# Filas de espera multiservidor con capacidad y redes

Rodrigo Maranzana

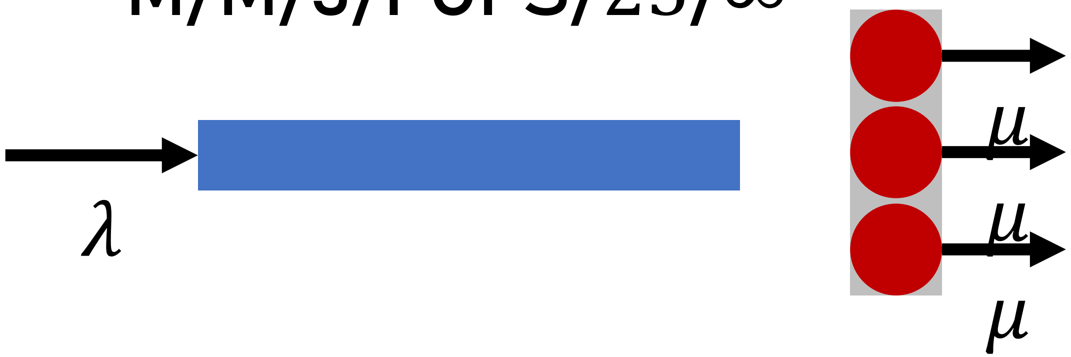
# Repaso: notación de Kendall

## M/M/1/FCFS/∞/∞


- Arribos $\sim Exp(\lambda)$
- Servicio $\sim Exp(\mu)$
- 1 servidor
- Primero llegado primero servido (FCFS)
- Capacidad infinita del sistema
- Fuente infinita

Se suele abreviar a:
M/M/1

## M/M/3/FCFS/25/∞




- Arribos $\sim Exp(\lambda)$
- Servicio $\sim Exp(\mu)$
- 3 servidores
- Primero llegado primero servido (FCFS)
- Capacidad de 25 personas
- Fuente infinita

Se suele abreviar a:
M/M/1/25

# Repaso: factor de tráfico

Es la relación entre la tasa de arribos y despachos.

Si “M” es la cantidad de servidores.

$$\rho = \frac{\lambda}{M\mu}$$

Casos:

$\rho \geq 1$ sistema inestable.

$\rho < 1$ sistema estable.

# Repaso: métricas y parámetros

## Cantidad de clientes promedio:

- En la fila: $L_q$ *[unidades o agentes]*
- En el sistema: $L_s$ o $L$ *[unidades o agentes]*

## Tiempo de espera promedio:

- En la fila: $W_q$ *[unidad de tiempo]*
- En el sistema: $W_s$ o $W$ *[unidad de tiempo]*

## Probabilidad de estado (que hayan “i” agentes):

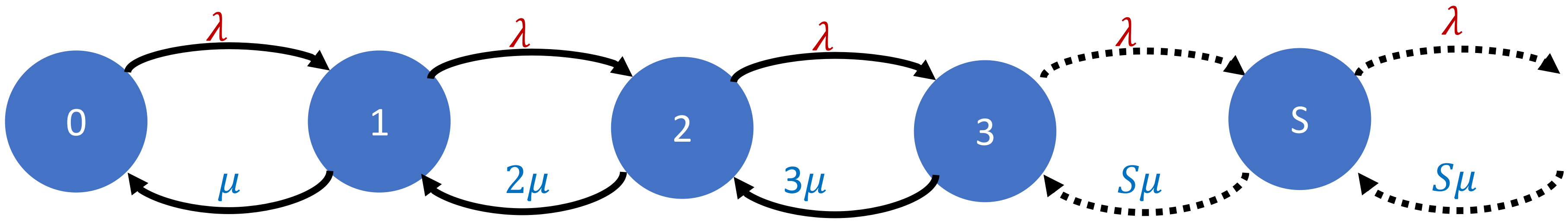
$P(X = i)$

# Filas de espera M/M/S

Representan el caso de una fila y múltiples servidores:

# Proceso de Nacimiento y Muerte M/M/S

## Matriz generadora:




$$
Q = \begin{bmatrix}
-\lambda & \lambda & 0 & 0 & \cdots & 0 & \cdots \\
\mu & -\lambda-\mu & \lambda & 0 & \cdots & 0 & \cdots \\
0 & 2\mu & -\lambda-2\mu & \lambda & \cdots & 0 & \cdots \\
0 & 0 & 3\mu & -\lambda-3\mu & \cdots & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \lambda & \cdots \\
0 & 0 & 0 & 0 & S\mu & -\lambda-S\mu & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
$$

# Filas de espera M/M/S

Factor de tráfico:

$$\rho = \frac{\lambda}{M\mu}$$

Probabilidad de sistema ocioso:

$$P_0 = \frac{1}{\left[\sum_{i=0}^{M-1} \frac{(\lambda/\mu)^i}{i!}\right] + \frac{(\lambda/\mu)^M}{M!\,(1-\rho)}}$$

Probabilidad de sistema con “n” agentes:

$$P_n = \begin{cases} \dfrac{P_0\left(\dfrac{\lambda^n}{\mu}\right)}{n!} & si\ 0 < n \leq M \\ \dfrac{P_0\left(\dfrac{\lambda^n}{\mu}\right)}{M!\,(M^{n-M})} & si\ n > M \end{cases}$$

# Filas de espera M/M/s

## Cantidad de clientes promedio

- En el sistema:

$$L = \lambda W$$

*(Ley de Little)*

$$L = L_q + \frac{\lambda}{\mu}$$

- En la fila:

$$L_q = \frac{P_0 \left(\frac{\lambda}{\mu}\right)^M \rho}{M!\ (1-\rho)^2}$$

## Tiempo de espera promedio

- En el sistema:

$$\mathrm{W} = \mathrm{W_q} + \frac{1}{\mu}$$

- En la fila:

$$\mathrm{W_q} = \frac{L_q}{\lambda}$$

# Ejemplo

Una línea automatizada tiene tres tornos CNC idénticos.

La materia forma una única fila de espera al pie de las 3 máquinas esperando ser procesada.

Las cantidades que arriban y se procesan siguen una distribución de Poisson.

Además se sabe que la tasa de procesamiento de los tornos es de $\mu = 6\ \mathrm{u/hora}$, y la materia prima llega con una tasa de $\lambda = 16\ u/hora$.

1. ¿El sistema es estable?
2. Largo de la fila promedio.
3. Tiempo que una unidad pasa en el sistema.

# 1. ¿El sistema es estable?

$$\rho = \frac{\lambda}{M\mu} = \frac{16\, u/h}{3 * 6\, u/h} = 0.88$$

Menor a 1, sistema estable.

# 2. Largo de la fila promedio

$$P_0 = \frac{1}{\left[\sum_{i=0}^{M-1} \frac{({}^{\lambda}/_{\mu})^i}{i!}\right] + \frac{({}^{\lambda}/_{\mu})^M}{M!\,(1-\rho)}} = \frac{1}{\left[\sum_{i=0}^{3-1} \frac{({}^{16}/_{6})^i}{i!}\right] + \frac{({}^{16}/_{6})^3}{3!\,(1-0.88)}} = 0.0311$$

$$P_0 = 3{,}11\%$$

$$L_q = \frac{P_0 \left(\frac{\lambda}{\mu}\right)^M \rho}{M!\ (1-\rho)^2} = \frac{0.028 \left(\frac{16}{6}\right)^3 0.88}{3!\ (1-0.88)^2} = 7.08\ unidades$$

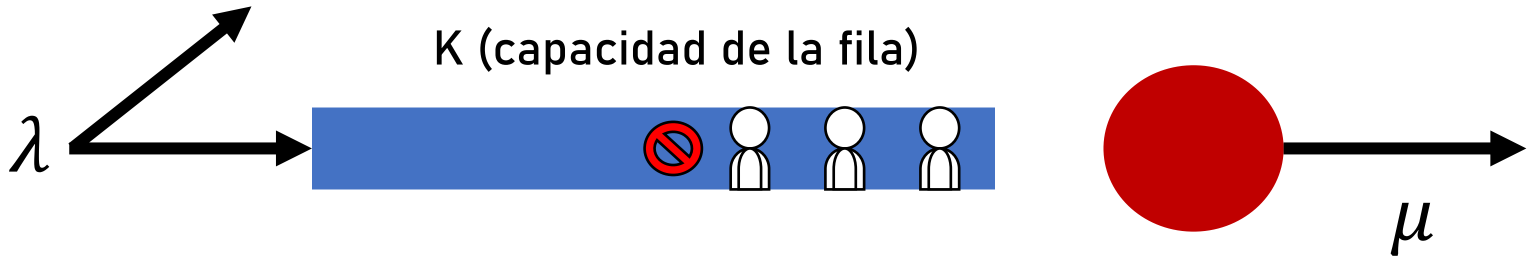
# 3. Tiempo en que un unidad pasa en el sistema

$$\mathrm{W_q} = \frac{L_q}{\lambda} = \frac{7.08\,u}{16\,u/h} = 0.442\,h = 26.55\,min$$

$$\mathrm{W} = \mathrm{W_q} + \frac{1}{\mu} = 0.442\,h/u + \frac{1}{6\,\frac{u}{h}} = 0.609\,h = 36.52\,min$$

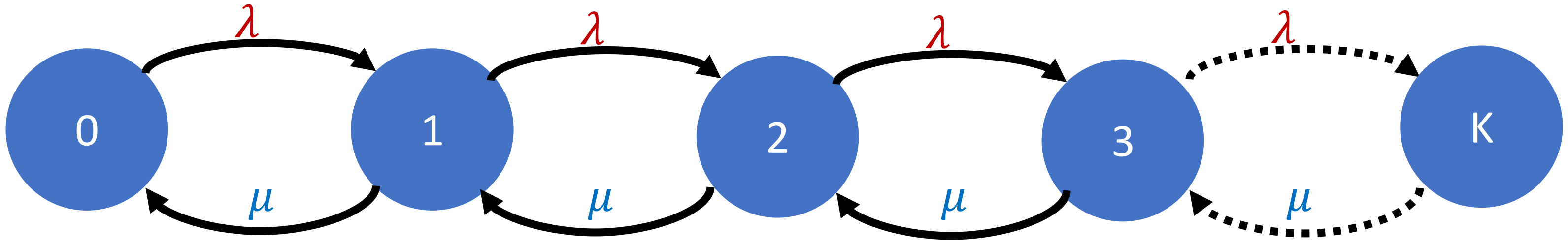
# Filas de espera M/M/1/K

Representan el caso de una fila con capacidad máxima:



# Proceso de Nacimiento y Muerte M/M/1/K

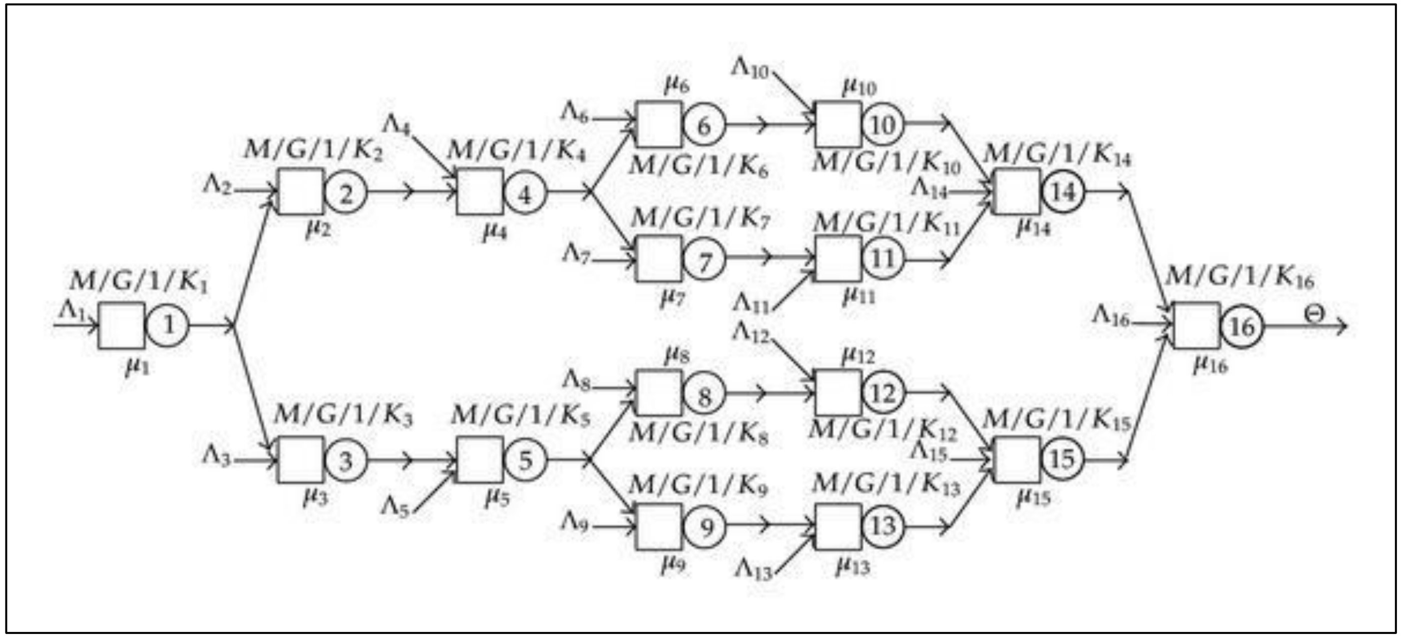
## Matriz generadora:




$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \cdots & 0 \\ \mu & -\lambda-\mu & \lambda & 0 & \cdots & 0 \\ 0 & \mu & -\lambda-\mu & \lambda & \cdots & 0 \\ 0 & 0 & \mu & -\lambda-\mu & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \lambda \\ 0 & 0 & 0 & 0 & \mu & -\lambda-\mu \end{bmatrix}$$

# Redes de filas de espera

Es una forma de modelizar un sistema mediante filas de espera conectadas.

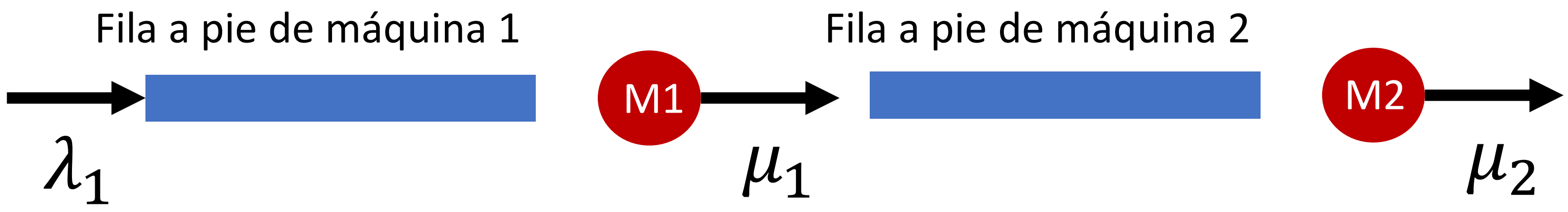



Ejemplo: red compleja modelizada con filas de espera

Cruz et Al. (2012), "Throughput Maximization of Queueing Networks with Simultaneous Minimization of Service Rates and Buffers" (https://www.hindawi.com/journals/mpe/2012/692593/ )

# Redes de filas de espera

Por ejemplo, una unidad productiva en serie se puede modelizar como:




La tasa de salida $\mu_1$ de M1 equivale a la de llegada de M2.