King Fahd University of Petroleum & Minerals Computer Engineering Dept

COE 543 – Mobile and Wireless Networks

Term 032

Dr. Ashraf S. Hasan Mahmoud

Rm 22-148-3

Ext. 1724

Email: ashraf@ccse.kfupm.edu.sa

Lecture Contents

1. Traffic Engineering - Erlang C and Erlang B models

This material is found in section 4.2.5 of Pahlavan's book (page 176)

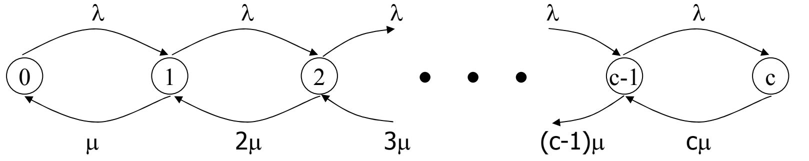
Performance of Fixed-Assignment Access Methods

- FDMA/TDMA provide a hard capacity limit (number of channels)
  - FDMA – maximum number of carriers per cell
  - TDMA – maximum number of slots per frame X number of carriers per cell
- CDMA-based also has a hard capacity limit dictated by the number of Walsh codes for example, but usually practical capacity is lower
  - Soft-capacity figure: Near the capacity boundary, the addition of one extra user degrades the link quality for all
  - Call admission control mechanism attempt to limit maximum number of ongoing calls before link quality degrades for all
- If you operate a maximum no of channels, then call blocking and call delay are the two important measures!

Erlang-B and Erlang-C Models

- More details to be provided in COE560
- Model designed to predict blocking probability (Erlang-B) and average call delay (Erlang-C) for a given number of channels and traffic intensity
- Valid for voice and traffic models conforming to the basic assumption (usually not applicable to data)
- Assumptions, Terminology and Parameters:
  - Channels ←→ Servers: c servers
  - Users ←→ Calls
  - Calls arrive according to a Poisson process with rate = λ
    - Inter-call arrival is an exponentially distributed r.v. with mean $1/\lambda$
  - Call duration is exponentially distributed r.v. with mean = $1/\mu$
  - Traffic intensity, $\rho = \lambda/\mu$

Erlang-B (M/M/c/c) Model – Call Blocking

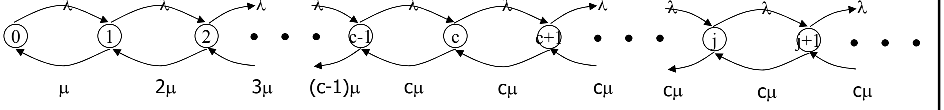


- As was shown earlier – review previous notes, the call blocking probability is given by

$$B(c,\rho) = \frac{\rho^c / c!}{\sum_{i=0}^{c} \rho^i / i!}$$

- ρ – is referred to as the offered load, while ρX[1-B(c,ρ)] is referred to as the carried load
- Note in this model – calls arriving while there are c calls are blocked – no buffering is employed

Erlang-C (M/M/c) Model – Call Delay



- The probability that an arriving call having to wait is given by

$$\Pr(delay > 0) = \frac{\rho^c}{\rho^c + c!\left(1 - \frac{\rho}{c}\right)\sum_{k=0}^{c-1} \frac{\rho^k}{k!}}$$

- The average delay is given by

$$D = \Pr(delay > 0) \times \frac{1}{\mu(c-\rho)}$$

- The probability of the delay exceeding t time units is given by

$$\Pr(delay > t) = \Pr(delay > 0) e^{-(N-\rho)\mu t}$$

Examples

- An IS-136 cellular provider owns 50 cell sites and 19 traffic carriers per carrier per cell each with bandwidth of 30 kHz. Assuming each user makes three calls per hour and the average holding time per call is 5 minutes. Determine the total number of subscribers that the service provider can support with a blocking rate less than 2%
- Solution:

c = 19X3 = 57 per cell

B(57,ρ) = 0.02 ➔ ρ = 45 Erlangs per cell

$(\lambda/\mu)_{sub}$ = 3/60*5 = 0.25 Erlangs per sub

Number of subs = total traffic / traffic per user

= 45 / 0.25

= 180 per cell

Number of subs for all sites = 180 X 50 = 8,000 subs

Note that $\rho_{all\_subs} = (\lambda/\mu)_{all\_subs} = \lambda_{all\_subs}/\mu$

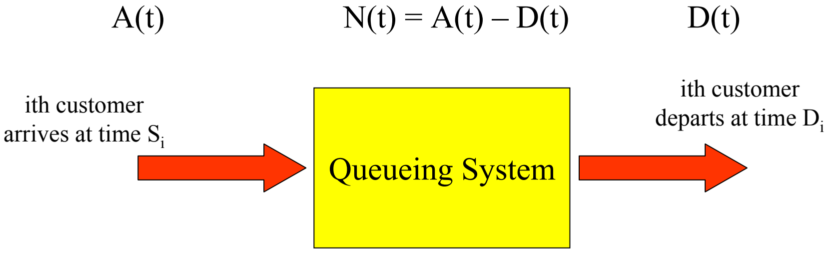
whereas, $\rho_{sub} = (\lambda/\mu)_{sub} = \lambda_{sub}/\mu$

λ_{all_subs} = no of subs X λ_{sub}

Background Slides

Queuing Model

- **Consider the following system:**

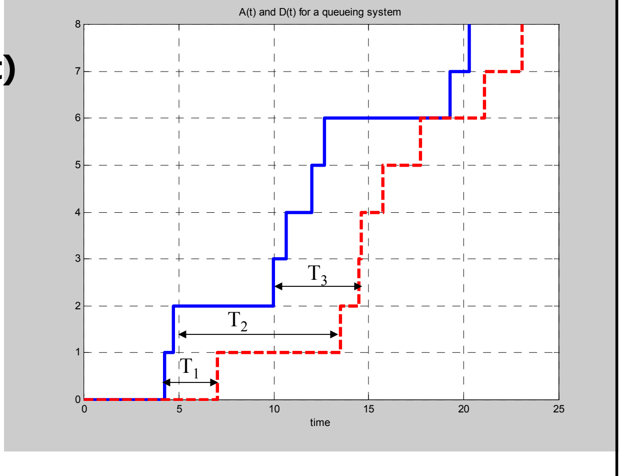


$$T_i = D_i - S_i$$

A(t) – number of arrivals in (0, t]
D(t) – number of departures in (0, t]
N(t) – number of customers in system in (0,t]
T_i – duration of time spent in system for ith customer

Number of Customers in System

- **Blue curve: A(t)**
- **Red curve: D(t)**
- **Total time spent in the system for all customers = area in between two curves**



Little's Formula – cont'd

- **Little's formula:**

  $$E[N] = \lambda E[T]$$

- **Which relates the average arrival rate (λ), the average number of customers in the system (E[N]), and the average time spent in the system (E[T])**

- **Holds for many service disciplines and for systems with arbitrary number of servers. It holds for many interpretations of the system as well**

Example 1:

- **<u>Problem</u>: Let Ns(t) be the number of customers being served at time t, and let τ denote the service time. If we designate the set of servers to be the "system"m then Little's formula becomes:**

$$E[Ns] = \lambda E[\tau]$$

Where E[Ns] is the average number of busy servers for a system in the steady state.

Example 1: cont'd

Note: for a single server Ns(t) can be either 0 or 1 ➔ E[Ns] represents the portion of time the server is busy. If p_0 = Prob[Ns(t) = 0], then we have

$$1 - p_0 = E[Ns] = \lambda E[\tau], \text{ Or}$$
$$p_0 = 1 - \lambda E[\tau]$$

The quantity $\lambda E[\tau]$ is defined as the utilization for a single server. Usually, it is given the symbol ρ

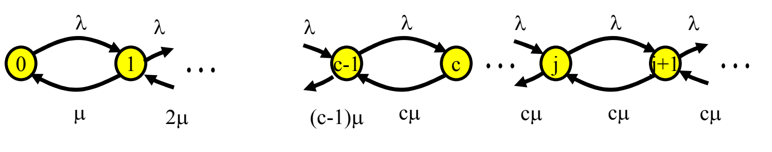
$$\rho = \lambda E[\tau]$$

For a c-server system, we define the utilization (the fraction of busy servers) to be

$$\rho = \lambda E[\tau] / c$$

Multi-Server Systems: M/M/c

- **The transition rate diagram for a multi-server M/M/c queue is as follows:**
  - **Departure rate = kμ when k servers are busy**



Multi-Server Systems: M/M/c – cont'd

- When k servers are busy, the time until the next departure is given by:
  $X = \min(\tau_1, \tau_2, \ldots, \tau_k)$

where τ_i are iid exponential r.v. with mean $1/\mu$

The CDF for X is given by (refer to definition)

$$\begin{aligned}
Prob[X > t] &= Prob[\min(\tau_1, \tau_2, \ldots, \tau_k) > t] \\
&= Prob[\tau_1>t, \tau_2>t, \ldots, \tau_k>t] \\
&= Prob[\tau_1>t]\ Prob[\tau_2>t] \ldots Prob[\tau_k>t] \\
&= e^{-\mu t}\, e^{-\mu t} \ldots e^{-\mu t} \\
&= e^{-k\mu t}
\end{aligned}$$

Therefore, the time till the next departure (X) is an exponentially distributed r.v. with mean $1/(k\mu)$

Multi-Server Systems: M/M/c – cont'd

- Writing the global balance equations:

$\lambda \quad p_0 = \mu p_1$

$j\mu \quad p_j = \lambda p_{j-1}$ for $j=1, 2, \ldots, c$

$c\mu \quad p_j = \lambda p_{j-1}$ for $j= c, c+1, \ldots$

➔

$p_j= a^j/j!\ p_0$ (for $j=1, 2, \ldots, c$) and

$p_j= \rho^{j-c}/c!\ a^c\ p_0$ (for $j=c, c+1, \ldots$)

where $a = \lambda/\mu$ and $\rho = a/c$

Multi-Server Systems: M/M/c – cont'd

- **To find p_0, we resort to the fact that $\Sigma p_j = 1$**

➔ $$p_0 = \left\{ \sum_{j=0}^{c-1} \frac{a^j}{j!} + \frac{a^c}{c!} \frac{1}{1-\rho} \right\}^{-1}$$

The probability that an arriving customer has to wait

Prob[W > 0] = Prob[N ≥ c]
= $p_c + p_{c+1} + p_{c+2} + \ldots$
= $p_c/(1-\rho)$

Erlang-C formula

Multi-Server Systems: M/M/c – cont'd

- **The mean number of customers in queue (waiting):**

$$E[N_q] = \sum_{j=c}^{\infty} (j-c) \Pr[N(t) = j]$$
$$= \sum_{j=c}^{\infty} (j-c) \rho^{j-c} p_c$$
$$= \frac{\rho}{(1-\rho)^2} p_c$$
$$= \frac{\rho}{1-\rho} \Pr[W > 0]$$

Multi-Server Systems: M/M/c – cont'd

- **The mean waiting time in queue:**

$$E[W] = E[N_q]/\lambda$$

- **The mean total delay in system:**

$$E[T] = E[W] + E[\tau]$$
$$= E[W] + 1/\mu$$

- **The mean number of customers in system:**

$$E[N] = \lambda E[T]$$
$$= E[N_q] + a$$

Why?

Example:

- **A company has a system with four private telephone lines connecting two of its sites. Suppose that requests for these lines arrive according to a Poisson process at rate of one call every 2 minutes, and suppose that call durations are exponentially distributed with mean 4 minutes. When all lines are busy, the system delays (i.e. queues) call requests until a line becomes available. Find the probability of having to wait for a line.**

Example: cont'd

- **Solution:**
  $\lambda = 1/2$, $1/\mu = 4$, $c = 4 \Rightarrow a = \lambda/\mu = 2$
  $\Rightarrow \rho = a/c = 1/2$

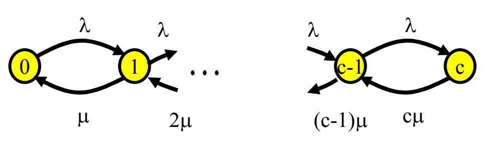
$p_0 = \{1+2+2^2/2!+2^3/3!+2^4/4!\ (1/(1-\rho))\}^{-1}$
$= 3/23$

$p_c = a^c/c!\ p0$
$= 2^4/4!\ X\ 3/23$

$Prob[W > 0] = p_c/(1-{}^{r})$
$= 2^4/4!\ X\ 3/23\ X\ 1/(1-1/2)$
$= 4/23$
≈ 0.17

Multi-Server Systems: M/M/c/c

- **The transition rate diagram for a multi-server with no waiting room (M/M/c/c) queue is as follows:**
  - **Departure rate = $k\mu$ when k servers are busy**



PMF for Number of Customers for M/M/c/c

- **Writing the global balance equations, one can show:**

  $p_j = a^j/j!\ p_0$ (for $j=0, 1, \ldots, c$)

where $a = \lambda/\mu$ (the offered load)

- **To find p_0, we resort to the fact that $\sum p_j = 1$**

$$p_0 = \left\{ \sum_{j=0}^{c} \frac{a^j}{j!} \right\}^{-1}$$

Erlang-B Formula

- **Erlang-B formula is defined as the probability that all servers are busy:**

$$\Pr[N = c] = p_c$$

$$= \frac{a_c / j!}{1 + a + a^2/2! + \ldots + a^c / c!}$$

Expected Number of customers in M/M/c/c

- **The actual arrival rate *into* the system:**

$$\lambda_a = \lambda(1 - p_c)$$

- **Average total delay figure:**

$$E[T] = E[\tau]$$

Why?

- **Average number of customers:**

$$E[N] = \lambda_a E[\tau]$$