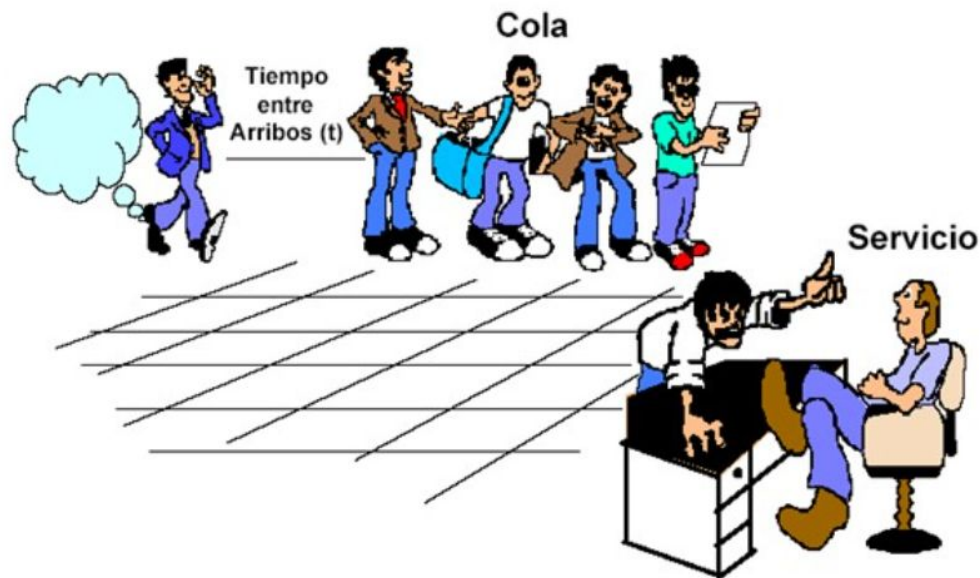


Modelos y Simulación

TEORIA DE COLAS y MODELOS DE ESPERA



UNIDAD 4: SIMULACIÓN DE EVENTOS DISCRETOS

Simulación de modelos de espera: Conceptos básicos de la Teoría de Colas. Modelado de Sistemas de Colas. Modelos de espera. Problemas de aplicación. Simulación de sistema de inventario: conceptos básicos del sistema de inventarios. simulación de montecarlo. problemas de aplicación. análisis y comparación de resultados

Cola o Línea de Espera



Se conoce como **línea de espera** a una hilera formada por uno o varios **clientes** que aguardan para recibir un servicio. Los clientes pueden ser personas, objetos, máquinas que requieren mantenimiento, contenedores con mercancías en espera de ser embarcados o elementos de inventarios a punto de ser utilizados.



Colas o Líneas de Espera

Las colas/filas de espera se encuentran en nuestra vida cotidiana.

Usos:

- la teoría de colas/líneas de espera es aplicable a empresas de servicios y manufactureras
- Personas esperando para realizar sus transacciones ante una caja en un banco, cajero automatico.
- Estudiantes esperando por obtener copias en la fotocopidora.
- Automóviles esperando pagar ante una estación de peaje o continuar su camino, ante un semáforo en rojo,
- Máquinas dañadas a la espera de ser reparadas.
- Los procesos en estado de "listos para ejecutar" esperando para la asignación del procesador.
- Los estudiantes haciendo cola para anotarse en las materias a cursar.



Esperar en un semáforo



Esperar en un cajero automático



Esperar en un supermercado

COLAS O LÍNEAS DE ESPERA

*Porque se forman las colas /
líneas de espera?*

Colas o Líneas de Espera

Debido a que el tiempo de llegada entre dos trabajos o clientes sucesivos varía y el tiempo de procesamiento varía de un consumidor a otro.

El fenómeno de las colas surge cuando la demanda real de un servicio es superior a la capacidad que existe para dar dicho servicio.

Cualquier sistema que se caracterice por elementos que posiblemente tengan que esperar en una fila para recibir algún servicio pueden conceptualizarse como un sistema de cola.

Colas o Líneas de Espera

Consideraciones:

- Los encargados de diseñar los modelos de línea de espera desean que las líneas de espera sean lo suficientemente corta.
- En La mayoría de los servicios tenemos que esperar un cierto tiempo. Para que exista un equilibrio de la satisfacción del consumidor y los costos de servicio de las empresas se mantengan en niveles adecuados.
- Las líneas de espera/cola tienen ciertos elementos comunes: consumidores que requieren adquirir el servicio, una fila donde se espera el servicio, las instalaciones de servicio y la parte de salida de servicios. Los cuales se conocen como estructuras de líneas de espera.

El estudio de las colas es importante porque proporciona tanto una base teórica del tipo de servicio que podemos esperar de un determinado recurso, como la forma en la cual dicho recurso puede ser diseñado para proporcionar un determinado grado de servicio a sus clientes.

Teoría de Colas o Líneas de Espera

QUE ES LA TEORIA DE COLA O DE LINEAS DE ESPERA

La **teoría de colas** es el estudio matemático del comportamiento de líneas de espera. Esta se presenta, cuando los "clientes" llegan a un "lugar" demandando un servicio a un "servidor", el cual tiene una cierta capacidad de atención. Si el servidor no está disponible inmediatamente y el cliente decide esperar, entonces se forma la línea de espera.

Una **cola** es una línea de espera y la **teoría de colas** es una colección de modelos matemáticos que describen sistemas de líneas de espera particulares o sistemas de colas. Los **modelos** sirven para encontrar un buen compromiso entre costes del sistema y los tiempos promedio de la línea de espera para un sistema dado.

Los **sistemas de colas** son modelos de sistemas que proporcionan servicio. Como modelo, pueden representar cualquier sistema en donde los trabajos o clientes llegan buscando un servicio de algún tipo y salen después de que dicho servicio haya sido atendido.

EN DEFINITIVA.....

- Una cola es una línea de espera
- La teoría de colas es un conjunto de modelos matemáticos que describen sistemas de líneas de espera particulares
- El objetivo es encontrar el estado estable del sistema y determinar una capacidad de servicio apropiada

Objetivos de un sistema basado en la Teoría de Colas

- Identificar el nivel óptimo de capacidad del sistema que minimiza el coste del mismo.
- Evaluar el impacto que las posibles alternativas de modificación de la capacidad del sistema tendrían en el coste total del mismo.
- Establecer un balance equilibrado (“óptimo”) entre las consideraciones cuantitativas de costes y las cualitativas de servicio.
- Prestar atención al tiempo de permanencia en el sistema o en la cola de espera.

ESTRUCTURA DE LOS PROBLEMAS DE COLAS/LÍNEAS DE ESPERA

El análisis de los problemas de líneas de espera comienza con una descripción de los elementos básicos de la situación. Cada situación específica tendrá características diferentes, **PERO CUATRO ELEMENTOS** son comunes en todas ellas.

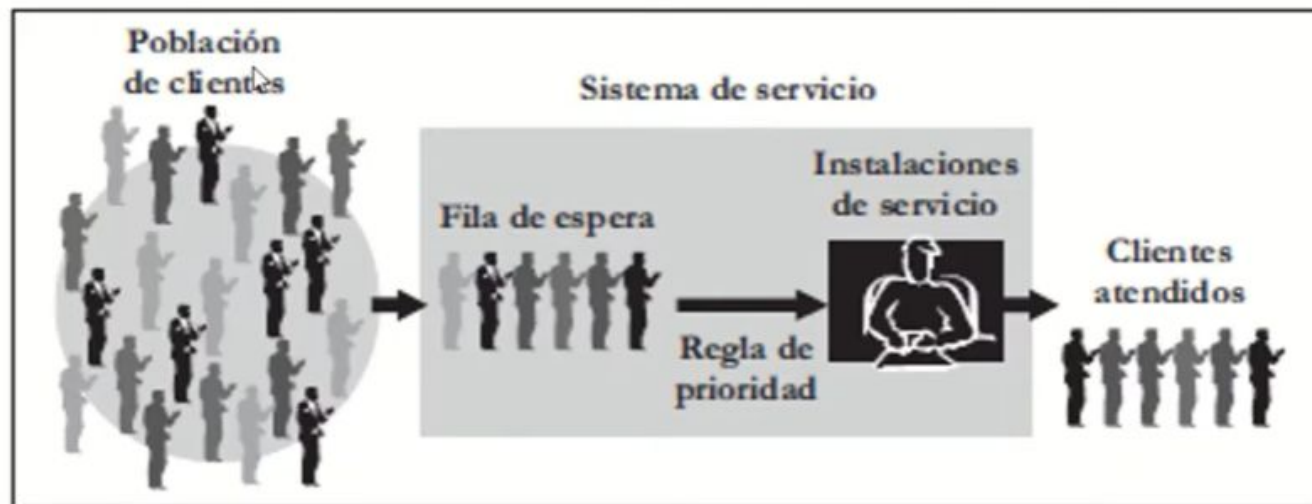
1. Un insumo o población de clientes, que genera clientes potenciales.

2. Una línea o fila de espera formada por los clientes.

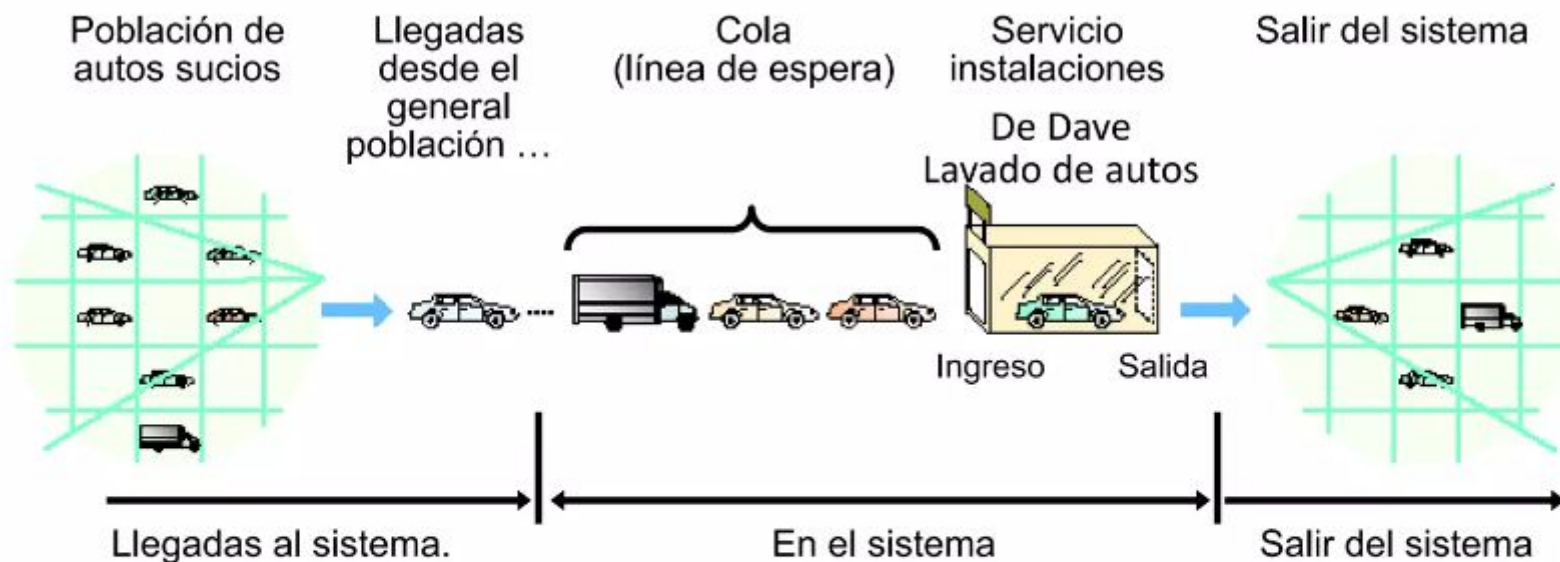
3. La instalación de servicio, constituida por una persona, una máquina o ambas.

4. Una regla de prioridad para seleccionar el siguiente cliente que será atendido.

El término 'cliente' se utiliza en un sentido general (una persona, piezas esperando su turno para ser procesadas, una lista de trabajo esperando para ser impresas en una impresora de red, etc.



- Posibles consumidores que pueden llevar a entrar. Estos clientes van a entrar y van a formar una fila de espera.
- Puede llegar a pasar que mientras esté en la fila de espera exista algo que se llama regla de prioridad. Generalmente se atiende al primer cliente que llega es el primero en ser atendido.
- Esto no se da, Por ejemplo, en los sanatorios esto no funciona. La prioridad cambia y está dada por la gravedad de la enfermedad.
- Las instalaciones del servicio puede estar formada por un máquina persona o ambas.



Características de llegada

- ▶ Tamaño de población
- ▶ Comportamiento de llegadas
- ▶ Distribución estadística de llegadas.

Características de la línea de espera

- ▶ Limitado vs. ilimitado
- ▶ Disciplina de la cola

Características de Servicio

- ▶ Diseño de servicio
- ▶ Distribución estadística del servicio

Figure D.1

ESTRUCTURA DE LOS PROBLEMAS DE COLAS/LÍNEAS DE ESPERA

En cada una de las siguientes situaciones identifica al cliente al servidor.

Situación	Cliente	Servidor
Aviones que llegan a un aeropuerto.		
Sitio de taxis que atiende a pasajeros que esperan.		
Operaciones en un estacionamiento.		

ESTRUCTURA DE LOS PROBLEMAS DE COLAS/LÍNEAS DE ESPERA

Situación	Cliente	Servidor
Aviones que llegan a un aeropuerto.	Aviones	Aeropuerto
Sitio de taxis que atiende a pasajeros que esperan.	Pasajero	Sitio de taxis
Operaciones en un estacionamiento.	Automóviles	Estacionamiento

TEORIA DE COLAS

Costos de Servicio y Costos de Espera

El diseño de cualquier sistema de colas debe tomar en consideración dos tipos de costos.

- El primero involucra la provisión del servicio. (Costo de Servicio)
- El segundo significa el tener elementos esperando en la cola. (Costo de Espera)

TEORIA DE COLAS

Costos de Servicio y Costos de Espera

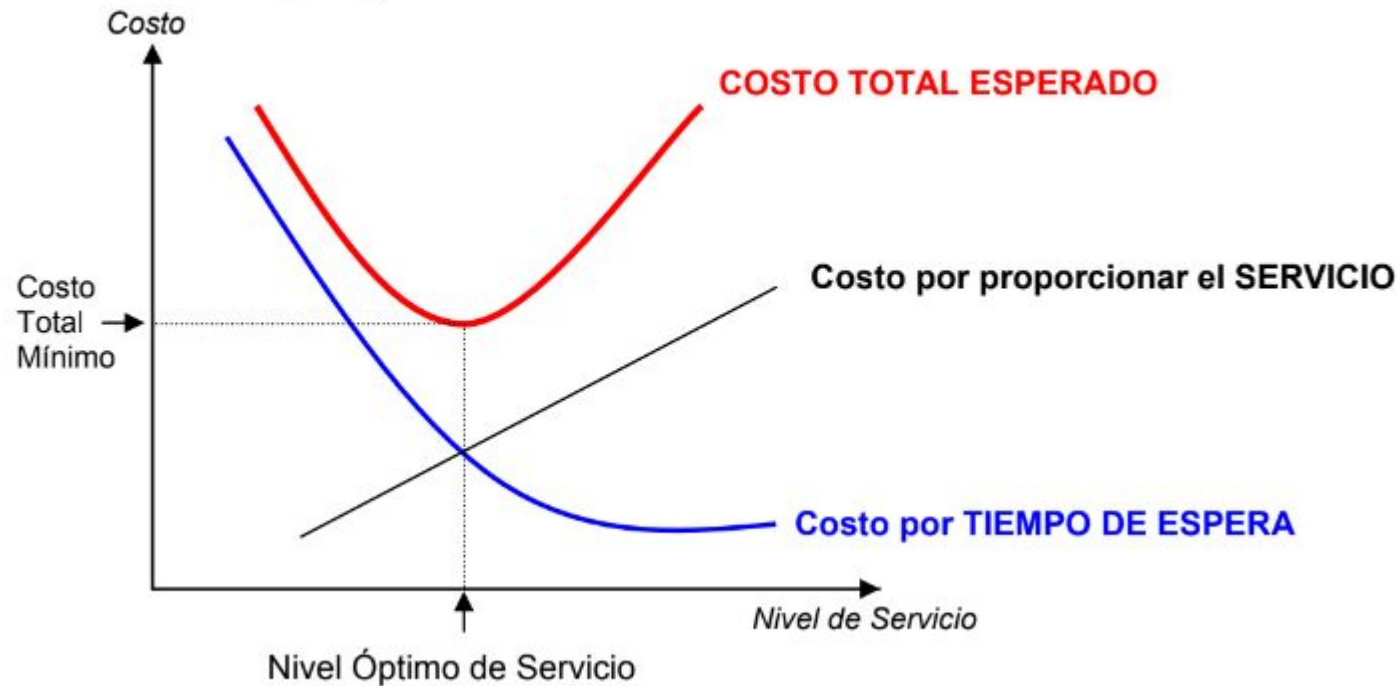
- Los Administradores reconocen el equilibrio que debe haber entre el **COSTO DE** proporcionar buen **SERVICIO** y el **COSTO** del tiempo **DE ESPERA** del cliente o de la máquina que deben ser atendidos.
- Los Administradores desean que las colas sean lo suficientemente cortas con la finalidad de que los clientes no se enojen e incluso se retiren sin llegar a utilizar el servicio o lo usen pero no retornen más.
- Sin embargo los Administradores contemplan tener una longitud de cola razonable en espera, que sea balanceada, para obtener ahorros significativos en el **COSTO DEL SERVICIO**

Un diseño óptimo del sistema debe considerar la minimización de ambos costos.

TEORIA DE COLAS

Costos de Servicio y Costos de Espera

Equilibrio entre Costos de Espera y Costos de Servicio



TEORIA DE COLAS

- Cuando el servicio mejora, disminuye el **costo** de tiempo perdido en las **líneas de espera**.
- Este costo puede reflejar pérdida de productividad de los operarios que están esperando que compongan sus equipos o puede ser simplemente un estimado de los clientes perdidos a causa de un mal **servicio** y **colas** muy largas.
- En ciertos **servicios** el **costo** de la espera puede ser intolerablemente alto.

TEORIA DE COLAS

COLAS MAS COMUNES

SITIO	ARRIBOS EN COLA	SERVICIO
Supermercado	Compradores	Pago en cajas
Peaje	Vehículos	Pago de peaje
Consultorio	Pacientes	Consulta
Sistema de Cómputo	Programas a ser corridos	Proceso de datos
Compañía de teléfonos	Llamadas	Efectuar comunicación
Banco	Clientes	Depósitos y Cobros
Mantenimiento	Máquinas dañadas	Reparación
Muelle	Barcos	Carga y descarga

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

- Una **cola de espera** está compuesta de tres elementos:
 1. Arribos o ingresos al sistema
 2. Línea de espera
 3. Servicio
- Estos tres componentes tienen *ciertas características* que deben ser examinadas antes de desarrollar el aspecto matemático de los modelos de cola.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

1. CARACTERÍSTICAS DE ARRIBO O INGRESOS DEL SISTEMA :

La fuente de ingreso que genera los arribos o clientes para el servicio tiene tres características principales:

- a. Tamaño de la población que arriba
- b. Patrón de llegada a la cola
- c. Comportamiento de las llegadas.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

1. CARACTERÍSTICAS DE ARRIBO O INGRESOS DEL SISTEMA

1.a. Tamaño de la Población: número máximo de clientes potenciales, que pueden solicitar el servicio.

- Infinito (ilimitado): Cuando el número de clientes o arribos en un momento dado es una pequeña parte de los arribos potenciales. Para propósitos prácticos poblaciones ilimitadas pueden considerarse: los vehículos que se *acercan* a una estación de peaje, los hinchas que van a una cancha, los clientes en un supermercado.

LA MAYORÍA DE LOS MODELOS ASUME ARRIBO INFINITO.

- Población de arribo limitada o finita: cuando se tienen muy pocos servidores y el servicio es restringido. Ej.: los pacientes en un consultorio médico.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

1. CARACTERÍSTICAS DE ARRIBO O INGRESOS DEL SISTEMA

1.b. Patrón de arribo a la cola:

1. Describe matemáticamente la manera en que se producen los arribos o llegadas de los clientes al sistema.
2. Tiempo entre llegadas, que es el tiempo que transcurre entre un cliente y otro sucesivo que llegan a demandar el servicio.
3. Existen dos clases básicas de tiempo entre llegadas: **determinístico** si se conoce el tiempo exacto que transcurre entre un arribo y otro, o **aleatorio**, cuya distribución probabilística se considera conocida.
 - Determinístico: se da donde los sistemas están operados por maquinarias. (empresas manufactureras). **Un paciente cada 15 minutos.**
 - Aleatorios: se da donde los sistemas están operados por personas (Confiterías, etc)
4. Se consideran que los arribos son aleatorios cuando éstos son independientes de otros y su ocurrencia no puede ser predecida exactamente.
5. Frecuentemente en problemas de colas, el número de arribos por unidad de tiempo pueden ser estimados por medio de la Distribución de Poisson que es una distribución discreta de probabilidad.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

1. CARACTERÍSTICAS DE ARRIBO O INGRESOS DEL SISTEMA

1.c. Comportamiento de los arribos:

La mayoría de los modelos de colas asume que los clientes son **pacientes** o sea que **esperan en la cola hasta ser servidos y no se pasan entre colas**. Desafortunadamente, la vida es complicada y la gente se reniega. Aquellos que se impacientan por la espera, **se retiran** de la cola sin completar su transacción.

Esta situación sirve para acentuar el estudio de la teoría de colas y el análisis de las líneas de espera, ya que un cliente no servido es un cliente perdido y hace mala propaganda de ese negocio.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

2. CARACTERÍSTICAS DE LA LÍNEA DE ESPERA

- La **LINEA DE ESPERA** es el segundo componente de un sistema de colas. La longitud de la cola puede ser también **LIMITADA** o **ILIMITADA**.
 - Cola **LIMITADA** es aquella que por aspectos físicos no puede incrementarse a tamaños infinitos. Puede ser el caso de una peluquería que tiene pocos peluqueros y sillas para atender.
 - Estudiaremos los modelos de colas asumiendo **colas de longitud infinita**. Una cola es **ILIMITADA** cuando su tamaño no tiene restricción como es el caso de una estación de peaje.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

2. CARACTERÍSTICAS DE LA LÍNEA DE ESPERA

- Una segunda característica de las líneas de espera se refiere a la **DISCIPLINA EN LA COLA** mediante la cual los clientes reciben el servicio. La mayoría de los sistemas usan la regla Primero En Entrar Primero En Salir (First In First Out).
- En las áreas de emergencia de hospitales sin embargo se omite esta regla dependiendo de la gravedad de las lesiones de las personas que arriban por auxilio médico.
- En supermercados, personas con menos de 10 artículos tienen la caja express que atiende a este tipo de clientes. Pero en la cola se les atiende con la política FIFO.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

2. CARACTERÍSTICAS DE LA LÍNEA DE ESPERA

FIFO (first in first out): Se atiende al cliente en el orden que llegan a la cola, el primero en llegar será el primero en ser atendido. En los modelos básicos de colas se supone como normal la disciplina de primero en entrar, primero en salir, a menos que se establezca de otra manera.

LIFO (last in first out): Consiste en atender primero al que ha llegado de último, también se le conoce como 'pila'.

RSS (random selection of service): Se selecciona a los clientes de una cola de forma aleatoria, con algún procedimiento de prioridad o algún otra preclasificación.

Processor Sharing: Todos los clientes experimentan con eficacia el mismo retraso, ya que comparten entre todos los clientes de la cola la capacidad del sistema atendiendo a todos por igual.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

2. CARACTERÍSTICAS DE LA LÍNEA DE ESPERA

La regla de prioridad determina a qué cliente se deberá atender a continuación.

1. A quien llega primero, se atiende primero.

2. Fecha prometida de vencimiento más próxima.

3. Tiempo de procesamiento más corto.

4. Cliente de alta prioridad.

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

3. CARACTERÍSTICAS DEL SERVICIO

El **tercer elemento** de un sistema de colas es el **SERVICIO**. En él son importantes dos propiedades básicas:

1. La **configuración** del sistema de servicio.
2. El **patrón de tiempos de servicio**

TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

3. CARACTERÍSTICAS DEL SERVICIO

3. Configuraciones básicas para el servicio

Los sistemas para el servicio son clasificados en función del número de **canales** (servidores) y el número de **fases** (número de paradas que deben hacerse durante el servicio).

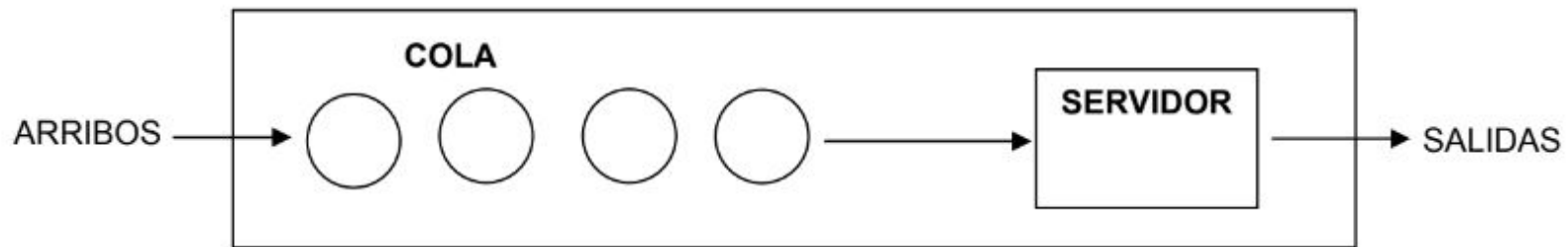
TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

3. CARACTERÍSTICAS DEL SERVICIO

3. 1. Configuraciones básicas para el servicio

**Sistemas de un solo canal y una sola fase: Todos los servicios solicitados por un cliente suelen impartirse por una instalación con un solo servidor.
Ej: Lavado automático de un auto.**



SISTEMA UN CANAL, UNA FASE

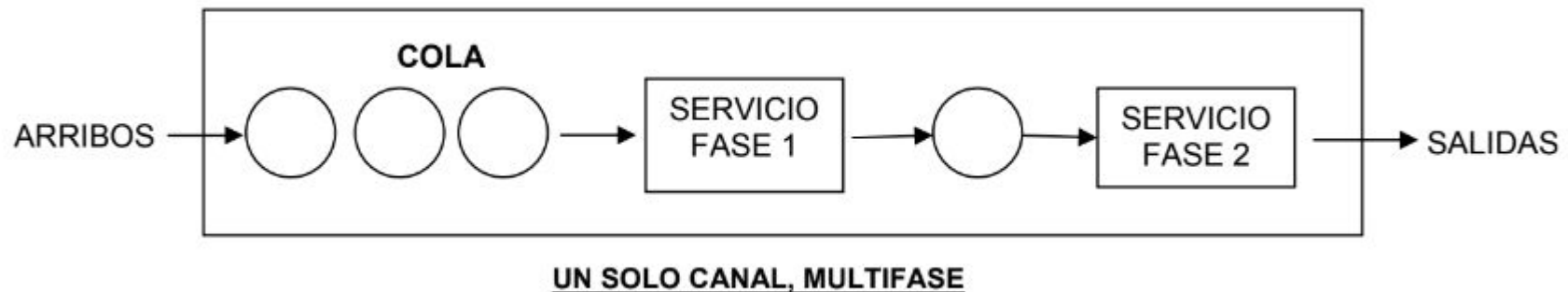
TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

3. CARACTERÍSTICAS DEL SERVICIO

3. 1. Configuraciones básicas para el servicio

Sistemas de un solo canal y múltiples fases: Se usa cuando es más conveniente que los servicios se imparten en secuencias por varias instalaciones. Pero el volumen de la clientela y otras restricciones limitan el diseño a un solo canal. Los clientes forman una sola fila y avanzan de forma secuencial, pasando de una instalación de servicio a la siguiente. **Ej: trámites para sacar pasaporte, DNI.**



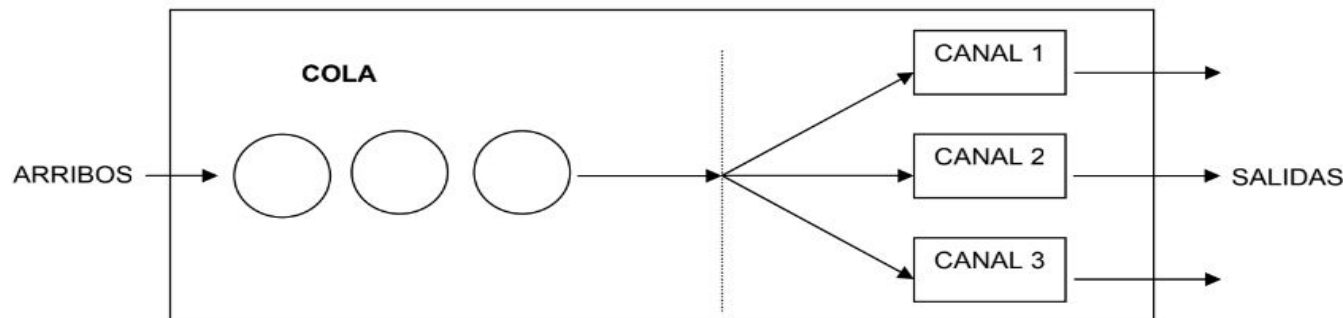
TEORIA DE COLAS

Características de una LÍNEA DE ESPERA

3. CARACTERÍSTICAS DEL SERVICIO

3. 1. Configuraciones básicas para el servicio

Sistema de cola multi-canal y una sola fase: Se usa cuando la demanda es suficientemente grande para justificar que se suministre el mismo servicio en más de una instalación (servidor) o bien cuando los servicios ofrecidos por las instalaciones son diferentes. Los clientes forman una fila o varias dependiendo del diseño. Son principalmente los cajeros de un banco en los cuales hay una sola cola y varias personas atendiendo a los clientes en diferentes cajas.



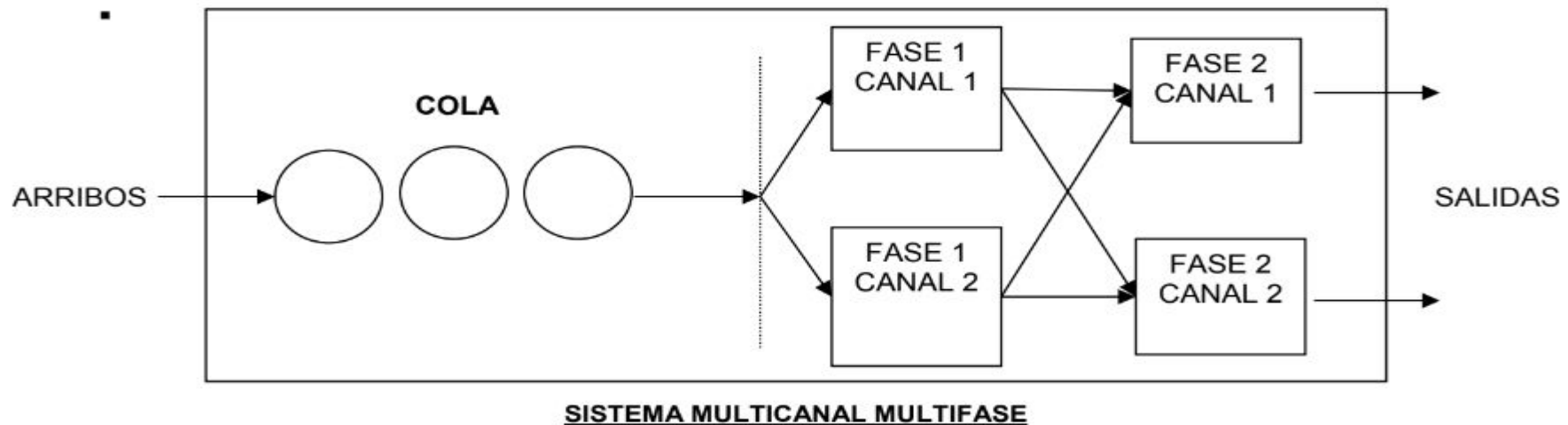
TEORIA DE COLAS

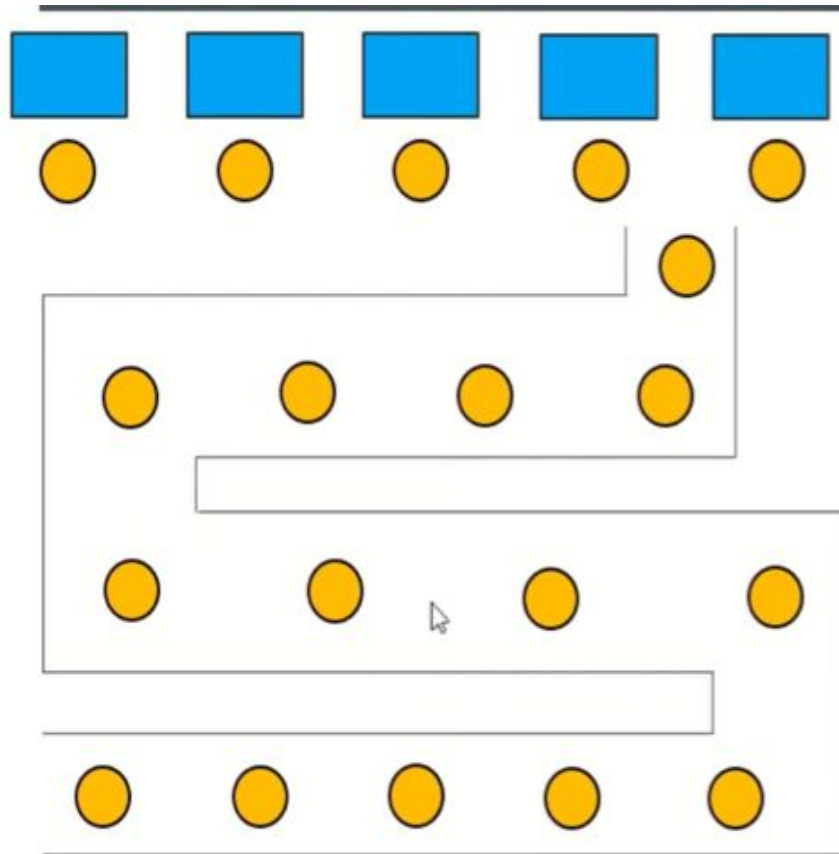
Características de una LÍNEA DE ESPERA

3. CARACTERÍSTICAS DEL SERVICIO

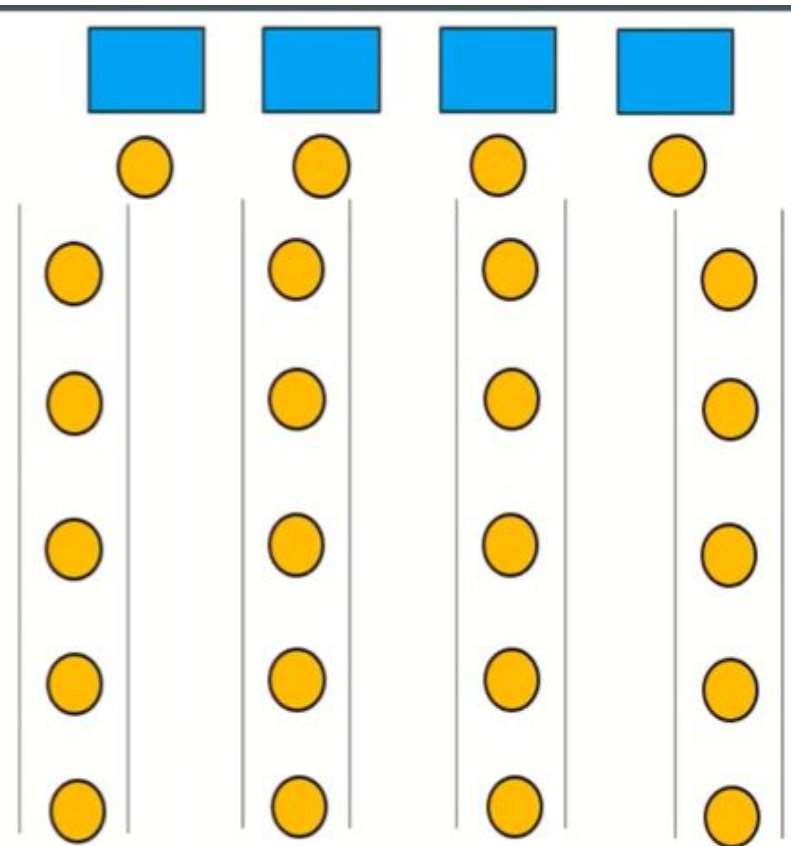
3. 1. Configuraciones básicas para el servicio

Sistema de cola multicanal multifase: Los clientes pueden ser atendidos por una de las instalaciones de la primera fase pero despues requieren los servicios de una instalación de la segunda fase y así sucesivamente. Ej: los negocios de servicios de lavado.





a) Una sola fila



b) Múltiples filas

El consultorio de un dentista familiar.



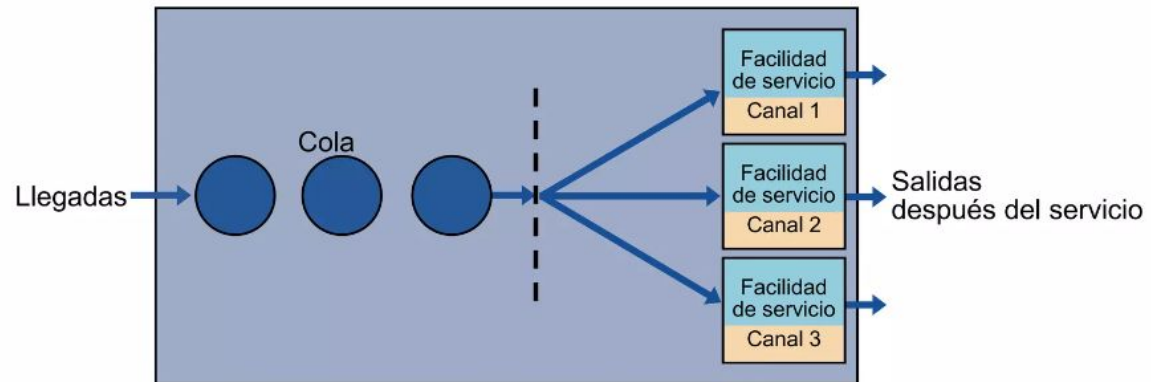
Servidor único, sistema monofásico

Un servicio al carro de doble ventana de McDonald's



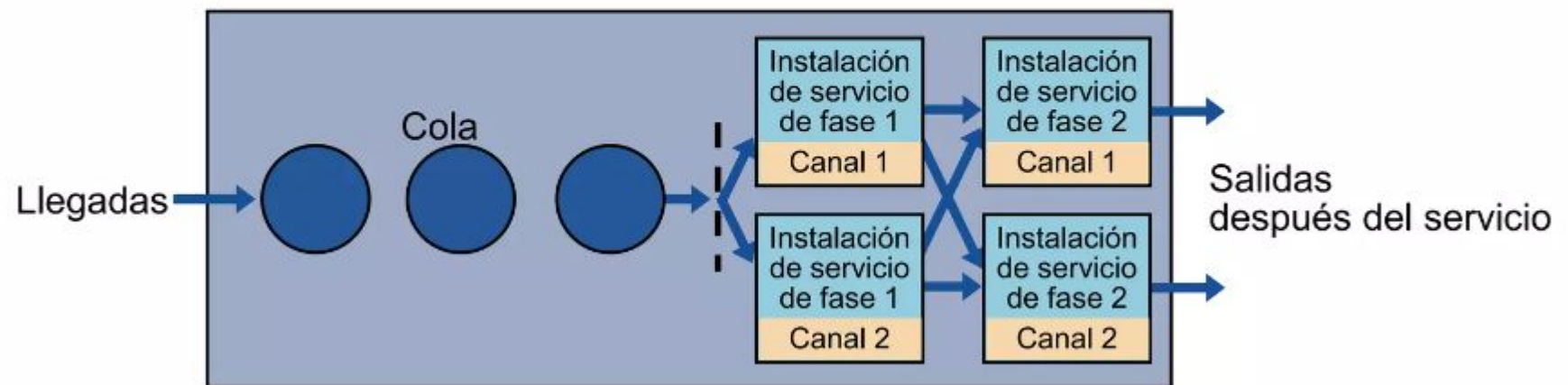
Servidor único, sistema multifásico

La mayoría de las ventanas de servicios bancarios y postales

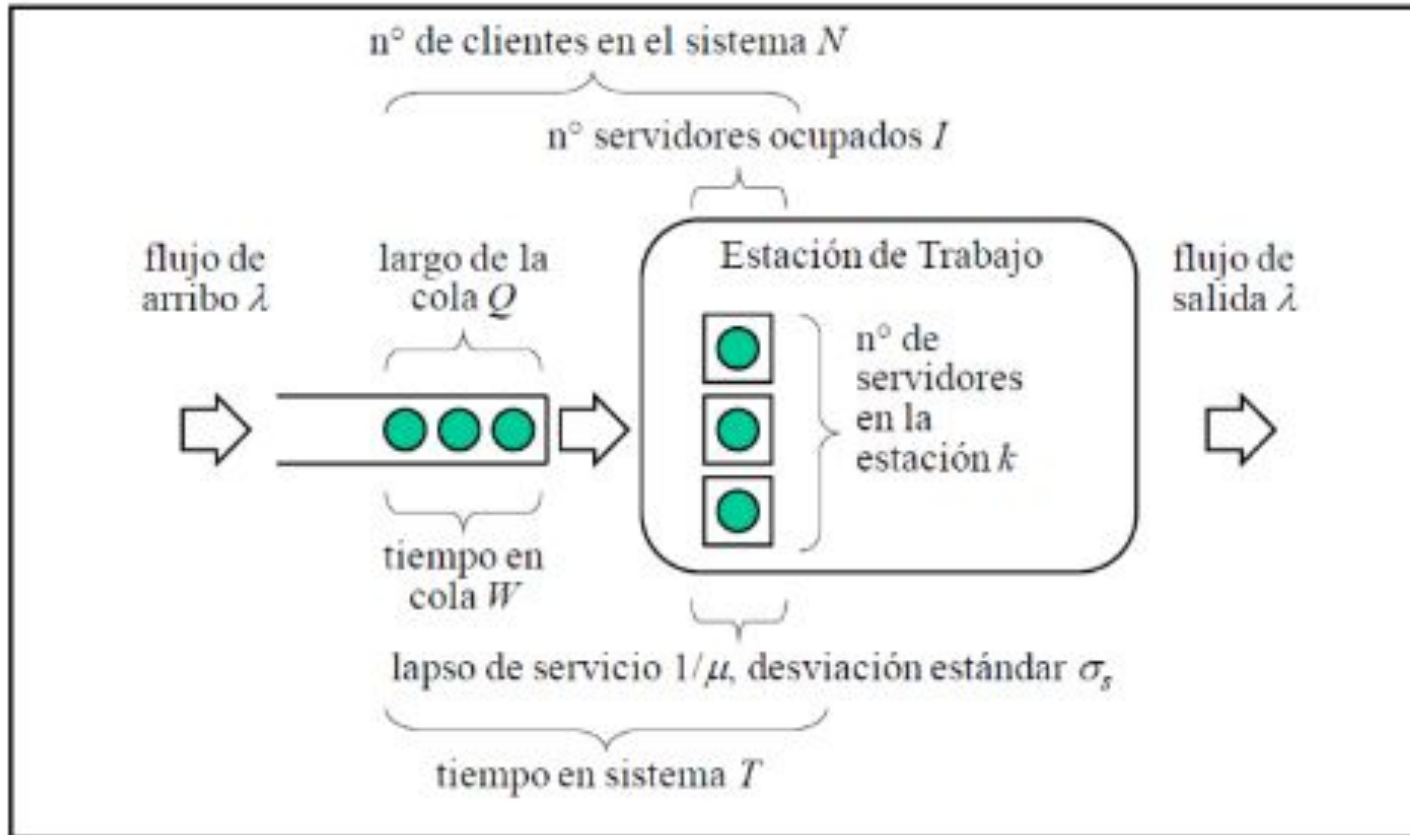


Multi-servidor, sistema monofásico

Algunas cafeterías universitarias



Modelo de Cola de Espera



Modelo de sistema de espera que consiste en un número de servidores y en una cola de clientes. Los valores N , λ y Q , entre otros, definen las características del sistema

Modelo de Cola de Espera

1. Los principales actores que participan en un modelo de Colas son los clientes y los servidores.
2. Entendiéndose por cliente una persona, una orden de servicio, un automóvil, entre otros y el servidor será aquella estación que pueda realizar la respectiva actividad de servicio sobre el cliente, cómo ser un cajero, una secretaria, una maquina, entre otros los clientes que requieran un servicio se generan en el tiempo en una fuente de entrada.
3. Éstos, ingresan al sistema y se unen a la cola. En determinado momento se selecciona a un miembro de la cola para proporcionar un servicio mediante alguna regla conocida como disciplina de la cola. Se realiza el servicio que el cliente requiere mediante un mecanismo de servicio y luego el cliente sale del sistema de cola.
4. La disciplina de cola se define como el proceso para decidir qué usuario será llamado de la cola o al que se le prestará servicio.

Las disciplinas básicas son:

- Forma en que los clientes llegan al punto de servicio: aleatoria o determinística.
- Forma como se realiza el servicio: aleatoria o determinística.
- Modo de elegir los clientes de la fila que espera el servicio. Se puede optar por la norma FIFO (primero que entra, primero que sale), la norma LIFO (último que entra, primero que sale) o la norma RSS (selecciona los clientes de manera aleatoria, de acuerdo a algún procedimiento de prioridad o a algún otro orden.). Cabe aclarar que si hay varios puntos de servicio, la descripción del fenómeno de espera necesita otras especificaciones.

TEORIA DE COLAS

Medición del Rendimiento de las Colas

- Los modelos de colas ayudan a los administradores a tomar **decisiones para balancear los costos de servicio deseables con los costos de espera en la línea.**
- **Existen muchas medidas de rendimiento diferentes que se utilizan para evaluar un sistema de colas en estado estable.**
-
- Los principales factores que se evalúan en estos modelos son:
 1. **Tiempo promedio** que cada cliente u objeto permanece en la cola
 2. **Longitud de cola promedio**
 3. **Tiempo promedio que cada cliente permanece en el sistema** (tiempo de espera + tiempo de servicio).
 4. **Número de clientes promedio en el sistema.**
 5. **Probabilidad** de que el servicio se quede vacío
 6. **Factor de utilización** del sistema
 7. **Probabilidad** de la presencia de un **específico número de clientes** en el sistema.

MODELOS DE COLAS

Notación de los Modelos de Colas

- Reconociendo la diversidad de los sistemas de colas, **Kendall** (1953) propuso un sistema de notación para sistemas de servidores paralelos que ha sido adoptado universalmente.
- La notación de **Kendall** está dada por

A/B/c/N/K

- A = Distribución de tiempo entre arribos.
- B = Distribución del tiempo de servicio.
- c = números de servidores paralelos número de servidores existentes en la estación, los que se suponen son idénticos y funcionan en paralelo
- N= Capacidad del sistema, o sea cantidad de clientes que pueden estar en el sistema (sea esperando o siendo atendidos). Cuando no hay límite para la cantidad de clientes que están en la cola, no se incluye notación alguna. Esto implica que la capacidad del sistema es infinita.
- K= Tamaño de la población. Cuando se omite, se asume infinito. Si el ingreso está limitado a K clientes, se escribe A/B/c/K/K implicando que la capacidad también se limita a K. Esto se hace incluso cuando la capacidad de la sala de espera no tenga límite, ya que no necesitamos una sala mayor que para K.

MODELOS DE COLAS

Notación de los Modelos de Colas

A

M = Distribución de Poisson.
D = Constante.
G = General, es decir tiene
varianza y medias conocidas.

B

M = Exponencial.
D = Constante.
G = General.

K

Los canales pueden ser 1, 2,
3, ..., etc.

MODELOS DE COLAS

Notación de los Modelos de Colas

Los siguientes son símbolos comunes para **A** y **B**:

- **M** = **Para distribución de tiempo entre llegadas**, la letra "M" significa un proceso markoviano de entradas, que indica que las llegadas se producen de manera aleatoria. Pueden provenir de una fuente infinita, en cuyo caso se debe especificar el número de llegadas por unidad de tiempo (average arrival rate). Si proceden de una fuente finita, se deberá establecer la tasa de llegadas para cada cliente.
Para distribución de tiempo de servicio, significa una distribución exponencial del tiempo de servicio (otra vez M proviene de markoviano). Ésta es la más común en la práctica, y representa en ese sentido el tiempo más aleatorio o impredecible.
- **D** = constante o determinística
- **Ek** = Erlang de orden k
- **PH** = Tipo fase
- **H** = Hiperexponencial
- **G** = Distribución de probabilidad no especificada o general. Ciertos resultados pueden obtenerse sin necesidad de algún supuesto respecto a la distribución, pero en este caso se deberá ingresar la desviación estándar de los tiempos de servicio.

MODELOS DE COLAS

Notación de los Modelos de Colas

- Por ejemplo: $M/M/1/\infty/\infty$ significa *un solo servidor*, capacidad de *cola ilimitada* y *población infinita* de arribos potenciales. Los tiempos entre arribos y los tiempos de servicio son **distribuidos exponencialmente**.
- Cuando N y K son *infinitos*, pueden ser *descartados* de la notación. $M/M/1/\infty/\infty$ es reducido a $M/M/1$.

MODELOS DE COLAS

Variedad de Modelos de Colas

- Existe una cantidad enorme de Modelos de Colas que pueden utilizarse.
- Vamos a concentrar en 4 de los modelos más usados.
- Los 4 modelos de colas a estudiar asumen:
 - o Arribos según la Distribución de Poisson
 - o Disciplina FIFO
 - o Una sola fase de servicio.

-

MODELOS DE COLAS

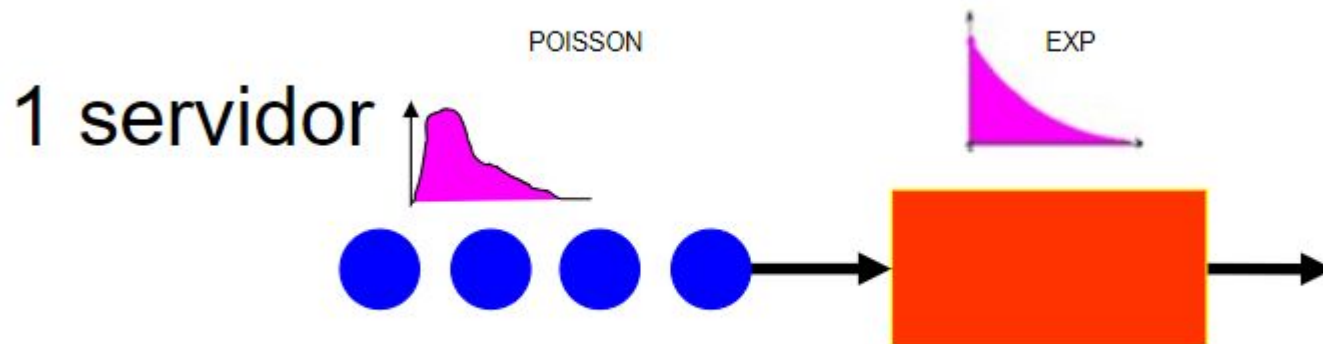
Variedad de Modelos de Colas

- **Modelo A: Un canal,**
- **Modelo B: Multicanal**
- **Modelo C: Tiempo de Servicio constante**
- **Modelo D: Población Limitada**

MODELOS DE COLAS

Modelo A: Modelo de Colas de un solo canal, (Modelo M/M/1)

- Arribos que siguen la distribución de Poisson
- Tiempos de Servicio Exponenciales
- Es el más sencillo, un solo servidor y una sola fila.



MODELOS DE COLAS

Modelo A: M/M/1

M=la forma en que llegan los consumidores

M= la forma que son atendidos

1= cuántos servidores están operando

Descripción del modelo

- Hay una sola cola, cuya capacidad es infinita, y un solo servidor, La disciplina será FIFO
- Las llegadas se producen según un proceso de Poisson de razón λ , donde λ es el número medio de llegadas por unidad de tiempo y $1/\lambda$ es el tiempo medio entre llegadas, Los tiempos entre llegadas se distribuirán exponencialmente, $\text{Exp}(\lambda)$
- Los tiempos entre servicios también se distribuirán exponencialmente, $\text{Exp}(\mu)$, de tal manera que μ es el número medio de clientes que el servidor es capaz de atender por unidad de tiempo y $1/\mu$ es el tiempo medio de servicio

Condición de no saturación

- Se demuestra que si $\lambda \geq \mu$, el sistema se satura, es decir, el número de clientes en la cola crece indefinidamente con el tiempo, Por consiguiente, la condición de no saturación será:

$$\rho < 1, \quad \text{donde} \quad \rho = \frac{\lambda}{\mu}$$

- Nosotros sólo estudiaremos las colas que no se saturan, Cuando una cola no se satura, también se dice que alcanza el estado estacionario,

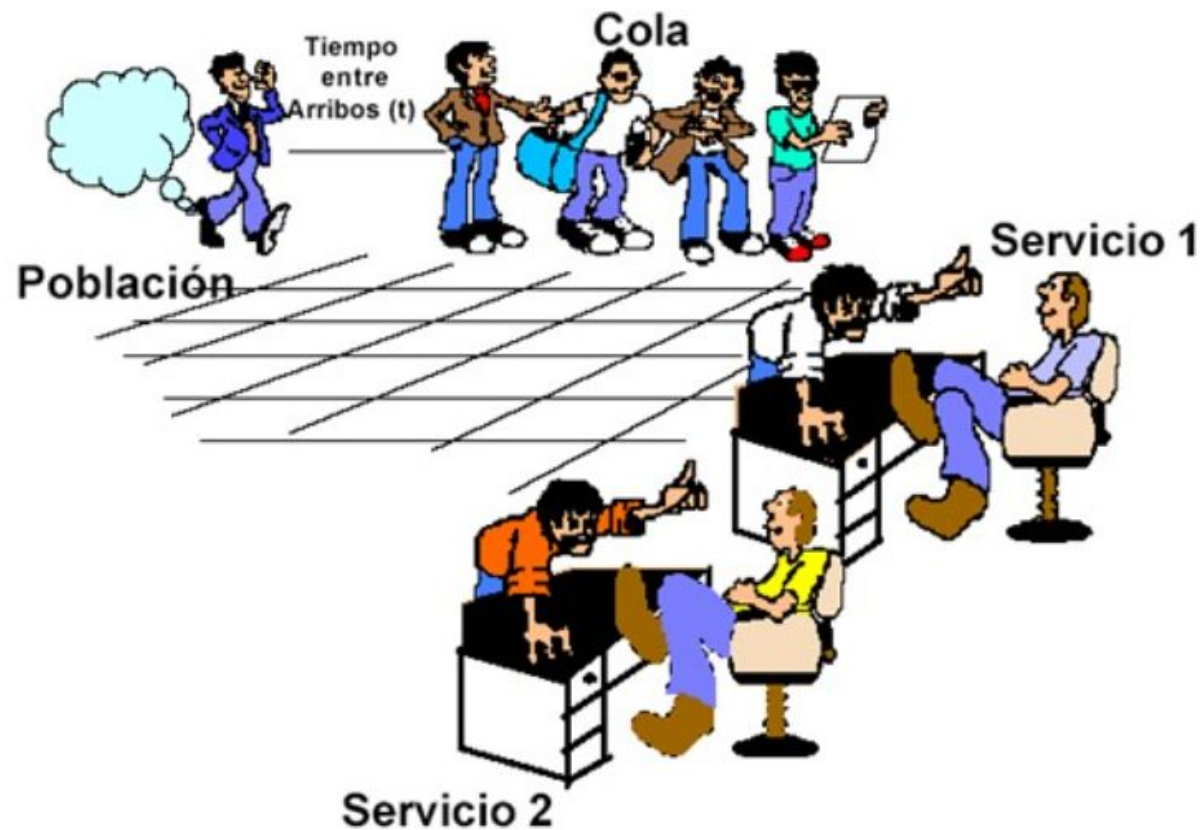
Cuando una cola no se satura se dice que alcanza el estado estacionario

Modelos de colas descritos		
MODELO	NOMBRE	EJEMPLO
A	Sistema de servidor único (M/M/1)	Mostrador de información en grandes almacenes

NÚMERO DE SERVIDORES (CANALES)	NÚMERO DE FASES	PATRÓN DE TASA DE LLEGADA	MODO DE TIEMPO DE SERVICIO	TAMAÑO DE LA POBLACION	DISCIPLINA DE LA COLA
Simple	Simple	Poisson	Exponencial negativo	Ilimitado	FIFO

MODELOS DE COLAS

Modelo B: (M/M/S)



MODELOS DE COLAS

Modelo B: (M/M/S)

- Dos o más servidores o canales están disponibles para atender a los clientes que arriban.
- Los clientes forman una sola cola y se los atiende de acuerdo al servidor que queda libre.
- Asumimos que los arribos siguen la distribución de probabilidad de Poisson y los tiempos de servicio son distribuidos exponencialmente.
- Los servicios se los hace de acuerdo a la política primero en llegar primero en ser servido (FIFO) y todos los servidores atienden a la misma tasa.

Modelos de colas descritos		
MODELO	NOMBRE	EJEMPLO
B	Servidor multiple (M/M/S)	Mostrador de boletos aéreos

NÚMERO DE SERVIDORES (CANALES)	NÚMERO DE FASES	PATRÓN DE TASA DE LLEGADA	MODO DE TIEMPO DE SERVICIO	TAMAÑO DE LA POBLACION	DISCIPLINA DE LA COLA
Servidor multiple	Simple	Poisson	Exponencial negativo	Ilimitado	FIFO

MODELOS DE COLAS

Modelo B: (M/M/S)

Cuando $\rho = \frac{\lambda}{\mu} \geq 1 \Rightarrow$

Tasa media de servicio menor que la tasa media de llegadas. Se forma una cola infinita

Cuando $\rho = \frac{\lambda}{\mu} < 1 \Rightarrow$

Tasa media de servicio mayor que la tasa media de llegadas

El sistema de colas alcanzará la condición de estado estable y podemos aplicar directamente los resultados de estado estable hallados anteriormente.

MODELOS DE COLAS

Modelo B: $(M/M/\infty)$

Número ilimitado de servidores

No hay cola. cada cliente que llega es servido directamente

MODELOS DE COLAS

Modelo C: (M/D/I)

- Algunos sistemas tienen tiempos de servicio constantes en lugar de exponencialmente distribuidos. Cuando los clientes son atendidos o equipos son procesados con un ciclo fijo como es el caso de una lavadora o ciertos entretenimientos en los parques de diversiones, el asumir servicio constante es adecuado.

MODELOS DE COLAS

Modelo C: (M/D/1)

Modelos de colas descritos		
MODELO	NOMBRE	EJEMPLO
C	Servicio constante (M / D / 1)	Lavado de autos automatizado

NÚMERO DE SERVIDORES (CANALES)	NÚMERO DE FASES	PATRÓN DE TASA DE LLEGADA	MODO DE TIEMPO DE SERVICIO	TAMAÑO DE LA POBLACION	DICIPLINA DE LA COLA
Simple	Simple	Poisson	Constante	Ilimitado	FIFO

MODELOS DE COLAS

Modelo D: Población limitada

► Supuestos

1. Solo hay un servidor
2. La población de unidades que buscan servicio es limitada.
3. Las llegadas siguen una distribución de Poisson, los tiempos de servicio son negativos distribuidos exponencialmente
4. Los clientes son atendidos por orden de llegada

► Supuestos

1. Solo hay un servidor
2. La población de unidades que buscan servicio es limitada.
3. Las llegadas siguen una distribución de Poisson, los tiempos de servicio son negativos distribuidos exponencialmente
4. Los clientes son atendidos por orden de llegada

RESUMEN DE LOS MODELOS DE COLAS

DESCRITOS

MODELO	NOMBRE	Nº DE CANAL ES	Nº DE FASES	PATRÓN DE ARRIBO	PATRÓN DE SERVICIO	TAMAÑO DE LA POBLACIÓN	DISCIPLINA DE COLA
A	SIMPLE M/M/I	UNO	UNA	POISSON	EXPONEN CIAL	INFINITA	FIFO
B	MULTI- CANAL M/M/S	MULTI CANAL	UNA	POISSON	EXPONEN CIAL	INFINITA	FIFO
C	SERVICIO CONSTANTE (M/D/I)	UNO	UNA	POISSON	CONSTAN TE	INFINITA	FIFO
D	POBLACION LIMITADA	UNO	UNA	POISSON	EXPONEN CIAL	FINITA	FIFO

FORMULAS PARA COLAS.

MODELO A: SISTEMA SIMPLE O M/M/1

λ = Número de llegadas por unidad de tiempo.

μ = Número de servicios por unidad de tiempo.

p = Utilización promedio del sistema.

$p_{(n)}$ = Probabilidad de que n clientes estén en el sistema

L_s = Número promedio de clientes en el sistema de servicio.

L_q = Número promedio de clientes en la fila.

w_s = Tiempo promedio transcurrido en el sistema, incluido en el servicio.

w_q = Tiempo promedio de espera en la fila.

FORMULAS PARA COLAS

MODELO A: SISTEMA SIMPLE O M/M/1

$$\lambda = \text{velocidad de llegadas} \frac{\text{clientes}}{\text{hora}}$$

$$\mu = \text{velocidad de servicio} \frac{\text{clientes}}{\text{hora}}$$

$$\frac{1}{\lambda} = \text{tiempo entre llegadas.}$$

$$\frac{1}{\mu} = \text{tiempo entre servicio.}$$

$$p = \frac{\lambda}{\mu}$$

$$p_{(n)} = (1 - p)p^n \quad \circ \quad p_{(n)} = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^n$$

$$L_s = \frac{\lambda}{\mu - \lambda}$$

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

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

$$w_q = pw_s \quad \circ \quad w_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Estas formulas solo son aplicables si y solo si:

$$\lambda < \mu \quad \circ \quad \mu > \lambda$$

FORMULAS PARA COLAS

MODELO A: SISTEMA SIMPLE O M/M/1

Fórmulas de cola para el modelo A: sistema de servidor único, también denominado M / M / 1

λ = número promedio de llegadas por período de tiempo

μ = número promedio de personas o artículos atendidos por período de tiempo (tasa de servicio promedio)

L_s = número promedio de unidades (clientes) en el sistema (esperando y siendo atendido)

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

W_s = tiempo promedio que una unidad pasa en el sistema (tiempo de espera más tiempo de servicio)

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

FORMULAS PARA COLAS

MODELO A: SISTEMA SIMPLE O M/M/1

Fórmulas de cola para el modelo A: sistema de servidor único, también denominado M / M / 1

$$\begin{aligned} L_q &= \text{número promedio de unidades que esperan en la cola} \\ &= \frac{\lambda^2}{\mu(\mu - \lambda)} \end{aligned}$$

$$\begin{aligned} W_q &= \text{tiempo promedio que una unidad pasa esperando en la cola} \\ &= \frac{\lambda}{\mu(\mu - \lambda)} = \frac{L_q}{\lambda} \end{aligned}$$

$$\begin{aligned} \rho &= \text{factor de utilización del sistema} \\ &= \frac{\lambda}{\mu} \end{aligned}$$

FORMULAS PARA COLAS

MODELO A: SISTEMA SIMPLE O M/M/1

Fórmulas de cola para el modelo A: sistema de servidor único, también denominado M / M / 1

P_0 = Probabilidad de 0 unidades en el sistema (es decir, la unidad de servicio está inactiva)

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

$P_{n>k}$ = probabilidad de más de k unidades en el sistema, donde n es el número de unidades en el sistema

$$= \left[\frac{\lambda}{\mu} \right]^{k+1}$$

EJEMPLO. MODELO A: SISTEMA SIMPLE O $M/M/1$

La gerente de un supermercado está interesada en brindar un buen servicio a las personas de mayor edad que compran en su local. Actualmente, el supermercado cuenta con una caja de salida reservada para los jubilados. Estas personas llegan a la caja a un ritmo promedio de 30 por hora, de acuerdo con una distribución de Poisson, y son atendidos a una tasa promedio de 35 clientes por hora, con tiempos de servicio exponenciales. Calcule los siguientes promedios:

- a) Utilización del empleado de la caja de salida.
- b) Número de clientes que entran al sistema.
- c) Número de clientes formados en la fila.
- d) Tiempo transcurrido dentro del sistema.
- e) Tiempo de espera en la fila.

EJEMPLO. MODELO A: SISTEMA SIMPLE O M/M/I

1. Identificar a λ (Velocidad de llegada clientes) y μ (Velocidad del servicio)

$$\lambda = 30 \text{ clientes/hr}$$

$$\mu = 35 \text{ clientes/hr}$$

a) Utilización del empleado de la caja de salida.

$$p = \frac{\lambda}{\mu} = \frac{30}{35} = 0.857 = 85.7\%$$

La utilización promedio del sistema

b) Número de clientes que entran al sistema.

$$L_s = \frac{\lambda}{\mu - \lambda} = \frac{30}{35 - 30} = \frac{30}{5} = 6 \text{ clientes}$$

Número promedio de clientes en el sistema es de 6.

EJEMPLO. MODELO A: SISTEMA SIMPLE O M/M/1

1. Identificar a λ (Velocidad de llegada clientes) y μ (Velocidad del servicio)

$$\lambda = 30 \text{ clientes/hr}$$

$$\mu = 35 \text{ clientes/hr}$$

c) Número de clientes formados en la fila.

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

$$L_q = (0.875)(6) = 5.14 \text{ clientes}$$

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{(30)^2}{35(35 - 30)} = \frac{900}{35(5)} = \frac{900}{175} = 5.14 \text{ clientes}$$

Número promedio de
clientes en la fila

d) Tiempo transcurrido dentro del sistema.

$$w_s = \frac{1}{\mu - \lambda} = \frac{1}{35 - 30} = \frac{1}{5} = 0.20 \text{ horas o } 12 \text{ minutos}$$

EJEMPLO. MODELO A: SISTEMA SIMPLE O M/M/1

1. Identificar a λ (Velocidad de llegada clientes) y μ (Velocidad del servicio)

$$\lambda = 30 \text{ clientes/hr}$$

$$\mu = 35 \text{ clientes/hr}$$

e) Tiempo de espera en la fila.

$$w_q = pws \text{ o } w_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

$$w_q = (0.875)(0.20) = 0.17 \text{ hora o } 10.28 \text{ minutos}$$

$$w_q = \frac{30}{35(35 - 30)} = \frac{30}{35(5)} = \frac{30}{175} = 0.17$$

EJEMPLO. MODELO A: SISTEMA SIMPLE O $M/M/1$

Solución: La caja de salida puede representarse como un sistema con un solo canal y una sola fase.

Porque???? La velocidad de llegada es menor a la velocidad de atención con un solo servidor. No se excede, el número promedio de clientes, la utilización del sistema, etc

UNA COLA CON UN SERVIDOR ÚNICO

Tom Jones, el mecánico de Golden Muffler Shop, está preparado para instalar nuevos tubos de escape a un ritmo medio de 3 por hora (o aproximadamente 1 cada 20 minutos), según una distribución exponencial negativa. Los clientes que buscan este servicio llegan al taller a un promedio de 2 por hora, siguiendo una distribución de Poisson. Se les atiende sobre la base del primero que entra, primero que sale, y proceden de una población de posibles clientes muy grande, casi infinita.

A partir de esta descripción, podemos obtener las características de funcionamiento del sistema de colas de Golden Muffler.

SOLUCIÓN ►

$\lambda = 2$ automóviles que llegan por hora

$\mu = 3$ automóviles atendidos por hora

$$L_s = \frac{\lambda}{\mu - \lambda} = \frac{2}{3 - 2} = \frac{2}{1}$$

= 2 automóviles en el sistema, como promedio

$$W_s = \frac{1}{\mu - \lambda} = \frac{1}{3 - 2} = 1$$

= tiempo medio de espera en el sistema de una hora

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{2^2}{3(3 - 2)} = \frac{4}{3(1)} = \frac{4}{3}$$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{2}{3(3 - 2)} = \frac{2}{3} \text{ horas}$$

= 40 minutos de tiempo medio de espera por automóvil

$$\rho = \frac{\lambda}{\mu} = \frac{2}{3}$$

= el mecánico está ocupado el 66,6 % del tiempo

$$P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{2}{3}$$

= probabilidad del 0,33 de que haya 0 automóviles en el sistema

Probabilidad de más de k automóviles en el sistema

K	$P_{n > k} = (2/3)^{k+1}$
0	0,667 ← Observe que esto es igual a $1 - P_0 = 1 - 0,33 = 0,667$
1	0,444
2	0,296
3	0,198 ← Implica que hay una posibilidad del 19,8 % de que haya más de 3 automóviles en el sistema.
4	0,132
5	0,088
6	0,058
7	0,039