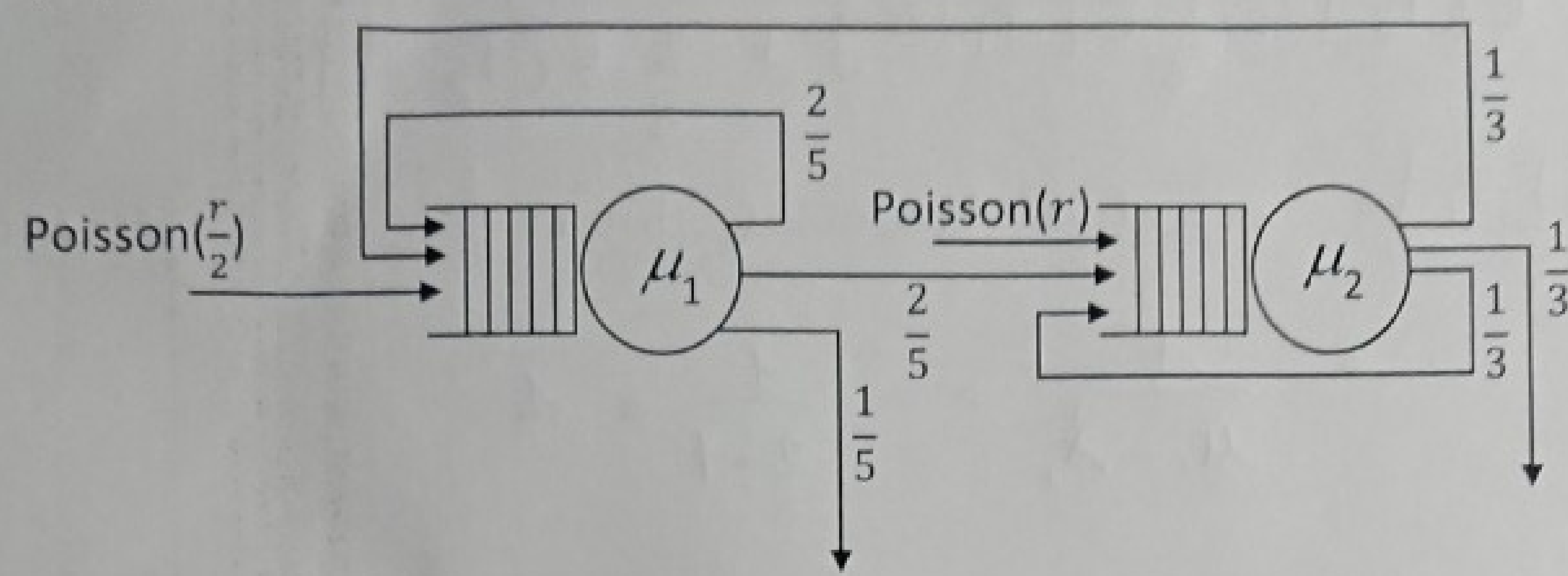
**Ex. 1 (10 pts)**




Consider the network given in the figure, where two routers have an exponential service rate of $\mu_1=2$ packets/second and $\mu_2=6$ packets/second, respectively. Each router has an infinite, FCFS buffer, and routing is probabilistic according to the probability values shown in the figure.
There are two entry points of average arrival rate r/2 and r, respectively, as shown in the figure, which are exponentially distributed.

Answer the following questions:

(a) What is the maximum allowable value $r_{MAX}$ of the parameter $r$ defining the two arrival rates, that the network can tolerate?.

(b) Set $r=0.5\ r_{MAX}$. What is the mean response time $T^1_{TOT}$ of a packet entering at router 1?

*Solve the problem in the space below, then write the answers in the box.*

**Answers:**

$r_{MAX} =$ 0,8

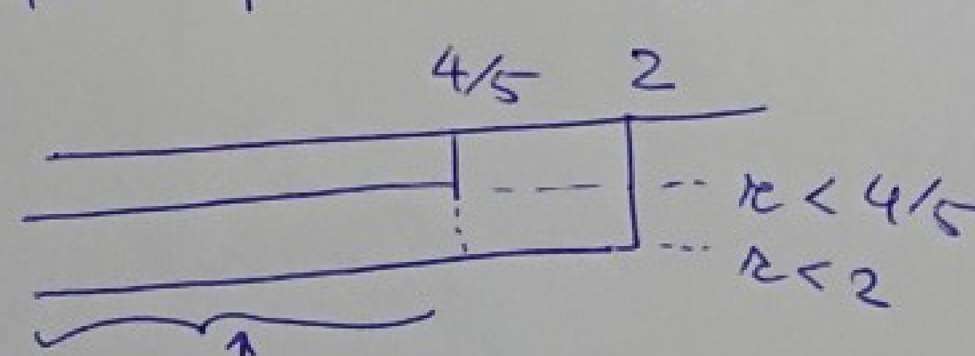
$T^1_{TOT} =$ 45/16

P

$$\left.\begin{aligned}\hat\lambda_1 &= \frac{r}{2} + \frac{2}{5}\hat\lambda_1 + \frac{1}{3}\hat\lambda_2 \\ \hat\lambda_2 &= r + \frac{2}{5}\hat\lambda_1 + \frac{1}{3}\hat\lambda_2\end{aligned}\right\} \Rightarrow \hat\lambda_1 = 5/2\, r \ ; \ \hat\lambda_2 = 3r$$

To ensure stability we impose $\rho_i < 1 \quad i = 1, 2$

$$\left.\begin{aligned}\rho_1 &= \frac{\hat\lambda_1}{\mu_1} = \frac{5}{4} r < 1 \\ \rho_2 &= \frac{\hat\lambda_2}{\mu_2} = \frac{\cancel{3}^{1}}{\cancel{6}_{2}} r < 1\end{aligned}\right\}$$




$r < \frac{4}{5} \rightarrow r_{MAX} = 0{,}8$ (a)

$r = 0.5\ r_{MAX} = 0{,}4$

$\Rightarrow \hat\lambda_1 = \frac{\cancel{5}^{1}}{\cancel{2}} \cdot \frac{\cancel{4}^{2}}{\cancel{10}_{2}} = 1 \ ; \ \hat\lambda_2 = 3 \cdot 0{,}4 = 1{,}2$

$$E[T_1^{TOT}] = E[T_1] + \frac{2}{5} E[T_1^{TOT}] + \frac{2}{5} E[T_2^{TOT}]$$

$$E[T_2^{TOT}] = E[T_2] + \frac{1}{3} E[T_1^{TOT}] + \frac{1}{3} E[T_2^{TOT}]$$

Where $E[T_1]$ = expected traversal time for a single visit =

$$= \frac{1}{\mu_1 - \hat{\lambda}_1} = \frac{1}{2-1} = 1$$

and $$E[T_2] = \frac{1}{\mu_2 - \hat{\lambda}_2} = \frac{1}{6-1.2} = \frac{5}{24}$$

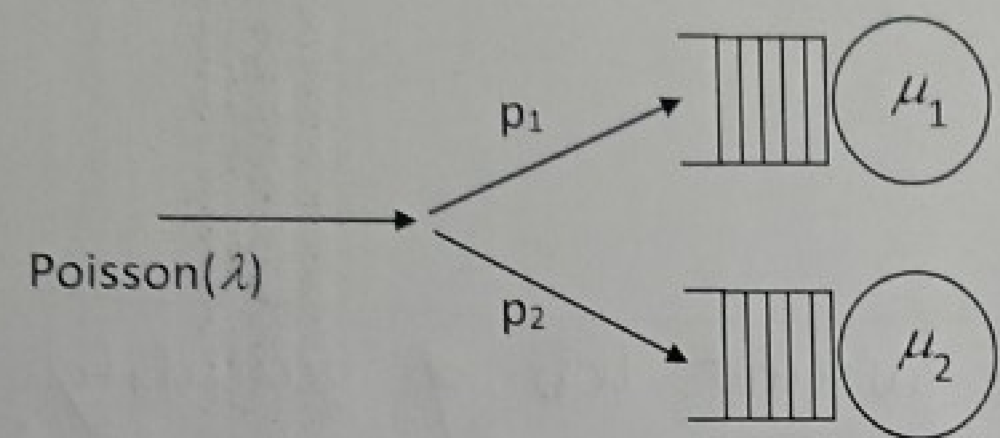
We set $x = E[T_1^{TOT}]$

$y = E[T_2^{TOT}]$

$$\left.\begin{aligned} x &= 1 + \frac{2}{5}x + \frac{2}{5}y \\ y &= \frac{5}{24} + \frac{1}{3}x + \frac{1}{3}y \end{aligned}\right\} \quad \left.\begin{aligned} 3x - 2y &= 1 \\ 16y - 8x &= 5 \end{aligned}\right\} \Rightarrow \begin{aligned} x &= 45/16 \\ y &= 55/32 \end{aligned}$$

$$\Rightarrow E[T_1^{TOT}] = x = 45/16$$

Ex. 2 (13 pts)




Consider the network given in the figure, where an arrival flow with Possonian distribution and rate $\lambda$ is routed probabilistically between two servers with exponential service rates of $\mu_1$ and $\mu_2$, respectively. Prove or disprove that, in the general case with $\mu_1 \neq \mu_2$, load balancing is the best choice in minimizing WAITING TIME and RESPONSE TIME of the above system.

**Answers:**

Load balancing gives optimal response time (Yes/No):
Load balancing gives optimal waiting time (Yes/No):

Case of load balancing: $E[V_1]\cdot E[S_1] = E[V_2]\cdot E[S_2] \Rightarrow \frac{p_1}{\mu_1} = \frac{p_2}{\mu_2}$

we set $p \triangleq p_1 \Rightarrow p_2 = 1-p \Rightarrow \frac{p}{\mu_1} = \frac{1-p}{\mu_2}$ from which

$$\boxed{\rho \triangleq \rho_1 = \frac{p\lambda}{\mu_1} = \frac{(1-p)\lambda}{\mu_2} = \rho_2}$$

$E[T] = p \cdot E[T_1] + (1-p) \cdot E[T_2]$

$E[T_Q] = p\, E[T_{Q1}] + (1-p)\, E[T_{Q2}] = E[T] - p E[S_1] - (1-p) E[S_2]$.

$$E[T_1] = \underbrace{\left(\frac{\rho}{1-\rho}\right)}_{E[N_1]} \cdot \frac{1}{p\lambda}$$

$$E[T_2] = \underbrace{\frac{\rho}{1-\rho}}_{E[N_2]} \cdot \frac{1}{(1-p)\cdot\lambda}$$

$$E[T] = \left(\frac{\rho}{1-\rho}\right)\cdot\left(\frac{1}{\lambda} + \frac{1}{\lambda}\right) = \frac{2\rho}{\lambda(1-\rho)}$$

$$E[T_Q] = \frac{2\rho}{\lambda(1-\rho)} - \frac{p}{\mu_1} - \frac{(1-p)}{\mu_2} =$$

$$= \frac{2\rho}{\lambda(1-\rho)} - \frac{1}{\lambda}(\rho + \rho) =$$

$$= \frac{2\rho}{\lambda(1-\rho)} - \frac{2\rho}{\lambda} = \frac{2\rho(1-1+\rho)}{\lambda(1-\rho)} =$$

$$= \frac{2\rho^2}{\lambda(1-\rho)}$$

In summary in case of load balancing we have

$\rho = \rho_1 = \rho_2$, and

$$E[T] = \frac{2\rho}{\lambda(1-\rho)}$$

$$E[T_Q] = \frac{2\rho^2}{\lambda(1-\rho)}$$

---

Let us consider $\mu_1 > \mu_2$ without loss of generality and set $p_1 = 1$ $p_2 = 0$ (we send all the jobs to the faster server $\Rightarrow$ UNBALANCED LOAD)

We have

$$E[T] = \frac{1}{\mu_1 - \lambda}$$

and $$E[T_Q] = \frac{1}{\mu_1 - \lambda} - \frac{1}{\mu_1}$$

---

Consider now a numerical example with $\lambda = 1$, $\mu_1 = 10$ and $\mu_2 = 1$

Balanced case

$$E[T]_B = \frac{2\rho}{\lambda(1-\rho)} = \frac{2 \cdot \frac{1}{11}}{1\,(10/11)} = 1/5$$

$$\rho = \frac{p_1 \lambda}{\mu_1}$$

$$\begin{cases} p_1 + p_2 = 1 \\ p_1 = \frac{\mu_1}{\mu_2} p_2 \end{cases} \Rightarrow \begin{cases} p_1 = \mu_1/\mu_1+\mu_2 = \frac{10}{11} \\ p_2 = \mu_2/\mu_1+\mu_2 = 1/11 \end{cases} \Rightarrow \rho = \frac{10/11}{10} = \frac{1}{11}$$

$$E[T_Q]_B = \frac{2 \cdot \frac{1}{121}}{(1 - 1/11)} = \frac{2}{\cancel{121}_{11}} \cdot \frac{\cancel{11}}{10} = \frac{1}{55}$$

---

Same numerical example but unbalanced case

$$E[T]_{UB} = \frac{1}{10-1} = \frac{1}{9} < E[T_B] = \frac{1}{5} \rightarrow \underline{\text{better}}$$ if unbalanced case if load is concentrated on the faster server.

$$E[T_Q]_{UB} = \frac{1}{9} - \frac{1}{10} = \frac{1}{90} < E[T_Q]_B = \frac{1}{55}$$