Modelos de Filas de Espera

La Teoría de Filas

Es el estudio de los comportamientos de los sistemas de atención, en el que los clientes esperan por un servicio

El Danés *Agner Erlang* (1878-1929) en 1909 aplicó la teoría de probabilidad para analizar el comportamiento de las llamadas telefónicas, para mejorar el servicio de espera.

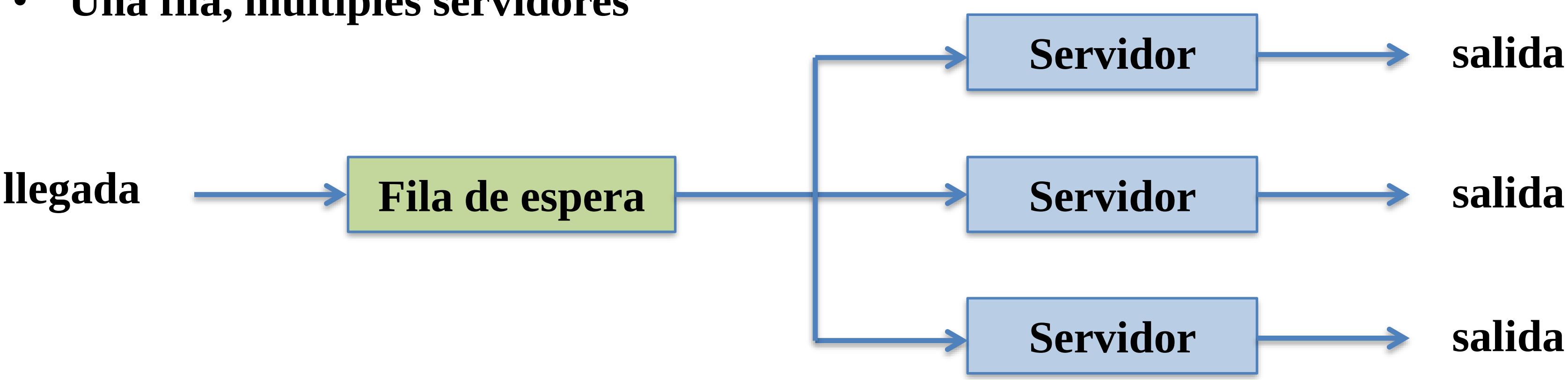
La aplicación de Teoría de Colas (Filas de Espera) busca la eficacia (dar un buen servicio) y la eficiencia (optimizar el uso de recursos) en procesos de atención a clientes

- **El tiempo de espera**
- **El tiempo de atención**

Tipos de Sistemas

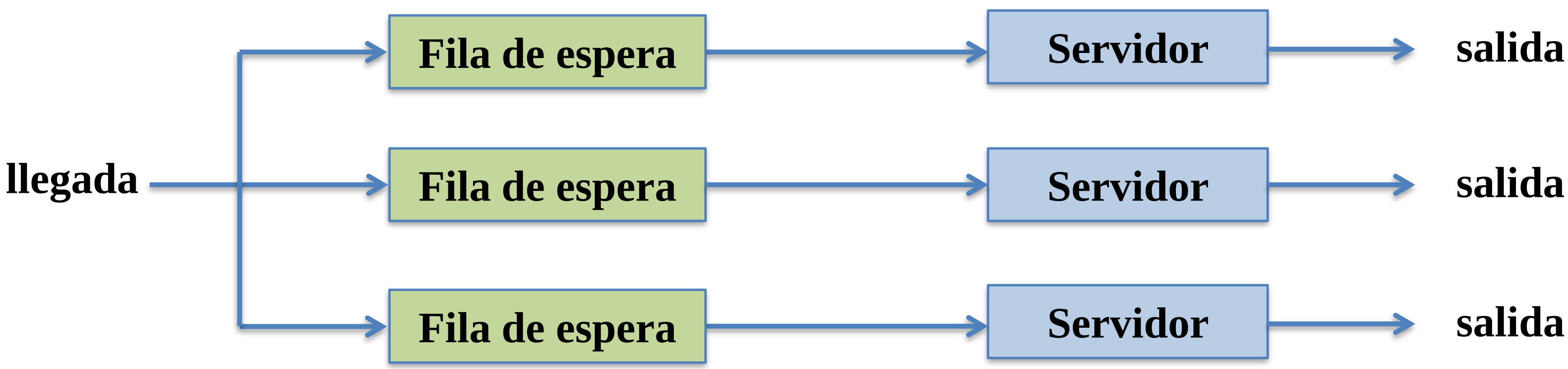
- **Una fila, un servidor**

- **Una fila, múltiples servidores**

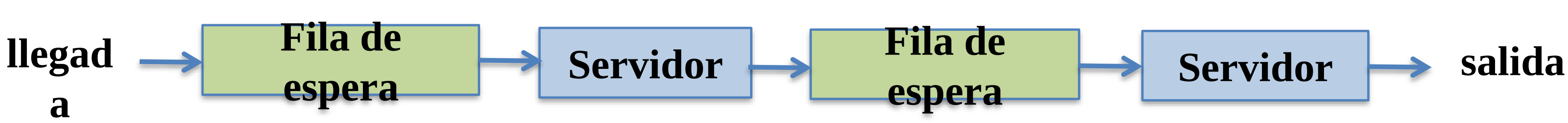


Tipos de Sistemas

- **Varias filas, múltiples servidores**



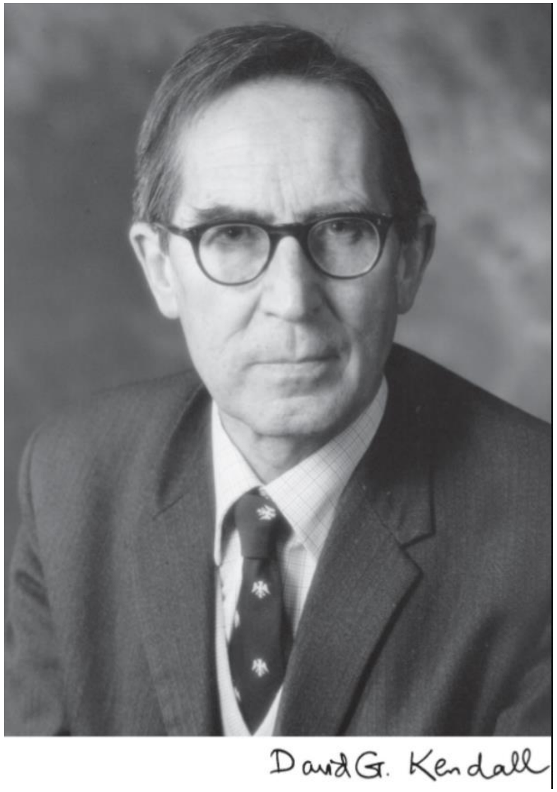
- **Una fila, servidores secuenciales**



Las características básicas que se deben utilizar para describir un sistema de filas

a) **Patrón de llegada de los clientes**
b) **Patrón de servicio de los servidores**
c) **Número de canales de servicio**
d) **Capacidad del sistema**
e) **Población de clientes posibles**
f) **Disciplina de fila**
g) **Número de etapas de servicio**

David G. Kendall introdujo la notación de filas de espera en 1953



A/S/m/B/K/SD

Notación de Colas (Kendall: A/S/m/B/K/SD)

A: Proceso de llegadas: La distribución de probabilidad de los periodos entre llegadas al centro de servicio.

S: Tiempo de servicio: La distribución de probabilidad de los periodos de servicio para cada petición en el centro de servicio

m: Número de servidores

B: Capacidad del sistema: El número máximo de clientes que pueden estar en el centro de servicio. Capacidad de almacenamiento.

K: Tamaño de la población: El número total de clientes potenciales que pueden llegar al sistema.

SD: Disciplina de servicio: La política de servicio (FIFS, LIFS, ...)

A/S/m/B/K/SD

Componente	Símbolo	Característica
Distribución de tiempos de llegada o Distribución de tiempos de servicio	M	Distribución Exponencial
	D	Distribución Determinística
	E_k	Distribución *Erlang* tipo k (0,1,2,..)
	G	Distribución General
Número de Servidores Capacidad del Sistema Tamaño de la Población	1,2,…,∞	
Disciplina de la cola	FIFS	Servir al primero que llega
	LIFS	El último que llega se sirve primero
	RSS	Selección aleatoria de servicio
	PR	Prioridad

Las variables aleatorias de las distribuciones de probabilidad deben ser *independientes e idénticamente distribuidas* (IID)

Modelos

M/M/1	Llegadas Distribución Exponencial Servicio Distribución Exponencial 1 Servidor
M/M/S	Llegadas Distribución Exponenciales Servicio Distribución Exponencial Múltiples Servidores (S)
M/M/1/k	Llegadas Distribución Exponenciales Servicio Distribución Exponencial 1 Servidor Capacidad limitada (k)
M/G/1	Llegadas Distribución Exponenciales Servicio con Distribución General 1 Servidor

Nomenclatura

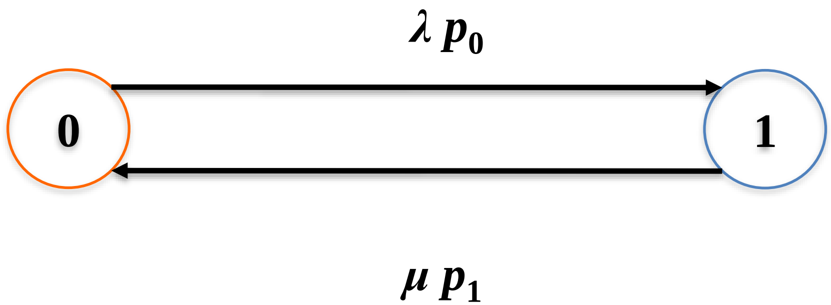
$\boldsymbol{\lambda}$	Número de llegadas por unidad de tiempo
$\boldsymbol{\mu}$	Número de servicios por unidad de tiempo si el servidor está ocupado
C, S, k	Número de servidores en paralelo
$\boldsymbol{\rho = \lambda/\mu}$	Congestión de un sistema con parámetros: *(λ, μ,)*
N(t)	Número de clientes en el sistema en el instante t
$\mathbf{N_q(t)}$	Número de clientes en la cola en el instante t
$\mathbf{N_s(t)}$	Número de clientes en servicio en el instante t
$\mathbf{P_n(t)}$	Probabilidad que haya *n* clientes en el sistema en el instante $P_r\{N(t)=n\}$
L	Número de clientes en el sistema en el estado estable
$\mathbf{P_n}$	Probabilidad de que haya *n* clientes en estado estable $P_n=P_r\{N=n\}$
$\mathbf{L_q}$	Número medio de clientes en la cola en el estado estable
$\mathbf{W_q = E[Tq]}$	Tiempo medio de espera de los clientes en la cola
$\mathbf{W_s = E[Ts]}$	Tiempo medio de servicio
$\mathbf{W = W_q + W_s}$	Tiempo total que un cliente invierte en el sistema
R	Número medio de clientes que se atienden por término medio
$\mathbf{P_b}$	Probabilidad de que cualquier servidor esté ocupado

M/M/1

1. Los clientes son atendidos con una política *FIFS* y cada llegada de un cliente es atendida sin importar la longitud de la fila de espera
2. Las llegadas son independientes
3. Las llegadas son descritas mediante la distribución de probabilidad de Poisson y con una tasa media de llegadas de λ.
4. La población potencial a atender es infinita o muy grande
5. Los tiempos de servicio al cliente varían y son independientes entre si
6. Los tiempos de servicio se representan mediante una distribución de probabilidad exponencial negativa , con una tasa media de servicio de μ.
7. La razón de servicio es mayor o igual que la razón de llegadas (estado estable.
8. La longitud de la fila de espera es ilimitada

M/M/1

Si tengo 2 posibles estados:



Es estado estable se define cuando el número de clientes atendidos es mayor o igual que el número de clientes que llegan (ley de Little).

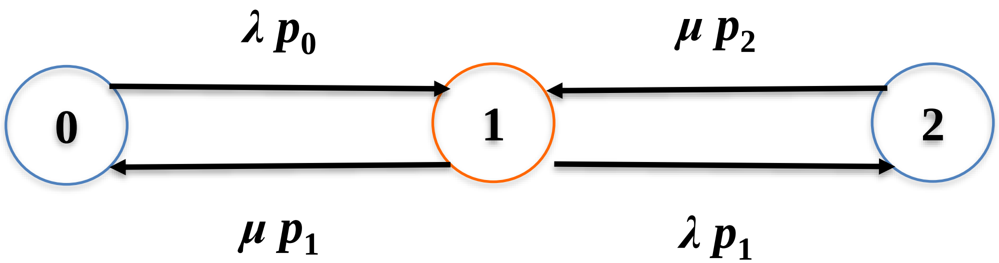
Si tenemos un estado estable, ¿cuál es el valor de p_1?

$$\mu p_1 = \lambda p_0$$

$$p_1 = (\lambda / \mu)\, p_0$$

M/M/1

$p_1 = (\lambda / \mu)\, p_0$

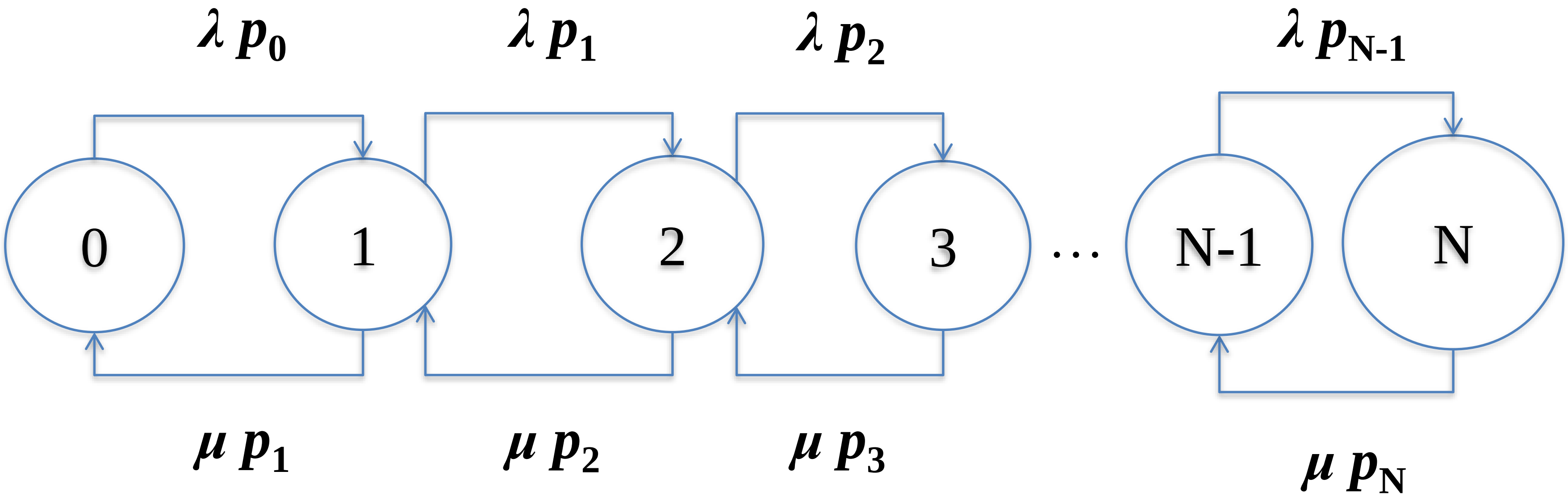


¿cuál es el valor de p_2?

$$\lambda p_0 + \mu p_2 = \mu p_1 + \lambda p_1$$

$$p_2 = \left(\frac{l}{m}\right)^2 p_0$$

M/M/1



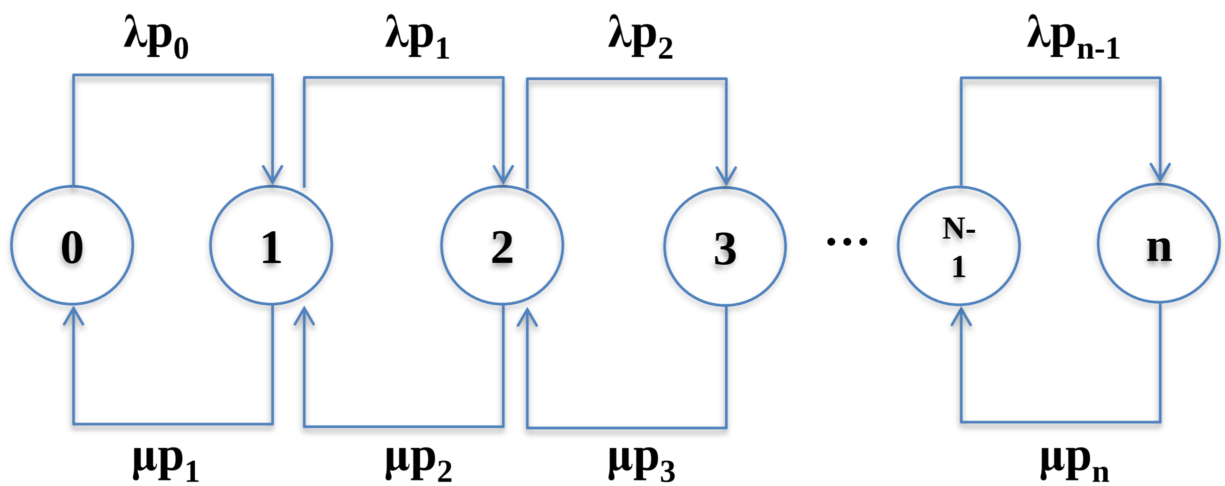
$$p_1 = (\lambda / \mu)\, p_0$$

$$p_2 = \left(\frac{l}{m}\right)^2 p_0$$

M/M/1

$$p_1 = (\lambda / \mu)\, p_0$$

$$p_2 = \left(\frac{l}{m}\right)^2 p_0$$



¿Cuál es el valor de p_n?

$$p_n = \left(\frac{l}{m}\right)^n p_0 \Rightarrow \boxed{p_n = r^n p_0}$$

¿Cuál es el valor de p_0?

M/M/1

Para calcular el valor de p_0

$$\sum_{n=0}^{\infty} p_n = 1 \qquad \text{si} \qquad p_n = \rho^n p_0$$

$$\sum_{n=0}^{\infty} \rho^n p_0 = 1$$

$$p_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n}$$

$$\sum_{n=0}^{\infty} \rho^n = \frac{1}{1-\rho}$$

ρ^n Corresponde a una serie geométrica

$$p_0 = \frac{1}{\frac{1}{1-\rho}}$$

$$\boxed{p_0 = 1-\rho}$$

M/M/1 — **Número de clientes esperando en el sistema**

$$L_s = \sum_{n=0}^{\infty} n p_n \qquad\qquad p_n = p_0 \rho^n$$

$$L_s = \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n p_0 \rho^n$$

$$L_s = \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n p_0 \rho^n = \sum_{n=0}^{\infty} n p_0 \rho^{n-1} \rho \;\; = p_0 \rho \sum_{n=0}^{\infty} n \rho^{n-1}$$

$$\sum_{n=0}^{\infty} n \rho^{n-1} = \frac{1}{(1-\rho)^2}$$

Obtener el valor de L_s en términos de λ y μ

$$L_s = p_0 \rho \frac{1}{(1-\rho)^2}$$

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

$$\boxed{L_s = \frac{\rho}{1-\rho}} \qquad \boxed{L_s = \frac{\lambda}{\mu-\lambda}}$$

M/M/1

Con el valor de L_s calculado $$L_s = \frac{\lambda}{\mu - \lambda}$$

Ley de Little L= λ W

Aplicando la Ley de Little, c*alcular el valor de W_s en términos de λ y μ*

$$W_s = \frac{1}{\mu - \lambda}$$

Si L_q está dado por la expresión indicada, calcular el valor de L_q en terminos de λ y μ

$L_q = \rho\, L_s$

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Si W_q está dado por la expresión indicada, calcular el valor de W_q en términos de λ y μ

$W_q = \rho\, W_s$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

$W_q = L_q/\lambda$

M/M/1 **Síntesis de medidas de desempeño**

$$p_n = \rho^n p_0 \qquad p_0 = 1 - \rho \qquad \rho = \frac{\lambda}{\mu}$$

$$L_s = \frac{\rho}{1-\rho} \qquad L_s = \frac{\lambda}{\mu - \lambda}$$

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

$$W_s = \frac{1}{\mu - \lambda}$$

$$W_q = \rho W_s \qquad Wq = \frac{L_q}{\lambda} \qquad W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Modelo M/M/1

Lavado de autos

Existe sólo un lugar para lavar autos, los autos llegan con una distribución de Poisson con un razón de 4 por hora. Si está ocupado los autos que llegan esperan, esto es no hay límite para el tamaño de la fila de espera . El número de servicios promedio es de 6 autos por hora.

El tiempo para lavar un auto sigue una distribución exponencial con 10 minutos en promedio.

λ = 4 autos /hr **μ = 6 autos / hr** **$\rho = \lambda/\mu = 0.6666 < 1$**

El sistema puede funcionar en condiciones de estado estable

a) $Lq= \lambda^2 / (\mu (\mu-\lambda)) = 1.33$ autos

$Ls= \lambda / (\mu-\lambda) = 2$ autos

b) $Ws= 1 / (\mu-\lambda) = 0.5$ hr

c) $Wq= \rho / (\mu-\lambda) = 0.333$ hr

d) $p_0= 1 - \rho = 0.333$

Modelo M/M/1

A un restaurante de comida rápida llegan en promedio 100 clientes por hora. Se tiene capacidad para atender en promedio a 150 clientes por hora, calcule las siguientes medidas de desempeño del sistema:

a) ¿Cuál es la probabilidad que el sistema este ocioso?
b) ¿Cuál es la probabilidad que un cliente llegue y tenga que esperar, porque el sistema está ocupado?
c) ¿Cuál es el número promedio de clientes en la fila de espera?
d) ¿Cuál es la probabilidad que hayan 10 clientes en la fila de espera?

Modelo M/M/1

¿Cuál es la probabilidad que el sistema este ocioso?

$P_0 = 1 - \rho = 0.333 \approx 33\%$

¿Cuál es la probabilidad que un cliente llegue y tenga que esperar, porque el sistema está ocupado?

$P(x \geq 1) = 1 - P_0 = 0.6666 \approx 67\%$

¿Cuál es el número promedio de clientes en la fila de espera?

$L_q = 1.33$

¿Cuál es la probabilidad que hayan 10 clientes en la fila de espera?

a) $P_{10} = 0.0057 \approx 0.57\%$

M/M/S

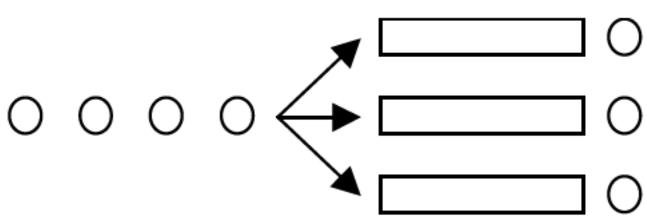
Modelos de colas multicanal

1. Dos o más servidores
2. Los clientes forman una sólo fila y son atendidos por el servidor desocupado
3. Las llegadas siguen una distribución de probabilidad de *Poisson*

4. Los tiempos de servicio siguen una distribución exponencial

5. Los servicios se atienden de acuerdo a la política de primeros en llegar primeros en atenderse (FIFS)
6. La capacidad de la fila de espera es infinita

$$r = \frac{l}{sm}$$

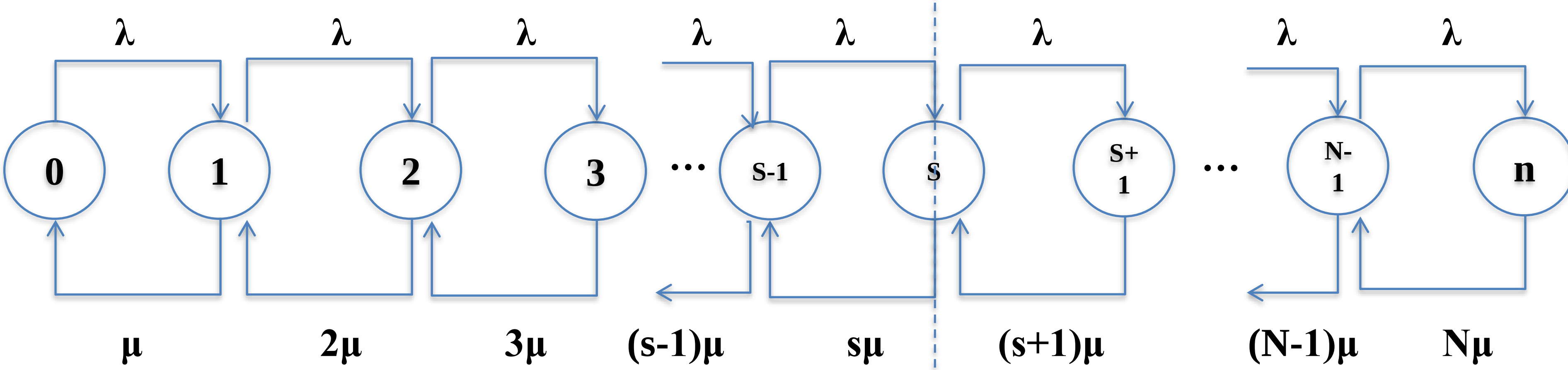


$$p_n = \rho^n p_0$$

M/M/S

X_t = número de clientes que hay en el sistema en el tiempo *t*

S= {*0,1,2,3,…, n*} conjuntos de estados



$$p_1 = \frac{\lambda}{\mu} p_0$$

$$p_2 = \frac{\lambda^2}{2\mu^2} p_0$$

$$p_3 = \frac{\lambda^3}{6\mu^3} p_0$$

¿Cuál es el valor de p_n, para los estados *s* al estado *s-1*?

$$p_s = \frac{\lambda^s}{s!\,\mu^s} p_0$$

$$p_n = \frac{\lambda^n}{n!\,\mu^n} p_0 \qquad \textbf{Si } n < s$$

¿Cuál es el valor de p_s, para los estados de *s* a *n*?

$$p_n = \frac{\lambda^n}{s!\,s^{n-s}\mu^n} p_0 \qquad \textbf{Si } n \geq s$$

M/M/S

$$
p_n = \begin{cases} \dfrac{\lambda^n}{n!\,\mu^n}\, p_0 & \text{Si } n < s \\[2ex] \dfrac{\lambda^n}{s!\, s^{n-s}\, \mu^n}\, p_0 & \text{Si } n \geq s \end{cases}
$$

¿cuál es el valor de p_0?

$$
\sum_{n=0}^{s-1} \frac{\lambda^n}{n!\,\mu^n} p_0 + \sum_{n=s}^{\infty} \frac{\lambda^n}{s!\, s^{n-s}\, \mu^n} p_0 = 1
$$

$$
\sum_{n=0}^{\infty} p_n = 1
$$

$$
p_0 = \left(\sum_{n=0}^{s-1} \frac{\lambda^n}{n!\,\mu^n} + \sum_{n=s}^{\infty} \frac{\lambda^n}{s!\, s^{n-s}\, \mu^n} \right)^{-1}
$$

M/M/S

$$p_n = \begin{cases} \dfrac{\lambda^n}{n!\,\mu^n} p_0 & \text{Si } n < s \\[2ex] \dfrac{\lambda^n}{s!\,s^{n-s}\mu^n} p_0 & \text{Si } n \geq s \end{cases}$$

¿cuál es el valor de p_0?

$$\sum_{n=0}^{s-1} \frac{\lambda^n}{n!\,\mu^n} p_0 + \sum_{n=s}^{\infty} \frac{\lambda^n}{s!\,s^{n-s}\mu^n} p_0 = 1 \qquad \sum_{n=0}^{\infty} p_n = 1$$

$$p_0 = \left(\sum_{n=0}^{s-1} \frac{\lambda^n}{n!\,\mu^n} + \sum_{n=s}^{\infty} \frac{\lambda^n}{s!\,s^{n-s}\mu^n} \right)^{-1}$$

$$\boxed{p_0 = \left(\sum_{n=0}^{s-1} \frac{\lambda^n}{n!\,\mu^n} + \frac{\lambda^s}{s!\,\mu^s\left(1 - \frac{\lambda}{s\mu}\right)} \right)^{-1}}$$

M/M/S

$$p_n = \begin{cases} \dfrac{\lambda^n}{n!\,\mu^n}\, p_0 & \textbf{Si n < s} \\[2ex] \dfrac{\lambda^n}{s!\, s^{n-s} \mu^n}\, p_0 & \textbf{Si n} \geq \textbf{s} \end{cases}$$

$$p_0 = \left(\sum_{n=0}^{s-1} \left(\frac{\lambda^n}{n!\,\mu^n}\right) + \frac{\lambda^s}{s!\,\mu^s \left(1 - \dfrac{\lambda}{s\mu}\right)} \right)^{-1}$$

$$L_q = \frac{\lambda^s}{s!\,\mu^s} \frac{\lambda}{s\mu} \frac{1}{\left(1 - \dfrac{\lambda}{s\mu}\right)^2} p_0$$

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

$$W_q = \frac{L_q}{\lambda} \qquad W_q = W - \frac{1}{\mu} \qquad W = \frac{L}{\lambda}$$

M/M/S

Los clientes llegan a un lavado de autos con una distribución de exponencial a una razón de 10 autos por hora y el tiempo de servicio es de 8 autos por hora. Si Se cuenta con 4 estaciones para el lavado de autos, calcular.

a) Número de autos en el sistema
b) Número de autos en la fila de espera
c) Tiempo invertido en el sistema
d) Tiempo invertido en la fila de espera

M/M/S

Un banco tiene 4 cajeros para atender a los clientes, llegan en promedio 80 personas sin cuenta en el banco por hora y 20 clientes con cuenta por hora. Los clientes al llegar hacen dos filas de espera al cajero desocupado: una para los que tienen cuenta y otra para los que no la tienen. Los clientes con cuenta son atendidos en cuanto un cajero se desocupa (preferencia).

El tiempo promedio que le toma al cajero atender un cliente es de 1.2 minutos.

Considera que los tiempos de llegada y tiempos de servicio siguen una distribución exponencial, calcule:

a) El número esperado de clientes en el banco
b) El tiempo esperado que cada tipo de cliente invierte en el banco
c) El tiempo esperado que cada tipo de cliente invierte en la fila de espera
d) El número esperado de clientes en cada fila de espera

Descargar y correr el ejemplo de Desmo-J

University of Hamburg
Department of Computer Science

DESMO-J Java

A Framework for Discrete-Event Modelling and Simulation