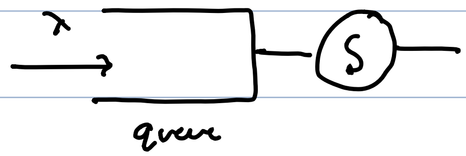
M/M/1 Queue

Markovian Process



Arrival rate

$\lambda < \mu$

$T_s = \frac{1}{\mu}$ ← service rate

$\rho = \lambda * T_s = \frac{\lambda}{\mu}$

Legend :

λ : average arrival rate (Poisson Distribution)

T_s : Average service time (Exp. Distribution)

μ : " " service rate ($\frac{1}{T_s}$)

ρ : " " utilization $0 \leq \rho \leq 1$

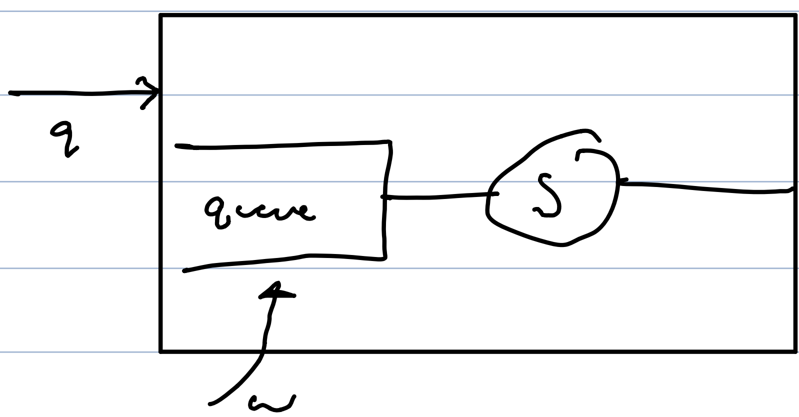
T_q : average turn around time

T_w : " " waiting time

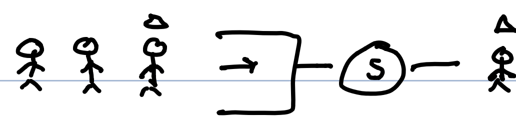
q : average number of request inside the system

w : " " queue

$$T_q = T_w + T_s$$



Littles Law



$$\lambda * T_q = q$$

$$\omega = \lambda T_\omega$$

$$T_q = T_\omega + T_s$$

$$\lambda T_q = \lambda T_\omega + \lambda T_s$$

$$q = \omega + \rho$$

$$q = \frac{\rho}{1-\rho}$$

①

Consider a very small interval of time h

During h
- → Nothing happens
- → One request arrives (birth)
- → One request departs (death).

$$\text{Prob}[\text{No arrivals in } h] = e^{-\lambda h}$$

$$\text{Prob}[\text{an arrival}] = 1 - e^{-\lambda h}$$

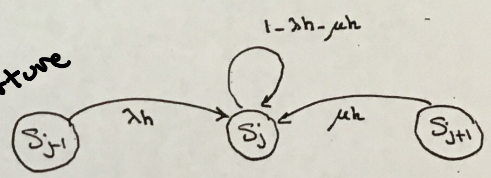
$$= 1 - \left(1 - \lambda h + \frac{(\lambda h)^2}{2!} - \dots\right)$$

→ ignore

$$= \lambda h$$

Similarly $\text{Prob}[\text{a departure}] = \mu h$

Let S_j denotes the state of the system with j customers



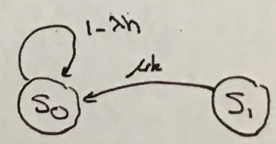
arrival ↓ departure ↓

$$Pr[S_j] = \lambda h \, Pr[S_{j-1}] + \mu h \, Pr[S_{j+1}] + (1-\lambda h-\mu h)\, Pr[S_j]$$ ← nothing happens

$$\mu \, Pr[S_{j+1}] = \mu + \lambda \; Pr[S_j] - \lambda \, Pr[S_{j-1}]$$

$$Pr[S_{j+1}] = 1+\rho \quad Pr[S_j] - \rho \, Pr[S_{j-1}] \dots\dots \quad \boxed{1}$$

S_0 is a special case:



$$Pr[S_0] = 1-\lambda h \;\; Pr[S_0] + \mu h \, Pr[S_1].$$

$$\lambda \, Pr[S_0] = \mu \, Pr[S_1]$$

$$Pr[S_1] = \rho \, Pr[S_0] \dots\dots \quad \boxed{2}$$

$$Pr[S_2] = 1+\rho \quad Pr[S_1] - \rho \, Pr[S_0]$$

$$= (1+\rho)(\rho \, Pr[S_0]) - \rho \, Pr[S_0] =$$

$$= \rho \, Pr[S_0] \, (1+\rho-1) = \rho^2 \, Pr[S_0]$$

Utilization ↗

$$Pr[S_j] = \rho^j \, Pr[S_0] \ldots\ldots\ldots \boxed{3}$$

All prob. must add up to 1, so

$$Pr[S_0] + Pr[S_1] + Pr[S_2] + \ldots\ldots = 1.$$

$$Pr[S_0]\left(1 + \rho + \rho^2 + \rho^3 + \ldots\ldots\right) = 1.$$

$$\frac{Pr[S_0]}{1-\rho} = 1. \qquad Pr[S_0] = 1 - \rho \ \ldots\ldots \boxed{4}$$

q = Expected number of Customers in the System.

$$q = 0 \, Pr[S_0] + 1 \, Pr[S_1] + 2 \, Pr[S_2] + \ldots\ldots$$

$$= 0 + \rho(1-\rho) + 2\rho^2(1-\rho) + \ldots\ldots$$

$$= (1-\rho)\left[\rho + 2\rho^2 + 3\rho^3 + \ldots\right]$$

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

$$= \frac{\rho}{1-\rho}$$

$$\boxed{q = \frac{\rho}{1-\rho} \ \ldots\ldots \boxed{5}}$$

Let $X = \rho + 2\rho^2 + 3\rho^3 + \ldots$

$$X = \rho\left(1 + 2\rho + 3\rho^2 + \ldots\right)$$

$$X = \rho\left(\begin{array}{l} 1 + \rho + \rho^2 + \ldots \\ \quad\ \ \rho + 2\rho^2 + \ldots \end{array}\right)$$

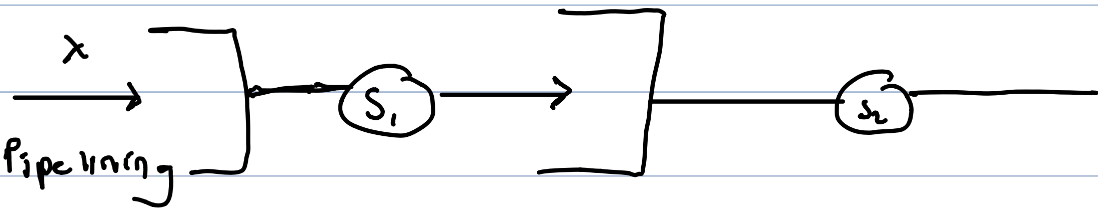
$$X = \rho\left(\frac{1}{1-\rho} + X\right)$$

$$(1-\rho)X = \frac{\rho}{1-\rho}$$

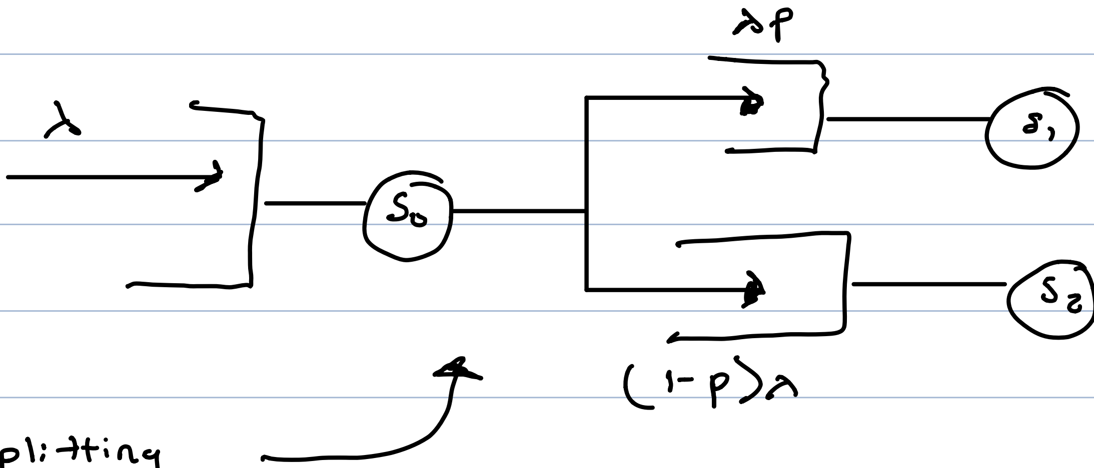
$$X = \frac{\rho}{(1-\rho)^2}$$

⋮

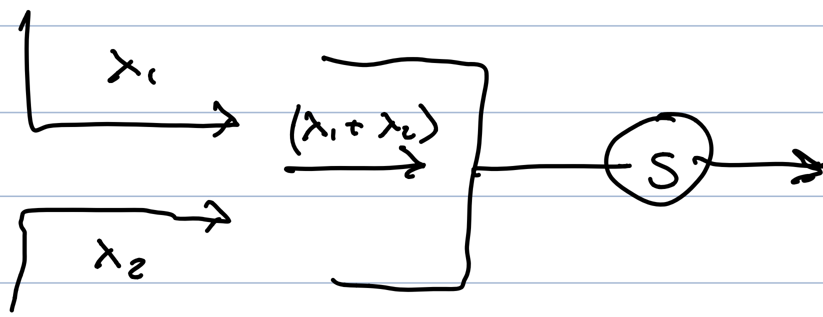
Network of queues



Pipelining

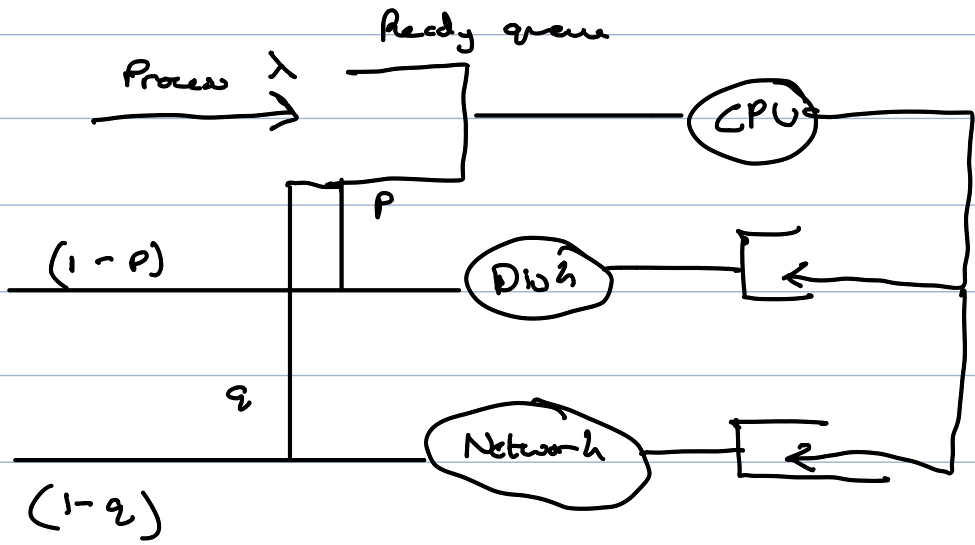


Splitting



Merging

Jackson's theorem



Assumptions

→ Arrivals following poisson
→ Service times follow Exp
→ Requests are either waiting or getting served
↳ each queue can be analyzed separately