Teoría de colas

Modelado de sistemas de colas

Modelado de Sistemas

Sistemas Discretos → Estáticos
→ Dinámicos

Sistemas Continuos

Ejemplos de utilización

- Una línea de cajas de un supermercado.
- Un aeropuerto.
- Una planta fabril.
- Pacientes que llegan a la guardia médica de un hospital para ser atendidos.

Sistemas de colas

Se deben identificar:

- **Objetos**
- **Eventos**
- **Colas**

Tipos de Objetos

- **Cliente**
- **Servidor**

Ambos pueden ser temporarios o permanentes.

Tipos de Eventos

- **Llegada**
- **Fin de Atención**
- **Eventos temporizados:**
  - Fin de la simulación
  - Interrupción de la atención
  - Interrupción de las llegadas
  - Tareas programadas. (Mantenimiento, Descanso)

Características de las Colas

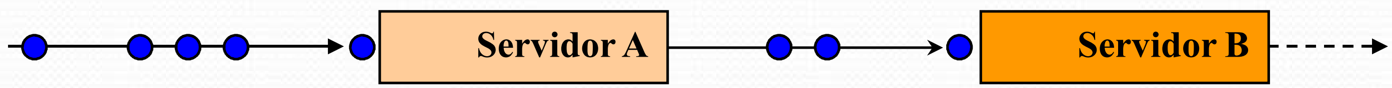
- ¿Tienen longitud máxima?
- FIFO/LIFO
- Impaciencia
- Prioridades

Es necesario establecer:

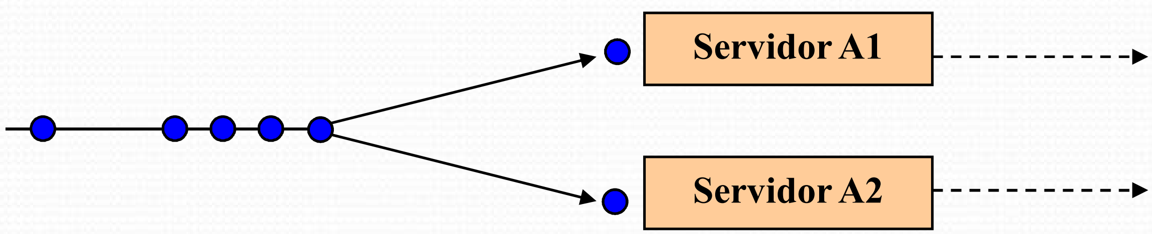
- Distribuciones estadísticas asociadas a los **tiempos entre llegadas**, y a los **tiempos de atención**
- Tipos de clientes, estados, atributos.
- Disciplina del servicio (interrupciones, prioridades).
- Distribución y cantidad de colas.
- Disposición de los servidores, estados, atributos.

Disposición de los Servidores

Serie

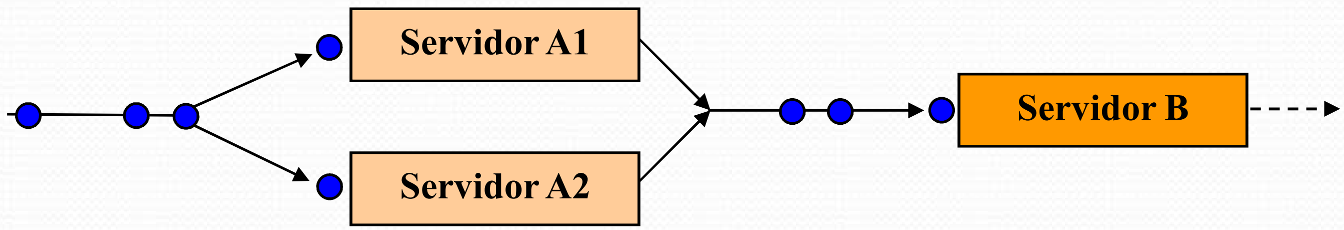


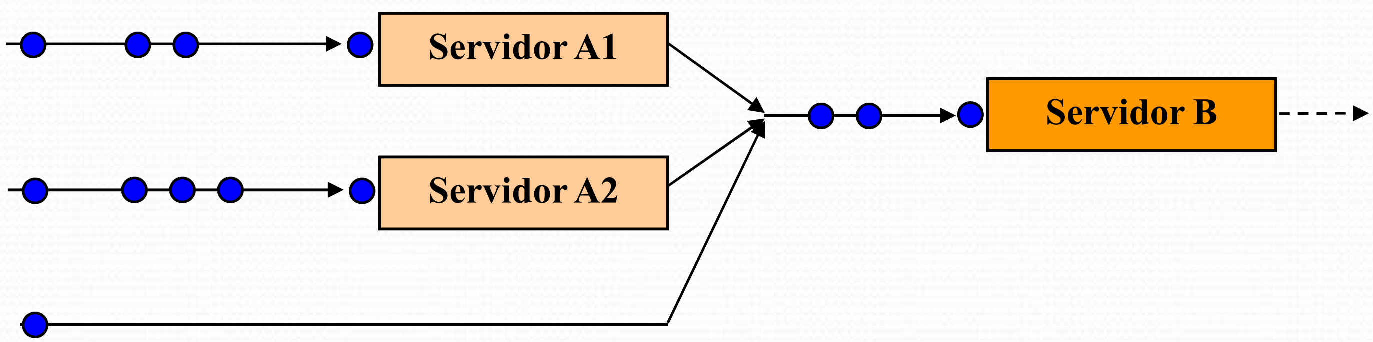
Paralelo



Disposición de los Servidores

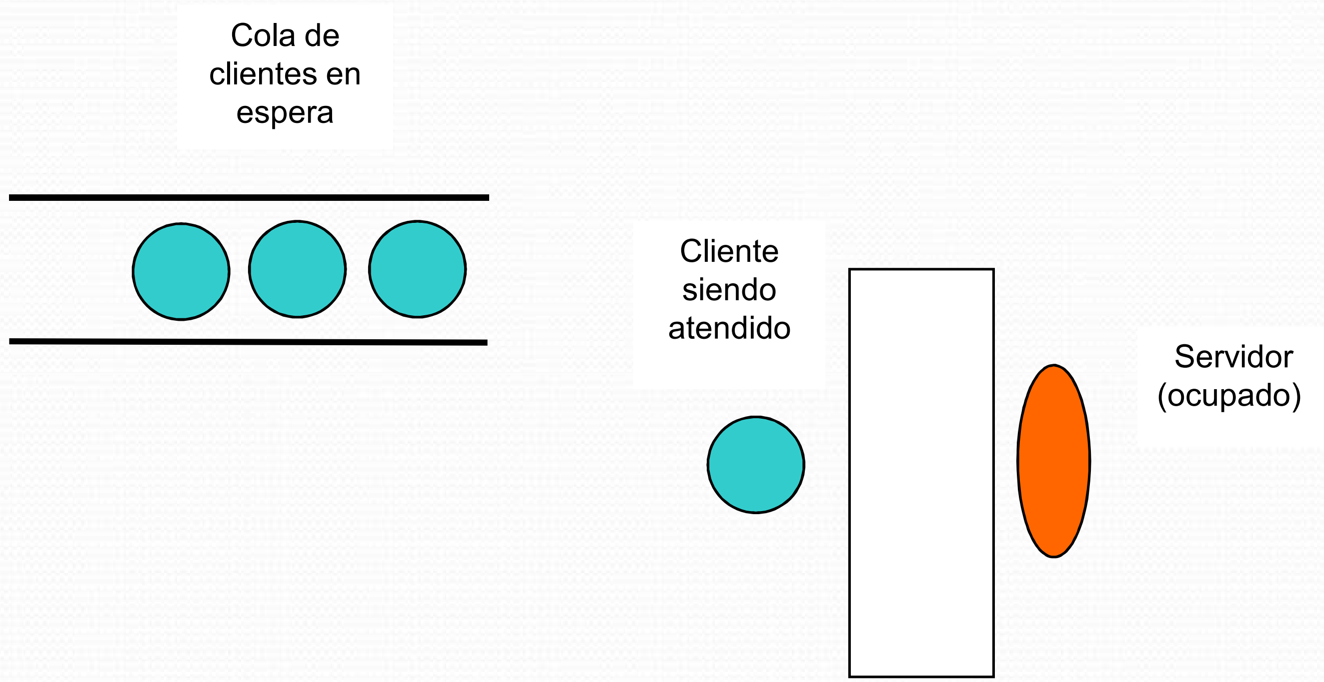
Combinados





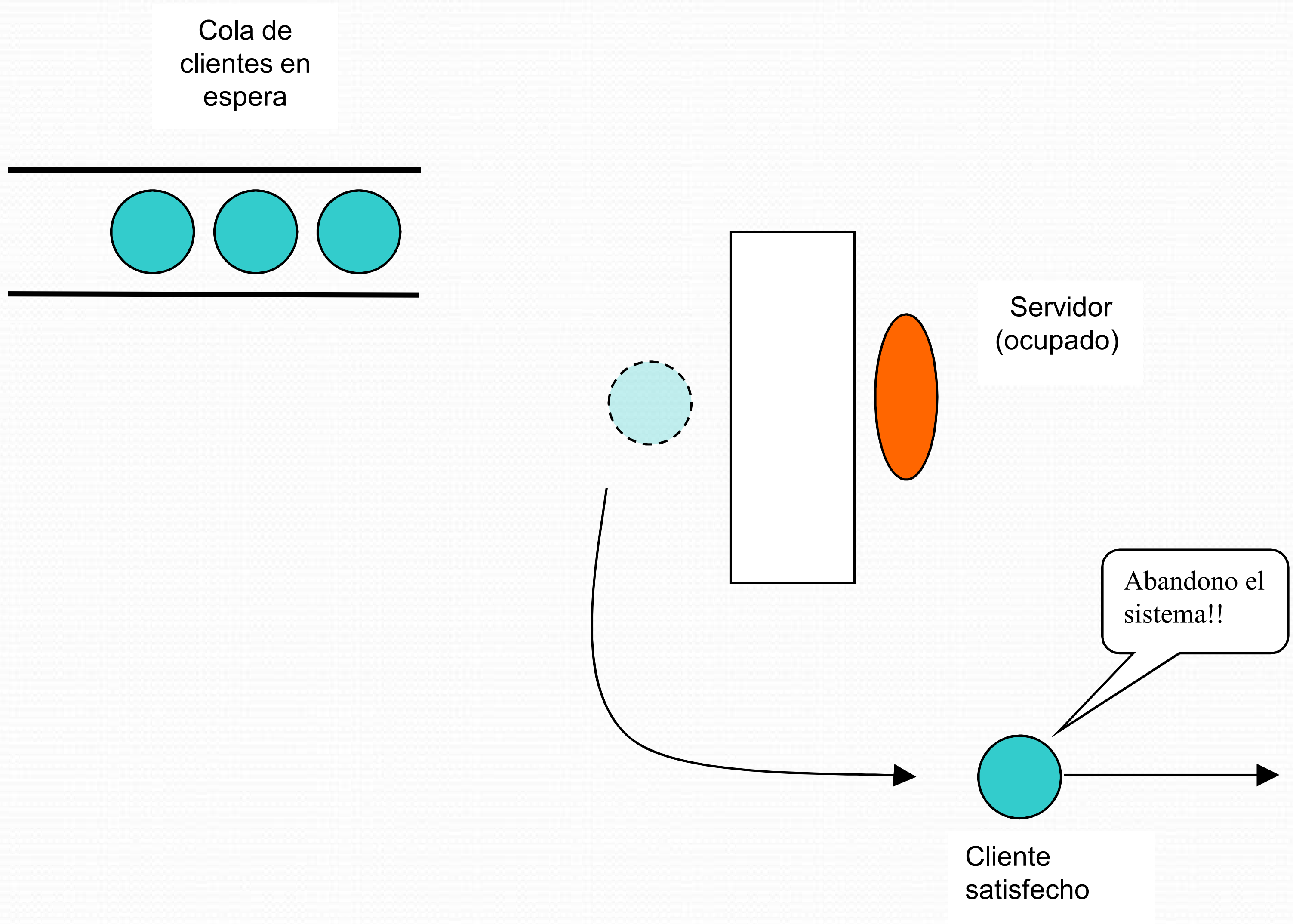
Sistema elemental:

un servidor , un solo tipo de clientes y una cola



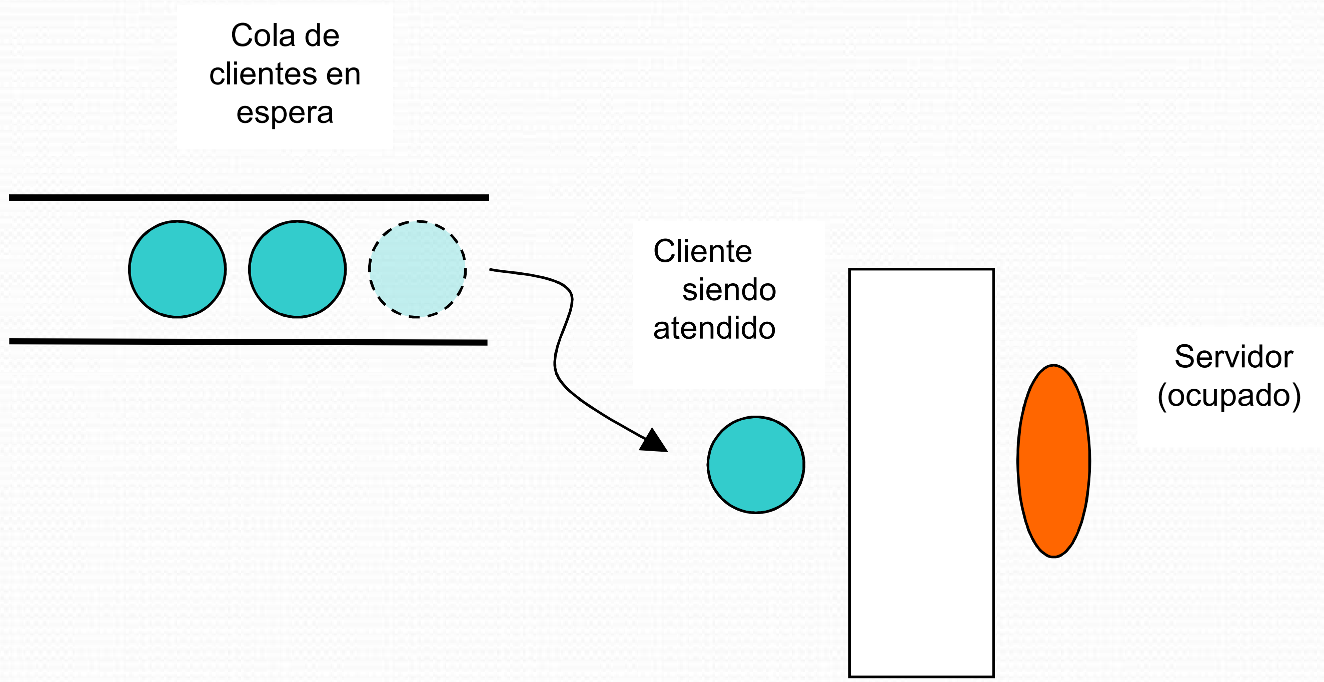
Sistema elemental:

un servidor , un solo tipo de clientes y una cola



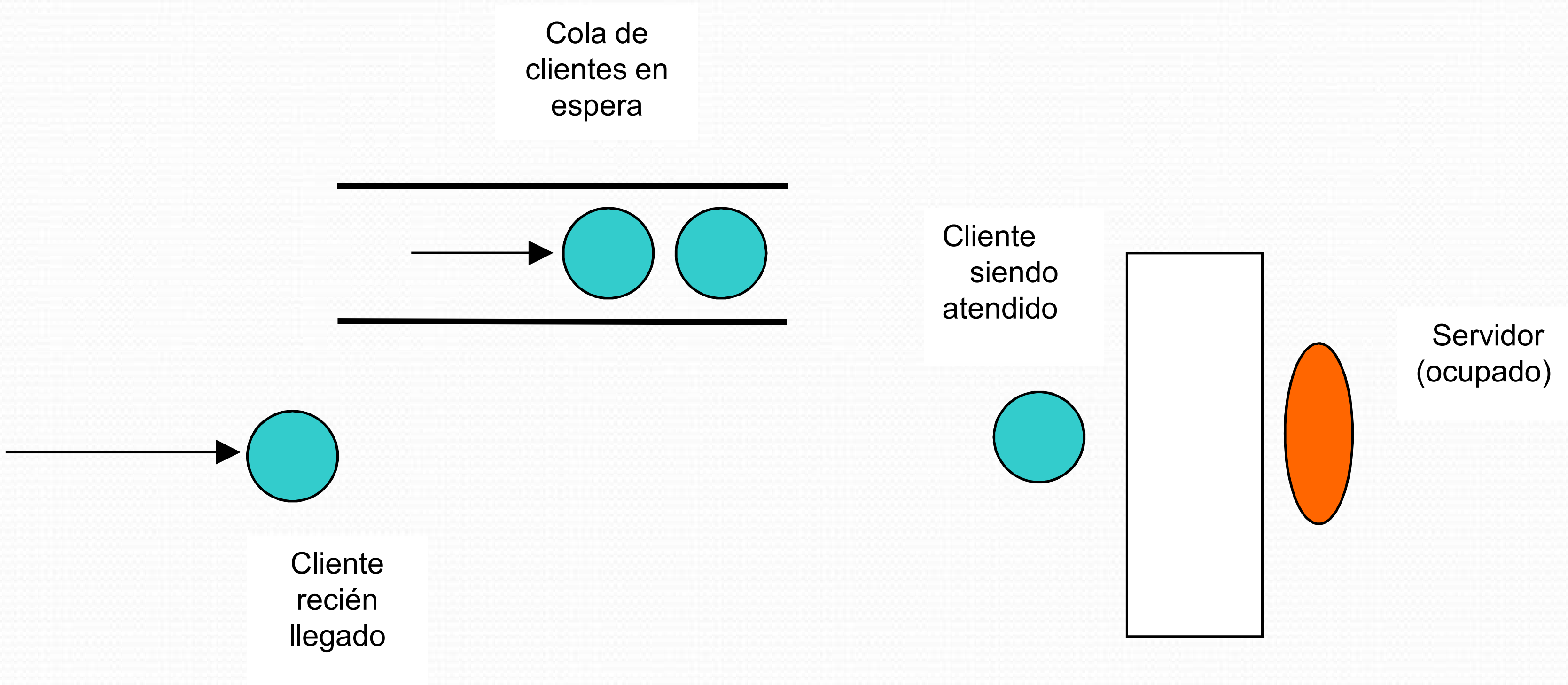
Sistema elemental:

un servidor , un solo tipo de clientes y una cola



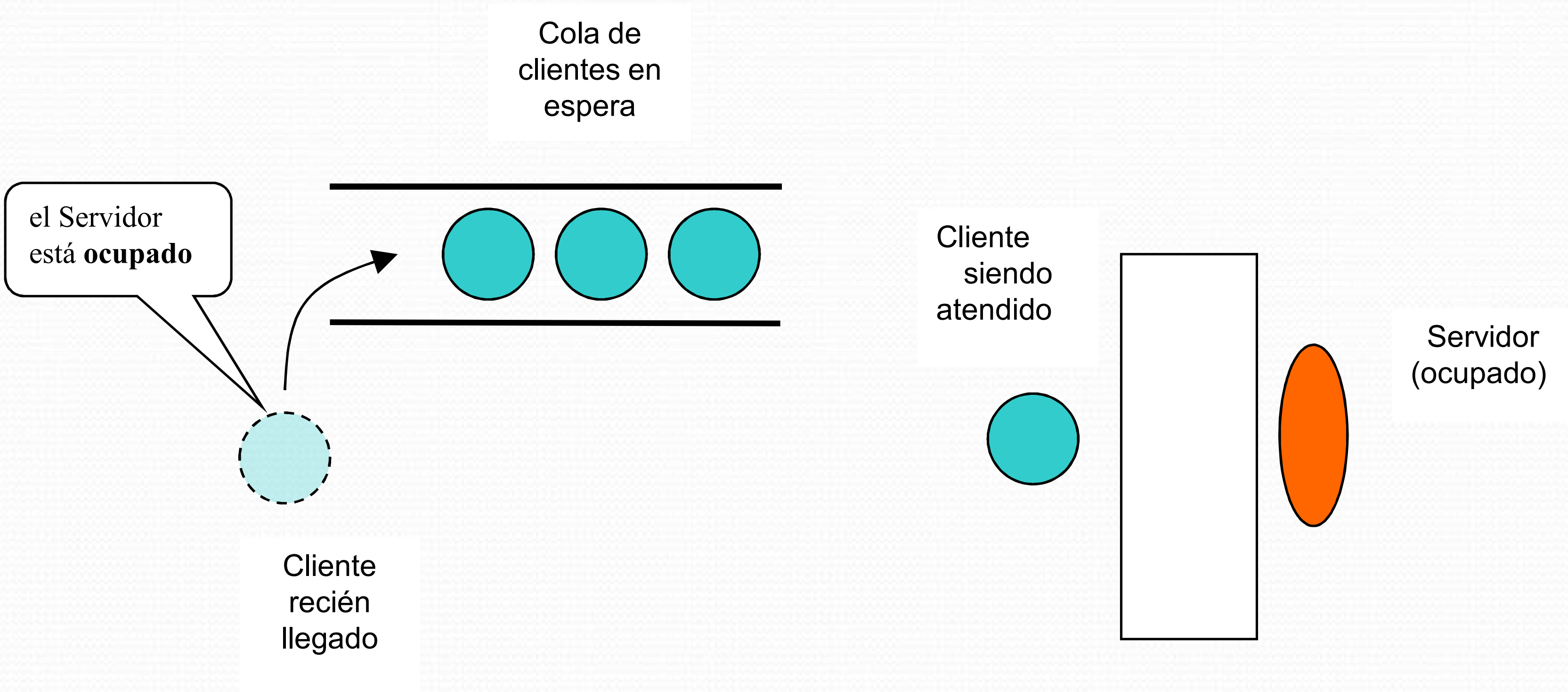
Sistema elemental:

un servidor , un solo tipo de clientes y una cola



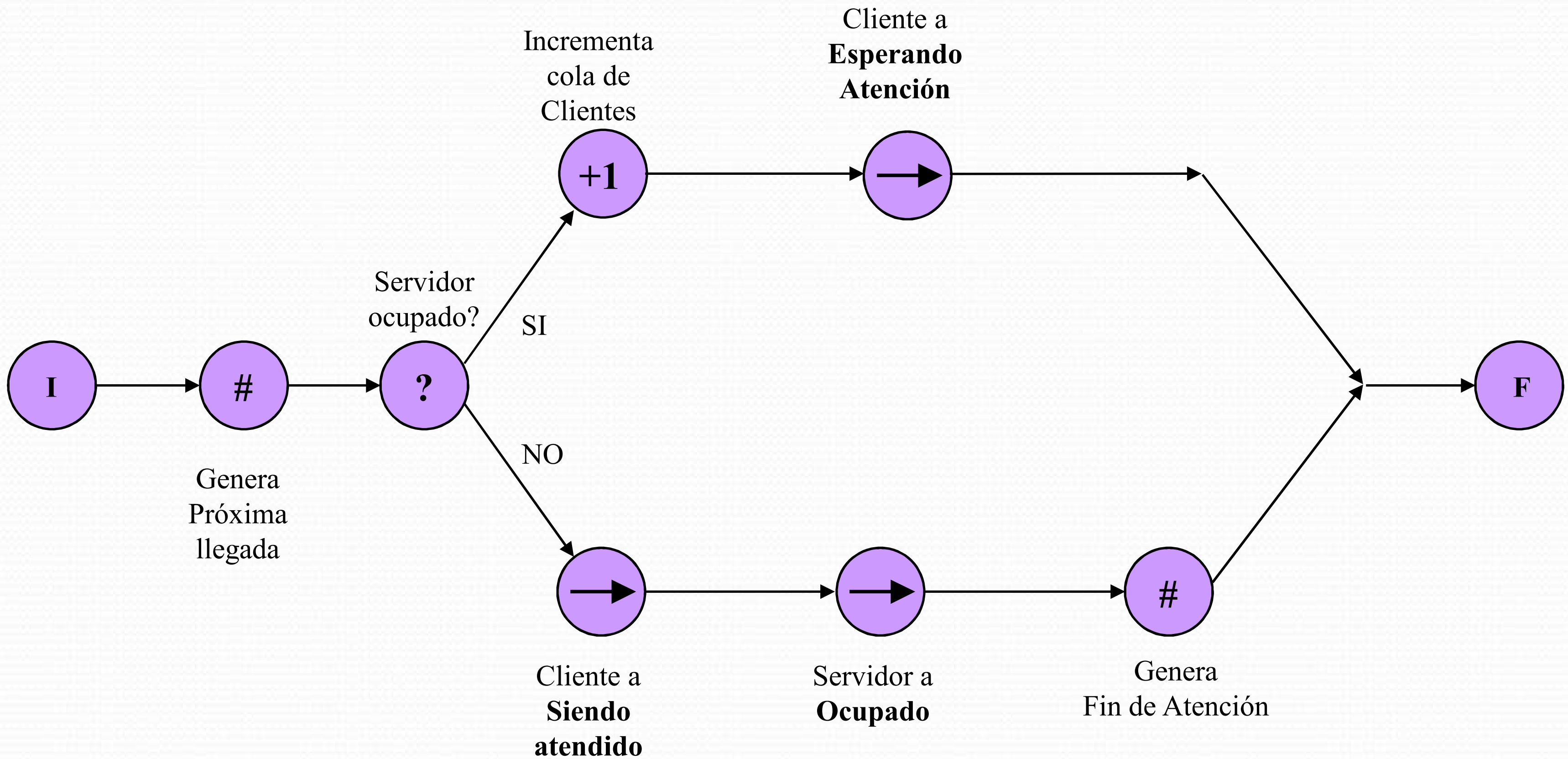
Sistema elemental:

un servidor , un solo tipo de clientes y una cola



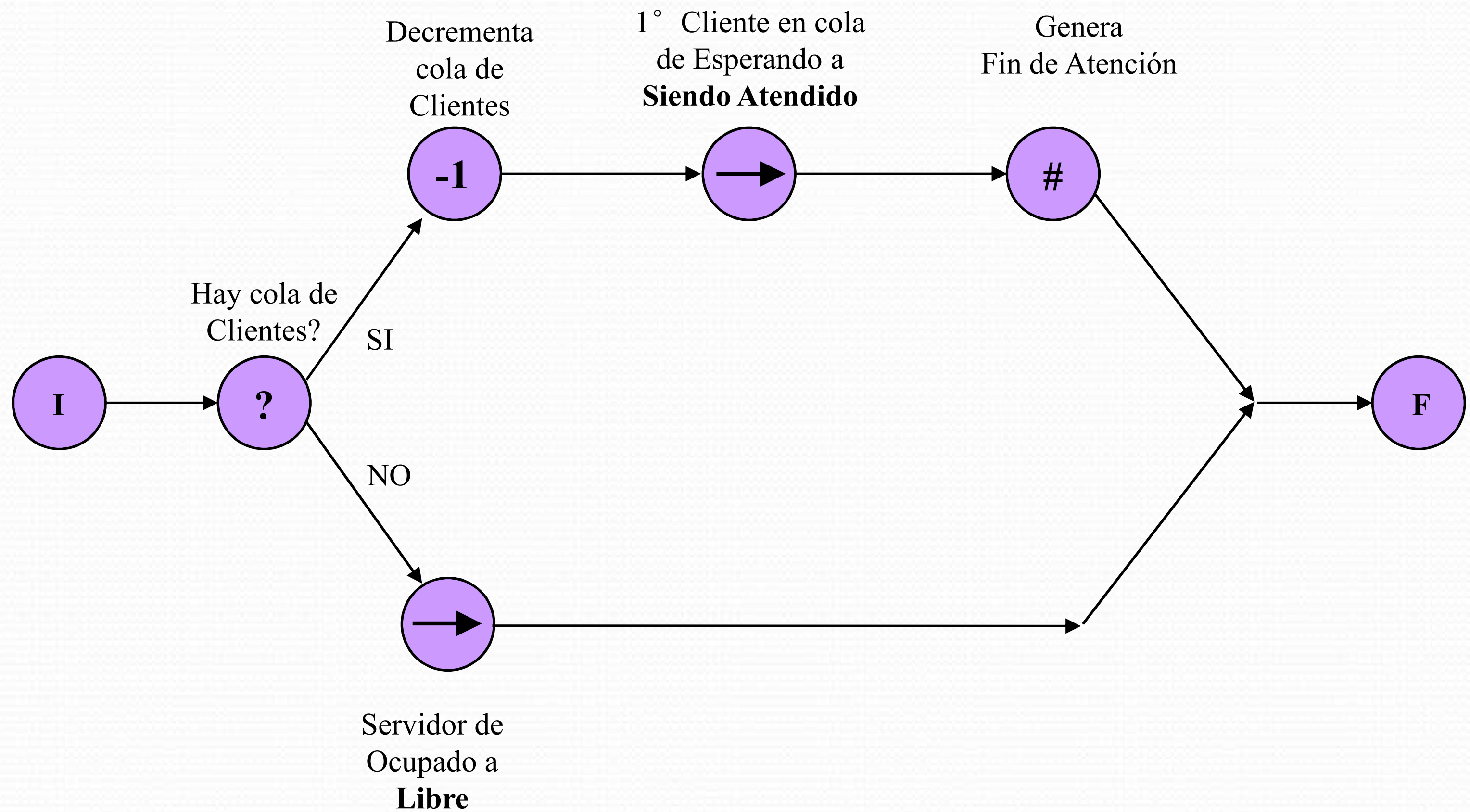
Llegada al sistema:

Cambios originados por el evento



Fin de servicio:

Cambios originados por el evento



Medidas de Desempeño

Tiempo promedio en cola (y en sistema) = $$\frac{\sum_{i=1}^{n} t_i}{n}$$

t_i: tiempo de permanencia del cliente i en cola

n: cantidad de clientes pasibles de entrar en cola

Medidas de Desempeño

Cantidad promedio de clientes en cola (y en sistema) $= \dfrac{\sum_{i=1}^{n} t_i}{t_{total}}$

t_i: tiempo de permanencia del cliente i en cola

t_{total}: tiempo total de la simulación

Medidas de Desempeño

Porcentaje de ocupación del Servidor = $\dfrac{\sum t_{ocup}}{t_{total}} \times 100$

t_{ocup}: intervalo de tempo en que el servidor está ocupado

t_{total}: tiempo total de la simulación

Otras Medidas de Desempeño

- **Porcentaje de clientes atendido**
- **Tiempo máximo de permanencia en cola**
- **Tiempo máximo de permanencia en el sistema**
- **Tiempo ocioso del servidor**
- **Probabilidad de que la espera sea mayor a…**
- **Probabilidad de que no sea atendido.**

Ejemplo de aplicación

Una librería con un empleado atiende la demanda de sus clientes.

Objetivos:

- Determinar el tiempo promedio de permanencia de clientes en cola.
- Conocer el porcentaje de ocupación del empleado.

Ejemplo de aplicación

<u>Objetos</u>

Cliente

- Temporario
- Estados:
  - Siendo atendido
  - Esperando en cola

Empleado(servidor)

- Permanente
- Estados:
  - Libre
  - Ocupado

Ejemplo: Una librería con un empleado atiende la demanda de sus clientes.

Eventos

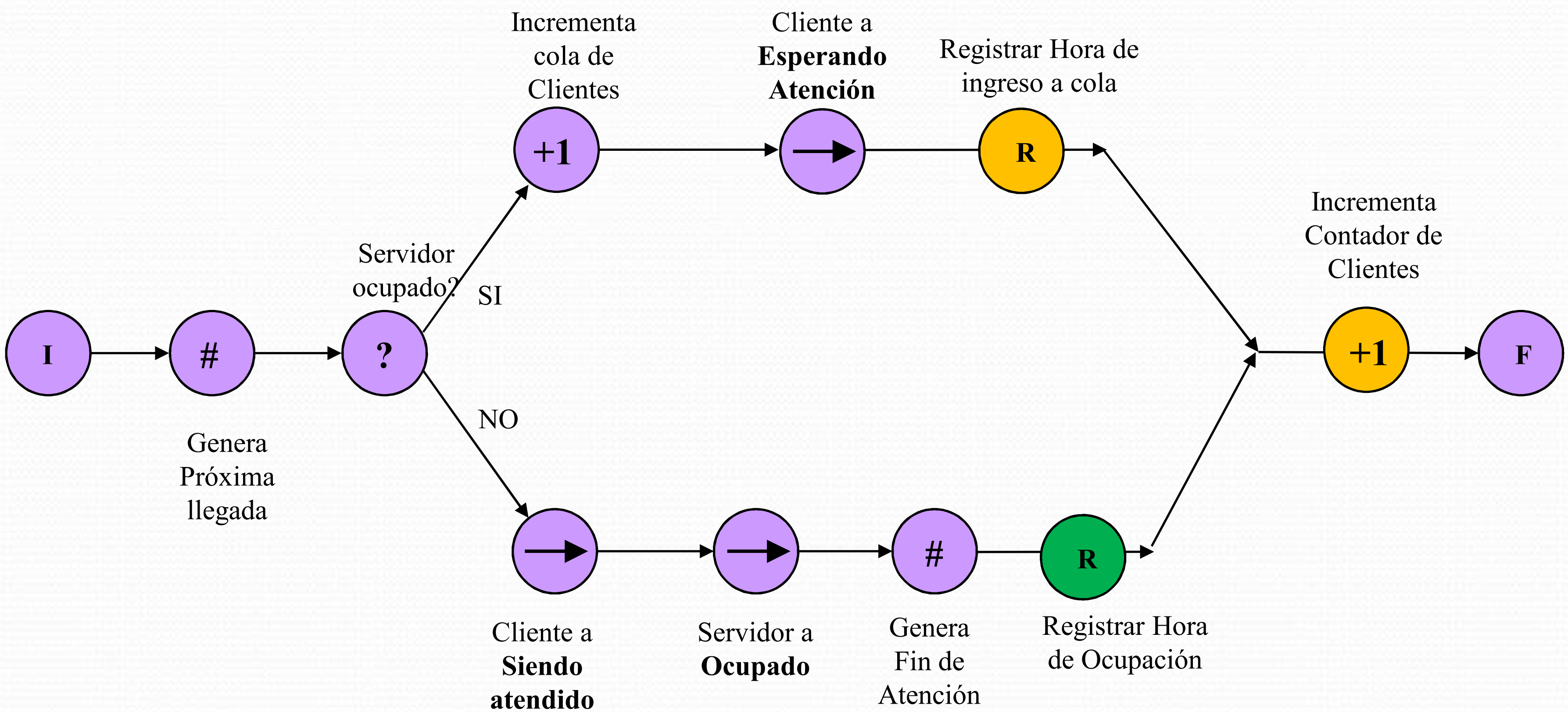
- Llegada de clientes al sistema
- Fin de atención al cliente
- Fin de la simulación (Temporizado)

Variables necesarias de acuerdo a los Objetivos:

- Determinar el tiempo promedio de permanencia de clientes en cola.
    - Acumulador de tiempos de permanencia en cola
    - Contador de clientes que llegan al sistema
- Conocer el porcentaje de ocupación del empleado.
    - Acumulador de tiempos de ocupación del empleado
    - Tiempo total de la simulación

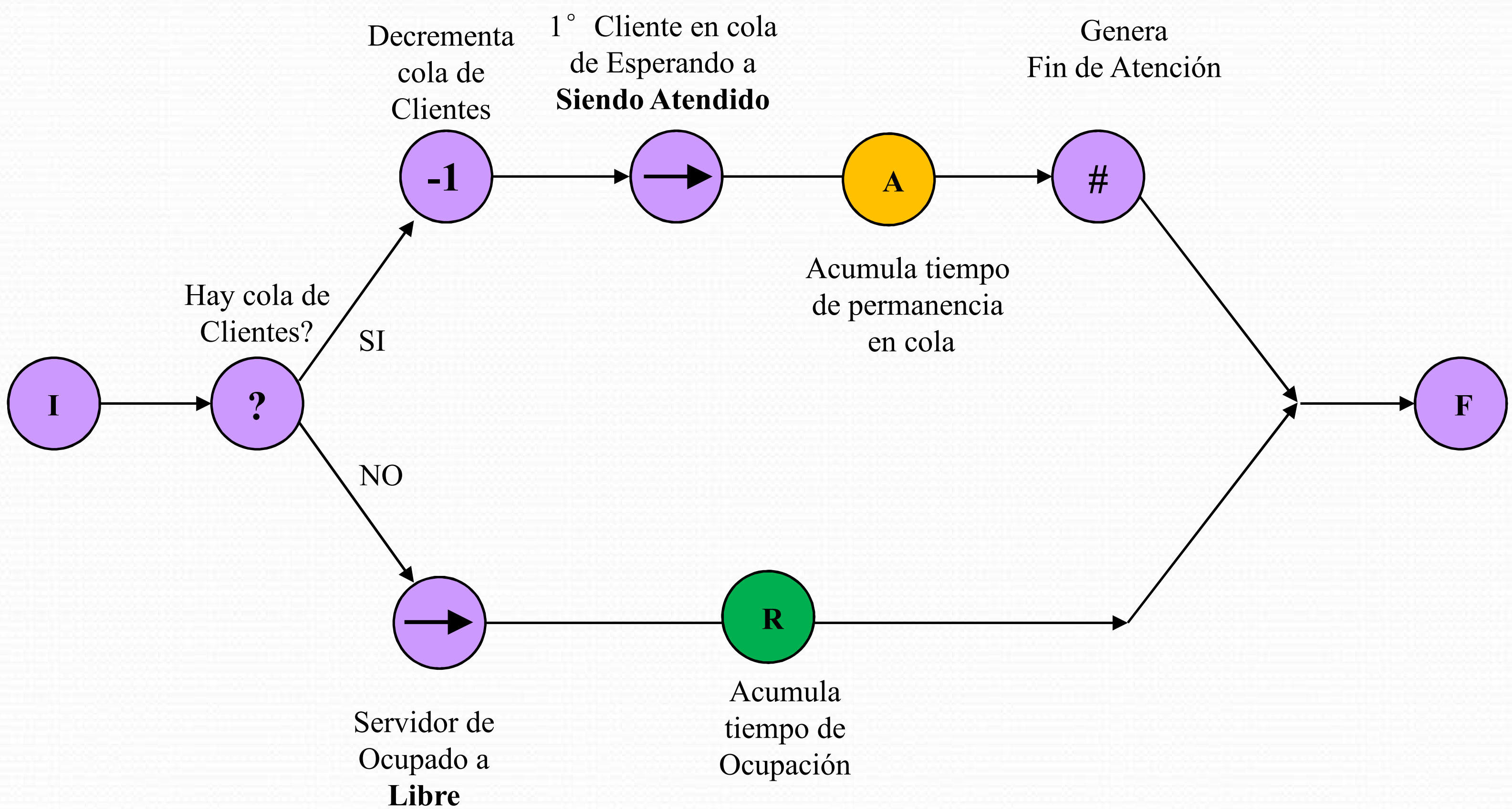
Llegada al sistema:

Cambios originados por el evento



Fin de servicio:

Cambios originados por el evento



Vector de estado

Vector estado: Una librería con un empleado atiende la demanda de sus clientes

		Llegada cliente					Fin de Atención					Acumuladores		Hora de llegada a la cola			
Evento	Reloj	Nº Random	Tiempo entre llegadas	**Próxima llegada**	Cola de clientes	Estado Servidor	Nº Random	Tiempo de Atención	**Fin de Atención**	Cont. Clientes	Inicio de ocup.	Tiempo de Ocup.	Tpo de Perman. en cola	Cte 2	Cte 3	Cte 5	Cte 6
Inicio	0			1,23	0	Libre			--	0							

Estadísticas del ejemplo

- Tiempo promedio de permanencia de clientes en cola:

  16,98 min / 6 clientes = **2,83 minutos**

- Tiempo de ocupación del empleado:

  (24,28 / 30) x 100 = **80,94 %**

Conclusiones

El tiempo medio de espera parece excesivo.

Sería conveniente agregar otro empleado que colabore con la venta, en momentos en que la cola tenga más de un cliente.

Resolución Analítica

Notación de Kendall:

A / B / s / k / t / d

A: distribución de tiempos entre llegadas
B: distribución de tiempos de servicio
s: número de servidores en paralelo (canales)
k: capacidad del sistema
t: tamaño de la fuente de entrada
d: disciplina de la cola

Modelos típicos

- M/M/1
- M/M/s
- M/M/1/k
- M/M/1/∞/∞/FCFS

Notación :

$$\lambda \equiv$$ Tasa de llegadas

$$\mu \equiv$$ Tasa de servicio

$$\rho = \frac{\lambda}{s\mu} \equiv$$ Factor de utilización del sistema

Si $\rho < 1$, entonces el sistema se estabiliza

Notación :

$L \equiv$ valor esperado del nro. de clientes en el sistema

$L_q \equiv$ valor esperado del nro. de clientes en cola

$W \equiv$ tiempo medio de respuesta (espera en sistema)

$W_q \equiv$ tiempo medio de espera en cola

$p_n \equiv$ Prob. de que “n” clientes estén en el sistema

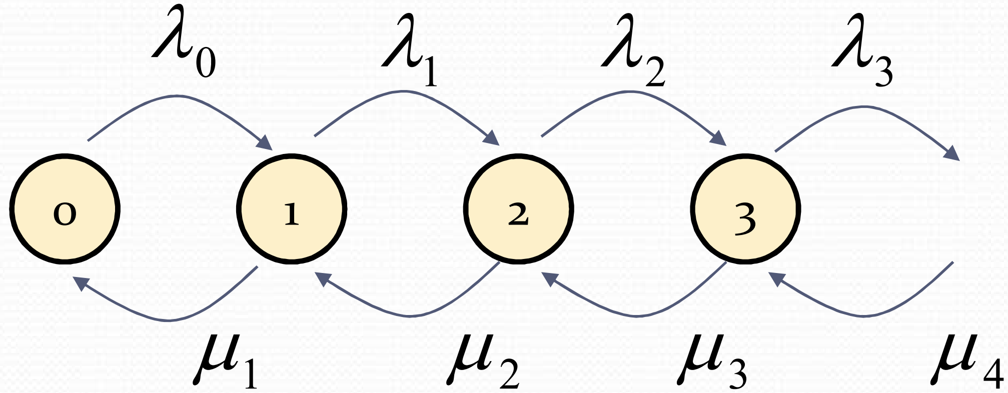
Relaciones básicas: Modelo general

Fórmula de Little: $L = \lambda W$ y $L_q = \lambda W_q$

además, $$W = W_q + \frac{1}{\mu}$$

de donde se deduce: $$L = L_q + \frac{\lambda}{\mu}$$

Estado “n”



$$p_n = ?$$

Ecuaciones de balance de flujo:

$$p_0\lambda_0 = p_1\mu_1$$

$$p_0\lambda_0 + p_2\mu_2 = p_1\lambda_1 + p_1\mu_1$$

$$p_1\lambda_1 + p_3\mu_3 = p_2\lambda_2 + p_2\mu_2$$

$$\cdots = \cdots$$

$$p_{n-1}\lambda_{n-1} + p_{n+1}\mu_{n+1} = p_n\lambda_n + p_n\mu_n$$

$$\cdots = \cdots$$

Si resolvemos la ecuaciones…

$$p_1 = \frac{\lambda_0}{\mu_1} p_0$$

$$p_2 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} p_0$$

$$p_3 = \frac{\lambda_2 \lambda_1 \lambda_0}{\mu_3 \mu_2 \mu_1} p_0$$

$$\cdots = \cdots$$

$$p_n = \frac{\lambda_{n-1} \cdots \lambda_1 \lambda_0}{\mu_n \cdots \mu_2 \mu_1} p_0$$

Para calcular p_0 , se utiliza:

$$p_0 + p_1 + p_2 + \cdots + p_n + \cdots = 1$$

Modelo M/M/1

En este caso, $\lambda_n = \lambda, \quad \mu_n = \mu, \quad \rho = \dfrac{\lambda}{\mu} < 1$

para todo n.

Entonces, $p_n = \rho^n p_0, \qquad p_0 = 1 - \rho,$

por lo que $p_n = \rho^n (1 - \rho)$

Por lo tanto, $L = \sum_{n=0}^{\infty} n p_n = \dfrac{\rho}{1 - \rho}$

y de la misma forma, $L_q = \sum_{n=1}^{\infty} (n-1) p_n = \dfrac{\rho^2}{1 - \rho}$

Modelo M/M/1

Por la fórmula de Little:

$$W = \frac{L}{\lambda} = \frac{1}{\mu\left(1-\rho\right)}$$

$$W_q = W - \frac{1}{\mu} = \frac{\rho}{\mu\left(1-\rho\right)}$$

La probabilidad de que haya k o más clientes en el sistema es:

$$P\left(N \geq k\right) = 1 - \sum_{n=0}^{k-1} p_k = 1 - \sum_{n=0}^{k-1} \rho^n \left(1-\rho\right) = 1 - \frac{\left(1-\rho\right)\left(1-\rho^k\right)}{\left(1-\rho\right)} = \rho^k$$

Por lo tanto,

$$P\left(N < k\right) = 1 - \rho^k$$

Modelo M/M/1

Resumiendo…

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

$$W = \frac{1}{\mu\left(1-\rho\right)}$$

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

$$W_q = \frac{\rho}{\mu\left(1-\rho\right)}$$

$$P\left(N < k\right) = 1-\rho^k$$

Modelo M/M/s (múltiples servidores)

La tasa de servicio depende del numero de servidores del sistema,
En este caso:

$$\rho = \frac{\lambda}{s\mu} < 1$$

y se puede probar que

$$p_0 = \frac{1}{\sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!(1-\rho)}}$$

y

$$p_n = \frac{(\lambda/\mu)^n p_0}{n!}, \quad si \quad 0 \leq n \leq s$$

$$p_n = \frac{(\lambda/\mu)^n p_0}{s!s^{n-s}}, \quad si \quad n > s$$

Modelo M/M/s (múltiples servidores)

Además:

$$L_q = \frac{(\lambda / \mu)^s p_0 \rho}{s!(1-\rho)^2}$$

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

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

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

Probabilidad de que un nuevo cliente tenga que esperar:

$$p_w = \left(\frac{\lambda}{\mu}\right)^s \frac{p_0}{s!\left(1-\rho\right)}$$