UNIVERSIDAD DE LA HABANA

Facultad de Matemática y Computación

Primer Proyecto de Simulación

Servidores Especializados vs Servidores Generalistas

Autor: Lidier Robaina Caraballo

Grupo: C-411

13 de abril de 2025

Índice general

Capítulo 1

Introducción

En el ámbito de la gestión de servicios, la elección entre estrategias especializadas o generalistas representa un dilema recurrente, donde la eficiencia operativa y la experiencia del usuario dependen de la configuración de los recursos disponibles. Este proyecto aborda dicho desafío mediante la simulación computacional basada en eventos discretos, un enfoque que permite modelar sistemas dinámicos bajo condiciones controladas.

El estudio se centra en comparar dos escenarios: uno con servidores dedicados a tareas específicas (estrategia especializada) y otro en el que los servidores son flexibles y pueden atender múltiples tipos de demandas (estrategia generalista). Aunque el análisis se contextualiza en un caso sencillo de una sucursal bancaria, los resultados se orientan a ofrecer conocimientos aplicables a sistemas de servicios más amplios y complejos.

1.1 Objetivos y metas

El proyecto tiene como objetivo principal analizar y comparar el desempeño de las dos configuraciones operativas mencionadas anteriormente. Para ello, se definen las siguientes metas:

1. Calcular la eficiencia global de cada estrategia, cuantificando métricas clave como: congestión del sistema, tiempo de espera del usuario, probabilidad de retrasos críticos y subutilización de recursos.

2. Evaluar trade-off entre la flexibilidad operativa (capacidad de atender múltiples tareas) y la velocidad de servicio, considerando posibles incrementos en los tiempos de atención al adoptar una estrategia generalista.

3. Proporcionar recomendaciones basadas en datos para la optimización de sistemas de servicios, extrapolables a contextos como logística, atención al cliente o salud.

1.2 Sistema específico a simular

> Una pequeña sucursal de un banco tiene dos empleados, uno para los pagos y otro para los cobros. Los clientes llegan a cada caja siguiendo una distribución de Poisson con una media de 20/hora (el total de llegada al banco es de 40/hora). El tiempo de servicio de cada empleado es una negativa exponencial de media 2 minutos. El encargado de la sección está pensando hacer un cambio en que los dos operarios puedan hacer tanto pagos como cobros para evitar situaciones en que una cola está llena y la otra parada. Sin embargo, se estima que cuando los empleados se encarguen de las dos cosas el tiempo de servicio aumentará a una media de 2,4 minutos. Compara el sistema que se emplea ahora con el propuesto, calculando el total de gente en el banco, el tiempo medio que pasaría un cliente en el banco hasta que es atendido, la probabilidad de que un cliente espere más de cinco minutos y el tiempo medio que están parados los empleados. [1]

De la definición del problema se obtienen las siguientes variables de interés:

- **avg_clients:** Media del total de clientes en el sistema en cada instante (congestión del sistema)
- **waiting_time:** Media del tiempo durante el que un cliente permanece en la cola (tiempo de espera)
- **p_over_tk:** Probabilidad de que que el tiempo de espera sea superior a un tiempo determinado (retrasos críticos)
- **idle_time:** Media del tiempo durante el cual los servidores no tienen clientes (subutilización de recursos)

1.3 Variables que describen el problema

- $\boldsymbol{\lambda_1}$**:** frecuencia de llegada de clientes para pagos (distribución Poisson con media λ_1)
- $\boldsymbol{\lambda_2}$**:** frecuencia de llegada de clientes para cobros (distribución Poisson con media λ_2)
- $\boldsymbol{\mu_1}$**:** tiempo de atención en el servidor 1 (distribución exponencial con media μ_1)
- $\boldsymbol{\mu_2}$**:** tiempo de atención en el servidor 2 (distribución exponencial con media μ_2)
- **strategy:** especializado vs generalista (variable categórica)
- $\mathbf{t_k}$**:** tiempo de espera crítico

En caso de estrategia especializada, el servidor 1 atiende los pagos y el servidor 2 los cobros. En caso de estrategia generalista, $\mu_1 = \mu_2$.

[1] Problema 6.5 de [1]

Capítulo 2

Implementación

2.1 Detalles de implementación

1. **Gestión de Eventos:** Se emplea una cola de prioridad (módulo `heapq`) para manejar la lista cronológica de eventos. Cada evento contiene:
   - Marca temporal de ejecución
   - Tipo (llegada o salida)
   - Metadatos específicos (índice de servidor/cola)
2. **Estructuras de Datos:**
   - **Colas de espera**: Arreglos separados para cada servicio en modo especializado vs cola única compartida en modo generalista
   - **Estado de servidores**: Arreglo booleano que indica disponibilidad
   - **Contadores de clientes**: Registro separado por colas (especializado) o contador único (generalista)
3. **Mecánica de Simulación:**
   - **Llegadas**: Generadas mediante proceso Poisson usando `random.expovariate()`
   - **Tiempos de servicio**: Modelados con distribución exponencial negativa
   - **Asignación de servidores**: Política FIFO con prioridad a servidores disponibles
4. **Recolección de Métricas:**
   - *Área acumulativa* para cálculos promediados en el tiempo
   - Lista de tiempos de espera individuales
   - Registro de ocupación de servidores
   - Cálculo final mediante integración temporal (método de área bajo la curva)

2.2 Pasos de la simulación

El flujo de ejecución sigue esta secuencia lógica:

1. **Inicialización:**
   - Crear estructura de colas según estrategia
   - Programar primeros eventos de llegada usando tasas λ_1 y λ_2
   - Inicializar contadores y registros estadísticos
2. **Bucle Principal de Eventos:**
```
    while time < sim_time:
        event = heappop(events)
        actualizar_estadisticas()
        procesar_evento(event)
```
3. **Procesamiento de Llegadas:**
   - Insertar cliente en la cola correspondiente
   - Si hay servidor disponible:
     - Iniciar servicio inmediato
     - Registrar tiempo de espera cero
     - Programar evento de salida
   - Generar próxima llegada según distribución Poisson
4. **Manejo de Salidas:**
   - Liberar servidor
   - Si existen clientes en cola:
     - Extraer siguiente cliente
     - Calcular tiempo de espera (current_time - arrival_time)
     - Programar nuevo evento de salida
5. **Actualización Estadística:**
   - Calcular tiempo transcurrido desde último evento
   - Acumular:
     - Clientes-tiempo en sistema
     - Tiempo ocupado de servidores
   - Mantener precisión temporal mediante integración continua

Capítulo 3

Resultados y experimentos

3.1 Hallazgos de la simulación

Se ejecutaron 1000 simulaciones independientes para cada configuración del sistema. Para cada métrica de interés, se comparan las distribuciones resultantes de ambas estrategias mediante histogramas normalizados con su correspondiente distribución normal teórica.

3.1.1. Clientes Promedio en el Sistema

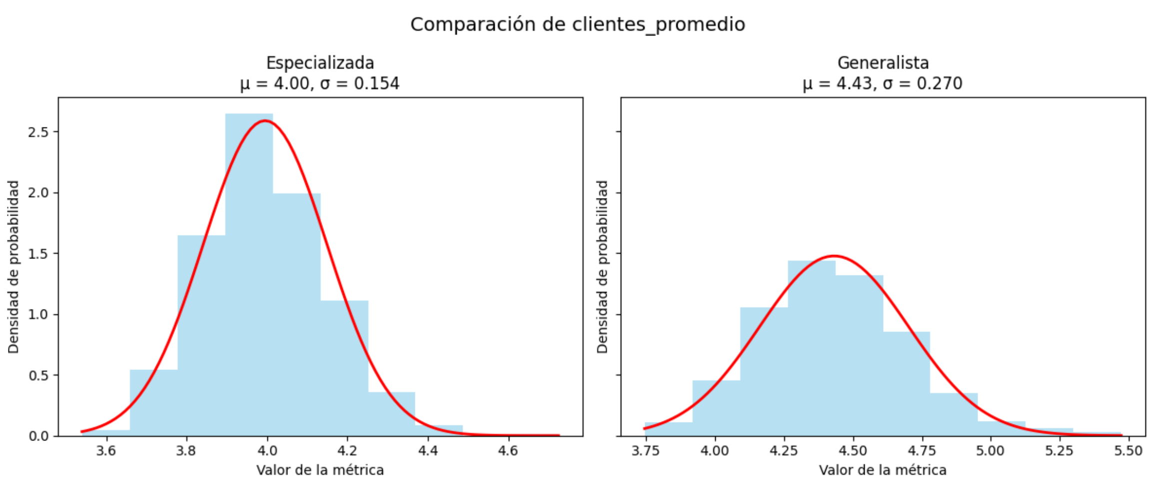

Figura 3.1: Comparación de clientes promedio en el sistema. La distribución normal teórica (roja) muestra el ajuste a los datos simulados (azul).

3.1.2. Tiempo Medio de Espera

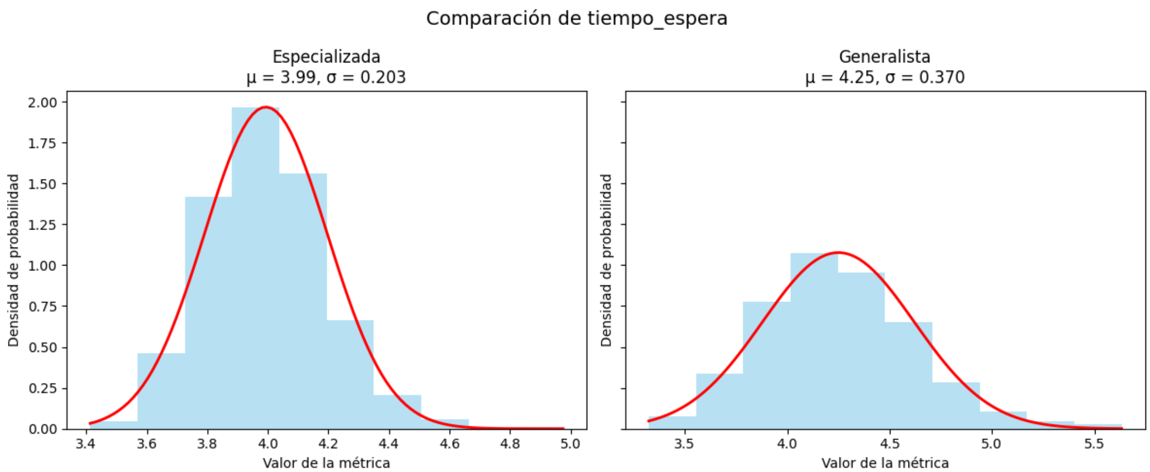

Figura 3.2: Distribución comparativa del tiempo de espera. Se observa mayor dispersión en la estrategia generalista debido al aumento del tiempo de servicio.

3.1.3. Probabilidad de Espera Crítica

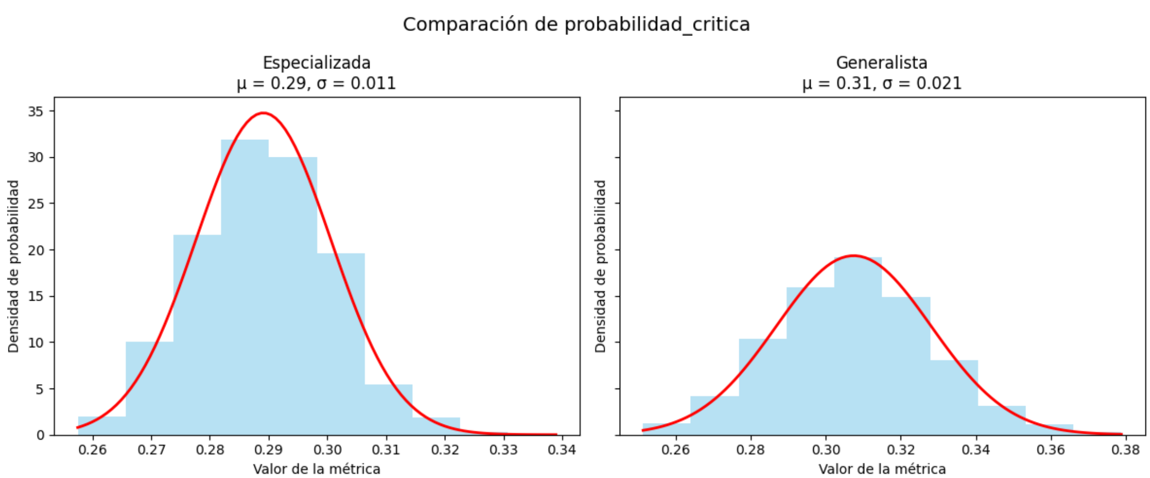

Figura 3.3: Probabilidad de superar el umbral de 5 minutos de espera. La estrategia especializada muestra menor varianza entre simulaciones.

3.1.4. Tiempo Inactivo de Servidores

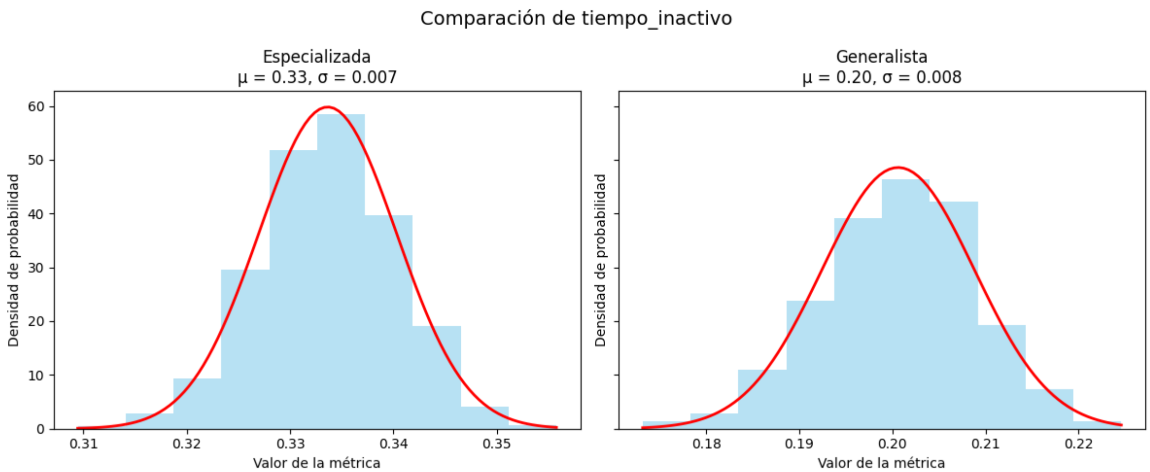

Figura 3.4: Comparación de la subutilización de recursos. La estrategia generalista mantiene menor tiempo inactivo promedio.

Capítulo 4

Modelo matemático

4.1 Modelos Probabilísticos y Estado Estacionario

4.1.1. Estado Estacionario en Teoría de Colas

El estado estacionario ocurre cuando el sistema se estabiliza con distribuciones de probabilidad constantes en el tiempo. Se define mediante:

- **Ecuaciones de Balance:** Tasa de entrada = Tasa de salida para cada estado

$$\lambda P_n = \mu P_{n+1} \quad \text{(Para M/M/1)}$$

- **Distribución Estacionaria:** Probabilidades P_n de tener n clientes en el sistema

4.1.2. Sistema Especializado (M/M/1 × 2)

Para cada caja individual:

Balance de Estados

$$\begin{aligned} \text{Estado 0:} \quad & \lambda P_0 = \mu P_1 \\ \text{Estado n:} \quad & (\lambda + \mu) P_n = \lambda P_{n-1} + \mu P_{n+1} \end{aligned}$$

Solución General

$$P_n = (1-\rho)\rho^n \quad \text{donde } \rho = \frac{\lambda}{\mu} < 1$$

Métricas Derivadas

$$
\begin{aligned}
L &= \sum_{n=0}^{\infty} nP_n = \frac{\rho}{1-\rho} \\
W &= \frac{L}{\lambda} = \frac{1}{\mu-\lambda} \\
P(W_q > t) &= \rho e^{-\mu(1-\rho)t} \\
\text{Tiempo ocioso} &= 1-\rho
\end{aligned}
$$

4.1.3. Sistema Generalista (M/M/2)

Para dos servidores compartidos:

Balance de Estados

$$
\begin{cases}
\lambda P_0 = \mu P_1 \\
(\lambda+\mu)P_1 = \lambda P_0 + 2\mu P_2 \\
(\lambda+2\mu)P_n = \lambda P_{n-1} + 2\mu P_{n+1} & n \geq 2
\end{cases}
$$

Solución General

$$
P_n = \begin{cases}
P_0 \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!} & 0 \leq n \leq 2 \\
P_0 \left(\frac{\lambda}{\mu}\right)^n \frac{1}{2^{n-2}2!} & n > 2
\end{cases}
$$

Probabilidad de Estado Vacío

$$
P_0 = \left[1 + \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^2}{2(1-\rho)}\right]^{-1} \qquad \rho = \frac{\lambda}{2\mu}
$$

Métricas Clave

$$
\begin{aligned}
L &= \frac{\lambda\mu(\lambda/\mu)^2}{(2\mu-\lambda)^2}P_0 + \frac{\lambda}{\mu} \\
W &= \frac{L}{\lambda} \\
P(W_q > t) &= \frac{(2\rho)^2}{2(1-\rho)}P_0 e^{-(2\mu-\lambda)t} \\
\text{Tiempo ocioso} &= 1-\rho
\end{aligned}
$$

4.2 Supuestos y Restricciones

- **Proceso de Poisson:** Llegadas independientes con tasa constante
- **Disciplina FIFO:** Primero en llegar, primero en ser atendido
- **Estado Estacionario:** $\rho < 1$ para convergencia
- **Servidores Homogéneos:** Misma capacidad de servicio
- **Independencia:** Tiempos entre llegadas y servicios independientes

4.3 Resultados Teóricos vs Simulación

4.3.1. Sistema Especializado

Métrica	Teórico	Simulado
Clientes totales	2.67	4.00
Tiempo sistema (min)	4.00	3.99
$P(W_q > 5 \text{ min})$	28.65 %	29 %
Tiempo ocioso	33.33 %	33 %

4.3.2. Sistema Generalista

Métrica	Teórico	Simulado
Clientes totales	3.56	4.43
Tiempo sistema (min)	5.34	4.25
$P(W_q > 5 \text{ min})$	35.12 %	31 %
Tiempo ocioso	20.00 %	20 %

4.4 Análisis de Discrepancias

- **Efectos Transitorios:** La simulación necesita tiempo para alcanzar el estado estacionario
- **Muestreo Finito:** Limitaciones en números pseudoaleatorios
- **Redondeos:** Precisión numérica en implementación
- **No idealidades:** En práctica, las llegadas pueden tener cierta correlación

Capítulo 5

Conclusiones

1. Servidores especializados provocan bloqueo parcial del sistema debido a las colas separadas, lo que lleva a una subutilización de recursos.

2. Servidores generalistas son más lentos debido a la falta de especialización, provocando una peor experiencia de usuario.

Bibliografía

[1] SABATER, J. P. G. Aplicando Teoría de Colas en Dirección de Operaciones. [S.l.]: Grupo ROGLE, Departamento de Organización de Empresas, Universidad Politécnica de Valencia, 2015/2016.