



**UNIVERSIDAD TECNOLÓGICA NACIONAL**  
FACULTAD REGIONAL ROSARIO

Cátedra: Simulación

**Trabajo Práctico Integrador**

Este trabajo práctico fue realizado por los alumnos:

Bassi, Danilo - legajo: 43725

Campitelli, Gabriel - legajo: 43677

Moreyra, Sebastian - legajo: 43684

---

Supervisado por:  
Juan Ignacio Torres

**Índice**

<b>1. Enunciado del Problema</b>	<b>2</b>
<b>2. Programación y Organización Lógica</b>	<b>2</b>
2.1. Entidades del Sistema . . . . .	2
2.2. Programa Principal . . . . .	2
<b>3. Comportamiento del sistema bajo distintas disciplinas.</b>	<b>3</b>
3.1. Disciplina FIFO . . . . .	3
3.2. Disciplina LIFO . . . . .	3
3.2.1. Cantidad promedio de clientes en cola .	3
3.3. Disciplina Prioridades . . . . .	4
<b>4. Cambios propuestos</b>	<b>5</b>
4.1. Cambio 1 . . . . .	5
4.1.1. Resultados . . . . .	5
4.2. Cambio 2 . . . . .	7
4.2.1. Resultados . . . . .	7
4.2.2. Cantidad promedio de clientes en cola .	9
4.2.3. Medidas de desempeño . . . . .	9
4.2.4. Tiempo promedio total del cliente en el sistema . . . . .	9
4.2.5. Utilización promedio del servidor . . . . .	9
<b>5. Conclusiones</b>	<b>10</b>
<b>6. Anexo</b>	<b>10</b>

# Simulación de Sistema de Espera

## Enunciado

En el siguiente trabajo práctico realizaremos una simulación de un sistema de espera compuesto por 3 colas y 5 servidores utilizando el lenguaje Python para su correspondiente programación.

### 1. Enunciado del Problema

Considere un Sistema de Espera compuesto por 2 servidores iniciales, los cuales tienen su propia cola cada uno. Luego de ser procesados en alguno de los dos servidores, los clientes pasan a una cola única en la cual esperan hasta ser atendidos por alguno de los 3 servidores restantes. Los tiempo entre arribos de clientes y los tiempos de servicio correspondientes a cada servidor son variables aleatorias IID. En la siguiente figura se representa el sistema expuesto junto con los parámetros estadísticos de las distribuciones de las variables aleatorias involucradas:

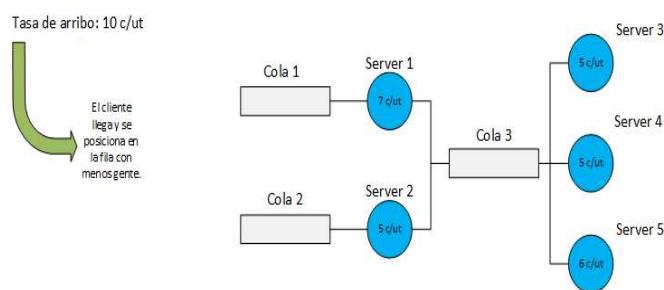


Figura 1: Sistema de Espera

Los eventos que se identifican en el sistema son:

- Arribo al Sistema.
- Partida desde Server 1.
- Partida desde Server 2.
- Partida desde Server 3.
- Partida desde Server 4.
- Partida desde Server 5.

### 2. Programación y Organización Lógica

#### 2.1. Entidades del Sistema

Para poder simular este sistema nos basamos en el paradigma de programación orientado a objetos, donde identificamos las clases que intervienen en el sistema y desarrollamos sus correspondientes métodos en base al rol que ocupan en el mismo.



Figura 2: Diagrama de Clases

#### 2.2. Programa Principal

Para llegar a cabo la simulación del sistema nuestro programa principal irá invocando una serie de rutinas. Al principio de la simulación se invocará una rutina de Inicialización, la cual se encargará de establecer las condiciones del sistema al tiempo cero. Una vez que el primer evento ocurre, avanza el reloj de simulación y se van invocando las rutinas correspondientes a cada tipo de evento, formando un ciclo hasta que se den las condiciones de fin de simulación.

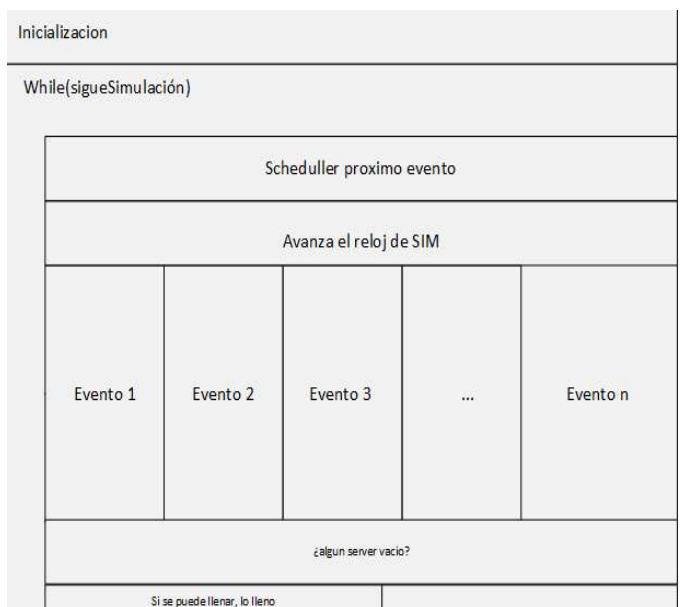


Figura 3: Diagrama de Chapin

### 3. Comportamiento del sistema bajo distintas disciplinas.

#### 3.1. Disciplina FIFO

A continuación realizaremos el análisis del comportamiento del sistema bajo la disciplina de cola FIFO.

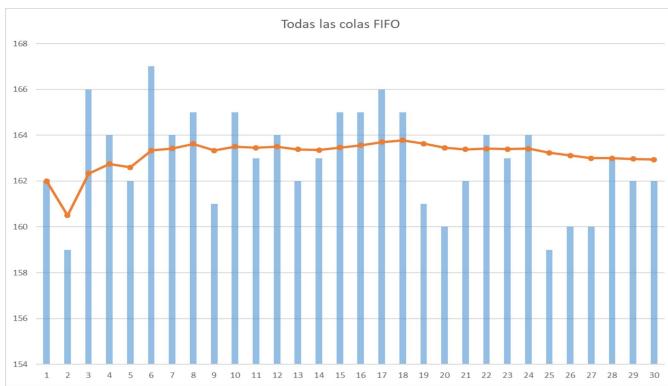


Figura 4: Clientes atendidos en cada corrida.

En el diagrama anterior se detalla la cantidad de clientes atendidos por el sistema a lo largo de 30 corridas del algoritmo de simulación. Es posible observar cómo, al utilizar números aleatorios para tanto las llegadas como para las partidas de los clientes a cada uno de los servidores, cambia la cantidad de clientes que completan servicio. Un claro ejemplo es la diferencia que se ve entre la corrida número 2 (159 clientes atendidos) y la corrida 17 (166 clientes atendidos).

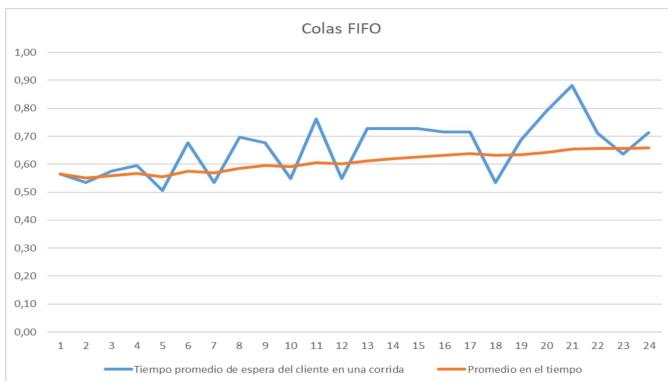


Figura 5: Promedio de espera de los clientes, disciplina FIFO.

En este segundo diagrama, es posible estudiar la "demora promedio de cada cliente en el sistema" a través de cada corrida. Podemos advertir una diferencia considerable entre la demora observada en la corrida 2 (aproximadamente 0,58 unidades de tiempo) y la corrida 21 (aproximadamente 0,89 unidades de tiempo).

Podemos también observar cómo la gráfica tiende a estabilizarse en un valor cercano a las 0,65 unidades de tiempo. Podemos decir que la media de las medias de espera de cada cliente en el sistema tiende a 0,65 unidades de tiempo, aunque para hacer una aproximación con un mayor grado de confianza

deberíamos realizar mas corridas.

#### 3.2. Disciplina LIFO

Al igual que para la disciplina FIFO, hacemos el análisis para la disciplina LIFO.

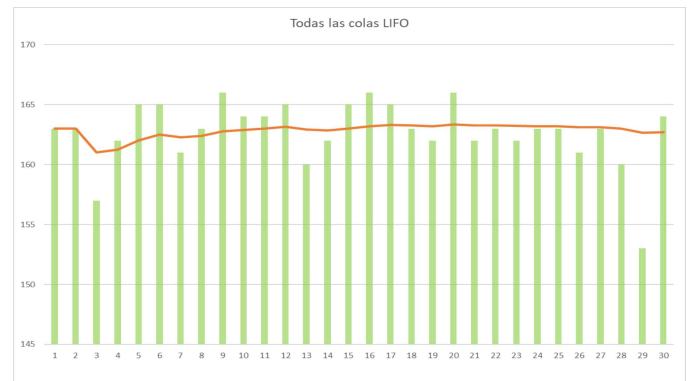


Figura 6: Clientes atendidos en cada corrida.

En este caso, vemos que la variable "cantidad de clientes atendidos" en cada corrida es mas estable que en el caso anterior, pues se mantiene alrededor del valor 163.

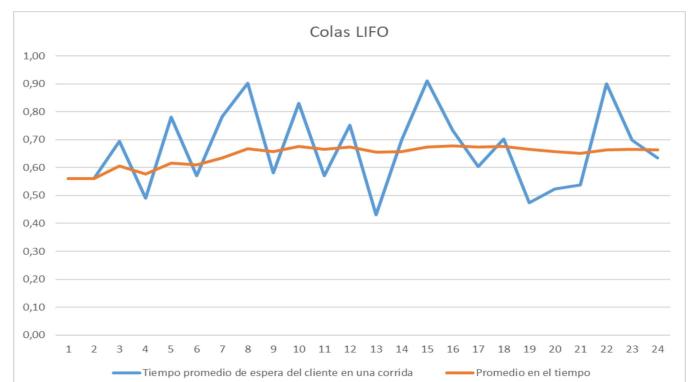


Figura 7: Promedio de espera de los clientes, disciplina LIFO.

En cuanto a la variable "demora promedio de clientes en cola", podemos observar que la variable no se comporta de manera tan estable como lo hacía en la disciplina FIFO.

La misma varía entre (aproximadamente) 0,41 y 0,90 unidades de tiempo.

#### 3.2.1. Cantidad promedio de clientes en cola

En la siguiente tabla observamos la cantidad promedio de cada cliente en cada cola:

Cola 1	1.4128606235431005
Cola 2	1.0238970387168544
Cola 3	0.1401929234074071

### 3 COMPORTAMIENTO DEL SISTEMA BAJO DISTINTAS DISCIPLINAS.

4

#### 3.3. Disciplina Prioridades

Al repetir el análisis para la disciplina de colas “Prioridades”, podemos obtener el siguiente gráfico de barras:

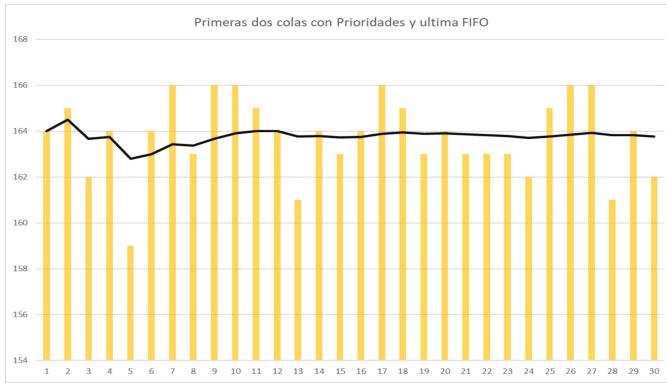


Figura 8: Clientes atendidos en cada corrida.

A partir del gráfico anterior observamos que, tal como ocurría para la disciplina LIFO, la “*cantidad de clientes atendidos*” en cada corrida se mantiene fluctuando, aunque en este caso, próxima al valor 164.

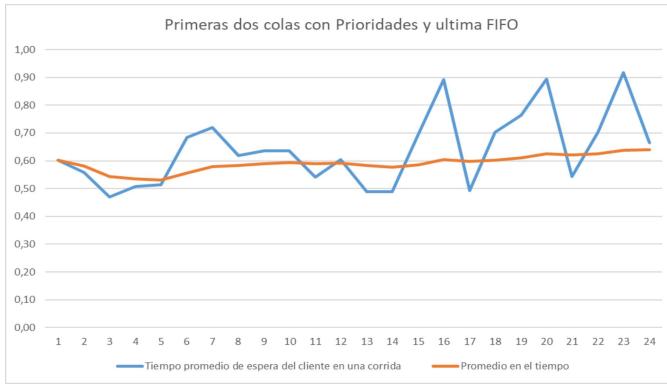


Figura 9: Promedio de espera de los clientes, Prioridades.

En cuanto a la variable “*tiempo promedio de espera del cliente*”, se observa que el valor de dicha variable oscila entre 0,50 y (aproximadamente) 0,82 unidades de tiempo.

En este caso la media de las medias de dicha variable aumenta con cada nueva corrida. Para obtener pruebas mas concluyentes sobre el valor de la variable en discusión se debe realizar un número mayor de corridas de la simulación.

En cuanto a la utilización de los servidores, en los 3 casos fue muy similar, por lo que decidimos no poner las gráficas y resumirlos en el siguiente cuadro:

Utilización promedio del Servidor 1	$\approx 84\%$
Utilización promedio del Servidor 2	$\approx 77\%$
Utilización promedio del Servidor 3	$\approx 81\%$
Utilización promedio del Servidor 4	$\approx 66\%$
Utilización promedio del Servidor 5	$\approx 55\%$

A continuación, se deja la  $B(t)$  del servidor 1 para las primeras 275 unidades de tiempo de la simulación. Por comodidad, y para una mejor obtención de los resultados estadísticos y para una correcta interpretación y comprensión de la información que se quiere exponer, decidimos no mostrar esta gráfica y sustituirla por un diagrama de bastones, en donde cada bastón en la posición  $i$  es el promedio de cada  $B(t)$  en la corrida  $i$ . Esto nos permite observar más datos en una misma gráfica, y también permite ver como converge el promedio a lo largo de las corridas.

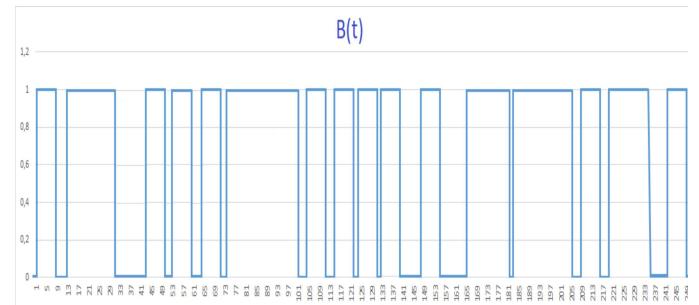


Figura 10:  $B(t)$  del servidor 1 en una corrida con todas las colas FIFO

Este tipo de gráfica será reemplazada para futuros análisis por una de el siguiente estilo:

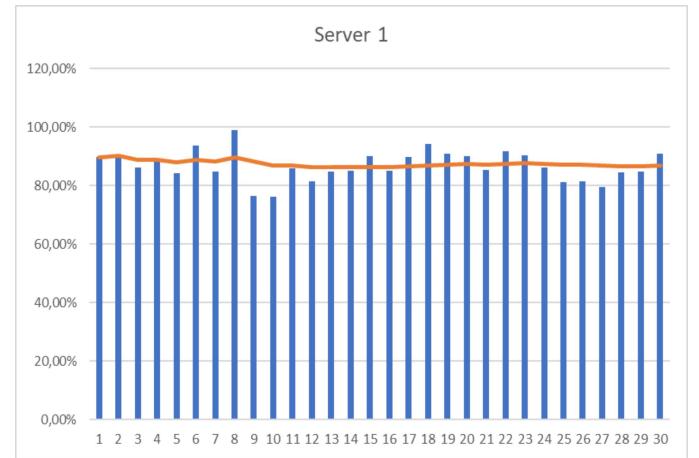


Figura 11: Gráfica que se va a utilizar para mostrar la utilizaciones de los servidores.

En donde cada bastón es el promedio de utilización de un servidor en una corrida, y la linea naranja es el promedio acumulado entre las diferentes corridas.

#### 4. Cambios propuestos

##### 4.1. Cambio 1

En base a los resultados obtenidos en la simulación 'raíz', por llamarlo de alguna manera, concluimos que el servidor 3 y 5 deberían ser cambiados de lugar. Esta conclusión se fundamenta en el hecho de que el servidor 3 es el que tiene mayor utilización, pero el servidor 5 es el que atiende mas rápido en promedio a los clientes, así que con ese pequeño cambio, se debería ver una pequeña mejora en las medidas de rendimiento.

También pensamos que podría llegar a ser beneficioso, incluir una cola exclusiva para gente con prioridad en la segunda linea de servidores que tenga un servidor exclusivo para dicha cola. Este servidor (5), atendería solo clientes con prioridad, pero en el caso de no existir gente con prioridad en dicha cola, dado que la probabilidad de que aparezca gente con prioridad es baja, atendería de manera igual a como lo venía haciendo anteriormente. Todas las políticas de las colas serían FIFO, dado que es la política más común y aceptada por los clientes.

El esquema del cambio 1 se vería en la figura 12.

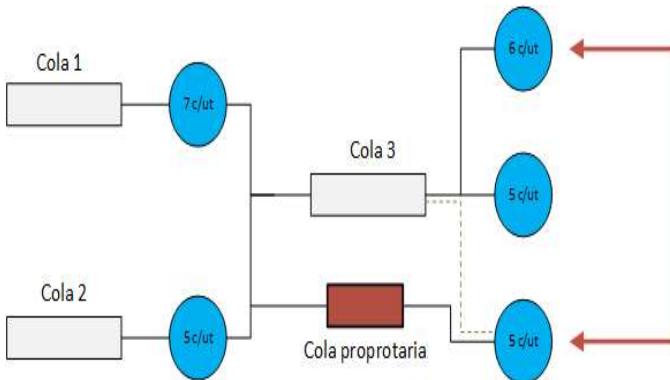


Figura 12: Esquema del cambio 1

De manera similar a lo descrito anteriormente, se procedió a realizar la codificación necesaria para adaptar el código al esquema planteado en el cambio 1, correr el algoritmo de simulación 30 veces y en base a esas corridas, se obtuvieron los resultados mostrados a continuación.

##### 4.1.1. Resultados

Al correr el algoritmo de simulación 30 veces, podemos observar que los clientes atendidos, fueron, en promedio similares a las políticas planteadas antes de realizar cambios. Ver figura 13.

Como en promedio, los clientes atendidos fueron muy similares a las otras políticas planteadas antes de hacer los cambios, el resto de conclusiones pueden ser interpretadas de la misma manera.

Si observamos la utilización de los servidores, en este caso, podemos observar que es muy similar a la utilización en la política FIFO. El servidor 1 se encuentra muy utilizado (Excediendo el 80 por ciento en muchos casos) y el servidor 3 también, aunque en este caso, el servidor 3 es un poco más rápido



Figura 13: Cantidad de clientes atendidos utilizando el esquema del cambio 1

que antes, lo cual debería verse reflejado en mejoras de rendimiento, aunque como veremos a continuación, debido a que no es significativamente más rápido, tampoco se obtiene una mejora considerable.

En la figura 15 se muestra la utilización de los 5 servidores.

Finalmente, se muestra una gráfica con el tiempo promedio en el que el cliente se encuentra en el sistema (ver fig. 14). El valor promedio de las 30 corridas fue aproximadamente 0.625 unidades de tiempo. Un valor un poco más bajo que las políticas anteriores, pero casi despreciable. Además, podría estar limitado a los distintos valores aleatorios de las 30 corridas del código. Aun así, creemos que la leve mejora se debe solamente al cambio de posición de el servidor 3 y 5 y que la cola prioritaria no influyó en los resultados finales de las medidas de rendimiento.



Figura 14: Tiempo promedio de clientes en el sistema para el Cambio 1

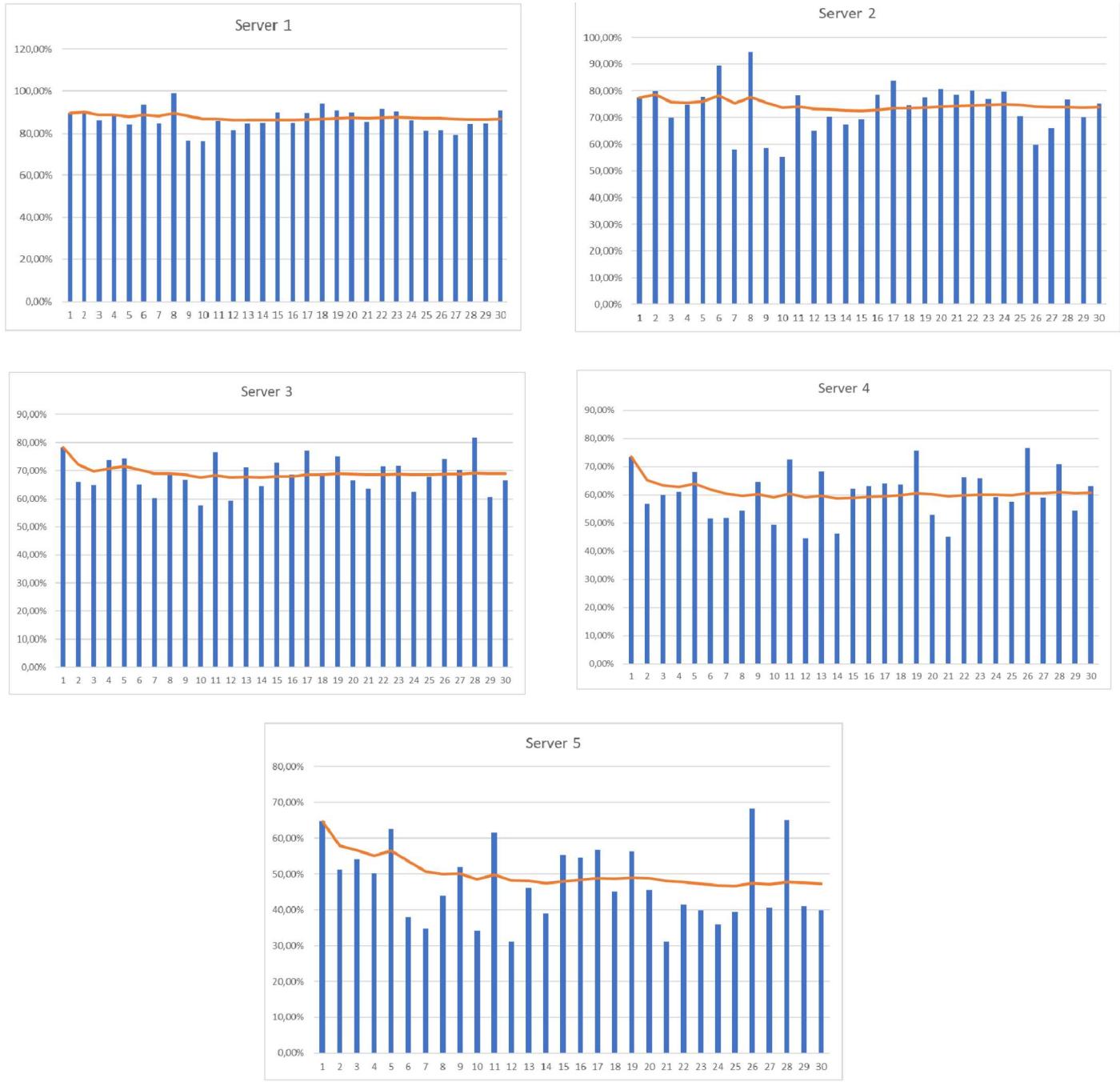


Figura 15: Tiempo promedio de clientes en el sistema para el Cambio 1. La linea naranja es el promedio de la utilización de cada servidor a lo largo de las 30 simulaciones

#### 4.2. Cambio 2

En base a lo visto en las políticas anteriores, incluyendo el cambio 1, observamos que el servidor 1 siempre se encuentra en un porcentaje mayor al 80 % de utilización. Varios autores remarcan que un servidor por encima de ese valor, indica un cuello de botella que limita la performance del sistema. Actualmente, en todas las pruebas, el servidor 1 se encuentra siempre por encima del 80 % de utilización (en algunas corridas superando el 90%) y el servidor 2, es el próximo con mayor utilización, lo que indicaría que el problema está en la primera línea de atención.

Como la política original de las dos primeras colas, es que el cliente elija siempre la cola con menos gente, y en caso de que no haya personas o que la cantidad de clientes en la cola 1 y 2 sea igual se posicione en la cola del servidor 1, es normal que este servidor siempre tenga utilización mayor al servidor 2, en promedio, la utilización del servidor 1 supera a la del servidor 2 en un 10 %.

Para intentar solventar este inconveniente, se podría haber realizado una cola única y que el servidor 1 y 2 tomen clientes de esa cola cuando se necesite, pero se tomó una perspectiva diferente, que mejora de manera mucho más significativa este inconveniente.

Actualmente se cuenta con 2 servidores en primera instancia; uno “rápido” que atiende en promedio 7 clientes por unidad de tiempo y uno “normal” que atiende 5 clientes por unidad de tiempo. Cuando realizamos la simulación, vemos que los servidores de segunda línea tienen utilización muy baja, en especial los servidores 4 y 5, lo que indicaría que la primera línea está limitando considerablemente a la segunda, haciendo que esta tenga un tiempo ocioso esperando a los servidores de primera línea. Lo que se realizó en este cambio, fue poner 3 servidores “normales” que atiendan 5 clientes/u.t cada uno en una primera línea, de esta manera, estos tres servidores van a poder atender mucho más rápido a los clientes que los 2 anteriores, y se puso a los dos servidores más rápidos en una segunda línea, de manera que estos puedan atender de manera veloz a los clientes despachados por los servidores de primera línea.

Ambas colas, tanto la de primera linea como la de segunda tienen una política de atención FIFO.

El esquema del cambio se observa en la figura 16.

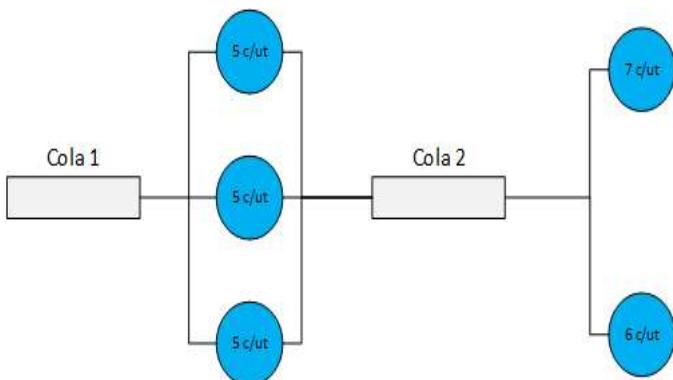


Figura 16: Tiempo promedio de clientes en el sistema

El numero de los servidores se mantienen con la misma nomenclatura de antes; 1,2 y 3 para los servidores de primera linea empezando del servidor que se encuentra en la parte de arriba y 4 y 5 para los de segunda linea, donde el servidor 4 es el que atiende 7 clientes por unidad de tiempo.

##### 4.2.1. Resultados

El primer dato que obtenemos, es el de los clientes atendidos en las 30 corridas de la simulación, en promedio, es similar a las anteriores políticas (aunque en promedio ligeramente más alto) por lo que las conclusiones del resto de datos pueden ser interpretadas de manera análoga. (ver fig. 17)

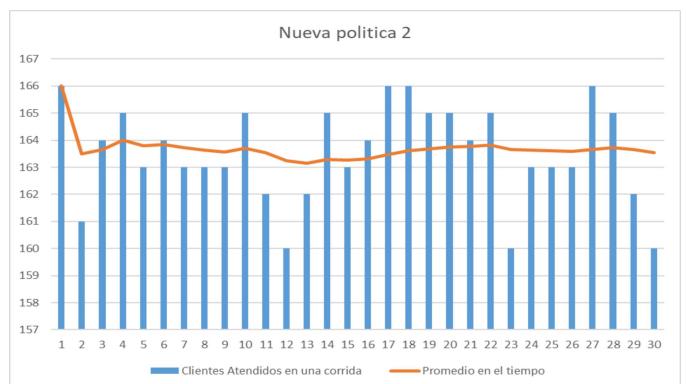


Figura 17: Tiempo promedio de clientes atendidos en el sistema con el Cambio 2

Cuando observamos la utilización de los servidores (ver fig. 19\*) podemos ver que en ningún caso los servidores exceden el 80 % de utilización en promedio. Esto significa que la utilización de los servidores se equiparó, de manera que el cuello de botella en la primera línea se eliminó, por lo que a priori, esto debería influir significativamente en los resultados de las medidas de rendimiento. Para estar seguro de esto, mostramos el tiempo promedio de los clientes en el sistema (ver fig. 18).

Recordamos que hasta ahora, el mejor resultado de este tiempo promedio es de aproximadamente 0.625 unidades de tiempo (en el cambio 1) y el peor de 0.665 (todas las colas LI-FO).

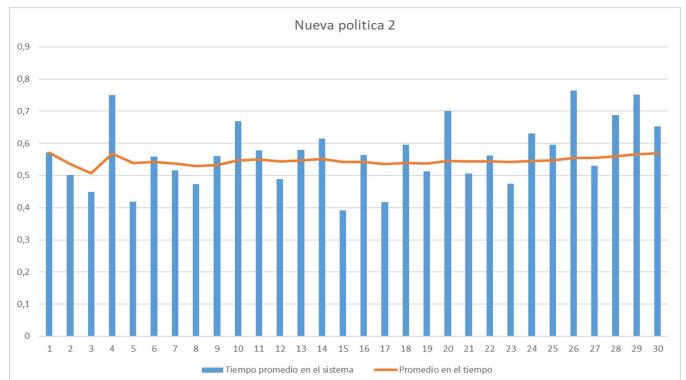


Figura 18: Tiempo promedio de clientes atendidos en el sistema con el Cambio 2

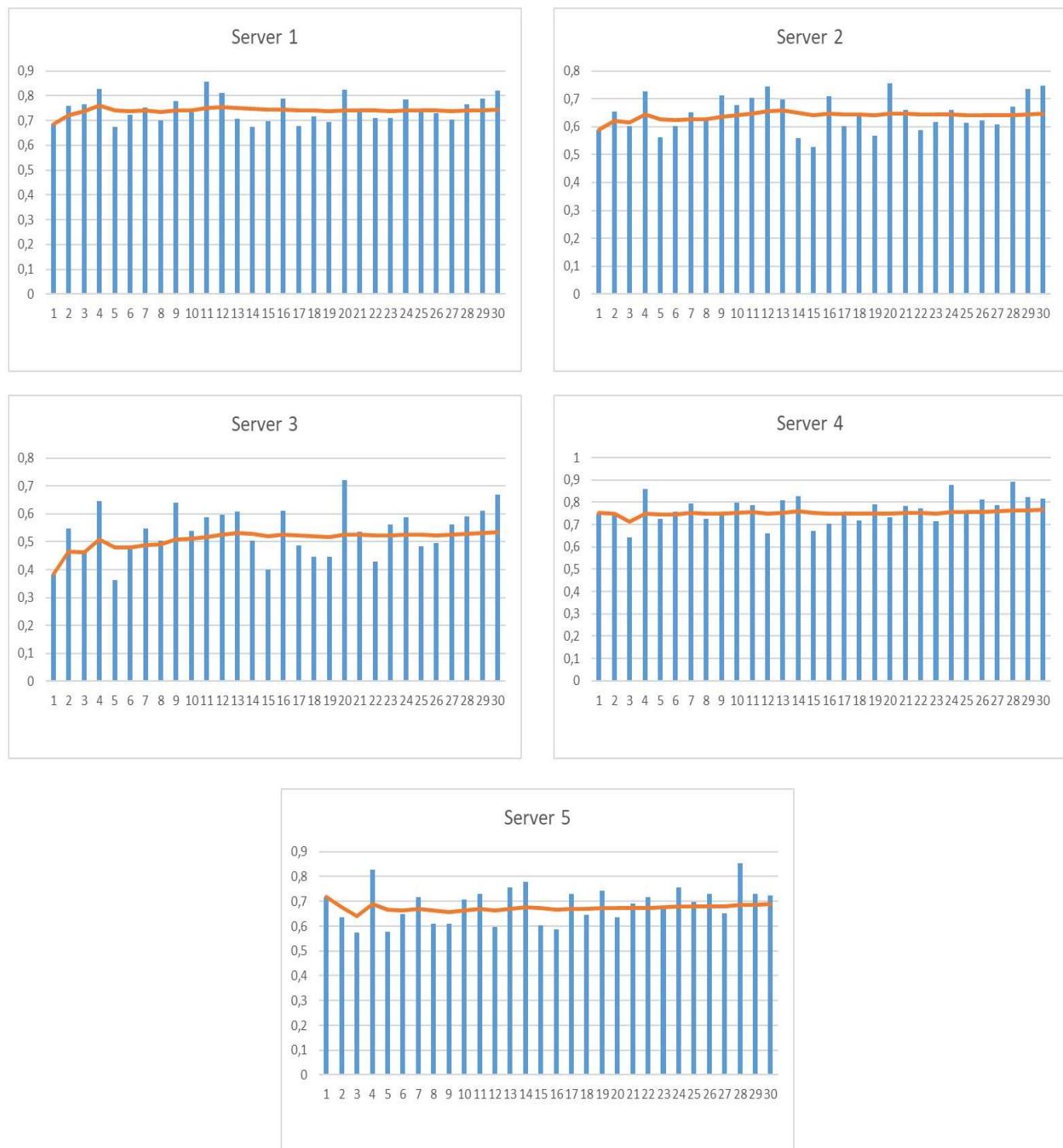


Figura 19: Tiempo promedio de utilización de los servidores en el Cambio 2

## 4 CAMBIOS PROPUESTOS

9

En las gráficas mostradas, puede observarse que el tiempo promedio del cliente en el sistema disminuyó considerablemente (0.568 unidades de tiempo) de esta manera, se mejoró mucho las medidas de rendimiento, esto es debido a que se eliminó el cuello de botella en la primera linea de atención.

### 4.2.2. Cantidad promedio de clientes en cola

Como se hizo en la sección anterior, se mostrarán en la siguiente tabla la cantidad promedio de cliente en las colas de cada servidor (en este caso en las Colas 1 y 2) en una corrida.

Cola 1	1.1916238817551392
Cola 2	1.6980385250779593

### 4.2.3. Medidas de desempeño

Para este cambio realizado pudieron obtenerse las siguientes gráficas que nos ayudan a ver el comportamiento de las medidas de desempeño y su convergencia: *tiempo promedio total del cliente en el sistema, utilización promedio de cada servidor y cantidad de clientes promedio en cola*.

### 4.2.4. Tiempo promedio total del cliente en el sistema

En la siguiente gráfica, se observa cómo, tras varias corridas, las esperas promedios convergen a un valor en común cuando alcanzan el estado estacionario.

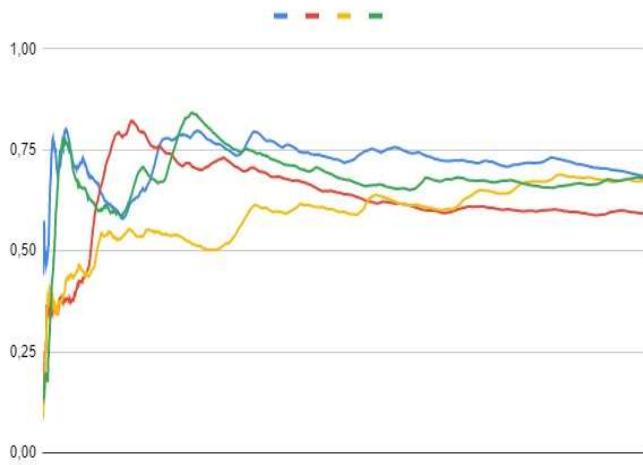


Figura 20: Espera promedio de cliente en la cola

### 4.2.5. Utilización promedio del servidor

Para verificar la convergencia de las utilizaciones promedio de los servidores, se realizó la figura 21. En ella, se muestran cómo avanza el promedio en el tiempo de la utilización del servidor a medida que avanza el reloj de la simulación en cada una de las corridas, representadas cada una por un color diferente.

En todas las gráficas que comparten el color naranja, puede observarse como es la utilización de los servidores en una corrida en particular. Precisamente en la corrida representada por este color, la llegada de clientes al sistema fue mas alta que

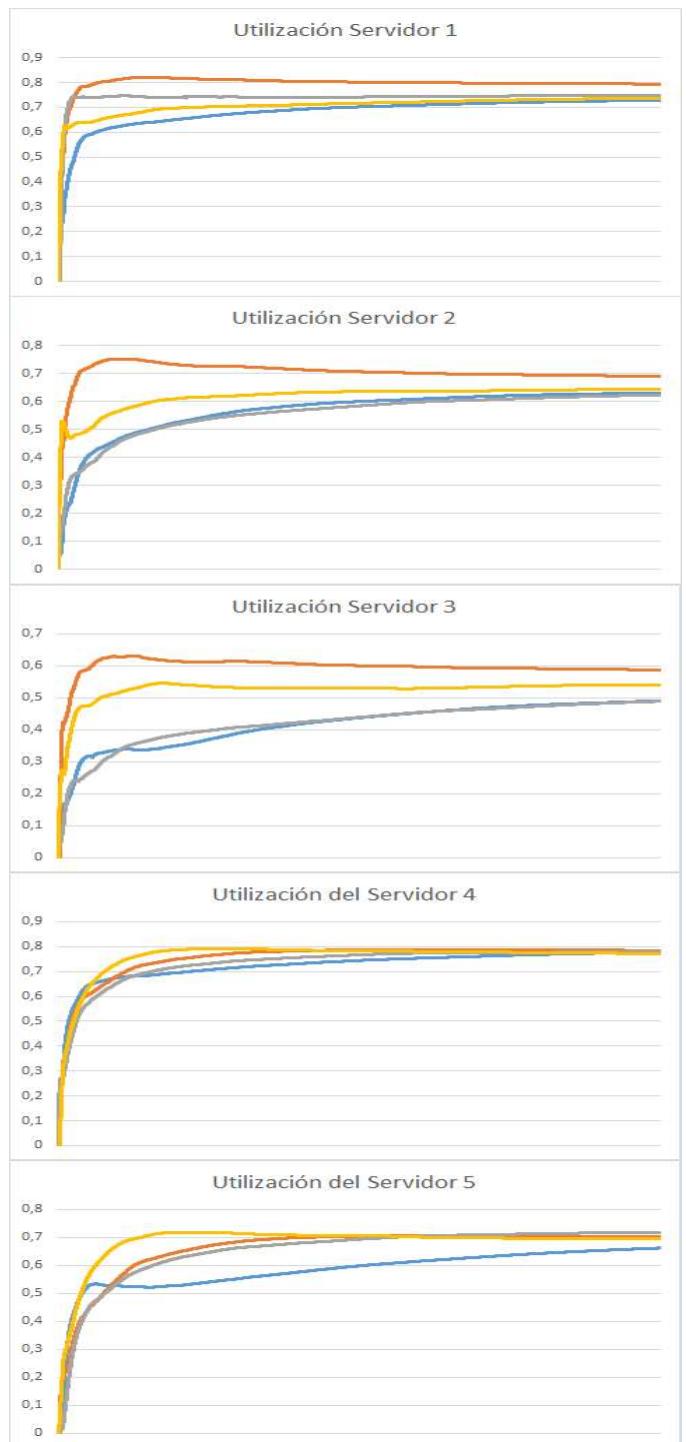


Figura 21: Utilización promedio de los servidores.

en las demás. De igual manera, el color celeste, representa la corrida que menos arribos tuvo.

Puede observarse que el servidor 3, es el que mayor holgura tiene en caso de tener pocos arribos, pero esto se debe pura y exclusivamente a como atienden los servidores en el algoritmo de simulación a los clientes en la cola; en caso de que estén todos los servidores desocupados, primero atiende el servidor 1, luego el 2 y finalmente el 3. Los valores a los que converge la

utilización promedio de los servidores, también pudo observarse en la figura 19.

## 5. Conclusiones

Según los datos expuestos, se vio que, con la primera estructura de 2 servidores en la primera línea de atención y 3 servidores en la segunda línea, se originaba un cuello de botella en los primeros servidores, de manera tal que los servidores de la segunda línea tenían un tiempo ocioso considerable. Este fenómeno se repitió para todas las políticas de atención. FIFO, LIFO, prioridades y la nueva cola prioritaria del cambio 1. Este cuello de botella fue eliminado completamente al realizar el esquema del cambio 2, en donde se pusieron 3 servidores en primera línea y 2 servidores un poco más rápidos en la segunda. En este caso, ningún servidor llegó a un 80 % de utilización promedio en las 30 corridas.

A continuación se deja una comparativa de la utilización promedio de los servidores en los diferentes casos probados.

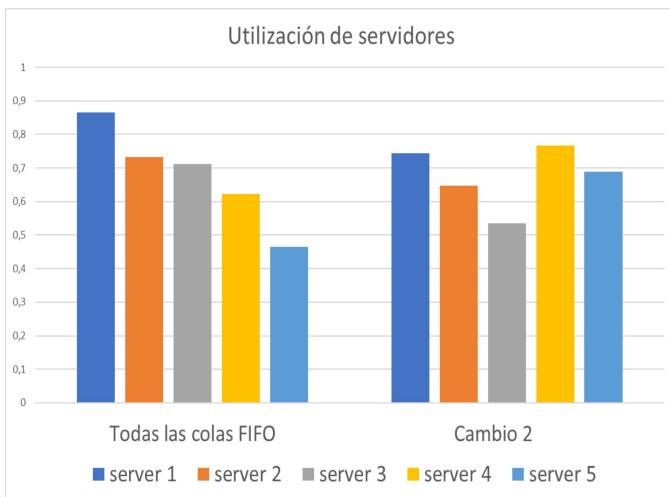


Figura 22: Utilización promedio de los servidores en FIFO vs Cambio 2

Como puede verse, se disminuyó considerablemente la utilización del servidor 1 y se equipararon las utilizaciones de los demás servidores. Esto último, influye de manera considerable en las medidas de rendimiento, como se observa a continuación, el tiempo de demora en el sistema disminuyó considerablemente. (fig. 23 ).

En el mejor de los casos anteriores, el tiempo que tarda el cliente desde que entra hasta que sale del sistema se ve disminuido en un 10 % aproximadamente (comparado con el cambio 1) y en el peor de los casos anteriores, el rendimiento se ve mejorado al disminuir en un 14.4 % el tiempo promedio en el sistema.

En base a los datos expuestos, concluimos que el cambio 2 es la mejor propuesta y es la que debería implementarse, no solo por el hecho de disminuir la cantidad de clientes en el sistema y el tiempo promedio del cliente en el sistema, sino porque también equilibra la utilización de los servidores, eliminando el

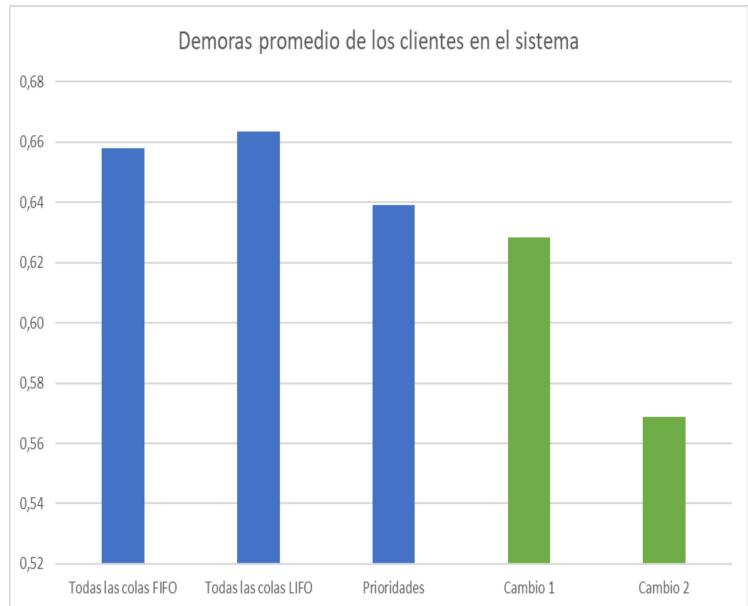


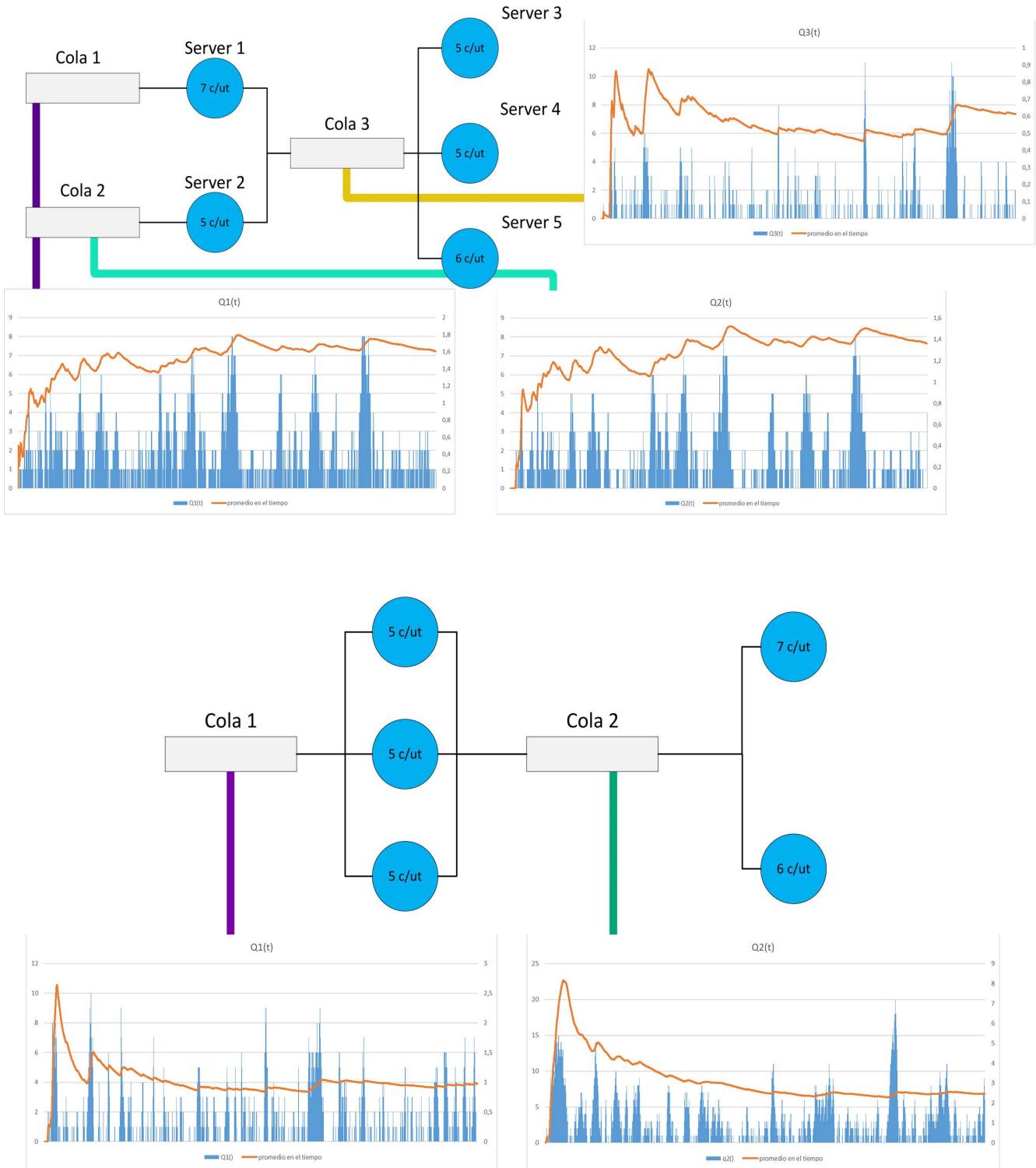
Figura 23: Tiempo promedio de los clientes en el sistema con todos los casos de prueba

tiempo ocioso producido por el cuello de botella inicial en la primera instancia.

## 6. Anexo

Se dejan como anexo a este trabajo práctico, tanto el link a GitHub del código utilizado para realizar las simulaciones así como también una gráfica que muestra  $Q(t)$  para cada una de las colas en el sistema original, y para el Cambio 2. (ver fig. 24)

<https://github.com/elcurco8/Simulacion-TP-Integrador>

Figura 24:  $Q(t)$  para cada una de las colas