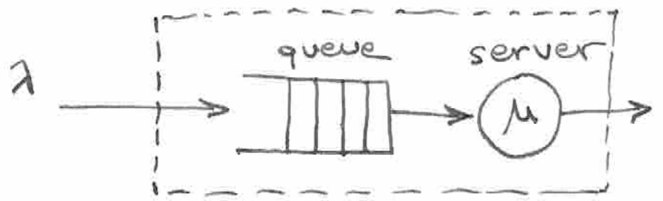
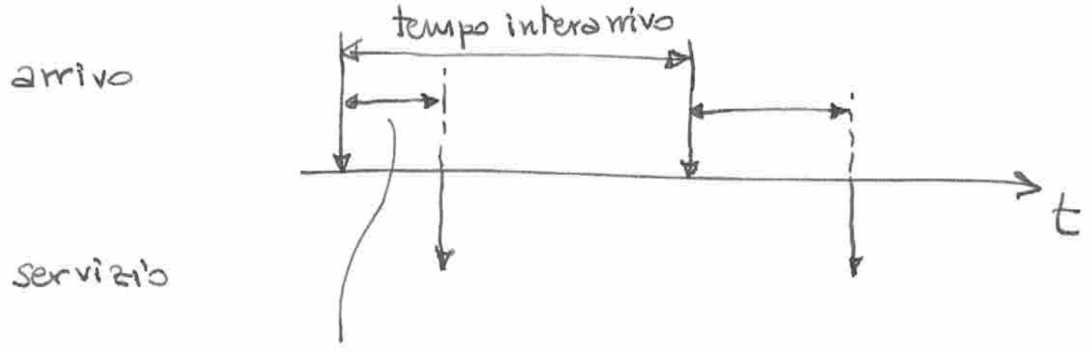
SISTEMI A CODA

In particolare ci focalizzeremo su quelli con popolazione ∞ in cui il server può essere singolo o multiplo e in cui la lungh. della coda è finita o infinita

Sul singolo servitore





Un parametro fondamentale per il progetto dei sistemi a coda è il <u>fattore di utilizzo</u> ρ.

$$\rho = \frac{\text{tasso ingresso}}{\text{max tasso smaltimento}}$$

partenze
uscite

A regime (steady state)

$$\rho = \mathbb{E}\{\lambda\}\,\mathbb{E}\{T_x\}$$

ritmo arrivi — tempo servizio

Se $\rho > 1$ allora il sistema a coda non riesce a smaltire tutte le richieste

Th. Per un qualsiasi sistema a coda con 1 servitore $(G/G/1)$ si ha

$$\boxed{\rho = 1 - P_\emptyset}$$ (prob. di almeno 1 pkt/ut nel sistema)

con $P_\emptyset$ prob. di servitore libero.

dim.

Osserviamo il sistema per un tempo t_{OBS} abbastanza grande $\Rightarrow$ numero di arrivi $E\{\lambda\} t_{OBS}$ (con prob. 1 se $t_{OBS} \to +\infty$)

Occupazione temporale del server $t_{OBS}(1 - P_\emptyset)$

Se $E\{T_x\}$ è il tempo medio di servizio, la quantità media di lavoro portato dagli utenti per unità di tempo è $E\{\lambda\} E\{T_x\}$

Con singolo servitore la capacità max di smaltire il lavoro per unità di tempo è 1 per cui

medio pkt/ut arrivati in t_{OBS} = # pkt/ut serviti in t_{OBS}

$$E\{\lambda\} t_{OBS} = \frac{t_{OBS}(1 - P_\emptyset)}{E\{T_x\}}$$

da cui

$$\rho = E\{\lambda\} \underbrace{E\{T_x\}}_{\frac{1}{\mu}} = 1 - P_\emptyset \quad \checkmark$$

Nota: $\rho < 1$

16 bis se disciplina di servizio Markoviana

Se la disciplina di servizio è markoviana significa che il tempo fra due servizi ha distribuzione esponenziale con parametro μ

$$f_{T_x}(\xi) = \mu e^{-\mu\xi} \quad ; \xi \geq 0 \text{ e zero altrove}$$

Per questa PDF

$$E\{T_x\} = \frac{1}{\mu}$$

$$E\{T_x^2\} = \frac{2}{\mu^2}$$

$$\sigma_{T_x}^2 = \frac{1}{\mu^2}$$

analogo al tempo di interarrivo

$$f_\tau(\xi) = \lambda e^{-\lambda\xi}$$

$$E\{\tau\} = \frac{1}{\lambda}$$

$$E\{\tau^2\} = \frac{2}{\lambda^2}$$

$$\sigma_\tau^2 = \frac{1}{\lambda^2}$$

$$E\{T_x\} = \int_{-\infty}^{+\infty} \xi \, f_{T_x}(\xi) \, dx$$

Parametri caratteristici sistema a coda

- descr. arrivi (distrib., λ)
- descr. servizio (distrib., μ)
- # servitori (c)

analysis $\Downarrow$ $\uparrow$ design

A regime (steady state)

coda o servitore

- prob. di stato (prob. k pkt/ut nel sist.) P_k
- # medio pkt/ut. nel sistema $L_s = E\{k\}$
- # medio pkt/ut in coda $L_q = E\{q\}$
- Tempo medio attesa nel sistema $W_s = E\{T_s\}$
- Tempo medio attesa in coda $W_q = E\{T_q\}$

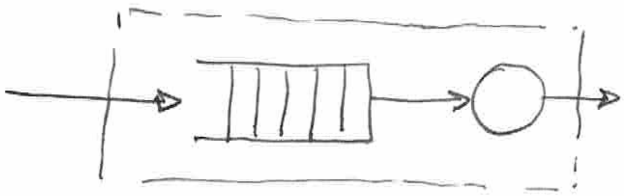
Vedremo che sono tutti in relazione fra loro.

Considereremo sist a coda su popolazioni infinite con server singoli o multipli e lungh. delle code (buffer) finite o infinite.

Ricorda: nei processi di nascita e morte un solo evento (arrivo or servizio)

partenza

Sistema M/M/1



$$\lambda_k = \lambda$$
$$\mu_k = \mu$$

Arrivi Markoviani (assenza di memoria) $\Leftrightarrow$ lo stato del sistema non dipende dalla storia precedente

* distribuzione degli arrivi Poissoniana $\lambda \left[\frac{pkt}{s.}, \frac{ut}{hr.}\right]$
* distribuzione del tempo di servizio esponenziale μ " "

Ricorda $\quad P_k = P_\emptyset \prod_{i=\emptyset}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \qquad k \geq 1$

da cui

$$\boxed{P_k = P_\emptyset \left(\frac{\lambda}{\mu}\right)^k}$$

e $P_\emptyset = \left[1 + \sum_{k=1}^{+\infty} \prod_{i=\emptyset}^{k-1} \frac{\lambda_i}{\mu_{i+1}}\right]^{-1}$

da cui

$$P_\emptyset = \left[1 + \sum_{k=1}^{+\infty} \left(\frac{\lambda}{\mu}\right)^k\right]^{-1} = \left[\sum_{k=\emptyset}^{+\infty} \left(\frac{\lambda}{\mu}\right)^k\right]^{-1} \underset{\frac{\lambda}{\mu}<1}{=} \left(\frac{1}{1-\frac{\lambda}{\mu}}\right)^{-1}$$

$$\Rightarrow \boxed{P_\emptyset = 1 - \frac{\lambda}{\mu}}$$

Quindi

$$P_k = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^k \; ; \; k \geq \emptyset$$

$$E\{\lambda\} = \sum_{k=\emptyset}^{+\infty} \lambda_k P_k = \lambda \underbrace{\sum_{k=\emptyset}^{+\infty} P_k}_{1} = \lambda$$

$$E\{T_x\} = \frac{1}{\mu} \qquad \Rightarrow \qquad \rho = \frac{\lambda}{\mu}$$

Markov (distrib. exp.)

quindi

$$\boxed{P_k = (1-\rho)\,\rho^k} \; ; k \geq \emptyset$$

$$= P_\emptyset \, (1-P_\emptyset)^k$$

Inoltre

$$L_S = E\{k\} = \sum_{k=\emptyset}^{+\infty} k\,P_k = \frac{\rho}{1-\rho}$$

$$E\{k^2\} = \frac{\rho^2+\rho}{(1-\rho)^2} \quad , \quad \sigma_k^2 = \frac{\rho}{(1-\rho)^2}$$

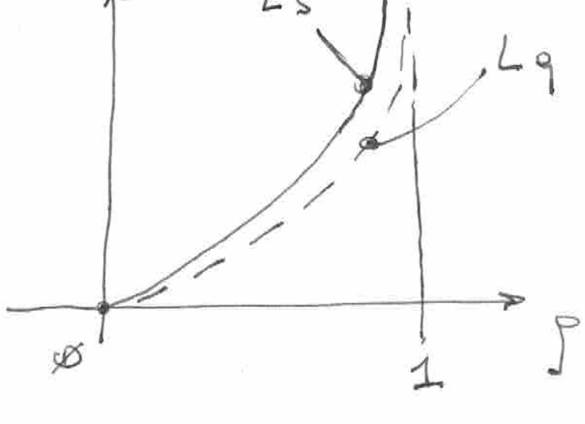
dalla relazione di Little

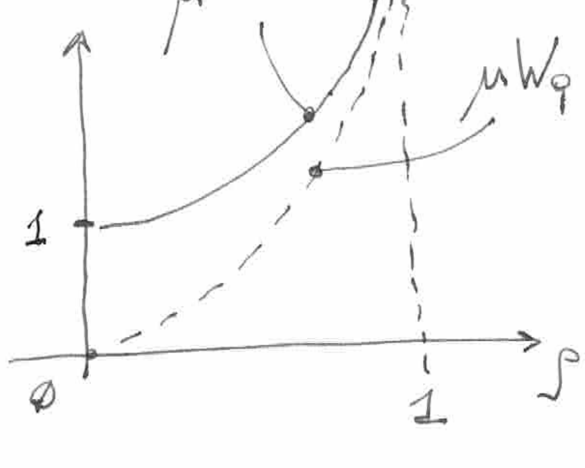
$$W_S = \frac{L_S}{\lambda} = \frac{\rho}{1-\rho}\frac{1}{\lambda} = \frac{1/\mu}{1-\rho} \qquad W_x = E\{T_x\}$$

$$W_q = W_S - W_x = W_S - \frac{1}{\mu} = \frac{1}{\mu}\;\frac{\rho}{1-\rho}$$

dalla relazione di Little

$$L_q = E\{q\} = \lambda W_q = \frac{\rho^2}{1-\rho}$$





Es. Sistema M/M/1 all'equilibrio con $\lambda = 8$ pkt/s., $\mu = 9$ pkt/s.

sono valori medi di var. aleatorie

$$\rho = \frac{8}{9} \simeq 0.89 \qquad P_0 = 1 - \rho = \frac{1}{9}$$

$\rho = 1 - P_0$ prob. di server in utilizzo

prob. di server libero (no pkt nel sistema)

$$\mathbb{P}\{\text{no queue}\} = P_0 + P_1 = \frac{1}{9} + \frac{8}{81} = \frac{17}{81} \simeq 0.21$$

$$P_1 = (1-\rho)\rho = \frac{8}{81}$$

nel 21% dei casi un pkt non entra in coda.

$$\mathbb{P}\{\text{queue}\} = P_2 + P_3 + \ldots = 1 - (P_0 + P_1) \simeq 0.79$$

nel 79% dei casi un pkt entra in coda e non viene servito subito

Prob. di 10 pkt nel sistema (1 servito e 9 in coda)

$$P_{10} = \rho^{10}(1-\rho) \simeq 0{,}034$$

nel 3.4% si trovano 10 pkt nel sistema.

$$L_s = \frac{\rho}{1-\rho} = \frac{8/9}{1-8/9} = 8$$

in media 8 pkt nel sistema di cui 1 servito e 7 in coda

$$L_q = \frac{\rho^2}{1-\rho} \simeq 7.11$$

$$W_s = \frac{L_s}{\lambda} = 1 \text{ s.}$$

un pkt aspetta mediamente 1 s. nel sistema (in coda o server)

$$W_q = \frac{L_q}{\lambda} \simeq 0.89 \text{ s.}$$

un pkt sta mediamente 0.89 s. in coda (prima di essere servito)

Osservazione: quando una risorsa Internet $(M/M/1)$ è condivisa fra pochi utenti $(\rho \to \emptyset)$ allora $W_s \to \frac{1}{\mu}$, $W_q \to \emptyset$, $L_s \to \emptyset$, $L_q \to \emptyset$.

Si può osservare che se la risorsa è condivisa fra pochi utenti l'arrivo dei pkt tende a essere deterministico con memoria.

Al crescere del numero di utenti l'arrivo dei pkt tende ad essere senza memoria (Markoviano)

Inoltre nei sist. M/M/1 la prob. di avere almeno N pkt/ut. nel sistema (coda o server) è

$$\mathbb{P}\{k \geq N\} = \sum_{k=N}^{+\infty} P_k = 1 - \sum_{k=\emptyset}^{N-1} P_k = 1 - \sum_{k=\emptyset}^{N-1} (1-\rho)\rho^k$$

$$= 1 - (1-\rho) \sum_{k=\emptyset}^{N-1} \rho^k = 1 - (1-\rho)\frac{1-\rho^N}{1-\rho} = \rho^N$$

In alcuni sistemi gli arrivi tendono ad essere ritardati tanto più quanto più il sistema è affollato

Sistema M/M/1 con arrivi rallentati

$$\lambda_k = \frac{\alpha}{k+1}$$

$$\mu_k = \mu$$

$$P_\phi \prod_{i=\phi}^{k-1} \frac{\alpha}{\mu} \frac{1}{i+1} = P_\phi \left(\frac{\alpha}{\mu}\right)^k \prod_{i=\phi}^{k-1} \frac{1}{i+1}$$

All'equilibrio $P_k = P_\phi \prod_{i=\phi}^{k-1} \frac{\lambda_i}{\mu_{i+1}} = P_\phi \left(\frac{\alpha}{\mu}\right)^k \frac{1}{k!}$

$$P_\phi = \left[1 + \sum_{k=1}^{+\infty} \left(\frac{\alpha}{\mu}\right)^k \frac{1}{k!}\right]^{-1} = \left[\sum_{k=\phi}^{+\infty} \left(\frac{\alpha}{\mu}\right)^k \frac{1}{k!}\right]^{-1} = e^{-\frac{\alpha}{\mu}}$$

da cui $\boxed{P_k = \frac{1}{k!} \left(\frac{\alpha}{\mu}\right)^k e^{-\frac{\alpha}{\mu}}}$; $k \geq \phi$

Il fattore di utilizzo è $\underset{\underset{E\{\lambda\} E\{T_x\}}{\parallel}}{\rho} = 1 - P_\phi = 1 - e^{-\alpha/\mu}$

da cui $E\{\lambda\} = \frac{\rho}{E\{T_x\}} = \frac{\rho}{1/\mu} = \rho\mu = \mu\left(1 - e^{-\alpha/\mu}\right)$

P_k Poissoniana per cui media $L_s = E\{k\} = \frac{\alpha}{\mu}$

$$W_s \underset{\text{Little}}{=} \frac{L_s}{E\{\lambda\}} = \frac{\alpha}{\mu} \frac{1}{\mu\left(1 - e^{-\alpha/\mu}\right)} = \frac{\alpha}{\mu^2 \rho}$$

$$W_q = W_s - E\{T_x\} = \frac{\alpha}{\mu^2 \rho} - \frac{1}{\mu} = \frac{\alpha - \rho\mu}{\mu^2 \rho}.$$

$$L_q \underset{\text{Little}}{=} W_q\, E\{\lambda\} = \frac{\alpha - \rho\mu}{\mu}.$$

Es.

Si consideri un sistema a coda M/M/1 (all'equilibrio) con arrivi rallentati

$$\lambda_k = \frac{\alpha}{k+1} \quad pkt/s. \qquad \alpha = 10 \; pkt/s.$$

Si determini il minimo tasso di servizio μ^* affinché il tempo medio nel sistema non superi $W_s^* = 500$ ms.

Rel. Little $\quad W_s = \frac{L_s}{E\{\lambda\}} = \frac{\alpha}{\mu^2 (1 - e^{-\alpha/\mu})}$

$$W_s \leq W_s^* \iff \frac{\alpha}{\mu^2 (1 - e^{-\alpha/\mu})} \leq W_s^*$$

$$\iff \mu^2 (1 - e^{-\alpha/\mu}) \geq \frac{\alpha}{W_s^*} = 20 \; pkt.$$

Sol. iterativa

$$g(\mu) = \mu^2 (1 - e^{-10/\mu}) \geq 20$$

$\mu = 5 \longrightarrow g(5) \simeq 21,62$

$\mu = 4 \longrightarrow g(4) \simeq 14,69$

$\mu = 4.7 \longrightarrow g(4.7) \simeq 19,46$

$\mu = 4.8 \longrightarrow g(4.8) \simeq 20,17$

$\mu^* \simeq 4.8$ pkt/s.