

PROGRAMA PARA LA SOLUCION DE MODELOS DE COLAS

Mauricio Zúñiga Castañeda
Departamento de Ingeniería de Sistemas y Computación
Facultad de Ingeniería
Universidad de Los Andes
Apartado Aéreo 4976, Bogotá

Fernando Palacios Gómez
Departamento de Ingeniería Industrial
Facultad de Ingeniería
Universidad de Los Andes
Apartado Aéreo 4976, Bogotá
Agosto de 1987

RESUMEN

Los modelos de colas han sido utilizados exitosamente para analizar el desempeño de sistemas de computación y de comunicaciones, como función de su topología, de las distribuciones y disciplinas de servicio de los dispositivos, y del volumen y variedad de su carga. Por esta razón, para llevar a cabo el diseño y evaluación de ese tipo de sistemas, es útil una herramienta de solución de modelos de colas.

Este documento presenta un programa desarrollado para solucionar modelos generales de redes de colas, el cual utiliza la técnica de agregación, el análisis del valor medio y los resultados para las colas M/M/m y M/G/∞. Estas técnicas dan resultados exactos cuando el modelo es de tipo BCMP. Para modelos generales dan soluciones que pueden ser aproximadas, con un error generalmente pequeño.

El programa puede ser utilizado en los microcomputadores de la familia IBM y sus compatibles. Además el programa es fácilmente transportable pues ha sido desarrollado en el lenguaje C estándar.

Este documento también muestra un ejemplo de solución de modelos de sistemas de computación utilizando el programa presentado, estableciendo que el diseño y evaluación de este tipo de sistemas puede facilitarse con el uso del programa.

1. INTRODUCCION

Los modelos de colas han sido utilizados exitosamente para analizar sistemas de computación y de comunicaciones (Baske72), (Buzen78), (Klein76), (Kobay78), (Laven83), (Reise76) (Sauer80), (Sauer81) y (Wong78). El análisis con modelos de colas permite establacer el desempeño del sistema como función de su topología, de las distribuciones y disciplinas de servicio de los dispositivos, y del volumen y variedad de su carga.

Por esta razón, para llevar a cabo el diseño y evaluación de sistemas de gran complejidad, es altamente útil una herramienta de análisis de modelos de colas, la cual permite establecer la conveniencia de configuraciones arbitrarias y la eficiencia del sistema en función de sus características.

Este documento presenta un programa desarrollado para el análisis de sistemas de computación y de comunicaciones basados en modelos de colas. Las características de los modelos que soluciona son las siguientes:

- [] Pueden ser modelos abiertos y cerrados
- [] Para los modelos abiertos la tasa de llegada de clientes es Poisson e independiente del estado del modelo
- [] Puede haber cualquier número de centros de servicio
- [] La disciplina de los centros de servicio es FCFS, PS, IS o LCFSPR
- [] Los centros de servicio pueden estar arbitrariamente interconectados por la trayectoria de los saltos de los clientes. Las probabilidades de salto entre los centros son independientes del estado del modelo
- [] En cada centro de servicio puede haber un número arbitrario de clases de clientes. Cada clase puede tener una probabilidad de salto particular a los diferentes centros de servicio
- [] Los tiempos de servicio de los centros siguen una distribución cuya transformada de Laplace es una función racional propia. Entre otras, la exponencial, hipereexponencial e hipoexponencial son distribuciones de este tipo
- [] Las tasas de servicio pueden ser función de la longitud de la cola en el centro, sin distinguir su composición
- [] Puede haber subredes con población limitada

El programa ofrece una gran libertad en la definición del modelo para incluir los detalles relevantes del sistema bajo estudio.

El programa permite la definición del modelo y la consulta de la solución en forma interactiva. Para esto utiliza un sistema de menús que guían al usuario a lo largo de estas actividades.

El programa soluciona los modelos de colas utilizando principalmente la técnica de agregación y el análisis del valor medio (Chand74), (Chand78), (Reise80), (Sauer81) y (Sauer83). Estas técnicas dan resultados exactos para el modelo cuando éste es de tipo BCMP (Baske75). Para modelos generales pueden dar soluciones aproximadas con un error generalmente pequeño (Sauer81).

El programa puede ser utilizado en cualquier computador con microprocesador 8086/8088 y sistema operacional MS-DOS. El computador debe tener por lo menos 128K y una unidad de diskette. Esto incluye los microcomputadores de la familia IBM y sus compatibles. El programa es fácilmente transportable a otros equipos pues ha sido desarrollado en lenguaje C, utilizando el conjunto de instrucciones y funciones descritas en el texto de Kernighan y Ritchie (Kerni78).

2. MODELOS DE SISTEMAS DE COMPUTACION

Un sistema de computación consta de varias clases de recursos como son el procesador central, los procesadores de E/S, la memoria, los procesadores de comunicación y las líneas de comunicación, entre otros. Los trabajos entran al sistema y compiten por los recursos haciendo colas de espera para ser atendidos. Cuando termina el proceso de servicio, el trabajo puede entrar nuevamente a una cola de espera para ser atendido por otro recurso. Por esta razón, un sistema de computación puede ser modelado como una red de centros de servicio interconectados.

Hay cuatro disciplinas de servicio importantes para el modelaje de los recursos del sistema: FCFS ("First Come First Served"), PS ("Processor Sharing"), IS ("Infinite Server") y LCFSPR ("Last Come First Served Preemptive Resume").

En la disciplina FCFS los clientes (trabajos, tareas, procesos o paquetes de datos, dependiendo del modelo) se atienden en el orden de llegada. El centro atiende el requerimiento total del cliente antes de atender el siguiente en la cola.

La disciplina PS es el caso límite de la disciplina "Round Robin" si el costo de la interrupción del servicio se asume

nulo y el quantum se hace tender a cero. En esta disciplina todos los clientes de la cola reciben servicio simultáneamente. Cuando hay n clientes en la cola, cada uno recibe servicio a la $1/n$ parte de la tasa a la cual recibiría el servicio si estuviera solo.

En la disciplina IS el cliente es atendido cuando llega y no hay demora en la cola. Esto corresponde a un centro donde hay más servidores que clientes.

En la disciplina LCFSPR la cola de clientes se comporta como una pila. El último cliente en llegar es atendido en el acto. El cliente es atendido hasta terminar o hasta que la llegada de otro cliente interrumpa el servicio. Una vez se finaliza el servicio de un cliente, se continuaría prestando el servicio a los demás en la pila, en el orden inverso al de llegada.

La figura 2-1 presenta el modelo de servidor central (Baske72), (Klein76), (Kobay78), (Laven83), (Reise76) y (Sauer81). Los clientes o trabajos viajan cíclicamente haciendo uso alternado del procesador central y de los procesadores de E/S. El modelo es cerrado debido a que el número de trabajos en el sistema, N , es constante. El valor de esta constante es el nivel de multiprogramación del sistema.

El modelo supone que el salto de los trabajos entre los recursos es instantáneo. Sin embargo, con la disciplina IS se pueden modelar demoras o tiempos de propagación, si ello se requiriera.

El tiempo de servicio se asume que es caracterizado por una distribución exponencial con tasa media μ_m , $m = 1, \dots, M$. Es posible suponer distribuciones más complejas para representar el tiempo de servicio.

El procesador central es modelado con disciplina de servicio PS y los procesadores de E/S con disciplina de servicio FCFS.

En un momento dado, el estado del modelo es un vector cuyos elementos son el número de trabajos en cada uno de los procesadores. Las suposiciones hechas permiten estudiar la evolución del estado del modelo en el tiempo como un proceso de Markov. De hecho, los procesos markovianos juegan un papel central en la solución de los modelos de colas considerados en este trabajo.

La solución de este modelo tiene utilidad múltiple. Por ejemplo permite conocer, para diferentes niveles de multiprogramación, el porcentaje de tiempo que los procesadores están inactivos y el tiempo que deben esperar los trabajos para hacer uso de los procesadores. También el modelo puede emplearse para establecer qué recursos del

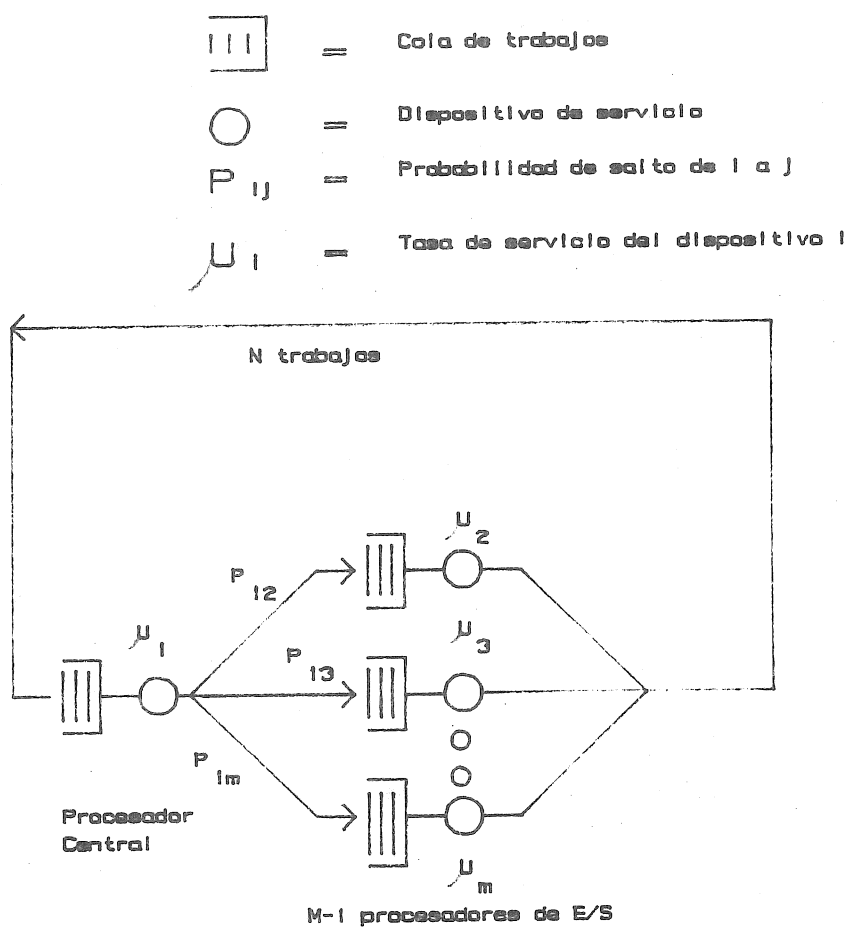


Figura 2 - 1 MODELO DE SERVIDOR CENTRAL. Los trabajos viajan cíclicamente alternando el uso del procesador central y de los procesadores de E/S.

sistema actúan como "cuello de botella".

Pueden formularse otros modelos con mayor nivel de sofisticación que reflejen detalles adicionales de la estructura y comportamiento de los sistemas reales. Igualmente pueden desarrollarse modelos que representen sistemas más complejos, tales como las redes de procesamiento centralizado y distribuido.

3. SOLUCION DE MODELOS POR AGREGACION

Una técnica de solución de modelos de colas es la agregación, llamada también descomposición. Mediante esta técnica se reemplaza una subred por un centro de servicio denominado compuesto. La figura 3-1 muestra un ejemplo de agregación.

Se ha demostrado que la agregación es exacta si el modelo tiene balance local (Chand74). Esto significa que cumpliéndose dicha condición, se asegura que el centro de servicio compuesto interactúa de la misma forma que la subred con el resto de los centros de servicio del modelo.

El balance local es una propiedad que tienen ciertos procesos markovianos y conduce a que las soluciones del estado del modelo sean en forma de producto (Baske75):

$$P_i = (P_{i1} \dots P_{iM}) / G, \quad 3-1$$

en donde P_i es la probabilidad del estado i del sistema; P_{im} es un factor determinado solamente por el estado del centro de servicio m , $m = 1, \dots, M$; y G es la constante de normalización.

Todos los modelos BCMP satisfacen balance local.

Para calcular la tasa de servicio del centro de servicio compuesto, es necesario aislar la subred a agregar del resto de los centros de servicio del modelo. Para esto se conecta su salida directamente a su entrada (las líneas punteadas de la figura 3-1). Y, para cada población posible de clientes en la subred, N , se calcula su tasa de flujo de clientes ("throughput"), $R(N)$. Esta tasa de flujo es, simplemente, la tasa de servicio del centro de servicio compuesto:

$$\mu(N) = R(N). \quad 3-2$$

Para calcular $R(N)$, existen varios algoritmos bastante eficientes que son aplicables si la subred satisface balance local (Buzen73), (Chand80) y (Reise80). Uno de los más importantes, en especial por sus buenas propiedades numéricas (Sauer83), es el análisis del valor medio.

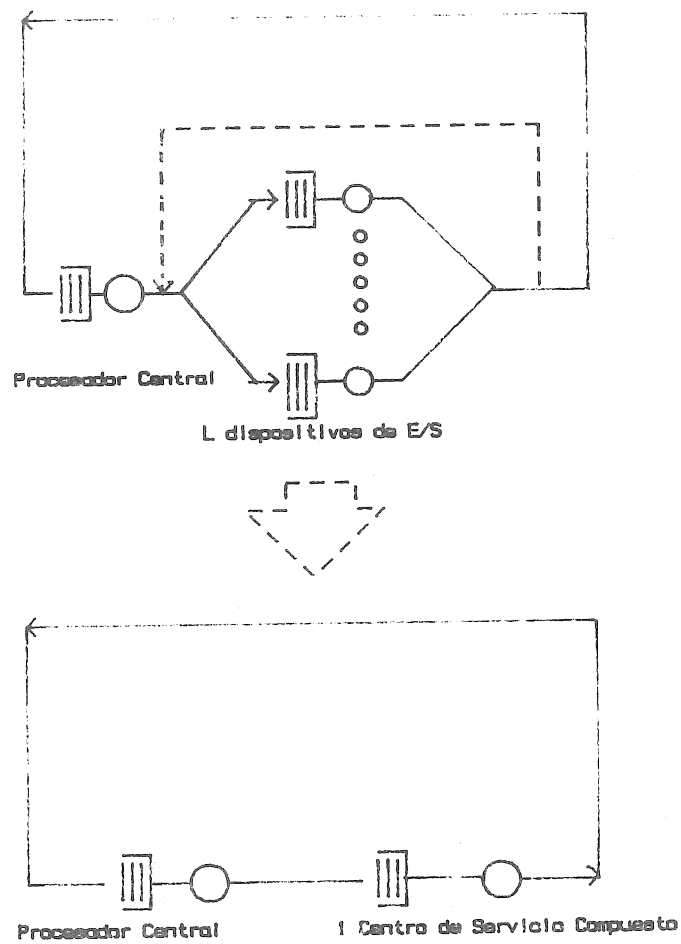


Figura 3 - 1 EJEMPLO DE AGREGACION. La subred de los dispositivos de E/S. es reemplazada por un centro de servicio compuesto con el mismo comportamiento.

El análisis del valor medio utiliza ecuaciones recursivas que relacionan el tiempo medio de espera en un centro de servicio, $Q_m(N)$, con la longitud media de su cola cuando el

modelo tiene un cliente menos, $L_m(N-1)$ (Reise80). Por ejemplo, para el modelo de la figura 2-1, se tiene:

$$Q_m(N) = \mu_m \cdot (1 + L_m(N-1)); \quad 3-3$$

y para centros de servicio con disciplina IS:

$$Q_m(N) = 1/\mu_m \quad 3-4$$

Luego, aplicando las ecuaciones de relación entre:

- I. La tasa de flujo en la red, $R(N)$, con su población, N , y su tiempo medio de travesía, $Q(N)$:

$$R(N) = N/Q(N) \quad 3-5$$

establecida por la ley de Little (Littl61);

- II. El tiempo medio de travesía de la red con el tiempo de espera en los centros de servicio:

$$Q(N) = \sum_{m=1}^M r_m Q_m, \quad 3-6$$

en donde r_m es la tasa relativa de flujo de clientes en el centro de servicio m , calculable con base en las probabilidades de salto de los clientes entre parejas de centros de servicio; y

- III. La tasa de flujo de la red con las tasas de flujo, $R_m(N)$, y relativa de flujo r_m de cualquier centro de servicio m :

$$R(N) = R_m(N)/r_m; \quad 3-7$$

se puede derivar la siguiente ecuación para el cálculo de la tasa de flujo de clientes en cada centro de servicio:

$$R_m(N) = r_m \frac{N}{\sum_{i=1}^M r_i Q_i(N)}, \quad 3-8$$

Las longitudes medias de las colas se calculan aplicando nuevamente el resultado de Little:

$$L_m(N) = R_m(N) \cdot Q_m(N). \quad 3-9$$

El análisis del valor medio implica un proceso iterativo y se inicia con $L_m(0) = 0$, $m = 1, \dots, M$. De esta manera se pueden determinar $Q_m(n)$, $R_m(n)$ y $L_m(n)$ para $n = 1, \dots, N$.

Una ventaja importante de la agregación es su economía en términos computacionales. Comparada con otras técnicas de solución de modelos de colas, su costo computacional es el menor.

Otra ventaja de la agregación es su aplicabilidad en modelos sin balance local. Tal es el caso de los modelos con centros de servicio pasivo. Un centro de servicio pasivo es aquel cuya tasa de servicio es determinada externamente por otros centros de servicio. La memoria es un caso típico de centro de servicio pasivo. En general la memoria permanece asignada en tanto el trabajo requiera servicio del procesador central y de las unidades de disco.

Para los modelos con centros de servicio pasivo, la agregación puede ser aproximada, siendo su error normalmente pequeño (Sauer81). Además, en muchas situaciones es superior a otras técnicas de solución de modelos de colas. El análisis numérico presenta problemas de control numérico y requiere la determinación exhaustiva de todos los estados del modelo, lo que es imposible cuando el número de estados no es finito. Finalmente y de otro lado, la simulación no produce necesariamente resultados más exactos que la agregación, además de que, en general, es computacionalmente más costosa.

4. PROGRAMA PARA LA SOLUCION DE MODELOS DE COLAS

Con la elaboración del programa se desea proveer una herramienta para el análisis de modelos de colas. Aunque se ha enmarcado la necesidad del programa en el diseño y evaluación de sistemas de computación, su utilidad se extiende a los sistemas en donde hay competencia por el uso de recursos.

El programa soluciona modelos abiertos y cerrados. Igualmente, no impone restricciones al número de centros de servicio que conforman el modelo ni a su modo de interconexión. Al mismo tiempo, permite poblaciones no homogéneas con lo cual el comportamiento probabilístico de los clientes del modelo puede ser diferente. Esto permite modelar sistemas de computación en donde hay trabajos que

hacen uso de los recursos como el procesador central y las unidades de E/S, en forma marcadamente diferente.

El programa puede ser utilizado en cualquier computador con microprocesador 8086/8088 y sistema operacional MS-DOS. El computador debe tener por lo menos 128K y una unidad de diskette. El programa ha sido escrito en lenguaje C utilizando el conjunto de instrucciones y funciones descritas en el libro de Kernigham y Ritchie (Kern178). Esto garantiza que el programa es transportable a otros equipos.

Para la solución del modelo el programa utiliza la técnica de agregación y el análisis del valor medio. La agregación de subredes se lleva a cabo con el fin de reducir la complejidad del modelo. Esto es indispensable cuando el modelo de colas no es de tipo BCMP. El análisis de subredes se hace con base en el análisis del valor medio. Además, para los modelos abiertos, el programa utiliza los resultados para las colas M/M/m y M/G/1.

La agregación, el análisis del valor medio y los resultados para las colas M/M/m y M/G/1 dan soluciones exactas cuando el modelo es de tipo BCMP. Para los modelos generales los resultados pueden ser aproximados; pero en general puede esperarse un error pequeño.

El programa permite la definición del modelo y la consulta de la solución en forma interactiva. Con este fin utiliza un sistema de menús que guían el uso de estas actividades.

Durante la definición del modelo de colas el programa verifica que las características de la red, dadas como entradas, sean válidas. Todos los errores se detectan en la fase de definición. La información de entrada que debe suministrarse al programa es la siguiente:

- [] nombre del modelo, entrada opcional
- [] número de centros de servicio del modelo
- [] disciplina de servicio de cada centro
- [] número de servidores de cada centro
- [] número de clases de clientes del modelo
- [] número de clientes o población del modelo, si éste es cerrado
- [] probabilidades de salto entre centros de servicio, para cada clase de clientes
- [] tasas de servicio de cada centro, para cada población de clientes posible en este

[] tasa de llegada de clientes, si el modelo es abierto

El programa crea un archivo en disco con las características del modelo, de tal manera que puede utilizarse en diferentes sesiones de análisis.

Los parámetros calculados por el programa son:

[] población media en los modelos abiertos

[] utilización de la red de los modelos abiertos

[] utilización de cada centro de servicio

[] tiempo medio de travesía en los modelos abiertos

[] tiempo medio de ciclo en los modelos cerrados cíclicos

[] tasa de flujo de clientes de los modelos abiertos

[] tasa de flujo de clientes en cada centro de servicio

[] tiempo medio de espera en cada centro

[] longitud media de la cola en cada centro

[] probabilidad de todas las posibles longitudes de la cola en los centros

Detalles adicionales pueden ser encontrados en el manual del usuario del programa (Zufig86).

5. USO DE LOS MODELOS DE COLAS

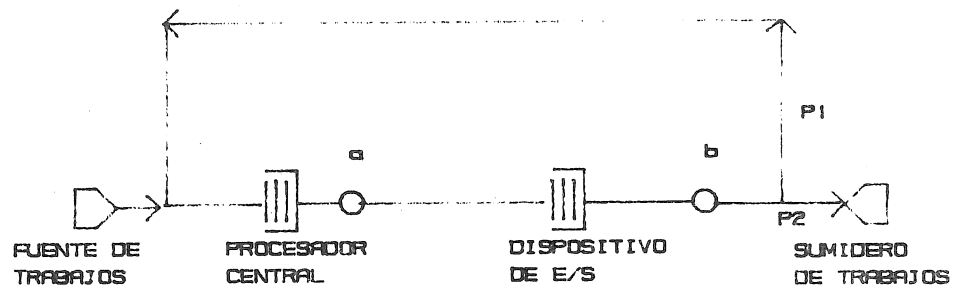
Para ilustrar el uso de los modelos de colas y del programa desarrollado, se ha calculado el desempeño de los modelos de sistemas de computación mostrados en la figura 5-1.

El primer modelo muestra un sistema al cual llegan trabajos con una tasa Poisson con media λ . Cada trabajo es atendido por el procesador central a una tasa de servicio exponencial con media a , y un dispositivo de E/S con tasa de servicio exponencial de media b . Los trabajos pueden requerir más de un ciclo de atención del procesador central y del dispositivo de E/S, ocurriendo esto con una probabilidad p_1 ; de otra

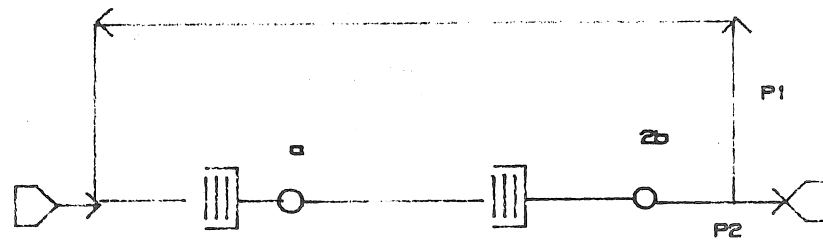
forma, los trabajos salen del sistema luego de recibir servicio en el dispositivo de E/S, con una probabilidad p_2 .

Supóngase que se desea aumentar la velocidad de servicio de E/S de este sistema de computación. Para esto se consideran las tres alternativas de la figura 5-1.

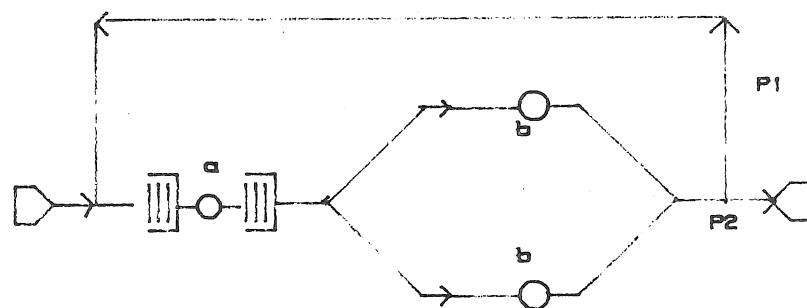
Sistema de computación



1ra alternativa



2da alternativa



3ra alternativa

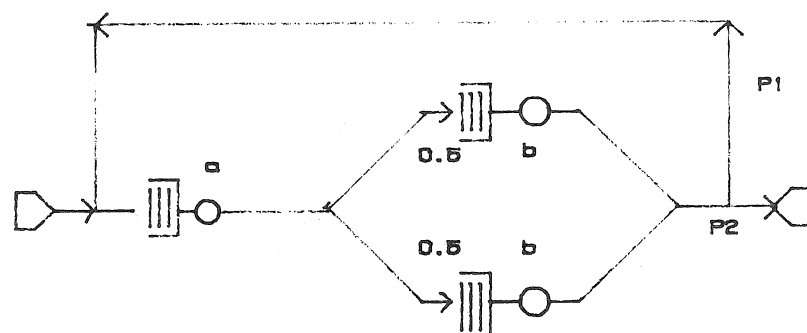


Figura 5 - 1 Alternativas para aumentar el desempeño de un sistema de computación

En la primera se reemplazaría el dispositivo de E/S por otro con tasa de servicio $2b$, dos veces más rápido que el dispositivo del sistema original.

En la segunda habría una cola de trabajos con dos dispositivos de E/S en paralelo, cada uno con tasa de servicio b . Los trabajos en la cola saltarían al primero de los dos dispositivos que se encuentre disponible.

En la tercera se tendrían en paralelo dos dispositivos con tasa de servicio b , los cuales serían escogidos por los trabajos en forma aleatoria con probabilidad $0,5$.

En las tres alternativas las restantes características del modelo original se mantienen.

Nótese que cada alternativa ofrece duplicar la velocidad del servicio de E/S; pero las dos últimas lo harían sólo en los casos en que se encuentren los dos dispositivos de E/S ocupados al mismo tiempo. Esta situación hace las alternativas similares.

Se desea saber:

- [] cuál de las tres alternativas daría el mejor aumento de productividad al sistema; y
- [] si el aumento de productividad justifica retener el procesador central.

Este problema ha sido estudiado en (Chand) y (Lipsk77).

El aumento de productividad de las alternativas de mejora del sistema, puede medirse de varias maneras, cada una de ellas expresando en forma diferente el hecho de que los trabajos serían atendidos más rápidamente. Algunas son las siguientes:

- [] la disminución del tiempo medio de espera de los trabajos para recibir el servicio de E/S;
- [] el aumento de la longitud de la cola en el procesador central; y
- [] el aumento de la utilización del procesador central.

Para el análisis se ha escogido la última pues, además de mostrar el aumento de la productividad del sistema, sirve de base para concluir la conveniencia de retener el procesador central.

La tabla 5-1 y la figura 5-2 muestran la utilización del procesador central para el modelo original y para las tres alternativas de mejora cuando $b/a = 0,3$. Esta relación corresponde a un sistema cuyo desempeño está limitado por el

$$b/a = 0.3$$

UTILIZACION DEL PROCESADOR CENTRAL

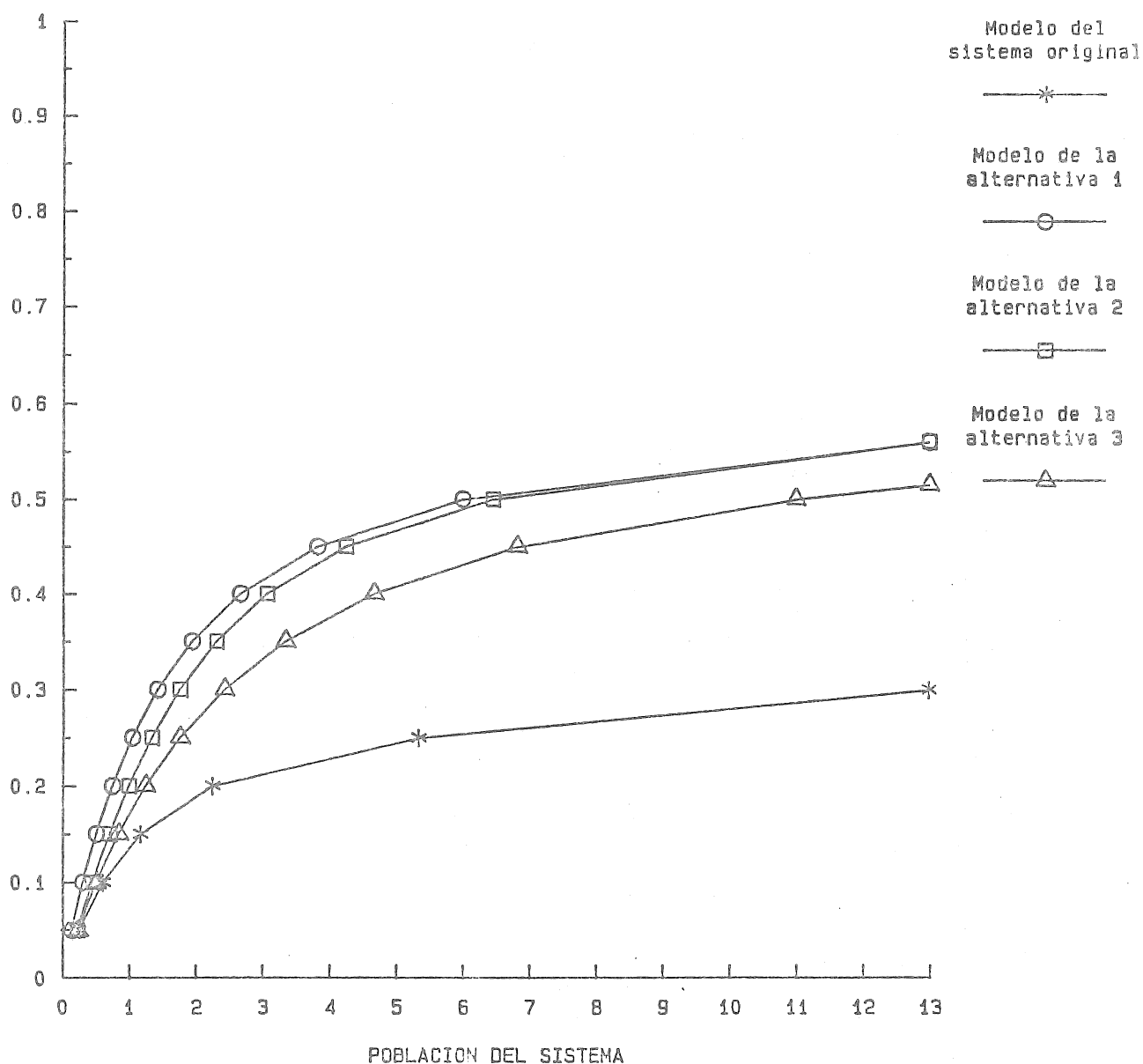


Figura 5-2. Utilizacion del procesador central como funcion de la poblacion del sistema

a = 1,0 b = 0,3

Tasa de Llegada	Utilización del Procesador Central	Población del Sistema	Población de la Alternativa 1	Población de la Alternativa 2	Población de la Alternativa 3
0.01	0.05	0.25263	0.14354	0.22046	0.23445
0.02	0.10	0.61111	0.31111	0.45397	0.51111
0.03	0.15	1.17647	0.50980	0.70980	0.94314
0.04	0.20	2.25000	0.75000	1.00000	1.25000
0.05	0.25	5.33331	1.04762	1.34174	1.76190
0.06	0.30	--	1.42857	1.76190	2.42857
0.07	0.35	--	1.93846	2.30688	3.33846
0.08	0.40	--	2.66666	3.06666	4.66666
0.09	0.45	--	3.81817	4.24674	6.81816
0.10	0.50	--	5.99998	6.45452	10.99995
0.11	0.55	--	12.22212	12.70038	22.22202
0.12	0.60	--	--	--	--

Tabla 5-1. Valor de la utilización del procesador central para poblaciones del modelo del sistema y de sus tres alternativas de mejora de la figura 5-1. En la derivación de estos resultados se varió el valor de la tasa de llegada y se tomó $p_1 = 0,8$, $p_2 = 0,2$, $a = 1,0$ y $b = 0,3$

servicio del dispositivo de E/S.

La utilización se ha calculado para diferentes poblaciones en cada modelo. Para esto se han fijado los valores de la probabilidad de que el trabajo dé un nuevo ciclo en el sistema, p , en 0,8; y la probabilidad de que el trabajo

salga del sistema, p_1 , en 0,2. Y, al mismo tiempo, se ha

variado el valor de la tasa de llegada de trabajos, λ , desde 0,01 hasta 0,12 con incrementos de 0,01. Esto ha permitido calcular la utilización del procesador central como función de una población variable en cada modelo.

Los mismos resultados podrían lograrse fijando el valor de λ y dando diferentes valores a p_1 y p_2 , con igual efecto de variar la población en cada modelo.

La tabla 5-1 muestra el valor de λ utilizado para derivar cada utilización y las poblaciones en el modelo del sistema y en las alternativas correspondientes.

La tabla 5-1 y la figura 5-2 permiten concluir que la primera alternativa da mayor utilización al procesador central; y que sólo cuando las poblaciones de los modelos son altas, se logra la misma utilización de 0,6.

Esto significa que cuando el nivel de multiprogramación del sistema (representado por la población del modelo) no es alto, la elección de la alternativa apropiada es de mayor importancia. Tal es el caso de los sistemas en donde el tamaño de la memoria limita el nivel de multiprogramación.

De la misma forma la tabla 5-2 y la figura 5-3 muestran los resultados para $b/a = 1$. En este caso el sistema tendría balanceados los servicios del procesador central y del dispositivo de E/S. De allí se concluye que nuevamente la mejor alternativa es la primera. También que sólo cuando las poblaciones de los modelos son altas, se logra igual utilización. Por último que el desempeño del sistema mejora un porcentaje pequeño lo que sugiere la necesidad de un nuevo procesador central de mayor velocidad.

En suma, estos ejemplos muestran que:

- [] para aumentar la velocidad de servicio de E/S en los sistemas de computación, es mejor utilizar dispositivos más rápidos que aumentar el número de dispositivos para una suma de velocidades igual;
- [] la elección de la alternativa apropiada es crucial en sistemas con bajo nivel de multiprogramación; y
- [] la mejora de la E/S puede conllevar el cambio del

a = 1,0 b = 1,0

Tasa de Llegada	Utilización del Procesador Central	Población del Sistema	Población de la Alternativa 1	Población de la Alternativa 2	Población de la Alternativa 3
0.01	0.05	0.10526	0.07827	0.10266	0.10391
0.02	0.10	0.22222	0.16374	0.21136	0.21637
0.03	0.15	0.35294	0.25755	0.32732	0.33863
0.04	0.20	0.50000	0.36111	0.45202	0.47222
0.05	0.25	0.66667	0.47619	0.58730	0.61905
0.06	0.30	0.85714	0.60504	0.73548	0.78151
0.07	0.35	1.07692	0.75058	0.89952	0.96270
0.08	0.40	1.33333	0.91667	1.08333	1.16667
0.09	0.45	1.63636	1.10850	1.29218	1.39883
0.10	0.50	2.00000	1.33333	1.53333	1.66666
0.11	0.55	2.44444	1.60153	1.81722	1.98084
0.12	0.60	2.99999	1.92857	2.15934	2.35714
0.13	0.65	3.71428	2.33862	2.58390	2.82010
0.14	0.70	4.66666	2.87179	3.13105	3.41025
0.15	0.75	5.99998	3.59999	3.87272	4.19999
0.16	0.80	7.99997	4.66665	4.95237	5.33332
0.17	0.85	11.33328	6.50577	6.70401	7.14490
0.18	0.90	17.99999	9.81812	10.12846	10.63630
0.19	0.95	27.99995	19.90449	20.22652	20.80925
0.20	1.00				

Tabla 5-2. Valor de la utilización del procesador central para poblaciones del modelo del sistema y de sus tres alternativas de mejora de la figura 5-1. En la derivación de estos resultados se varió el valor de la tasa de llegada y se tomó $p = 0,8$, $p = 0,2$, $a = 1,0$ y $b = 1,0$

$$b/a = 1,0$$

UTILIZACION DEL PROCESADOR CENTRAL

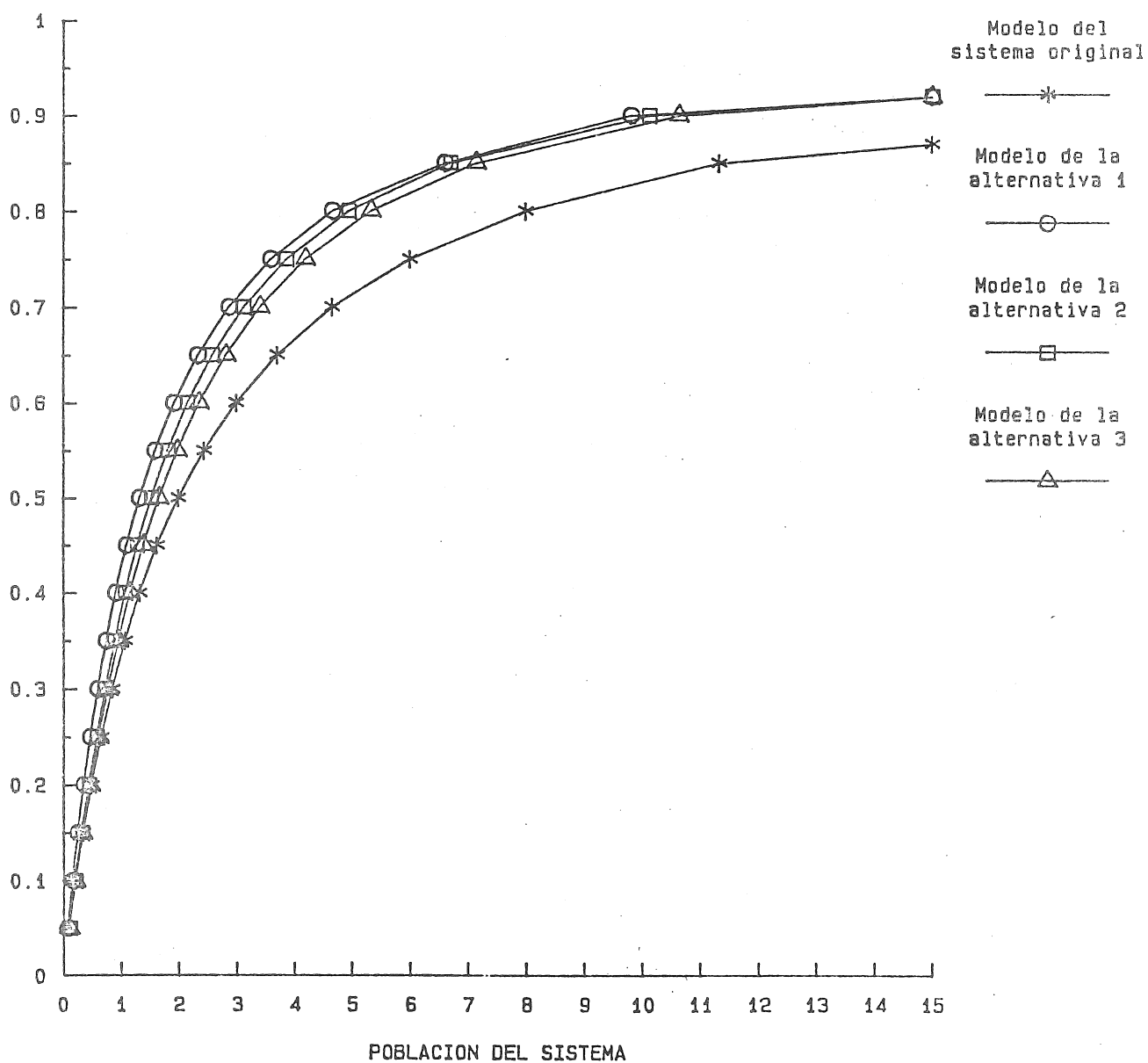


Figura 5-3. Utilizacion del procesador central como funcion de la poblacion del sistema

procesador central si el sistema tiene balanceadas las velocidades de estos servicios.

Los supuestos hechos para los modelos permiten generalizar los resultados a cualquier sistema de computación y a otros recursos como el procesador central.

En el análisis del procesador central se puede decir que en un sistema de computación con bajo nivel de multiprogramación, es mejor colocar un procesador más rápido que hacer multiprocesamiento con varios procesadores con una suma de velocidades equivalente, sin importar si hay sólo una cola de trabajos o si cada uno atiende su propia cola. Esta conclusión se debe entender sólo en términos del desempeño del sistema y sin considerar otros aspectos como la confiabilidad de este.

Por último debe observarse que éste análisis no ha requerido hacer medición de parámetros de ningún sistema real. Sólo se ha necesitado un conjunto de supuestos que mantiene la generalidad de los resultados. Pero, de otro lado, en otras aplicaciones de esta metodología de análisis, podría ser necesaria la estimación de parámetros de sistemas reales para permitir su modelaje mediante redes de colas.

6. CONCLUSIONES

Se ha presentado un programa para llevar a cabo el análisis de sistemas de computación y de comunicaciones basados en modelos de colas. Al mismo tiempo, se ha dado un ejemplo de evaluación de modelos de sistemas de computación utilizando el programa presentado. Finalmente, se ha mostrado que el programa puede ser de gran utilidad en el diseño y evaluación de sistemas, en particular de los sistemas de computación.

7. AGRADECIMIENTOS

Los autores quieren agradecer de una manera muy especial a la Facultad de Ingeniería de la Universidad de Los Andes, por el apoyo dado a este trabajo.

8. BIBLIOGRAFIA

- (Baske72) Baskett, F. y Palacios Gómez, F. Processor Sharing in a Central Server Queueing Model of Multiprogramming with Applications. Proc. Sixth Annual Princeton Conference on Information Science and Systems,

Princeton U., Princeton, N.J. Marzo 1972, 598-603.

- (Baske75) Baskett, F.; Chandy, K.M.; Muntz, R.R. y Palacios-Gómez, F. Open, Closed and Mixed Networks of Queues whit Different Classes of Customers. J.ACM 22,2 (abril 1975), 248-260.
- (Buzen73) Buzen, J.P. Computational Algorithms for Closed Queueing Networks whit Exponential Servers. C.ACM 16,9 (septiembre 1973), 527-531.
- (Buzen78) Buzen, J.P. A Queueing Network Model of MVS. Comp. Surveys 10,3 (septiembre 1978), 319-331.
- (Chand) Chandy, K.M. Queueing Models in Computer Systems. Fotocopia. Sin fecha.
- (Chand75) Chandy, K.M.; Herzog, V. y Woo, L. Parametrics Analysis of Queueing Network Models. RC4730, IBM Reseach, Yorktown Heights, New York. Febrero, 1974. IBM J. Res. Dev. 19,1 (enero 1975), 36-42.
- (Chand78) Chandy, K.M. y Sauer, C.H. Approximate Analysis of Queueing Networks Models of Computing Systems. Comp. Surveys 10,3 (septiembre 1978), 281-317.
- (Chand80) Chandy, K.M. y Sauer, C.H. Computational Algorithms for Product Form Queueing Networks. C.ACM 23,10 (octubre 1980), 573-583.
- (Kerni78) Kernigham, B.W. y Ritchie, D.M. The C Programming Language. Prentice Hall, Englewood Cliffs, N.J. 1978.
- (Klein76) Kleinrock, L. Queueing Systems. Vol. II: Computer Applications. Wiley. New York. 1976.
- (Kobay78) Kobayashi, H. System Design and Perfomance Analysis Using Analytic Models. En Current Trends in Programming Methodology, Vol III: Software Modeling and its Impact on Perfomance. K.M. Chandy y R.T. Yeh (Eds.), Prentice Hall, Inc., Englewood Cliffs, N.J., 1978, 72-114.
- (Laven83) Notes and Reports in Computer Science and Applied Mathematics, Vol IV: Computer Perfomance Modeling Handbook. S.S. Lavenberg (Ed.), Academic Press, Inc., New York, N.Y., 1983.
- (Lipsk77) Lipsky, L. y Church, J.D. Applications of a Queueing Network Model for a Computer System. Comp. Surveys 9,3 (septiembre 1977), 205-221.
- (Litt161) Little, J.D.C. A Proof of the Formula $L = \lambda W$. Op. Res. 9 (1961), 383-387.

- (Reise76) Interactive Modeling of Computer Systems. IBM Syst. J. 12,4 (abril 1976), 309-327.
- (Reise80) Reiser, M. y Lavenberg S.S. Mean Value Analysis of Closed Multi Chains Queueing Networks. J.ACM 27,2 (abril 1980), 313-322. Corrigendum J.ACM 28,3 (julio 1981), 629.
- (Sauer80) Sauer, C.H.; MacNair, E.A. y Salza, S. A Language for Extended Queueing Network Models. IBM J. Res. Dev. 24,6 (noviembre 1980), 747-755.
- (Sauer81) Sauer, C.H. y Chandy K.M. Computer Systems Performance Modeling. Prentice Hall, Englewood Cliffs, N.J., 1981.
- (Sauer83) Sauer, C.H. Computational Algorithms for State Dependent Queueing Networks. ACM Trans. on Comp. Syst. 1,1 (febrero 1983), 67-92. Corrigendum ACM Trans. on Comp. Syst. 1,4 (noviembre 1983), 369.
- (Wong78) Wong, J.W. Queueing Network Modeling of Computer Communication Networks. Comp. Surveys 10,3 (septiembre 1978), 343-351.
- (Zunig86) Zuniga, M. Modelos de Redes de Colas. Tesis de Master. Depto. Ing. Sist. y Comp. U. de Los Andes, Bogotá. 1986.