



# TÉCNICAS DE SIMULACIÓN

ING. JORGE MOYA

ESTUDIO DE TIEMPO DE ESPERA CON  
MONTECARLO EN ATENCION AL CLIENTE PARA  
SOPORTE TÉCNICO EN LA EMPRESA CELL  
TUNNING DE MANTA

SÁNCHEZ COBEÑA CRISTÓBAL ANDRÉS  
OCTAVO NIVEL "B"

8 DE SEPTIEMBRE DE 2017

## Contenido

Resumen .....	3
Introducción .....	4
Objetivos.....	5
Objetivo General .....	5
Objetivos Específicos .....	5
Generalidades .....	6
Acerca de la empresa .....	6
Diversificación de productos.....	6
Organigrama .....	6
Figura 1. Organigrama EMPRESA CELL TUNNING.....	6
Marco teórico.....	7
PROCESOS ESTOCÁSTICOS .....	7
TEORÍA DE COLAS.....	7
Figura 2. Estructura de sistemas de colas .....	8
Sistema de clasificación .....	8
Figura 3. Clasificación Kendall y Lee .....	8
Figura 4. Tipos de distribuciones .....	8
Medidas de desempeño .....	9
REDES DE COLAS .....	9
Figura 5. Redes de colas .....	9
Clasificación de las redes de colas.....	10
Figura 6. Clasificación de las redes de colas. ....	10
Medidas de desempeño .....	10
APLICACIONES DE TEORÍA DE COLAS .....	11
SIMULACIÓN.....	11
OPTIMIZACIÓN.....	11
Tipos de Optimización.....	11
REDES DE COLAS EN CENTROS DE SERVICIO.....	12
DEFINICIÓN DE LA RED DE COLAS .....	12
Figura 7. Tipos de redes .....	13
Figura 7. Estructura de la red de colas definida .....	14
FACTORES DE INFLUENCIA.....	14
Distribuciones de probabilidad de los tiempos entre llegadas.....	14
Distribuciones de probabilidad de tiempos entre llegadas a contrastar.....	15

Tabla 1. Resumen parámetros de distribuciones de llegada .....	17
Modelos de simulación .....	17
Figura 8. Modelos de simulación. ....	17
RESULTADOS DEL EXPERIMENTO .....	17
GENERALIDADES: INTERACCIONES ENTRE LOS FACTORES.....	18
ANÁLISIS DEL NÚMERO PROMEDIO DE PERSONAS EN COLA .....	18
Relación entre número esperado de personas en cola y nodo de atención .....	19
Figura 9. Redes de colas con capacidad finita .....	19
Resumen de interacción de factores sobre el número esperado de personas en cola .....	19
Tabla 9. Resumen relación entre número esperado de personas en cola y nodo de atención .	19
Tabla 10. Resumen relación número esperado de personas en cola y factor de utilización .....	20
Tabla 11. Resumen relación número esperado de personas en cola y número de servidores .	20
Tabla 12. Resumen relación número esperado de personas en cola y capacidad del nodo ....	21
Resumen de interacción de factores sobre el tiempo esperado en cola.....	22
Tabla 18. Resumen relación tiempo esperado en cola y nodo de atención.....	22
Tabla 19. Resumen relación tiempo esperado en cola y el factor de utilización .....	22
Tabla 20. Resumen relación tiempo esperado en cola y número de servidores.....	23
Tabla 21. Resumen relación tiempo esperado en cola y capacidad del nodo .....	23
Tabla 22. Resumen relación número esperado de personas en cola, capacidad del nodo y cantidad de servidores.....	24
ANÁLISIS GENERAL DEL SISTEMA .....	24
RESUMEN MEDIDAS DE DESEMPEÑO .....	24
Costos de utilización del sistema .....	25
Costos de espera .....	25
Otros Costos .....	26
MODELO DE OPTIMIZACIÓN.....	26
Definición de variables .....	26
Funciones objetivo .....	26
Restricciones .....	27
Restricciones asociadas con los servidores .....	27
Restricciones asociadas con la capacidad .....	27
Restricciones asociadas con los costos .....	28
Modelo Final .....	28
Conclusión.....	29
Bibliografía.....	29
Linkografía.....	29

## Resumen

---

En los sitios de atención en donde se forman colas, se hace necesaria la gestión eficiente de los recursos con el propósito de brindar un mejor servicio. Para ello, se debe conocer el comportamiento de dichos sistemas a través de un estudio de las variables que intervienen en él y cómo estas afectan las medidas de desempeño, buscando llegar a conclusiones acerca de diseño óptimo de un sistema en las complejas estructuras de colas que existen actualmente.

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. La teoría de colas estudia modelos matemáticos que describen sistemas de línea de espera con diferentes características. Los modelos nos ayudan a encontrar un equilibrio entre los costos del sistema y los tiempos promedio de la línea de espera para un sistema dado. Algunos estudios han llegado a la conclusión que, por término medio, un ciudadano promedio pasa cinco años de su vida esperando en distintas colas, y de estos cinco años casi seis meses esperando que cambie la luz en los semáforos (claro aquellos que respetan los semáforos).

Para cumplir con dicho objetivo en este estudio se pretende realizar un análisis descriptivo y estadístico mediante la simulación de escenarios tradicionales y otros estresados que brinden la inteligencia a un algoritmo heurístico multi-objetivo concebido desde la teoría de juegos, desarrollado con procesos de búsqueda local y evaluado mediante principios

económicos de igualdad, que sea capaz de generar el diseño óptimo de un sistema y a la vez ser una base conceptual para la toma de decisiones.

## Introducción

---

Los centros de servicio con atención al cliente viven día a día problemas de saturación, al momento de ir a la empresa en busca de algún servicio. Tanto su concepción como rediseño implica un estudio basado en medidas de desempeño que se ven afectadas constantemente por muchas variables dentro y fuera del proceso.

Cada una de estas variables impacta de alguna forma el modo como el sistema se comporta y cómo éste es capaz de satisfacer las necesidades del cliente eficientemente. Especialmente el recurso humano es el factor clave en un buen diseño, ya que tiene mayor impacto en las medidas de evaluación y por lo tanto representan el mejor punto de partida para aplicar procesos de mejora, su estudio y análisis solo es posible mediante pruebas estadísticas que evidencien su efecto en los indicadores alrededor de las interacciones con los demás factores que inciden en un modelo de colas.

Este estudio se encarga de realizar dicho análisis combinando diferentes tipos de factores como tasas de atención y llegada, distribuciones de probabilidad, longitud de una de una red en serie, número de servidores y capacidad del sistema para modelar problemas de colas en diferentes aplicaciones y con resultados significativamente diferentes.

Además a partir del análisis obtenido, es posible construir un algoritmo de optimización que involucre el conocimiento de los factores y su influencia en los indicadores, de tal manera que sea inteligente en la escogencia del mejor conjunto de soluciones para el problema, es decir, la cantidad de servidores y la capacidad necesaria para satisfacer las necesidades del cliente. Las técnicas de optimización multi-objetivo resultan muy útiles al momento de minimizar dos objetivos opuestos en donde no existen parámetros que optimicen ambos objetivos simultáneamente.

La inclusión de técnicas conocidas en la investigación de operaciones a un algoritmo de optimización concebido desde la teoría de juegos, desarrollado con procesos de búsqueda local y evaluado mediante principios económicos de igualdad, brinda una herramienta de análisis poco tradicional e innovadora con diferentes conclusiones acerca del estado óptimo

de un sistema. De esta manera se concibe un estudio de investigación y análisis de redes de colas que brinde criterios de mejora y eficiencia en la asignación eficiente del recurso humano

## Objetivos

### Objetivo General

Analizar el comportamiento y desempeño de línea de espera en procesos de servicio para que minimizando los costos asociados con la espera y los tiempos de permanencia en el sistema se asigne eficientemente el recurso humano en organizaciones con estructuras de este tipo.

### Objetivos Específicos

- Definir una red de colas sobre la cual sea posible realizar modificaciones de ciertos factores de las líneas de espera, como tasas de llegada al servicio, distribuciones de probabilidad de los tiempos entre llegadas, cantidad de servidores disponibles y carga de trabajo, entre otros.
- Determinar vía simulación el desempeño de los posibles sistemas que se pueden generar en la red de colas al combinar los tratamientos definidos en el experimento diseñado y evaluar cómo la variación de dichos parámetros en el sistema impacta su eficiencia.
- Construir una herramienta computacional orientada a organizaciones del sector servicios que determine el desempeño del sistema a partir de los resultados de la simulación y asigne el recurso en la red a partir de la solución del modelo de optimización.

## Generalidades

### Acerca de la empresa

CELL TUNNING es una empresa que se encarga de la venta de accesorios para móvil y soporte técnico, como empresa pretende dar un buen servicio que satisfaga las expectativas del cliente.

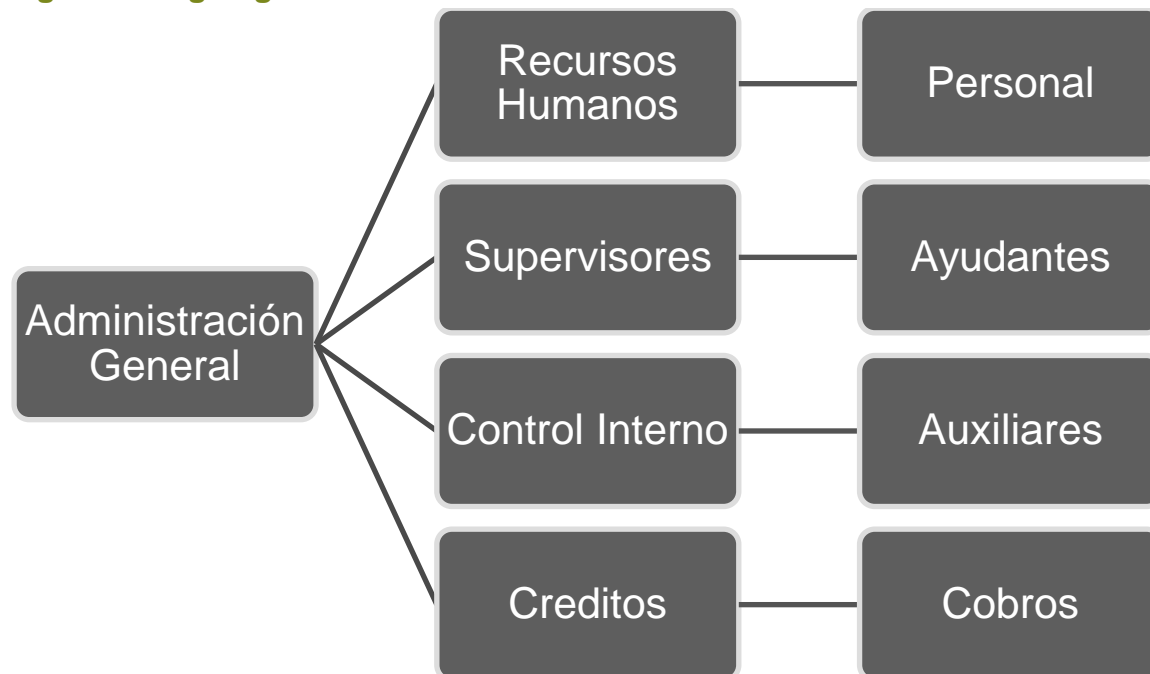
Esta empresa se encuentra ubicada en la calle 13 de Manta.

### Diversificación de productos

Entre los productos que se manejan en la empresa está toda la variedad que posee los accesorios para los diferentes dispositivos móviles, además de dar al cliente soporte técnico.

### Organigrama

Figura 1. Organigrama EMPRESA CELL TUNNING



## Marco teórico

### PROCESOS ESTOCÁSTICOS

Un proceso estocástico se define como una colección indexada de variables aleatorias  $\{X_t\}$ , donde el índice  $t$  toma valores de un conjunto  $T$  dado. Con frecuencia  $T$  se toma como el conjunto de enteros no negativos y  $X_t$  representa una característica de interés medible en el tiempo  $t$ . Por ejemplo,  $X_t$  puede representar el número de personas en un sistema al final del periodo  $t$ .

El interés de los procesos estocásticos es describir el comportamiento de un sistema en operación durante algunos periodos. Para saber su comportamiento, existen muchas aplicaciones y es usado también como sustento para muchas teorías probabilísticas una de ellas es la teoría de colas.

### TEORÍA DE COLAS

Muchas industrias de servicios tienen un sistema de colas, en el que los productos o clientes llegan a una estación y esperan en una fila, obtienen algún tipo de servicio y luego salen del sistema. Este tipo de fenómeno se origina cuando los usuarios de un determinado servicio llegan con mayor rapidez a la que este tiene capacidad de despachar, y por tal motivo se acumula personas u objetos que deben esperar para ser atendidos.

La teoría de colas es el estudio de dicha espera en las sus diferentes formas, usando los modelos de colas se representan los sistemas de líneas de espera que surgen en la práctica. Las formulas para cada modelo indican cual debería ser desempeño del sistema correspondiente y señalan la cantidad esperada de tiempo y personas en una cola, en una gama de circunstancias.

El carácter aleatorio de los fenómenos de colas implica evidentemente que el arma principal de este análisis es el cálculo de probabilidades. El conjunto de formulaciones y relaciones que usa a los datos a una determinada distribución probabilística constituye un modelo matemático que se denomina proceso estocástico en el cual una o varias magnitudes varían en forma aleatoria en función del tiempo. Estructura básica de los modelos de colas.

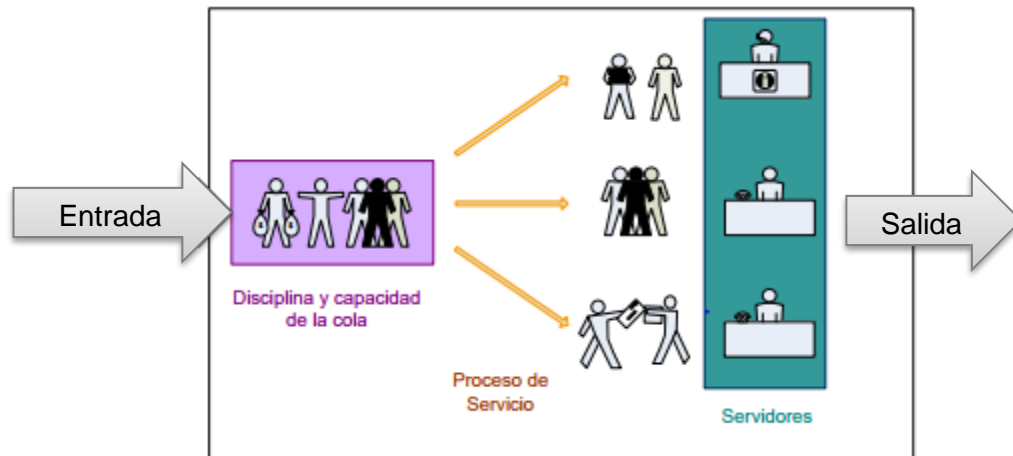
Una línea de espera está constituida por un cliente que requiere de un servicio que es proporcionado por un servidor en un determinado periodo. Los clientes entran aleatoriamente al sistema y forman una o varias colas para ser atendidos. Si el servidor está desocupado, de acuerdo a ciertas reglas preestablecidas con el nombre de disciplina del servicio, se proporciona el servicio a los elementos de la cola en un periodo determinado de tiempo, llamado tiempo de servicio y luego abandonan el sistema. Las líneas de espera se pueden clasificar de acuerdo a:

- a) El número de clientes que pueden esperar en la cola y población.
- b) Intervalo de tiempo transcurrido entre un cliente y otro.
- c) El tiempo de servicio.
- d) La disciplina de la cola. (FIFO, LIFO, etc.)



- e) El número de servidores.
- f) La estructura de las estaciones de servicio.
- g) La estabilidad del sistema.

**Figura 2. Estructura de sistemas de colas**



### Sistema de clasificación

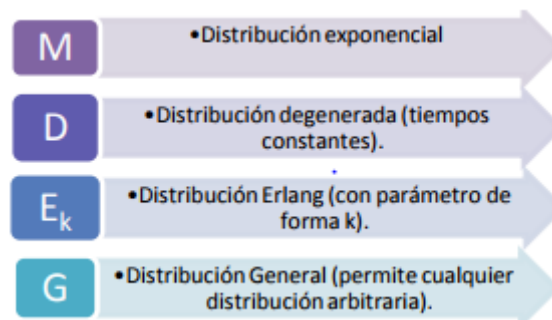
El sistema de clasificación más conocido es el de Kendall y Lee que en 1953 elaboraron con el fin de estandarizar características similares en algunos modelos y de esta forma analizarlos por igual. Esto no solo es benéfico en términos de la notación sino que abre la posibilidad de asociar cierto tipo de fórmulas ya calculadas a sistemas con características similares en el caso de las distribuciones exponenciales

**Figura 3. Clasificación Kendall y Lee**



1. Distribución de probabilidad del tiempo entre llegadas de las entidades
2. Distribución de probabilidad del tiempo de servicio
3. Número de servidores
4. Disciplina de atención
5. Número máximo de entidades que soporta el sistema en un periodo  $t$
6. Población

**Figura 4. Tipos de distribuciones**



Por ejemplo un sistema M/M/1 significa unos tiempos de llegada y atención con distribución exponencial y un solo servidor atendiendo dicha colas; como no se especifican las demás características se asume la disciplina es general y los otros parámetros infinitos. En el caso de un G/M/S estaríamos hablando entonces de una llegada con una distribución general, atención exponencial y S servidores. Debido al complejo manejo de las distribuciones no exponenciales en las colas para encontrar los valores a las incógnitas es necesaria la simulación.

### Medidas de desempeño

En la revisión analítica de las colas hay dos tipos de medidas de desempeño que son las más importantes y usadas, las dos medidas por lo común se expresan en términos de sus valores esperados y son:

**L** = número esperado de clientes en el sistema

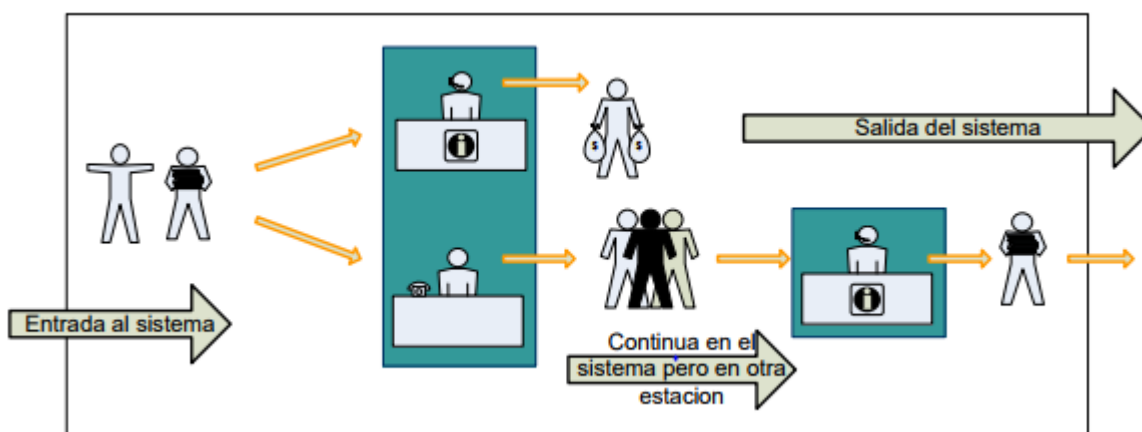
**W** = tiempo de espera esperado en el sistema

Ahora, si no se cuenta a las personas que se atienden sino solo las de la cola se tiene las mismas dos fórmulas pero con subíndice q ( $L_q$ ,  $W_q$ ).

## REDES DE COLAS

Hasta ahora se han tomado en cuenta nada mas los sistemas de colas que tienen una instalación de servicio, pero los sistemas de colas que se encuentran en los estudios de investigación de operaciones, en realidad a veces son redes de colas, es decir, redes de instalaciones de servicio en las que los clientes solicitan un servicio de algunas o todas ellas.

**Figura 5. Redes de colas**



Una red de colas es un conjunto de sistemas de colas conectados entre sí. En este tipo de estructuras los clientes llegan un nodo o punto de servicio y esperan su turno hasta ser atendidos como en un sistema de colas convencional; la diferencia es que al instante mismo de abandonar un nodo pueden o no dirigirse a otro para recibir un servicio adicional o

continuar un servicio en una dependencia diferente, esto sucede indefinidamente hasta que todos los requerimientos del servicio queden atendidos.

### Clasificación de las redes de colas

**Redes abiertas:** Reciben entrada de clientes procedentes de una o varias poblaciones externas y tienen salidas procedentes hacia el exterior. Son las usadas en centros de servicio y otras más aplicaciones.

**Redes cerradas:** no reciben entradas de poblaciones externas ni tienen salidas al exterior. Número constante de clientes circulando dentro de la red.

Figura 6. Clasificación de las redes de colas.



### Medidas de desempeño

Número medio de trabajos en el sistema,  $L_{red}$ , que es la suma de los números medios de trabajos en cada uno de los nodos:

$$L_{red} = \sum_{i=1}^K L_i$$

Tiempo medio en el sistema,  $W_{red}$ , que es el tiempo medio que pasa una tarea desde que entra en la red hasta que sale de ella:

$$W_{red} = \frac{L_{red}}{\lambda_{red}}$$

Propiedad de Equivalencia: Supóngase que una instalación de servicio tiene  $S$  servidores, un proceso de entrada Poisson con parámetro  $\lambda$ , y la misma distribución de los tiempos de servicio para cada servidor con parámetro  $\mu$ , en donde  $\rho$

## APLICACIONES DE TEORÍA DE COLAS

La teoría de colas ha tenido un lugar sobresaliente en las técnicas analíticas modernas, pero hasta aquí el enfoque se ha limitado a la formulación de una teoría matemática descriptiva. Así, no concierne en forma directa a la teoría de colas alcanzar la meta de la investigación de operaciones: la toma de decisiones óptimas. No obstante, en un insumo necesario para este fin que puede alcanzarse con la simulación y optimización.

### SIMULACIÓN

Simulación es una técnica numérica para conducir experimentos que comprenden ciertos tipos de relaciones matemáticas y lógicas, las cuales son necesarias para describir el comportamiento y la estructura de sistemas complejos del mundo real a través de largos periodos de tiempo.

Dentro de la gran variedad de paquetes de simulación disponibles en el mercado, los más conocidos y usados en la actualidad son:

**Arena:** Simulador que se usa para modelar procesos, informar y documentar. Simular la respuesta futura del sistema de cara a comprender relaciones complejas e identificar posibilidades de mejora. Visualizar las operaciones con gráficos dinámicos animados. Analizar como el sistema funciona en su configuración “tal cual” y bajo un conjunto de posibles alternativas posibles de forma que se pueda elegir con certeza la mejor de ellas.

### OPTIMIZACIÓN

Es el uso de modelos matemáticos, con el objetivo de brindar un conocimiento para la realización de la toma de decisiones. Busca mejorar el rendimiento de los procesos, para así determinar cómo se pueden maximizar o minimizar los beneficios.

#### Tipos de Optimización

Existen varios tipos de optimización usados para diferentes situaciones que por su complejidad se constituyen en un problema. Algunos de esos métodos son:

- Programación lineal: Procedimiento de resolución de problemas mediante formulación a través de ecuaciones lineales, optimizando la función objetivo teniendo que está restringida por ecuaciones lineales.
- Programación entera: Al igual que la lineal utiliza un procedimiento o algoritmo para resolver problemas. En estos modelos algunos o todas sus variables deben ser valores enteros.
- Programación dinámica: Es un enfoque general para la solución de problemas en los que es necesario tomar decisiones en etapas sucesivas. Las decisiones tomadas en una etapa condicionan la evolución futura del sistema y afectan a las situaciones en las que el sistema se encontrará en el futuro (denominadas estados), y a las decisiones que se plantearán en ese momento

- Metaheurísticas: Es un método para resolver un tipo de problema general, usando los parámetros dados por el usuario sobre unos procedimientos genéricos y abstractos de una manera que se espera eficiente. Las metaheurísticas generalmente se aplican a problemas que no tienen un algoritmo o heurística específica que dé una solución satisfactoria; o bien cuando no es posible implementar ese método óptimo.

Algunas metas-heurísticas de mayor utilización y por consiguiente más conocidas son:

- Optimización aleatoria
- Búsqueda local
- Búsqueda primero el mejor
- Optimización basada en colonias de hormigas
- Búsqueda Tabú
- Algoritmos Genéticos
- Algoritmos Matemáticos
- Inteligencia enjambre
- Búsqueda por difusión estocástica
- Optimización extrema

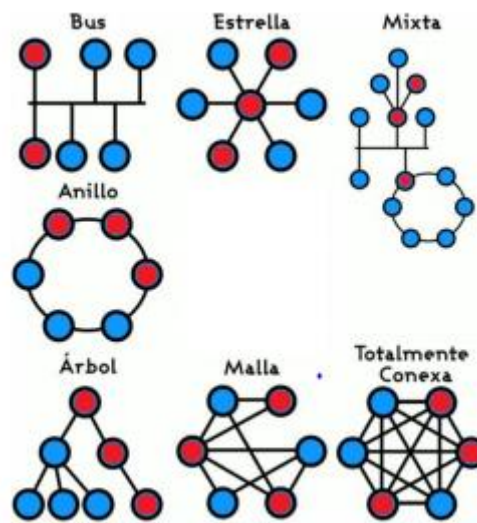
## REDES DE COLAS EN CENTROS DE SERVICIO

Los centros de servicios han evolucionado y, hoy en día, poseen características físicas, tecnológicas y procedimentales muy específicas que van más allá de lo que la teoría de colas ha intentado modelar y explicar hasta el momento. La gran mayoría de esfuerzos y hallazgos en esta teoría se han centrado en el análisis de líneas de espera sencillas, dejando a un lado el estudio de estructuras más complejas como las redes de colas, las cuales representan de manera más acertada centros de servicio actuales. En este capítulo se diseñará un experimento que permitirá analizar y determinar cómo la variación de los factores propios de las líneas de espera impacta el comportamiento y desempeño de redes de colas en procesos de servicio.

### DEFINICIÓN DE LA RED DE COLAS

Los procesos de servicio en general son estructuras complicadas cuya composición depende de los flujos de llegada o salida hacia los puntos de atención o hacia los nodos de servicio. Debido a esto, cada centro de atención dispone según su conveniencia de una distribución en forma de red para atender al cliente. Dentro de este tipo de estructuras se encuentran las redes tipo estrella, bus, cíclica, mixta, anillo, árbol, maya y totalmente conexa.

Figura 7. Tipos de redes



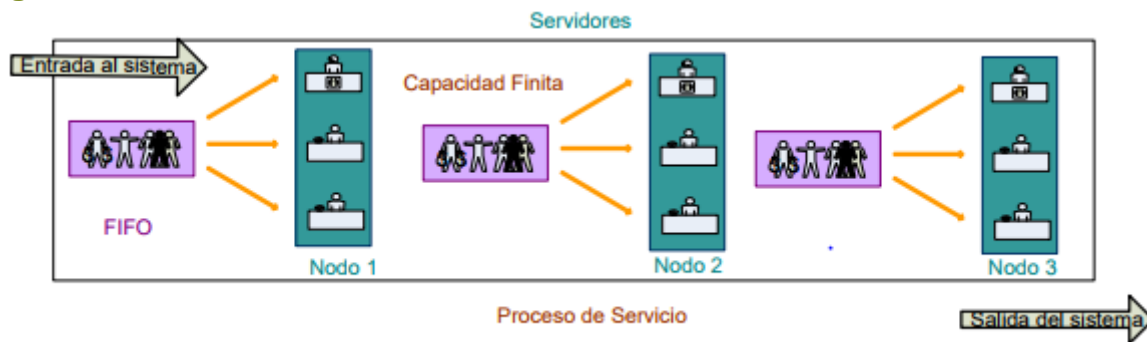
La figura muestra algunas de las estructuras en red más usadas y comunes en el sector empresarial. Cada nodo o círculo azul representa un punto de atención al interior de la red; los nodos de color rojo son puntos de atención que identifican nodos secuenciales o redes en serie al interior de las demás estructuras.

Los nodos en serie son el tipo de red más simple y ocurrente; representan una base estructural dentro de todas las demás redes y por ello la importancia que cobra su uso en el presente estudio. El uso de redes en serie es la mejor manera de construir procesos consecutivos sin limitarse a casos particulares y con la posibilidad de modelar diversas estructuras de atención, mediante procesos de expansión, unión o combinación. Debido a esto, se trabajará sobre una red en serie con cantidad de nodos constantes y sin ciclos entre uno o más nodos.

La selección de red acíclica en serie permite simplificar el funcionamiento de sistemas reales mediante el fraccionamiento de todo el proceso de atención en subprocesos en serie. De acuerdo a esto, los clientes podrán entrar por un único punto, ser atendidos en cada nodo según la capacidad de los servidores dispuestos y finalmente salir del sistema.

Adicionalmente, se asume que a dicha red no llegarán entidades de forma masiva, no habrá prioridades percibidas para la atención de entidades, la disciplina de la fila será FIFO, se generará una cola diferente en cada nodo, la cantidad de servidores cambiará por nodo y cada nodo tendrá una capacidad restringida distinta.

Figura 7. Estructura de la red de colas definida



## FACTORES DE INFLUENCIA

La distribución de probabilidad de los tiempos entre llegadas y de los tiempos de servicio, el número de servidores, la capacidad del sistema, la disciplina de la cola y el tamaño de la población o de la fuente de origen, hacen parte de dicha notación que permiten construir diferentes modelos y determinar el comportamiento de las medidas de desempeño de los sistemas mediante la variación de sus características.

### Distribuciones de probabilidad de los tiempos entre llegadas

Las redes de colas, y en general la teoría de colas, han sido estudiadas en su mayoría bajo la distribución de probabilidad exponencial debido a su facilidad matemática. Su aplicación se volvió muy popular, con las herramientas computacionales que hoy en día existen es posible realizar estudios mucho más complejos de los que en su momento Jackson y Burke concebían, cambiando el enfoque hacia nuevas distribuciones de probabilidad para modelar los tiempos entre llegadas. Aunque en el pasado la distribución exponencial resultó muy útil, en el presente trabajo su uso significa una simplificación matemática de lo que realmente ocurre en un sistema, generando resultados incongruentes en la gran mayoría de casos.

La distribución exponencial descarta la posibilidad de entradas al sistema mediante el muestreo de tiempos entre llegadas, pues es una distribución equidispersa cuya media y desviación de los tiempos de arribo es la misma, suposición que se cumple en muy pocas oportunidades en el entorno empresarial. Adicional a esto, el hecho de poseer un único parámetro no le permite modelar eficazmente el comportamiento de un sistema en operación, pues entre más parámetros contenga una distribución es más probable que el modelo pueda explicar la realidad.

Debido a lo expuesto anteriormente, se hace necesario el uso de otras distribuciones de probabilidad diferentes a la exponencial para modelar los tiempos entre llegadas, con el fin de suplir sus deficiencias y tener representaciones cada vez más aproximadas a la realidad. Se deben buscar distribuciones de probabilidad más versátiles, más robustas, con más aplicabilidad al contexto de atención empresarial que es donde se necesitan.

La distribución Híper-Erlang es una distribución de probabilidad continua con parámetros  $\lambda > 0$ ,  $l > 0$  (número de fases), y  $C \geq 1$  (parámetro de forma).

Su función de densidad es:

$$f(x) = \sum_{i=1}^I \alpha_i \frac{\lambda_i^{c_i}}{\Gamma(c_i)} x^{c_i-1} e^{-\lambda_i x}$$

Donde  $\alpha_i$  es la probabilidad de ocurrencia de cada fase ( $\sum_{i=1}^I \alpha_i = 1, \alpha_i \geq 0$ )

Su función acumulada de probabilidad es:

$$F(x) = 1 - \sum_{i=1}^I \alpha_i e^{-\lambda_i x} \sum_{j=0}^{c_i-1} \frac{(\lambda_i x)^j}{j!}$$

La función generadora de momentos de una variable aleatoria X con distribución HíperErlang es:

$$E(X^n) = \sum_{i=1}^I \alpha_i \lambda_i^{-n} \frac{\Gamma(c_i + n)}{\Gamma(c_i)}$$

### Distribuciones de probabilidad de tiempos entre llegadas a contrastar

De acuerdo a lo anterior, se van a contrastar dos distribuciones de probabilidad para modelar los tiempos entre llegadas: la distribución exponencial y la distribución HíperErlang. Para hacer comparables sus resultados es necesario tomar medidas equivalentes entre ellas y estimarse parámetros semejantes de las distribuciones de probabilidad seleccionadas. Para ello se supone que el parámetro lambda de la exponencial y de la Híper-Erlang es el mismo.

El Lambda o tasa de llegadas al sistema es el número esperado de llegadas por unidad de tiempo y se determinará a partir de la fórmula de utilización del sistema de la siguiente manera:

$$\rho = \frac{E(X)}{E(t)}$$

- Para la Distribución Exponencial

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

$$\lambda = \rho s \mu$$

- Para la Híper-Erlang



$$\rho = \frac{1/s\mu}{Ic/\lambda}$$

$$\lambda = \rho cs\mu$$

En donde  $\mu$  es la tasa de atención,  $s$  el número de servidores,  $c$  el orden de la HíperErlang e  $I$  el número de sus fases.

La distribución de probabilidad Híper-Erlang posee un parámetro de forma que debe ser estimado, así como un número de fases y una probabilidad de ocurrencia de cada una de éstas. Para fijarle un valor a estos parámetros, se toma el valor esperado de la distribución de probabilidad y se suponen lambdas iguales para cada una de las fases, como se muestra a continuación:

$$E[X] = \sum_{i=1}^I \alpha_i \lambda_i^{-1} \frac{(C_i - 1 + 1)!}{(C_i - 1)!}$$

$$E[X] = \alpha_i \lambda^{-1} \frac{(C)!}{(C-1)!} + \alpha_{i+1} \lambda^{-1} \frac{(C)!}{(C-1)!}$$

$$E[X] = (\alpha_i \lambda^{-1} + \alpha_{i+1} \lambda^{-1}) \frac{(C)!}{(C-1)!}$$

$$E[X] = \left(\frac{1}{\lambda}\right) \frac{(C)!}{(C-1)!}$$

$$E[X] = \frac{C}{\lambda}$$

Como resultado de este procedimiento se obtiene que si los parámetros en cada fase de la distribución Híper-Erlang son iguales y  $C=1$ , se generará una distribución exponencial; y si  $C=2$  se tendrá una distribución Erlang. Esto evidencia, que a partir de valores iguales o superiores a 3 se tendrían variables aleatorias Híper-Erlang.

Tomando como referencia este hallazgo, se fijan los parámetros faltantes de la siguiente forma:

- $C$  (orden de la distribución Erlang) = 3. A partir de este valor se garantiza la existencia de una distribución Híper-Erlang.
- $I$  (número de fases) = 2. Con este parámetro se puede ver la interacción entre las fases y aprovechar la propiedad tipo fase, la cual permite determinar un modelamiento con mayor precisión y con un rango más amplio de sucesos
- $\text{Alfa} = 1/I = 1/2$ . De esta manera se asigna la misma probabilidad de ocurrencia para ambas fases.

En conclusión, se contrastará el desempeño del sistema con tiempos de llegada exponenciales con parámetro lambda; y tiempos de llegada Híper-Erlang con parámetros lambda, C=3 y dos fases con la misma probabilidad de ocurrencia ( $\frac{1}{2}$ ).

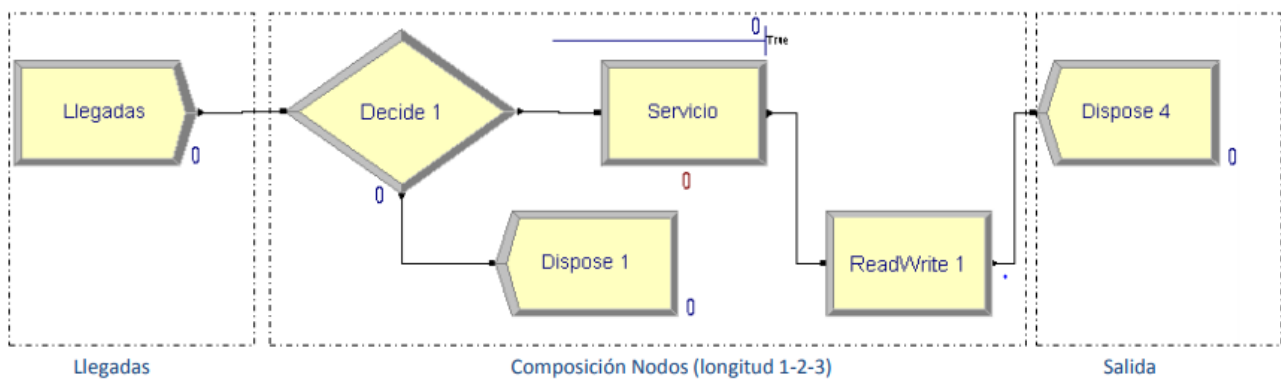
**Tabla 1. Resumen parámetros de distribuciones de llegada**

Distribución	Exponencial	Híper-Erlang
Lambda	$\lambda = \rho s \mu$	$\lambda = I \rho c s \mu$
Orden (forma)	1	C = 3
Fases	1	I = 2
Probabilidad (fases)	1	$\alpha = \frac{1}{2}$

## Modelos de simulación

Cada modelo de simulación está compuesto por la llegada, los nodos según la longitud del modelo (1,2 y 3) y finalmente una salida

**Figura 8. Modelos de simulación.**



## RESULTADOS DEL EXPERIMENTO

Después de la realización de la simulación y el cálculo de los indicadores de desempeño de los modelos planteados, es necesario realizar un análisis de los sistemas y determinar cómo la variación de las características de la red de colas aquí consideradas permite lograr mejores resultados. Con este análisis se podrán vislumbrar las diferencias principales entre las dos distribuciones seleccionadas y determinar el cambio y los efectos entre los escenarios estudiados.

Para este análisis se procede a aplicar pruebas estadísticas con la ayuda del programa SPSS sobre la información consolidada de los 49.393 modelos del experimento. Por medio de este análisis se podrá comprobar las relaciones evidenciadas mediante un análisis descriptivo.

Dentro de las pruebas a realizarse esta la homogeneidad de varianzas que mide la influencia de las variables categóricas (factores) sobre la variable dependiente cuantitativa<sup>101</sup>, en este caso sobre  $L_q$  y  $W_q$ ; y valora los efectos principales de los factores e interrelaciones de éstos entre sí<sup>102</sup>. Adicional a esta prueba, se complementa el análisis con un estudio de los conjuntos homogéneos y los intervalos de confianza entre las interacciones de los factores.

A continuación se presenta el análisis de la simulación, el cual busca determinar mediante procedimientos estadísticos cómo se comporta un modelo de colas en red según los factores expuestos en el diseño del experimento y extraer de sus resultados información clave en busca de mejorar la eficiencia general de los sistemas aquí desarrollados.

### GENERALIDADES: INTERACCIONES ENTRE LOS FACTORES

Con los resultados es posible ver la diferencia entre las medias marginales de los factores y dentro de ellas ver la relaciones individuales (primer grado) y combinadas (segundo y tercer grado) de los factores sobre los índices de desempeño. Resultados inferiores al 5% del nivel de significancia o p value confirmará la interrelación de factores.

Dentro de dicha interacción las relaciones de segundo orden para  $L_q$  y  $W_q$  son:

- Factor de utilización y número de servidores (únicamente para  $W_q$ )
- Factor de utilización y capacidad del sistema
- Número de servidores y capacidad del sistema
- Distribución de los tiempos entre llegadas y factor de utilización
- Distribución de los tiempos entre llegadas y capacidad del sistema
- Distribución de los tiempos entre llegadas y número de servidores

Las interacciones de tercer orden para  $L_q$  y  $W_q$  son:

- Factor de utilización, distribución y capacidad
- Factor de utilización, capacidad y servidores

Interacciones entre los factores como niveles de utilización, distribución y servidores; o nivel de utilización, distribución, servidores y capacidad no tienen una relación significativa con los resultados, esto quiere decir que si bien individualmente dichos factores impactan los resultados de las variables dependientes ( $L_q$  y  $W_q$ ), cuando se aumenta el nivel de interacción con la combinación de mas factores su efecto pierde importancia.

### ANÁLISIS DEL NÚMERO PROMEDIO DE PERSONAS EN COLA

El número esperado de personas en cola es uno de los dos indicadores de desempeño seleccionados para medir el comportamiento de las redes de colas. Dentro del diseño del experimento se señalan todas las variables que serán estudiadas por afectar la fluctuación de este indicador. A continuación, mediante la organización y estudio de la información, se realizará un análisis descriptivo del comportamiento evidenciado.

## Relación entre número esperado de personas en cola y nodo de atención

El número esperado de personas bajo parámetros de entrada y atención semejantes y capacidades infinitas tendría que ser igual en cada uno de los nodos según el teorema de Jackson y Burke en sistemas con tiempos de llegada y servicios exponenciales; sin embargo, en sistemas con capacidades finitas la tasa media de salida de un nodo a otro difiere de la tasa de entrada. Este comportamiento se presenta debido a que no todas las entidades que quisieron acceder al servicio pudieron hacerlo, y en consecuencia, abandonaron el sistema disminuyendo la tasa al nodo siguiente.

**Figura 9. Redes de colas con capacidad finita**




Esta situación hace que la tasa de llegada a los nodos siguientes en la red sea menor y el número esperado de personas en cola disminuya.

## Resumen de interacción de factores sobre el número esperado de personas en cola

**Tabla 9. Resumen relación entre número esperado de personas en cola y nodo de atención**

Relación	Distribución Exponencial	Distribución Hiper-Erlang
<div style="display: flex; align-items: center;"> <div style="margin-right: 10px;"> </div> <div> <p>Nodos</p> <p>Lq</p> </div> </div>	<p>Nodos:</p> <p>Disminución del Lq reduciendo la velocidad conforme el nodo está más alejado del origen de entrada. Variación del nodo uno a dos menor que del dos al tercer nodo.</p>	<p>Nodos:</p> <p>Disminución del Lq reduciendo la velocidad conforme el nodo está más alejado del origen de entrada. Mayor variación entre los nodos y especialmente del primer al segundo nodo</p>

**Tabla 10. Resumen relación número esperado de personas en cola y factor de utilización**

Relación	Distribución Exponencial	Distribución Híper-Erlang
	<p>Nodo 1: Aumento lineal del Lq con tendencia a acentuarse cuando el Rho se acerca a 1. Número esperado máximo promedio de personas en fila de 8.</p>	<p>Nodo 1: Aumento del Lq reduciendo la velocidad conforme el Rho se acerca a 1. Número mínimo de personas en fila 10.</p>

**Tabla 11. Resumen relación número esperado de personas en cola y número de servidores**


Relación	Distribución Exponencial	Distribución Híper-Erlang
	<p>Nodo 1: Disminución del Lq conforme existan más servidores, en especial el cambio de 1 a 2 servidores. Lq similar con Rho 0,7 de 2 a 10 servidores.</p> <p>Nodo 2 y 3: Resultados similares con menor Lq en cada uno de los puntos.</p>	<p>Nodo 1: Disminución del Lq conforme existan más servidores, en especial el cambio de 1 a 2 servidores. Se producen anomalías en la tendencia cuando el Lq está cercano a la capacidad máxima del sistema</p> <p>Nodo 2: Lq menores con aumentos a menor velocidad que en el nodo uno y aumentando conforme el Rho se acerca a 1. Ver anexo D. Figura 57, 58, 59, 60.</p> <p>Nodo 3: Lq menores a los anteriores nodos con un crecimiento totalmente lineal. Anexo C figuras 61, 62, 63.</p>

Tabla 12. Resumen relación número esperado de personas en cola y capacidad del nodo



Relación	Distribución Exponencial	Distribución Híper-Erlang
 <p>Capacidad</p> <p>Lq</p>	<p>Nodo 1: Lq aumenta conforme la capacidad lo hace, con excepción de la utilización de 0,7. El crecimiento se torna convexo en las capacidades de 20 y 30; y totalmente lineal en la capacidad de 10 personas.</p> <p>Nodo 2 y 3: Lq con aumentos lineales poco pronunciados y con tendencia a mantenerse constante. Ver Anexo D.</p>	<p>Nodo 1: Lq aumenta conforme la capacidad también lo hace según una curva logarítmica.</p> <p>Nodo 2 y 3: Comportamiento lineal de Lq en todas las capacidades con inflexiones drásticas cuando el sistema llega a su límite de atención. Ver Anexo D. figuras 53, 57 y 61.</p>

Tabla 13. Resumen Relación número esperado de personas en cola, capacidad del nodo y cantidad de servidores

Relación	Distribución Exponencial	Distribución Híper-Erlang
 <p>Lq</p> <p>Capacidad</p> <p>Servidores</p>	<p>Nodo 1: Disminución acelerada del Lq al pasar de uno a dos servidores. La capacidad genera mayor impacto a menor sea la cantidad de servidores hasta volverse indiferentes a partir de 7 servidores.</p> <p>Nodo 2 y 3: Lq con aumentos lineales poco pronunciados y con tendencia a mantenerse constante.</p>	<p>Nodo 1: Disminución del Lq al aumentar los servidores. La capacidad afecta el indicador menos cuando los servidores son diez.</p> <p>Nodo 2: Aumento lineal del Lq entre menos servidores existan. Las capacidades son indiferentes cuando hay 10 servidores.</p> <p>Nodo 3: Aumento acelerado del Lq entre menos servidores existan. Las capacidades con más de 7 servidores son indiferentes</p>

## Resumen de interacción de factores sobre el tiempo esperado en cola

Tabla 18. Resumen relación tiempo esperado en cola y nodo de atención


Relación	Distribución Exponencial	Distribución Híper-Erlang
	<p>Nodos:</p> <p>Disminución del <math>Wq</math> reduciendo la velocidad conforme el nodo está más alejado del origen de entrada.</p>	<p>Nodos:</p> <p>Disminución del <math>Wq</math> reduciendo la velocidad conforme el nodo está más alejado del origen de entrada. Mayor variación entre los nodos y especialmente del primer al segundo nodo</p>

Tabla 19. Resumen relación tiempo esperado en cola y el factor de utilización


Relación	Distribución Exponencial	Distribución Híper-Erlang
	<p>Nodo 1:</p> <p>Aumento lineal de <math>Wq</math> que tiende a acelerarse después de <math>\rho</math> igual a 0,8 en todos los nodos.</p> <p>Nodo 2 y 3:</p> <p>Estabilidad en los resultados del <math>Wq</math> con pequeños movimientos según el factor de utilización.</p>	<p>Nodo 1:</p> <p>Aumento del <math>Wq</math> aumentando con cambios drásticos a partir de <math>\rho</math> 0.8.</p> <p>Nodo 2 y 3:</p> <p><math>Wq</math> inferiores al primer nodo debido a la saturación de este y la caída de la tasa de salida al interior de los nodos.</p>



Tabla 20. Resumen relación tiempo esperado en cola y número de servidores



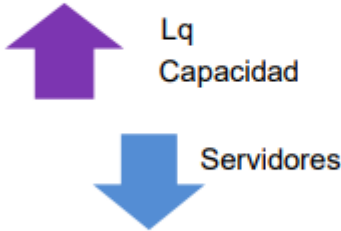
Relación	Distribución Exponencial	Distribución Híper-Erlang
	<p>Nodo 1,2,3: Disminución del Wq conforme existan más servidores, en especial el cambio de 1 a 2 servidores. Wq similar con Rho 0,7 de 2 a 10 servidores.</p> <p>Nodos 2 y 3: Resultados similares con menor Wq en cada uno de los puntos. Ver anexo E. Figuras 65, 66, 67.</p>	<p>Nodo 1: Aumento del Wq reduciendo la velocidad conforme el Rho se acerca a 1. Entre menos servidores existan el cambio del Wq es más evidente.</p> <p>Nodo 2: Wq menores con aumentos a menor velocidad que en el nodo uno y reduciéndose conforme el Rho se acerca a 1.</p> <p>Nodo 3: Wq menores a los anteriores nodos con un crecimiento lineal y comportamientos atípicos debido los resultados sobre el límite de la capacidad del sistema</p>

Tabla 21. Resumen relación tiempo esperado en cola y capacidad del nodo

Relación	Distribución Exponencial	Distribución Híper-Erlang
	<p>Nodo 1: Wq aumenta conforme la capacidad lo hace, con excepción de la utilización de 0,7. El crecimiento se torna convexo en las capacidades de 20 y 30; y totalmente lineal en la capacidad de 10 personas.</p> <p>Nodos 2 y 3: Resultados similares con menor Wq en cada uno de los puntos.</p>	<p>Nodo 1: Wq aumenta conforme la capacidad también lo hace según una curva cóncava, con inflexiones cuando el sistema llega a su capacidad máxima.</p> <p>Nodo 2: Comportamiento lineal de Wq en todas las capacidades.</p> <p>Nodo 3: Comportamiento aproximado a lineal mucho menor que en los anteriores nodos.</p>



**Tabla 22. Resumen relación número esperado de personas en cola, capacidad del nodo y cantidad de servidores**

Relación	Distribución Exponencial	Distribución Híper-Erlang
	<p><b>Nodo 1,2,3:</b> Aumento acelerado de <math>Wq</math> entre menos servidores existan. Las capacidades generan mayor impacto en el <math>Wq</math> a menor número de servidores y son indiferentes a partir de 7 servidores. La brecha del resultado de <math>Wq</math> de uno a dos servidores es mayor que la diferencia de dos a más servidores.</p> <p><b>Nodos 2 y 3:</b> Resultados similares con menor <math>Wq</math> en cada uno de los puntos. Ver anexo E. 65 – 71.</p>	<p><b>Nodo 1:</b> Aumento acelerado del <math>Wq</math> entre menos servidores existan. Las curvas que representan las capacidades son similares cuando poseen el valor de 20 y 30.</p> <p><b>Nodo 2:</b> Las capacidades son indiferentes cuando hay 10 servidores.</p> <p><b>Nodo 3:</b> Efecto similar al del nodo dos, pero con curvas menos convexas. Ver anexo E. 59, 60, 61.</p>

## ANÁLISIS GENERAL DEL SISTEMA

Hasta el momento se han analizado las colas nodo a nodo, sin embargo cobra aun mas sentido analizar el comportamiento de las variables de respuesta,  $Lq$  y  $Wq$ , en el sistema completo. Este análisis permite estudiar el efecto de los factores anidados, ya que es posible ver configuraciones de un nodo de acuerdo a un nivel diferente al de los nodos siguientes; además se puede observar la relación integral de los factores dentro de la sinergia existente en una red y comprobar su efecto e incidencia a nivel individual en los nodos y general en toda la red.

## RESUMEN MEDIDAS DE DESEMPEÑO

El análisis que hasta aquí se ha desarrollado se basa en los resultados que arrojó la simulación después de la combinación de los factores y niveles definidos para este estudio. Las distribuciones de probabilidad definidas para los tiempos entre llegadas permitieron realizar comparaciones sobre las medidas de desempeño de las interacciones del sistema, siendo siempre superiores las medidas de desempeño para los tiempos Híper-Erlang. Adicionalmente, fue posible evidenciar similitudes con respecto a los cambios, patrones y tendencias en los dos escenarios analizados, lo cual hace pensar que las conclusiones de este experimento pueden extenderse a otras distribuciones de probabilidad, ya que el comportamiento para ambas distribuciones fue homogéneo respecto al efecto de los factores.

A pesar de esto, queda abierta la posibilidad de modificar y/o adicionar otra distribución de probabilidad al experimento para evidenciar y corroborar la hipótesis aquí planteada.

## Costos de utilización del sistema

Uno de los costos más importante a la hora de definir cómo prestar un servicio son las instalaciones físicas. Ray Croc, promotor de McDonalds, afirma que la ubicación de un establecimiento no solo es importante sino que es el factor que más influye en el éxito de un negocio. Junto con el lugar se asocian muchos más costos que son derivados del mismo y que son necesarios para mantener el lugar de atención en las condiciones pertinentes y aptas para desarrollar la actividad que éste prevista.

Debido a la no caracterización de una oficina o punto en específico de atención en este estudio, se tomará un costo total mensual, que puede ser consultado en recibos, facturas etc., definido de la siguiente forma:

$$C_{tot} = C_{ar} + C_{adm} + C_{man} + C_{rh} + C_{as} + C_{im} + C_{sp}$$

Donde,  $C_{ar}$  es el costo de arrendamiento del lugar,  $C_{adm}$  representa el costo de administración,  $C_{man}$  el costo de mantenimiento,  $C_{rh}$  el costo de seguridad,  $C_{as}$  aseo,  $C_{im}$  impuestos y  $C_{sp}$  servicios públicos. Dicho costo está asociado con el número de personas posibles dentro del sistema (servidores (S) y capacidad de clientes (K) y es sensible a la variación de ellos.

Luego, para ver los costos que implica cambiar la capacidad o el número de servidores expresamos dicho costo por persona así:

$$C_u = \frac{C_{tot}}{(s+k)}$$

## Costos de espera

La espera de clientes en el sistema está asociada con el tiempo ocioso que éstos pierden mientras esperan por ser atendidos. Evaluar este costo es un factor clave al momento de cuantificar el efecto que éste tiene en la eficiencia de un negocio. Existen empresas como la Aerolínea Jet Blue que como política de calidad paga a sus clientes que esperen un bono en su próximo vuelo sin embargo, existen otras compañías en donde es muy difícil llegar a calcular dicho valor o costo debido a su naturaleza intangible.

La caracterización de estos costos dependerá de cada empresa y está definida según la función lineal expuesta por Hamdy Taha en donde el costo de espera es proporcional al tamaño de la espera.

$$C_{op} = C_o Ab$$

Donde  $C_p$  representa el costo de oportunidad que tiene un cliente que abandona el sistema, y  $Ab$  los abandonos por nodo.

### Otros Costos

Finalmente, existen otros costos particulares según la empresa que pueden ser especificados. Por lo general, estos costos dependen de los clientes y de los servidores

$$C_{ot} = C_{il}(S+K)$$

Donde  $C_{il}$  representa los costos indirectos por persona y otros costos relacionados con la prestación del servicio, por ejemplo costos indirectos de saturación de un local comercial como multas, daños materiales, desgaste de propiedad planta y equipo, etc.

## MODELO DE OPTIMIZACIÓN

Con este modelo se estará desarrollando una herramienta de apoyo para la toma de decisiones en la dirección de empresas de servicios que buscan mejorar la eficiencia y la calidad en sus sistemas de atención.

### Definición de variables

Para establecer el modelo de optimización es necesario definir todas las variables y parámetros que interactúan en el sistema y que afectan su desempeño. Dentro de la configuración de los modelos y los factores que los afectan están: las tasas de llegada y de atención de los servidores, sus distribuciones de probabilidad, la capacidad de cada uno de los nodos en el sistema.

### Funciones objetivo

Las empresas dentro de su proceso de mejora continua están atentas y ávidas de propuestas de cambio que permitan optimizar su operación, y en un centro de atención la mejor optimización esta en brindar el mejor servicio al menor costo posible.

Para cuantificar este objetivo, se proceden a formar ecuaciones que sirvan de guía para llegar a las soluciones óptimas del sistema.

El resultado es la multiplicación del costo salarial de cada servidor por el total de servidores en cada uno de los nodos, además la multiplicación de los costos de utilización por la capacidad de cada nodo y finalmente el costo de espera por el número de personas en cola en cada uno de los nodos.

$$\min Z = \underbrace{(C_s + C_{pt}) \sum_{n=1}^{N_o} S_n}_{\text{Costo de servidores y puesto de trabajo}} + \underbrace{(C_u + C_{tl})(\sum_{n=1}^{N_o} K_n + \sum_{n=1}^{N_o} S_n)}_{\text{Costo de Capacidad y otros costos}} + \underbrace{C_e \sum_{n=1}^{N_o} Lq_n}_{\text{Costo de Espera y Oportunidad}} + \underbrace{C_o \sum_{n=1}^{N_o} Ab_n}_{\text{Costo de Espera y Oportunidad}}$$

No obstante, aunque esta función objetivo es funcional no contempla situaciones en donde los costos de espera no son tangibles; de ser este el caso el resultado descontaría por completo el valor del  $Lq$ , cuyo resultado es vital en la optimización de un sistema de colas.

Lo ideal y recomendable en las compañías es realizar un ejercicio de evaluación de factores que permitan volver tangibles costos que no lo sean. A pesar de esto, si el último factor no es posible, adicional a la función objetivo actual se suma una función exclusiva para la minimización del tiempo de espera en cola que se deriva del número de personas en el sistema.

$$\min Z = \sum_{n=1}^{N_o} Lq_n$$

Esta función objetivo se utilizará independientemente de si el costo de espera es tangible o no, y busca por un lado evitar que se elimine el efecto de  $Lq$  dentro de la optimización; y en el caso de que el costo de espera sea conocido, busca incorporar un peso mayor al número de personas en cola.

## Restricciones

Las restricciones delimitan un marco de acción lógico y permisible de acuerdo con las necesidades de un problema. Para el caso de estudio éstas deben estar asociadas a las variables y el resultado de los objetivos.

### Restricciones asociadas con los servidores

La primera de estas restricciones ve encaminada a restringir la cantidad de servidores máximos dentro del sistema. Esta restricción está relacionada con el máximo costo permitido por las organizaciones para la contratación de empleados y/o por las instalaciones físicas de atención que sólo permitan un número determinado de servidores.

$$S_n < S_{tn} \quad \forall n, \quad (1)$$

$$\sum_{n=1}^{N_o} S_n < \sum_{n=1}^{N_o} S_{tn} \quad (2)$$

### Restricciones asociadas con la capacidad

La restricción de capacidad delimita un marco de referencia para evaluar hasta qué punto es posible ampliar la capacidad de un nodo.

$$K_n < K_{tn} \quad \forall n, \quad (3)$$

$$\sum_{n=1}^{N_o} K_n < \sum_{n=1}^{N_o} K_{tn} \quad (4)$$

### Restricciones asociadas con los costos

Cada negocio posee un punto de equilibrio en donde el resultado del ejercicio periódico de producir un bien o un servicio no representa pérdidas ni ganancias. Mediante el cálculo de este indicador es posible identificar cuál es el costo máximo que pueden soportar sin incurrir en pérdidas. Dicho valor está compuesto por la suma total de la nómina de los servidores, las cargas prestacionales, los aportes parafiscales, aportes a seguridad, el costo total de espera, el costo de mantenimiento del sistema, los costos del puesto de trabajo, tecnología y los otros costos referenciados.

$$C_{mo} + C_{utot} + C_w + C_{ot} + C_{wt} + C_o < C \quad (5)$$

## Modelo Final

Luego de definir las variables y parámetros, construir las funciones objetivo y las restricciones, el modelo de optimización en su versión compacta es:

<p><u>Conjuntos:</u>  <b>N</b> = Número de nodos</p> <p><u>Parámetros:</u>  <b><math>\lambda</math></b> = Tasa de llegada  <b><math>\mu</math></b> = Tasa de atención por servidor  <b><math>K_n</math></b> = Capacidad en el nodo n.  <b><math>S_t</math></b> = Número total de servidores  <b><math>N_o</math></b> = Número de nodos del sistema  <b>C</b> = Costo máximo del sistema</p> <p><u>Variables:</u>  <b><math>S_n</math></b> = Número de servidores en el nodo</p>	<p><u>Funciones objetivo:</u>  <math>\min_{\mathbb{Z}} Z =</math>  <math>(C_s + C_{pt}) \sum_{n=1}^{N_o} S_n + (C_u + C_{tl})(\sum_{n=1}^{N_o} K_n + \sum_{n=1}^{N_o} S_n) + C_e \sum_{n=1}^{N_o} Lq_n + C_o \sum_{n=1}^{N_o} Ab_n</math>  <math>\min Y = \sum_{n=1}^{N_o} Lq_n</math></p>
	<p><u>Restricciones:</u>  <math>S_n &lt; S_{tn} \quad \forall n,</math>  <math>\sum_{n=1}^{N_o} S_n &lt; \sum_{n=1}^{N_o} S_{tn}</math>  <math>K_n &lt; K_{tn} \quad \forall n,</math>  <math>\sum_{n=1}^{N_o} K_n &lt; \sum_{n=1}^{N_o} K_{tn}</math>  <math>C_{mo} + C_{utot} + C_w + C_{ot} + C_{wt} + C_o &lt; C</math>  <math>N, K, S_n, S_t, N_o \in \mathbb{Z}</math>  <math>\lambda, \mu, C \in \mathbb{R}</math></p>

## Conclusión

---

- En base a lo que aprendí al desarrollar el tema de simulación, me pude percatar que es un modelo que se emplea en muchísimas partes ya sea de la industria, comercio. La simulación como yo la había conocido era más en el ambiente laboral como en simulación de pilotos de aviones, o para los cohetes de la luna. Pero no me había puesto a pensar que en la gran mayoría lo utilizamos para análisis cuantitativos.
- En sistemas de colas con Montecarlo gobernados por tiempos entre llegadas con alta variabilidad, permite realizar un modelamiento más acertado al mostrar los datos de la simulación.
- Con el desarrollo de la simulación es posible conocer el comportamiento de las redes de colas según la combinación de los factores que influyen en ella y a partir de ese conocimiento alimentar un modelo de optimización con el ánimo de asignar eficientemente los recursos.
- La mayoría de las empresas utilizan este método para calcular sus ventas futuras, o si un nuevo producto se va a lanzar al mercado como es que su mercado reaccionario.
- Es un buen método para tomar decisiones y al desarrollarlo y concluir podemos llegar a la conclusión de que si fue una buena decisión o no.

## Bibliografía

---

1. Hamdy A. Taha **Investigación de operaciones 7ª. Edición** Editorial Pearson, México, 2004.
2. PRAWDA, Juan. **Métodos y modelos de investigación de operaciones.** (volumen 2) México: Editorial Limusa, 1984, 1023pp.

## Linkografía

---

1. <https://prezi.com/0swyzzvgxwd8/estructura-basica-de-los-modelos-de-linea-de-espera/>
2. [https://es.wikipedia.org/wiki/Teor%C3%ADa\\_de\\_colas](https://es.wikipedia.org/wiki/Teor%C3%ADa_de_colas)