Homework 6

$\mathrm{CS}~547$

Due Friday, Mar. 9, in class

Hyperexponential Residual Service Times

Consider the two-stage hyperexponential service time distribution with parameters p, μ_1 , and μ_2 . When a new customer enters service, it receives exponentially distributed service at rate μ_1 with probability p, and exponentially distributed service at rate μ_2 with probability 1 - p.

It's easy to see that the average service time is simply

$$\overline{s} = p\frac{1}{\mu_1} + (1-p)\frac{1}{\mu_2}$$

Now, consider the following argument for the average residual service time.

A customer in service at an arrival instant must be receiving exponentially distributed service at either rate μ_1 or rate μ_2 . I can use PASTA to reason about the service rate of the customer currently in service. With probability p, the customer is receiving service at rate μ_1 , and with probability 1 - p it is receiving its service at rate μ_2 . Exponential service times are memoryless, so the expected remaining service time is simply

$$\overline{z} = p \frac{1}{\mu_1} + (1-p) \frac{1}{\mu_2}$$

Therefore, hyperexponential service times must be memoryless!

This argument is incorrect. Hyperexponential service times are definitely not memoryless, and $\overline{z} \neq \overline{s}$ in general. What is my mistake?

Scheduling Comparisons

Place the following queues in order of increasing average residence time. You may assume that all of the queues have the same Poisson arrival rate and same average service time. Identify any queues with identical average residence times.

- deterministic service times and FCFS scheduling
- deterministic service times and LCFSPR scheduling
- exponential service times and FCFS scheduling
- hyperexponential service times and FCFS scheduling
- hyperexponential service times and PS scheduling
- Erlang-3 service times and FCFS scheduling
- Pareto service times and FCFS scheduling
- deterministic service times and SJF scheduling

Tandem Servers

Consider an open production line composed of 3 stations with Poisson arrivals. The stations are arranged in series, so that the output of one station feeds directly into the next, and customers exit the system after the third station. This model is very important in manufacturing, where it can represent many types of assembly lines.

Here are the measured characteristics of the three stations:

1.
$$\overline{s} = 1.0, c_s^2 = 2.5$$

2. $\overline{s} = 5.0, c_s^2 = .5$
3. $\overline{s} = 0.5, c_s^2 = 5.0$

Which station is the bottleneck?

What value of the arrival rate λ is required to keep the bottleneck at 80% utilization?

Calculate the average residence time in the entire production line using this value of λ .

A Network with Random Routing

The following model represents a CPU with three attached disks.



Figure 1: A CPU and three disks

Suppose we measured the following attributes for the system.

- arrival rate: $\lambda = 100$ requests/sec
- CPU: $\overline{s} = 2$ ms, PS scheduling discipline
- Disk 1: $\overline{s} = 12 \text{ ms}, c_s^2 = .75, \overline{V}_1 = .4$
- Disk 2: $\overline{s} = 15 \text{ ms}, c_s^2 = .30, \overline{V}_2 = .25$
- Disk 3: $\bar{s} = 11 \text{ ms}, c_s^2 = 1.2, \overline{V}_3 = .35$

The disks can be approximated as M/G/1 queues with FCFS scheduling.

If a request visits the CPU, accesses one of the disks, then leaves the system, what is its average residence time?

Calculate the average number of customers in the entire system.