EE549: Problem Set #2 Solutions

I. DEPARTURES

a) (See Fig. 1). The total departures before time t is defined as $D(t)$. These departures represent packets that were fully served (in FIFO order), and the service times thus expended a total time of $\sum_{i=1}^{D(t)} S_i$ throughout the interval $[0, t]$. It follows that this time duration is less than or equal to t.

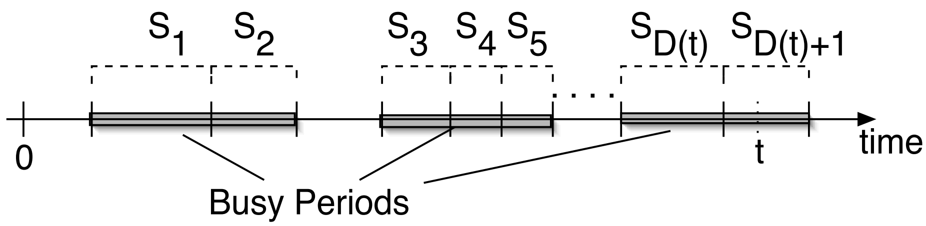

Fig. 1. A timeline illustrating the inequality of part 1(a).

b) If $D(t)$ does not increase to infinity as $t \to \infty$, then the departure rate is 0, which is certainly less than or equal to $\min[1/\mathbb{E}\{S\}, \lambda]$. Here we consider the case when $D(t) \to \infty$ as $t \to \infty$. From the inequality $\sum_{i=1}^{D(t)} S_i \leq t$ (which holds for all $t \geq 0$), we have:

$$\frac{1}{D(t)} \sum_{i=1}^{D(t)} S_i \leq \frac{t}{D(t)}$$

Taking a limit as $t \to \infty$ and using the law of large numbers yields:

$$\mathbb{E}\{S\} \leq \lim_{t \to \infty} \frac{t}{D(t)}$$

and so the departure rate is less than or equal to $1/\mathbb{E}\{S\}$. However, because $D(t) \leq N(t)$ for all t (as the departures cannot be more than the arrivals), we know that:

$$\lim_{t \to \infty} D(t)/t \leq \lim_{t \to \infty} N(t)/t = \lambda$$

because $N(t)$ has rate λ. Therefore, the departure rate is less than or equal to λ and it is *also* less than or equal to $1/\mathbb{E}\{S\}$. Thus, the departure rate is less than or equal to $\min[1/\mathbb{E}\{S\}, \lambda]$.

c) If $\lambda > 1/\mathbb{E}\{S\}$, we know that the departure rate satisfies $\lim_{t\to\infty} D(t)/t \leq \min[1/\mathbb{E}\{S\}, \lambda] = 1/\mathbb{E}\{S\}$ with prob. 1. We have that $L(t) = N(t) - D(t)$. Thus:

$$\begin{aligned} \lim_{t \to \infty} \frac{L(t)}{t} &= \lim_{t \to \infty} \left[\frac{N(t)}{t} - \frac{D(t)}{t} \right] \\ &= \lambda - \lim_{t \to \infty} D(t)/t \\ &\geq \lambda - 1/\mathbb{E}\{S\} \quad \text{with prob. 1} \end{aligned}$$

Because the limit of $L(t)/t$ is greater than or equal to the positive value $\lambda - 1/\mathbb{E}\{S\}$ with probability 1, it follows that $L(t) \to \infty$ with probability 1.[1]

[1] In fact, for any sample path $L(t)$ that satisfies $L(t)/t \to c > 0$, for any $\epsilon > 0$ there exists a time t_ϵ such that $L(t)/t \geq c - \epsilon$ for all $t \geq t_\epsilon$. Thus, $L(t) \geq (c - \epsilon)t$ for all $t \geq t_\epsilon$.

II. Fixed Length Packets and the Multiplexing Inequality

a) Proof for $L_{single}(t)$: Fix a time t. Suppose that $U_{single}(t) = 0$. Then the single-server system is empty and so we also have $L_{single}(t) = 0$ (so that $L_{single}(t) = \left\lceil \frac{U_{single}(t)}{B} \right\rceil$ holds for this case). Now suppose that $U_{single}(t) > 0$. Then $L_{single}(t) > 0$. Also, because of FIFO service, there are exactly $[L_{single}(t) - 1]$ packets in the buffer (none of which have started their transmission), and there is exactly 1 packet in the server. Thus: $U_{single}(t) = (L_{single}(t) - 1)B + R(t)$, where $R(t)$ is the residual amount of bits to be processed in the server, and satisfies $0 < R(t) \leq B$. Thus:

$$L_{single}(t) = \frac{U_{single}(t) - R(t)}{B} + 1 = \frac{U_{single}(t)}{B} + (1 - \frac{R(t)}{B})$$

Note that $1 > 1 - \frac{R(t)}{B} \geq 0$. Because $L_{single}(t)$ is an integer and it is greater than or equal to $U_{single}(t)/B$ and less than $U_{single}(t)/B + 1$, it follows that it is the smallest integer greater than or equal to $U_{single}(t)/B$. $\Box$

Proof for $L_{multi}(t)$: Fix a time t. Note that $U_{multi}(t) = \sum_{i=1}^{L(t)} R_i(t)$, where $R_i(t)$ is defined as the residual amount of unprocessed bits of packet i in the system at time t. We thus have:

$$U_{multi}(t) = \sum_{i=1}^{L_{multi}(t)} R_i(t) \leq L_{multi}(t)B$$

where the final inequality follows because $R_i(t) \leq B$ for all i. Thus, $L_{multi}(t)$ is an integer and $L_{multi}(t) \geq U_{multi}(t)/B$. Therefore, $L_{multi}(t) \geq \lceil U_{multi}(t)/B \rceil$. $\Box$

b) We know by the multiplexing inequality that $U_{single}(t) \leq U_{multi}(t)$ for all t. Therefore $\lceil U_{single}(t)/B \rceil \leq \lceil U_{multi}(t)/B \rceil$ for all t. From part (a) we also have that $L_{single}(t) = \lceil U_{single}(t)/B \rceil$ and $L_{multi}(t) \geq \lceil U_{multi}(t)/B \rceil$ for all t. Thus, $L_{single}(t) \leq L_{multi}(t)$ for all t.

c)

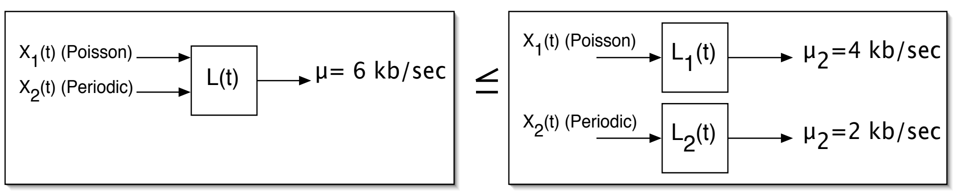

Fig. 2. The packet version of the multiplexing inequality for problem 2c.

From part (b), we know that $L(t) \leq L_1(t) + L_2(t)$, where $L_1(t)$ and $L_2(t)$ are the total number of packets in the multi-queue system shown in Fig. 2, with constant transmission rates μ_1 and μ_2 such that $\mu_1 + \mu_2 = 6$. Therefore:

$$\overline{L} \leq \overline{L}_1 + \overline{L}_2 \tag{1}$$

Let $\mu_1 = 4$ and $\mu_2 = 2$. The queue $L_1(t)$ is thus $M/D/1$ with loading $\rho = \lambda_1 B/\mu_1 = 3/4$. Thus, by the $M/D/1$ formula, we have:

$$\overline{L}_1 = \frac{(3/4)^2}{2(1 - 3/4)} + 3/4 = 15/8 = 1.875$$

The queue $L_2(t)$ has a periodic packet of size 2 kb arriving every second, and has a service time of 1 second, and so there is never more than one packet in queue 2 (i.e., $L_2(t) = 1$ for all t, and so $\overline{L}_2 = 1$). Combining these expressions for $\overline{L}_1$ and $\overline{L}_2$ in the inequality (1) yields:

$$\overline{L} \leq 23/8 = 2.875 \ packets$$

III. Broadcasting with the Law of Large Numbers

a) Each state of the Markov chain represents a number of users that have the current packet (so that no users have the packet in state 0, and all users have the packet in state 3). The first equation states that $\overline{T}_0$ is equal to the sum of 1 (as we must transmit the packet at least once), plus the expected remaining time (conditioning on all possible states that we can be in after the first transmission). No users receive the first transmission with probability $(1-p)^3$, and hence with this probability we are still in state 0 after the first transmission, so the expected remaining time is still $\overline{T}_0$. One user successfully receives the packet after the first transmisison with probability $3p(1-p)^2$, and hence with this probability the expected remaining time is $\overline{T}_1$, etc.

b) Solving the equations gives: $\overline{S}_2 = 1/p = 1.11111$, $\overline{S}_1 = 1.21212$, $\overline{S}_0 = 1.3040$.

c) Let Z_i be the *service time* of packet i, defined as the time required for all 3 users to receive packet i, starting from when it is first transmitted by the satellite. Clearly $\{Z_i\}_{i=1}^{\infty}$ are i.i.d. and have mean $\mathbb{E}\{Z_i\} = \overline{S}_0 = 1.3040$. Thus, $D(t)$ is a renewal process with i.i.d. inter-renewal times Z_i, and so the time average departure rate is equal to $1/\overline{S}_0 = 0.7669$ packets/sec.

IV. Multiplexing Inequality, Slow Trucks, and Delay

a) If the multi-server system never empties, then we are done, as we trivially have that the single-server system empties before the multi-server system. Else, let t^* be the time when the multi-server system empties. We know by the multiplexing inequality that $U_{single}(t) \leq U_{multi}(t)$ for all t, and hence:

$$U_{single}(t^*) \leq U_{multi}(t^*) = 0$$

It follows that $U_{single}(t^*) = 0$, and so the single-server system is also empty at time t^*. Thus, it must have emptied either on or before the time t^*, the multi-server system emptied.

b) Let $B_1 = B$ kb. Let $B_2 = B_3 = \ldots = B_{100} = 1$ kb.

In the single-server system, we have:

$$W_1^{single} = B/2, W_2^{single} = B/2 + 1/2, W_3^{single} = B/2 + 2/2, \ldots, W_{100}^{single} = B/2 + 99/2$$

That is, $W_i^{single} = B/2 + (i-1)/2$ for $i \in \{2, 3, \ldots, 100\}$. Therefore:

$$\begin{aligned}
\overline{W}_{single} &= \frac{1}{100}\sum_{i=1}^{100}[B/2 + (i-1)/2] \\
&= B/2 + \frac{1}{200}\sum_{i=0}^{99} i \\
&= B/2 + \frac{1}{200}(99)(100)/2 \\
&= B/2 + \frac{99}{4}
\end{aligned}$$

In the multi-server system: Suppose that $B/1 \geq 99/1$, so that the first packet takes longer to serve in the multi-server system than all 99 other packets. Thus, we have:

$$W_1^{single} = B, W_2^{single} = 1, W_3^{single} = 2, \ldots, W_{100}^{single} = 99$$

and so:

$$\begin{aligned}
\overline{W}_{multi} &= \frac{B}{100} + \frac{1}{100}\sum_{i=2}^{100} W_i^{multi} \\
&= \frac{B}{100} + \frac{1}{100}\sum_{i=2}^{100}(i-1) \\
&= \frac{B}{100} + \frac{1}{100}(99)(100)/2 \\
&= \frac{B}{100} + \frac{99}{2}
\end{aligned}$$

Therefore:

$$\frac{\overline{W}_{multi}}{\overline{W}_{single}} = \frac{B/100 + 99/2}{B/2 + 99/4} = \frac{1/50 + 99/B}{1 + 99/(2B)}$$

Let $B = 9900$. Then $99/B = 1/100$, and we have:

$$\frac{\overline{W}_{multi}}{\overline{W}_{single}} = \frac{1/50 + 1/100}{1 + 1/200} = 6/201 < 1/30$$

V. Dynamic Routing to Parallel Servers

a) The JSQ strategy routes to the queue $i \in \{1, \ldots, K\}$ such that:

$$i = \arg \min_{i \in \{1,\ldots,K\}} \left[U_i(t^-)/\mu_i\right] \quad \text{(JSQ)}$$

The Greedy strategy routes to the queue $i \in \{1, \ldots, K\}$ such that:

$$i = \arg \min_{i \in \{1,\ldots,K\}} \left[(U_i(t^-) + B)/\mu_i\right] \quad \text{(Greedy)}$$

If $\mu_i = \mu$ (equal rates) for all $i \in \{1, \ldots, K\}$, then the Greedy strategy chooses the queue i with the smallest value of $U_i(t^-)/\mu + B/\mu$, which is the same as choosing the queue with the smallest value of $U_i(t^-)/\mu$, which is the same as the JSQ strategy.

b) Due to a typo, there was no part (b) for this problem.

c) Suppose that we use the Greedy policy for all time, and that $U_i(t) = 0$. Consider any queue j. We want to show that $U_j(t) \leq B_{max}\mu_j/\mu_i$. If $U_j(t) = 0$, then we are done. Else, let B be the size of the packet in queue j that is last in line (so that it is either last in the buffer, or there are no packets in the buffer and it is being processed in the server. The time that this packet will exit queue j is exactly $t_1 = t + U_j(t)/\mu_j$. Because this packet was routed according to the Greedy strategy, this time t_1 must be less than or equal to the time it would have emptied queue i if it was routed there, which is less than or equal to $t_2 = t + B/\mu_i$ (because it would have started its service in queue i at least by time t). Therefore, $t_1 \leq t_2$, and so:

$$t + U_j(t)/\mu_j \leq t + B/\mu_i \leq t + B_{max}/\mu_i$$

Rearranging terms in the above inequality yields:

$$U_j(t) \leq B_{max}\frac{\mu_j}{\mu_i}$$

d) Let $U_{greedy}(t)$ be the total unfinished work under the greedy strategy, and let $U_{other}(t)$ be the total unfinished work under any other routing policy. Fix a particular time t.

Case 1: Suppose time t is *not* in a fully loaded interval of the Greedy system.

In this case, the Greedy system has at least one idle server (so that $U_i(t) = 0$ for at least one queue i), and so $U_j(t) \leq B_{max}\mu_j/\mu_i$ for all $j \neq i$. It follows that:

$$U_{greedy}(t) = \sum_{j \neq i} U_j(t) \leq \sum_{j \neq i}\left(B_{max}\frac{\mu_j}{\mu_i}\right) \leq B_{max}\sum_{j \neq i}\frac{\mu_j}{\mu_1} \leq B_{max}\sum_{j=1}^{K}\frac{\mu_j}{\mu_1} \leq B_{max}\sum_{j=1}^{K}\frac{\mu_j}{\mu_1} + U_{other}(t)$$

Thus, the result holds for Case 1.

Case 2: Suppose time t *is* in a fully loaded interval of the Greedy system.

In this case, let t_f be the start of the current fully loaded interval, and note that $t_f \leq t$ and that all servers of the Greedy system transmit at the full offered transmission rate during the interval from t_f to t. Thus:

$$\begin{aligned}
U_{greedy}(t) &= U_{greedy}(t_f^-) + X[t_f, t] - Y_{greedy}[t_f, t] &(2)\\
&= U_{greedy}(t_f^-) + X[t_f, t] - \int_{t_f}^{t} \mu(\tau)d\tau &(3)\\
&\leq B_{max}\sum_{j=1}^{K}\frac{\mu_j}{\mu_1} + X[t_f, t] - \int_{t_f}^{t}\mu(\tau)d\tau &(4)\\
&\leq B_{max}\sum_{j=1}^{K}\frac{\mu_j}{\mu_1} + X[t_f, t] - Y_{other}[t_f, t] &(5)\\
&\leq B_{max}\sum_{j=1}^{K}\frac{\mu_j}{\mu_1} + U_{other}(t_f^-) + X[t_f, t] - Y_{other}[t_f, t] &(6)\\
&= B_{max}\sum_{j=1}^{K}\frac{\mu_j}{\mu_1} + U_{other}(t) &(7)
\end{aligned}$$

where (2) holds by the I-O equation, (3) holds because the Greedy strategy is fully loaded during $[t_f, t]$, (4) holds because the time just before t_f is not a fully loaded time, so that the total unfinished work is upper bounded by $B_{max}\sum_{j=1}^{K}\frac{\mu_j}{\mu_1}$ as shown in Case 1, and (5) holds because the other policy cannot have more bit departures over the interval $[t_f, t]$ than the sum of the integrals of the offered transmission rates over that interval. Finally, (6) and (7) follow by the I-O equation. Thus, the result also holds for Case 2, and we are done.

e) The JSQ strategy ensures that it keeps unfinished work is no more than $(K-1)B_{max}$ beyond that of any other policy. The Greedy strategy ensures that unfinished work is no more than $B_{max}\sum_{i=1}^{K}(\mu_i/\mu_1)$ beyond any other strategy. This bound can be arbitrarily large (depending on the ratio of the largest rate to the smallest rate). Thus, the JSQ bound is better. However, this does not mean that one policy is "better" than the other. Indeed, the greedy policy can perform better in terms of delay in many cases (such as when only one packet arrives), even if it has a worse bound. The Greedy and JSQ policies are exactly the same if $\mu_i = \mu$ for all i.

VI. Rate Stability for Multi-Server, Shared Buffer Systems

a) Note by the tracking inequality we have:

$$U_{single}(t) \leq U_{multi}(t) \leq U_{single}(t) + (K-1)B_{max} \quad \text{for all } t$$

where $U_{single}(t)$ represents a single-sever queue with the same input process but with a transmission rate process $\mu(t) = \mu_1(t) + \ldots + \mu_K(t)$. Therefore:

$$\lim_{t\to\infty}\frac{U_{single}(t)}{t} \leq \lim_{t\to\infty}\frac{U_{multi}(t)}{t} \leq \lim_{t\to\infty}\frac{U_{single}(t)}{t} + \lim_{t\to\infty}\frac{(K-1)B_{max}}{t}$$

Because $(K-1)B_{max}$ is a constant, we have $(K-1)B_{max}/t \to 0$ as $t \to \infty$, and so:

$$\lim_{t\to\infty}\frac{U_{single}(t)}{t} \leq \lim_{t\to\infty}\frac{U_{multi}(t)}{t} \leq \lim_{t\to\infty}\frac{U_{single}(t)}{t}$$

Now note that the time average rate of $\mu(t)$ is given by $\overline{\mu}_1 + \ldots \overline{\mu}_K$. Thus, by the rate stability theorem, $U_{single}(t)$ is rate stable if and only if $r \leq \overline{\mu}_1 + \ldots \overline{\mu}_K$. Thus, if $r \leq \sum_{i=1}^{K}\overline{\mu}_i$, we have $\lim_{t\to\infty} U_{single}(t)/t = 0$ with prob. 1, and so with prob. 1 we have:

$$0 \leq \lim_{t\to\infty} U_{multi}(t)/t \leq 0$$

which means that $\lim_{t\to\infty} U_{multi}(t)/t = 0$ with prob. 1, and so the multi-server system is rate stable.

b) Suppose that $r > \sum_{i=1}^{K}\overline{\mu}_i$. Note that:

$$U_{multi}(t) \geq X(t) - \int_0^t \sum_{i=1}^{K}\mu_i(\tau)d\tau \quad (8)$$

That is, $U_{multi}(t)$ is at least as large as the total bit arrivals minus the maximum possible bit departures. Dividing (8) by t and taking limits yields (with prob. 1):

$$\begin{aligned}\lim_{t\to\infty}\frac{U_{multi}(t)}{t} &\geq \lim_{t\to\infty}\frac{X(t)}{t} - \sum_{i=1}^{K}\lim_{t\to\infty}\frac{1}{t}\int_0^t \mu_i(\tau)d\tau \\ &= r - \sum_{i=1}^{K}\overline{\mu}_i > 0\end{aligned}$$

Therefore, $U_{multi}(t)/t$ does not tend to 0 as $t \to \infty$, and so $U_{multi}(t)$ is not rate stable.

VII. Rate Stability for a Server Scheduling Problem

Suppose we can design a server allocation rule that yields well defined time average server rates $\overline{\mu}_i$ for all queues i. Note by the rate stability theorem that $U_i(t)$ is then rate stable if and only if $r_i \leq \overline{\mu}_i$. It follows that we just need to design the policy to achieve time average server rates $\overline{\mu}_i$ such that $\overline{\mu}_i \geq r_i$ for all i. The following lemma about probabilistic server allocation is useful:

Lemma 1: Suppose that every slot, a server is placed to queue i i.i.d. with probability p. Then $\overline{\mu}_i = \mu_i p$.

Proof: Let $t_k = kT$ be the time at which the kth timeslot ends (for $k \in \{1, 2, \ldots\}$). Then:

$$\begin{aligned}\frac{1}{t_k}\int_0^{t_k}\mu_i(\tau)d\tau &= \frac{1}{kT}\sum_{m=1}^{k}\int_{(m-1)T}^{mT}\mu_i(\tau)d\tau \\ &= \frac{1}{kT}\sum_{m=1}^{k}\mu_i T 1_m \\ &= \mu_i\left(\frac{1}{k}\sum_{m=1}^{k}1_m\right)\end{aligned}$$

where 1_m is an indicator function that is 1 if queue i is selected on slot m, and zero else. Note that $\{1_m\}_{m=1}^{\infty}$ are i.i.d. with $\mathbb{E}\{1_m\} = p$. Then:

$$\begin{aligned}\lim_{k\to\infty}\frac{1}{t_k}\int_0^{t_k}\mu_i(\tau)d\tau &= \mu_i\left(\lim_{k\to\infty}\frac{1}{k}\sum_{m=1}^{k}1_m\right) \\ &= \mu_i p \quad \text{with prob. 1}\end{aligned}$$

and thus $\mu_i(t)$ has a well defined time average rate that is given by $\mu_i p$. □

a) Let $r_1 = .9$ kb/sec, $r_2 = .4$ kb/sec, and $r_3 = .5$ kb/sec. Suppose that $\mu_1 = \mu_2 = \mu_3 = 1$ kb/sec.

Independetly every slot, flip a coin and make the following decisions:

- Serve queues 1 and 2 with probability 0.5.
- Serve queues 1 and 3 with probability 0.5.

Let $p_i = Pr[\text{Serve queue } i]$. Then clearly $p_1 = 1$, $p_2 = p_3 = 0.5$. Therefore: $\overline{\mu}_1 = p_1 = 1$ kb/sec, $\overline{\mu}_2 = p_2 = 0.5$ kb/sec, $\overline{\mu}_3 = p_3 = 0.5$ kb/sec. Then $r_i \leq \overline{\mu}_i$ for $i \in \{1, 2, 3\}$, and so all queues are rate stable.

b) Let $r_1 = .9$ kb/sec, $r_2 = .45$ kb/sec, and $r_3 = .6$ kb/sec. Supppose that we have the same system as part (a), so that $\mu_1 = \mu_2 = \mu_3 = 1$ kb/sec.

Independently every slot, flip a coin and make the following decisions:

- Serve queues 2 and 3 with prob. 0.1.
- Serve queues 1 and 2 with prob. 0.4.
- Serve queues 1 and 3 with prob. 0.5.

Then:

$$\begin{aligned}p_1 = 0.9 &\implies \overline{\mu}_1 = 0.9 \geq r_1 \\ p_2 = 0.5 &\implies \overline{\mu}_2 = 0.5 \geq r_2 \\ p_3 = 0.6 &\implies \overline{\mu}_3 = 0.6 \geq r_3\end{aligned}$$

Thus, by the rate stability theorem, all queues are rate stable.

c) For all time $t \geq 0$ and for any $i \in \{1, 2, 3\}$, we have:

$$\begin{aligned} U_i(t) &\geq X_i(t) - \int_0^t \mu_i(\tau)d\tau \\ &\geq X_i(t) - t \end{aligned}$$

where the final inequality holds because $\mu_i(\tau) \leq 1$ for all τ. Dividing by t yields:

$$\frac{U_i(t)}{t} \geq \frac{X_i(t)}{t} - 1$$

Taking a limit yields $\lim_{t\rightarrow\infty} U_i(t)/t \geq r_i - 1$. Therefore, if $r_i > 1$, it is impossible for queue i to be rate stable. Thus, if the system is rate stable, then all queues are rate stable and the inequalities $r_1 \leq 1$, $r_2 \leq 1$, $r_3 \leq 1$ cannot be violated.

It remains only to prove that $r_1 + r_2 + r_3 \leq 2$ cannot be violated. We have for all $t \geq 0$:

$$\begin{aligned} U_1(t) + U_2(t) + U_3(t) &\geq X_1(t) + X_2(t) + X_3(t) - \int_0^t \sum_{i=1}^3 \mu_i(\tau)d\tau \\ &\geq X_1(t) + X_2(t) + X_3(t) - 2t \end{aligned}$$

where the final inequality follows because $\sum_{i=1}^3 \mu_i(\tau) \leq 2$ for all τ. Dividing by t and taking a limit yields:

$$\lim_{t\rightarrow\infty} \left[\frac{U_1(t)}{t} + \frac{U_2(t)}{t} + \frac{U_3(t)}{t}\right] \geq r_1 + r_2 + r_3 - 2$$

Therefore, if $r_1 + r_2 + r_3 > 2$, the left hand side cannot be zero, and so the system cannot be rate stable. It follows that $r_1 + r_2 + r_3 \leq 2$ cannot be violated.

d) Independently every slot, flip a coin and make the following decisions:

- Serve queues 4 and 3 with prob. 0.3.
- Serve queues 1 and 2 with prob. 0.5.
- Serve queues 1 and 3 with prob. 0.2.

Then:

$$\begin{aligned} p_1 = 0.7 &\implies \overline{\mu}_1 = 0.7 \geq r_1 \\ p_2 = 0.5 &\implies \overline{\mu}_2 = 0.5 \geq r_2 \\ p_3 = 0.5 &\implies \overline{\mu}_3 = 0.5 \geq r_3 \\ p_4 = 0.3 &\implies \overline{\mu}_4 = 2(0.3) = 0.6 \geq r_4 \end{aligned}$$

Thus, by the rate stability theorem, all queues are rate stable.

e)

Step 1: (Choosing new rates $\tilde{r}_i$): Suppose that the rate vector (r_1, r_2, r_3) satisfies the constraints in (c). Since $r_1 + r_2 + r_3 \leq 2$, we can always choose new rates $(\tilde{r}_1, \tilde{r}_2, \tilde{r}_3)$ such that $r_i \leq \tilde{r}_i \leq 1$ for $i \in \{1, 2, 3\}$, and such that $\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3 = 2$. This can be done as follows: Let $x = 2 - r_1 + r_2 + r_3$. Note that $x \geq 0$.

- If $x = 0$ let $\tilde{r}_1 = r_1, \tilde{r}_2 = r_2, \tilde{r}_3 = r_3$. (Note that $r_i \leq \tilde{r}_i \leq 1$ for all i, and $\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3 = 2$).
- If $x > 0$ and $r_1 + x \leq 1$, then let $\tilde{r}_1 = r_1 + x, \tilde{r}_2 = r_2, \tilde{r}_3 = r_3$. (Note that $r_i \leq \tilde{r}_i \leq 1$ for all i, and $\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3 = 2$).
- If $x > 0$ and $r_1 + x > 1$, then let $\tilde{r}_1 = 1, \tilde{r}_2 = r_2 + [x - (1 - r_1)], \tilde{r}_2 = r_3$. (Note again that $r_i \leq \tilde{r}_i$ for all i, that $\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3 = 2$, and that $\tilde{r}_1 \leq 1, \tilde{r}_3 \leq 1$, and $\tilde{r}_2 \leq 1$. The fact that $r_2 \leq 1$ follows because $\tilde{r}_2 = r_2 + (2 - r_1 - r_2 - r_3) - 1 + r_1 = 1 - r_3 \leq 1$.)

Step 2: (Implementing a server scheduling algorithm based on the $\tilde{r}_i$ rates): Independently every slot, flip a coin and make the following decisions:

- Serve queues 1 and 2 with prob. α.
- Serve queues 1 and 3 with prob. β.
- Serve queues 2 and 3 with prob. γ.

where α, β, γ are probabilities such that $\alpha + \beta + \gamma = 1$ and $\alpha \geq 0, \beta \geq 0, \gamma \geq 0$. These probabilities are to be designed to satisfy certain properties. First note that if we have such probabilities, then we have:

$$\begin{aligned} p_1 = \alpha + \beta &\implies \overline{\mu}_1 = \alpha + \beta \\ p_2 = \alpha + \gamma &\implies \overline{\mu}_2 = \alpha + \gamma \\ p_3 = \beta + \gamma &\implies \overline{\mu}_3 = \beta + \gamma \end{aligned}$$

Thus, it suffices to choose the probabilities α, β, γ to satisfy:

$$\alpha + \beta = \tilde{r}_1 \quad (9)$$

$$\alpha + \gamma = \tilde{r}_2 \quad (10)$$

$$\beta + \gamma = \tilde{r}_3 \quad (11)$$

Solving the above system of three equations with three unknowns yields:

$$\begin{aligned} \alpha &= \frac{\tilde{r}_1 + \tilde{r}_2 - \tilde{r}_3}{2} \\ \beta &= \frac{\tilde{r}_3 + \tilde{r}_1 - \tilde{r}_2}{2} \\ \gamma &= \frac{\tilde{r}_3 + \tilde{r}_2 - \tilde{r}_1}{2} \end{aligned}$$

It is easily verified that this solution satisfies the additional required properties that $\alpha + \beta + \gamma = 1$, because:

$$\alpha + \beta + \gamma = (\tilde{r}_1 - \tilde{r}_1/2) + (\tilde{r}_2 - \tilde{r}_2/2) + (\tilde{r}_3 - \tilde{r}_3/2) = \frac{\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3}{2} = 1$$

Furthermore, these are *valid probabilities* as they are non-negative.[2] Indeed:

$$\begin{aligned} \alpha &= \frac{\tilde{r}_1 + \tilde{r}_2 - \tilde{r}_3}{2} = \frac{2 - \tilde{r}_3 - \tilde{r}_3}{2} = 1 - \tilde{r}_3 \geq 0 \\ \beta &= \frac{\tilde{r}_3 + \tilde{r}_1 - \tilde{r}_2}{2} = \frac{2 - \tilde{r}_2 - \tilde{r}_2}{2} = 1 - \tilde{r}_2 \geq 0 \\ \gamma &= \frac{\tilde{r}_3 + \tilde{r}_2 - \tilde{r}_1}{2} = \frac{2 - \tilde{r}_1 - \tilde{r}_1}{2} = 1 - \tilde{r}_1 \geq 0 \end{aligned}$$

Thus, implementing this randomized policy with these probabilities α, β, γ ensures that:

$$\overline{\mu}_1 = \alpha + \beta = \tilde{r}_1 \geq r_1 \quad (12)$$

$$\overline{\mu}_2 = \alpha + \gamma = \tilde{r}_2 \geq r_2 \quad (13)$$

$$\overline{\mu}_3 = \beta + \gamma = \tilde{r}_3 \geq r_3 \quad (14)$$

and hence all queues are rate stable.

Note: The reason I chose to define new rates $\tilde{r}_i$ is that it is often easier to find values α, β, γ that meet linear constraints with *equality* (as in constraints (9)-(11)), rather than with *inequality* (as in (12)-(14)).

[2]Note that since the are non-negative, the additional property that $\alpha + \beta + \gamma = 1$ ensures that $\alpha \leq 1, \beta \leq 1, \gamma \leq 1$, and so they are indeed valid probabilities.