Quiz No. 1

Problem 1 Consider a single server system with two types of jobs. Customer of type 1 arrive at a rate of λ_1 jobs/sec. Customer of type 2 arrive at a rate of λ_2 jobs/sec. Both types of jobs have an average service time of $E[S]$. We know that $(\lambda_1 + \lambda_2)E[S] < 1$. The system has an infinite waiting facility were jobs can wait to access the server.

1. Provide expressions, function of λ_1, λ_2 and $E[S]$, for the fraction of time the server is busy serving jobs of types 1 and 2, respectively.
   Answer: The average time a job spends in service is $E[S]$, whether the job is of type 1 or 2. Applying Little's Law to the server, we therefore, get the following expressions for the expected number of type i jobs in the server:
   $$E[N_i] = \lambda_i E[S], i = 1, 2$$
   Because $(\lambda_1 + \lambda_2)E[S] < 1$, the system is ergodic and the stationary distribution exists, so that $E[N_i]$ is also the fraction of time the server is busy serving a job of type i, *i.e.,*
   $$\begin{aligned} E[N_i] &= p(\text{server busy serving type } i \text{ job}) \cdot 1 + p(\text{server not serving type } i \text{ job}) \cdot 0 \\ &= p(\text{server busy serving type } i \text{ job}) \end{aligned}$$

2. Assume that you are told that there are on average 50 jobs in the *system* (queue + server). What is, as a function of λ_1, λ_2 and $E[S]$, the average number of jobs in the *queue*?
   Answer: Applying Little's Law to the entire system, we have $E[N] = \lambda E[T]$, where $\lambda = \lambda_1 + \lambda_2$ and $E[N] = 50$. The total number of jobs in the system is equal to the number of jobs in the queue plus the number of jobs in service. As in the previous question, applying Little's Law to the server tells us that the average number of jobs in service is $\lambda E[S]$. Hence the average number of jobs in the queue is of the form
   $$E[Q] = 50 - (\lambda_1 + \lambda_2)E[S]$$

Problem 2 Consider the closed system depicted in Fig. 1, where the number inside each "server" circle (CPU and disk) denotes the corresponding service rate. Assume that the number of jobs circulating in the system is $N = 10$. Give the best possible upper bound you can for the system throughput X.

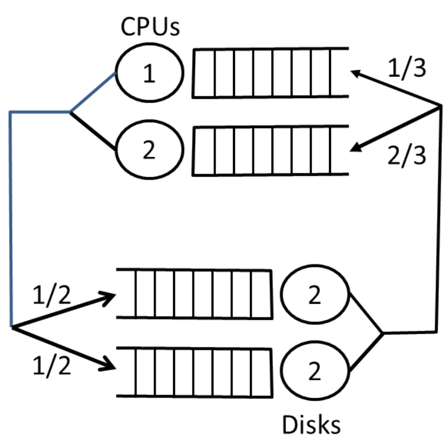

Figure 1: Closed CPU+disk system.

Answser: The expected service times to the two CPUs are $E[S_{CPU1}] = 1$ and $[S_{CPU2}] = \frac{1}{2}$. Because jobs visit a CPU only once and the split ratio between the two CPUs is $\frac{1}{3}, \frac{2}{3}$, the expected number of visits to each

CPU per job are also of the form $E[V_{CPU1}] = \frac{1}{3}$ and $E[V_{CPU2}] = \frac{2}{3}$, so that $E[D_{CPU1}] = \frac{1}{3} \cdot 1 = \frac{1}{3}$ and $E[D_{CPU2}] = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$. Similarly, for the disks, $E[S_{disk1}] = E[S_{disk2}] = \frac{1}{2}$, $E[V_{disk1}] = E[V_{disk2}] = \frac{1}{2}$, and therefore $E[D_{disk1}] = E[D_{disk2}] = \frac{1}{4}$.

This then implies that $D_{\max} = \frac{1}{3}$ and $D = \frac{1}{3} + \frac{1}{3} + \frac{1}{4} + \frac{1}{4} = \frac{7}{6}$. This in turn gives $N^* = \frac{D}{D_{\max}} = \frac{7}{2} = 3.5$. Hence, given that $N = 10 > N^*$, we have the following upper bound for the throughput X:

$$X \leq \frac{1}{D_{\max}} = 3 \text{ jobs/sec}$$