 - e^{-t/Ws}$

Similarly,

$W(t) = P\{W \leq t\} = 1 - e^{-t/W}$

Also,

$W_q(t) = P\{q \leq t\} = 1 - \rho e^{-t/W}$ (When queue discipline is FCFS)

PASTA (Poisson Arrivals See Time Averages) Theorem

- The state of an M/M/1 queue is the number of customers in the system.

- More general queueing systems have a more general state that may include how much service each customer has already received.

- For Poisson arrivals, the arrivals in any future increment of time is independent of those in past increments and for many systems of interest, independent of the present state S(t) (true for M/M/1, M/M/m, and M/G/1).

- In steady state, arrivals see steady state probabilities.

M/M/1 Analysis – Example Problem

A computer system has tasks arriving on average every 0.4 sec and requires service of 0.3 sec CPU time to process each job.

Assume an exponential distribution to process each job and their random arrivals. Find out:

a) *Utilization of CPU*

b) *Avg. System Time*

c) *Avg. Waiting Time*

d) *Avg. number of jobs in the queue*

e) *Avg. number of jobs inside the system*

f) *Probability that there are 5 or more tasks waiting for service.*

g) *If inter-arrival time is reduced to 0.25 sec, what will happen to the queuing system ?*

Answers:

a) 0.75
b) 1.2 sec
c) 0.9 sec
d) 2.25
e) 3 tasks
f) 0.1779
g) Unstable system

Task

The time between requests to a Web server is found to approximately follow an exponential distribution with a mean time between requests of 8 msec. The time required by the server to process each request is also found to roughly follow an exponential distribution with an average service time of approximately 5 msec.

(Use M/M/1 queueing analysis)

a) What is the average response time observed by the users making requests on this server?

b) How much faster must the server process requests to halve this average response time?

Answers:

a) 13.33 msec

b) 275 / sec