CSE 538 – Fall 2015 Midterm

4 Problems – 80 points total

Your Name:

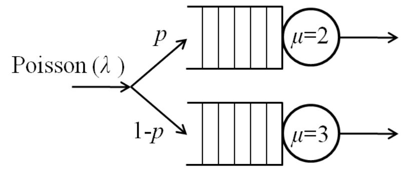

Figure 1: Open network model.

Problem 1 [15 points] Consider the open network of Fig. 1 with two servers of rates $\mu_1 = 2$ and $\mu_2 = 3$. Jobs arrive according to a Poisson process of rate λ, with jobs assigned to server 1 with probability p and to server 2 with probability $(1-p)$.

1. **[5 points]** Assuming a stable system and a corresponding λ value, give the system's maximum throughput and the value(s) of p for which it is realized.

Write your answer in the space below

2. **[5 points]** What is the value of p that maximizes the system's stability region, *i.e.,* will allow the highest possible value of λ while ensuring stability? Justify your answer.

Write your answer in the space below

3. **[5 points]** Assume that $\lambda = 2$ and compute, as a function of p, the probability $P\{3, 0\}$ that there are three (3) jobs in the top system **and** that the bottom system is empty (0 jobs). Explain what happens when $p = 1$.

Write your answer in the space below

Problem 2 [35 points] Consider the interactive system of Fig. 2 that consists of a CPU and two disks, a fast one and a slow one. Jobs visit the CPU once, but can visit the disks multiple times (multiple R/W operations). Note that p may be a design parameter, *i.e.,* we may be able to influence what fraction of R/W operations go to the fast disk, but β is outside our control, *i.e.,* is a function of the data footprint of the instructions being executed.

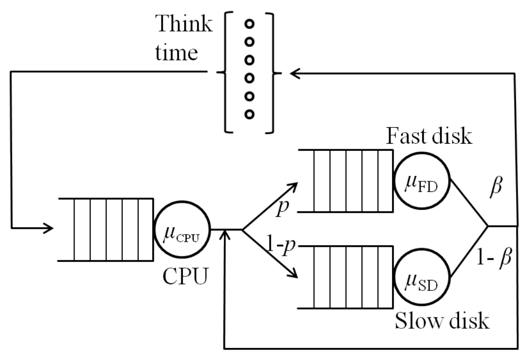

Figure 2: Closed system model.

We make the following measurements to assess the system's performance:

- Measurement duration: 20 minutes
- Average think time: 10 seconds
- Number of completed transactions in measurement interval: 1,500
- Number of CPU visits: 1,500
- Number of fast disk accesses: 30,000
- Number of slow disk accesses: 10,000
- CPU busy time: 1,000 seconds
- Fast disk busy time: 500 seconds
- Slow disk busy time: 600 seconds

1. **[10 points]** What is the average total service time of an individual transaction?

Write your answer in the space below

2. **[5 points]** Give asymptotic bounds for the system's throughput X and response time $E[R]$, as a function of N, the number of terminals using the interactive system.

Write your answer in the space below

3. **[20 points]** We are considering making the following changes to the system.

   (a) Turn the slow disk off;
   (b) Add a second fast disk and load-balance across all three disks;
   (c) Replace the CPU by one that is 50% faster;
   (d) All of the above, *i.e.,* faster CPU, slow disk off and load-balancing across two fast disks.

   Provide expressions, function of N, for the system's throughput X and response time $E[R]$ in each of the four configurations **[5 points each]**.

Write your answer in the space below

Write your answer in the space below

Problem 3 [20 points] Consider the two priority system of Fig. 3. High priority jobs arrive according to a Poisson process of rate λ_H, while low priority jobs arrive according to an independent Poisson process of rate λ_L. High and low priority jobs are assigned to separate queues, both of infinite capacity, but share a common server of unit service rate. Both high and low priority jobs have exponentially distributed service times with the same mean $1/\mu$. The system is assumed to be stable, *i.e.,* $\frac{\lambda_H+\lambda_L}{\mu} < 1$.

The system operates according to a *preemptive* priority policy. In other words, the server only serves jobs from the low priority queue if the high priority queue is empty. In particular, if a high priority job arrives (to an empty high priority queue) and finds the server busy serving a low priority job, the service of the low priority job is immediately interrupted and the server begins serving the high priority job. The low priority job resumes service only once the high priority queue is empty.

Note 1: The memoryless property of the exponential distribution ensures that a low priority job that resumes service is indistinguishable from one that just starts service.

Note 2: None of the questions below require solving for the probability distribution of the Markov chain representing the overall system.

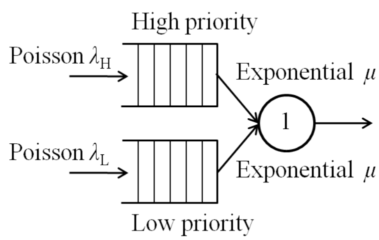

Figure 3: Priority system.

1. **[5 points]** Find an expression, function of the system parameters, for the probability $P^{(L)}_{\text{server busy}}$ that the server is busy serving a *low priority* job.

Write your answer in the space below

2. **[5 points]** Find an expression, function of the system parameters, for the probability π_0 that the system is empty (both queues are empty and the server is idle).

Write your answer in the space below

3. **[5 points]** Find an expression, function of the system parameters, for the average number of high priority jobs in the system, $E[N_H]$.

Write your answer in the space below

4. **[5 points]** Now, find an expression, function of the system parameters, for the average number of low priority jobs in the system, $E[N_L]$. (**Hint**: Unlike the result of the previous question that can be derived directly, this requires an intermediate step.)

Write your answer in the space below

Problem 4 [10 points] Consider the discrete time Markov chain (DTMC) of Fig. 4.

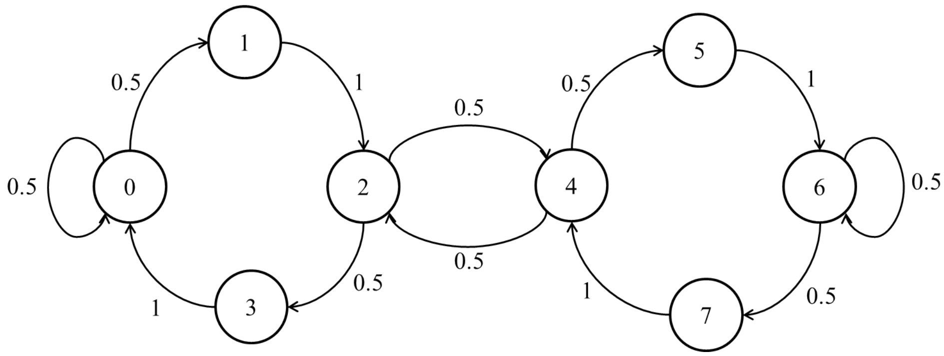

Figure 4: Discrete time Markov chain.

1. **[2 points]** Does the chain admit a limiting distribution? Justify your answer.

Write your answer in the space below

2. **[8 points]** Compute the stationary probabilities $\pi_0, \pi_1, \ldots, \pi_7$ for the chain.

Write your answer in the space below