

Dimensionado y Planificación de Redes

Tema 5 – Sistemas con fuentes finitas

Ramón Agüero Calvo

ramon.agueroc@unican.es

Contenido

- Introducción
- Modelos de fuente
- Sistemas de espera pura (1 servidor)
- Sistemas de pérdida pura – Fórmula de Engset

Contenido

- **Introducción**
- Modelos de fuente
- Sistemas de espera pura (1 servidor)
- Sistemas de pérdida pura – Fórmula de Engset

Motivación

- Estos sistemas se dan cuando el número de usuarios del sistema es pequeño
- No se puede hablar de una tasa de llegadas constante, ya que un usuario en el sistema no puede generar más peticiones
- Distinción entre diferentes tipos de terminal
 - Un único terminal puede comportarse como fuente “*infinita*”, cuando genera paquetes según un proceso de Poisson (de manera continuada)
 - Si una petición solo puede llevarse a cabo tras haber finalizado la anterior (protocolo *stop&wait*, servicios de compra on-line, etc) el modelado sería diferente
- En función del número de fuentes (M), de servidores (S) y del tamaño del subsistema de espera (W) se podrían distinguir varios supuestos
 - Si $S \geq M$: no hay ni pérdida ni espera, ya que una petición siempre encontrará un recurso disponible
 - Si $S+W \geq M$ y $M \geq S$: no se pierden peticiones – sistema de espera pura
 - Si $W = 0$, $M \geq S$: no hay espera – sistema de pérdida pura

Contenido

- Introducción
- Modelos de fuente
- Sistemas de espera pura (1 servidor)
- Sistemas de pérdida pura – Fórmula de Engset

Modelos de fuente

- A la hora de modelar el comportamiento de las fuentes de tráfico hay que distinguir el nivel en el que se está trabajando: (1) sesión, (2) flujo, (3) paquete
- Una aproximación sencilla y bastante empleada es utilizar el modelo ON-OFF
- La fuente puede estar en dos estados diferentes
 - ON: genera llamadas a una tasa determinada
 - OFF: no genera tráfico
- A partir de las probabilidades de los estados ON y OFF y de la velocidad, se puede obtener la velocidad media de la fuente

Modelos de fuente

- Otra aproximación, complementaria con la anterior, es asumir que las fuentes tienen dos “estados”
 - [OFF] Thinking time (reposo)
 - [ON] Holding time (activo)
- En el estado de reposo se establece el instante en el que se generará la siguiente llamada (habitualmente, a través de una va exponencial negativa (T_0) de media $E(T_0) = \alpha^{-1}$)
- En el estado activo se establece la duración de la llamada en curso, utilizando (habitualmente) una va exponencial negativa (T_S), de media $E(T_S) = \mu^{-1}$
- A partir de esos valores se puede establecer el “tráfico” por fuente, como $a = \frac{\alpha}{\mu}$
- Como se verá posteriormente, el problema se da cuando la fuente se encuentra en otra situación (por ejemplo, esperando)
 - Notar que las vas asumen que la fuente se encuentra en un estado concreto

Contenido

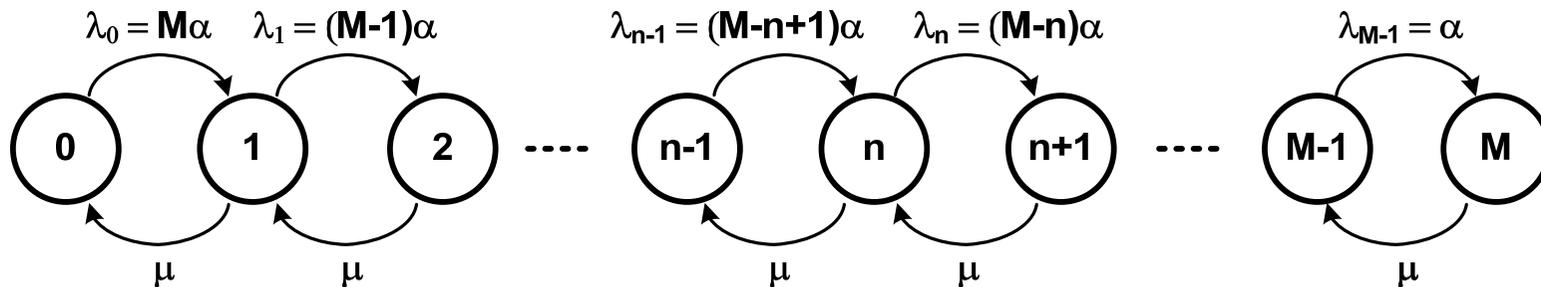
- Introducción
- Modelos de fuente
- Sistemas de espera pura (1 servidor)
- Sistemas de pérdida pura – Fórmula de Engset

Introducción

- Se considera un sistema con un único recurso, M fuentes y capacidad suficiente de espera (esto es $K+1 \geq M$)
- Las fuentes (libres) generan peticiones a una tasa α
 - [IMPORTANTE] La va correspondiente (exponencial negativa de media α^{-1}) está condicionada al hecho de que la fuente esté en el estado libre (*Thinking*)
 - Cuando una fuente está en el sistema (ya sea en el servicio o esperando) **no** puede generar más peticiones
- La duración media del servicio sigue una distribución exponencial negativa, con media μ^{-1}
- Como el sistema tiene capacidad suficiente para todas las posibles peticiones no hay pérdidas, por lo que se trata de un sistema de espera pura

Modelo

- La cadena de Markov correspondiente tendrá $M+1$ estados
- La tasa de nacimiento no es constante, ya que depende del número de fuentes libres
 - $\lambda_n = (M - n)\alpha$
- La tasa de muerte es igual para todos los estados (μ), ya que únicamente hay un recurso disponible



- Aplicando balance de flujos y resolviendo para p_n , se llega a...

$$p_n = \frac{M!}{(M - n)!} a^n \cdot p_0 \qquad p_0 = \frac{1}{\sum_{i=0}^M \frac{M!}{(M - i)!} a^i}$$

Parámetros de rendimiento

- Número medio de unidades en el sistema

$$\bar{N}_t = \sum_{i=0}^M i \cdot p_i = \dots = M - \frac{1 - p_0}{a}$$

- Número medio de unidades en espera

$$\bar{N}_Q = \sum_{i=1}^M (i - 1) \cdot p_i = \dots = M - \frac{(1 - p_0)(1 + a)}{a}$$

- Número medio de unidades en el servidor (coincide con el tráfico cursado)

$$\bar{N}_S = \sum_{i=1}^M p_i = 1 - p_0 = \bar{N}_t - \bar{N}_Q$$

- Número medio de fuentes libres

$$\bar{N}_m = \sum_{i=0}^M (M - i) \cdot p_i = \dots = \frac{1 - p_0}{a}$$

Parámetros de rendimiento

- Al aplicar la relación de Little se tiene que tener en cuenta que la tasa de llegadas no es constante, por lo que se tiene que obtener su valor medio

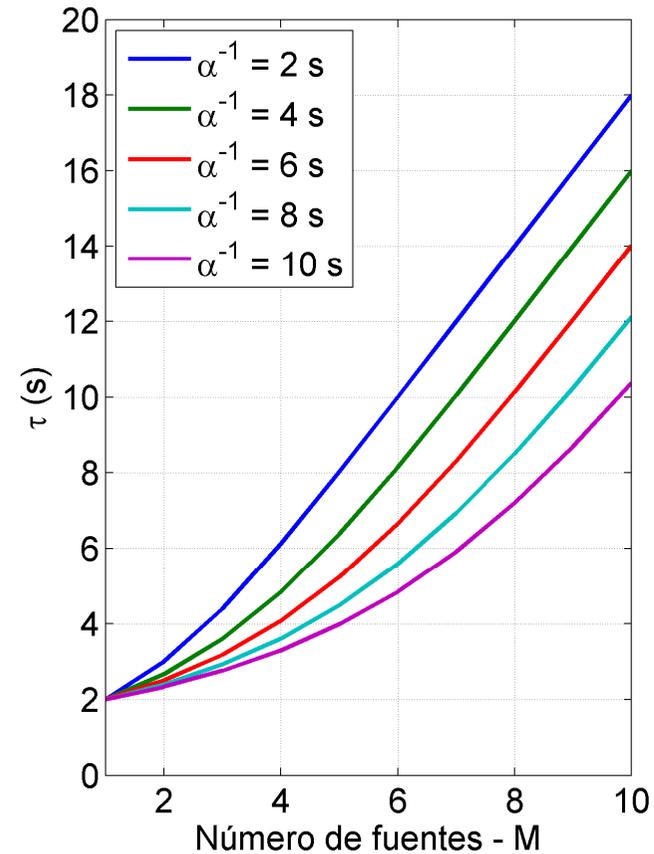
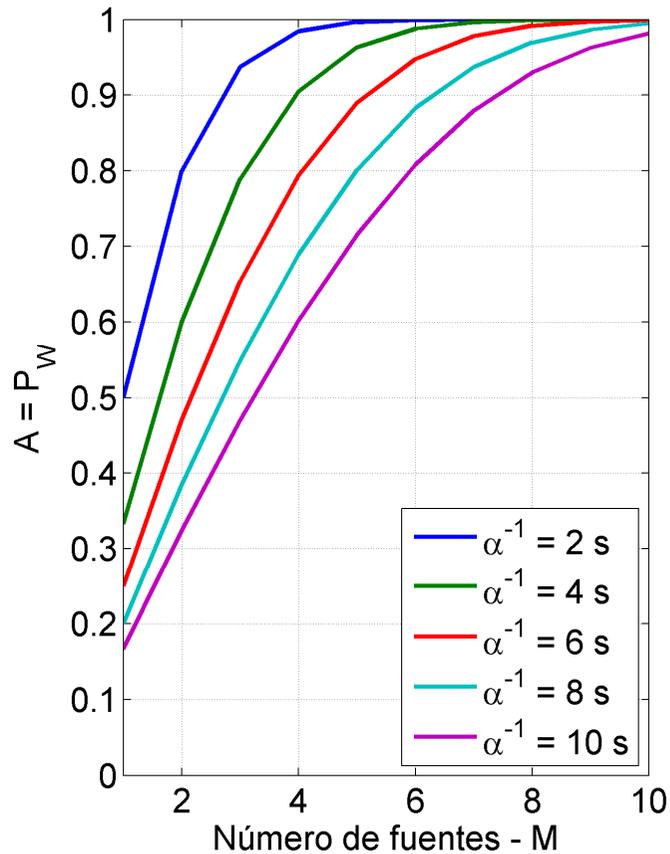
$$\bar{\lambda} = \sum_{i=0}^M \lambda_i \cdot p_i = \sum_{i=0}^M (M - i) \cdot \alpha \cdot p_i = \dots = \mu \cdot (1 - p_0)$$

- A partir de la que se pueden obtener los tiempos medios de permanencia en el sistema, en espera y en el servidor

$$\begin{aligned} \bar{T}_t &= \frac{\bar{N}_t}{\bar{\lambda}} = \dots = T_s \left(\frac{M}{1 - p_0} - \frac{1}{a} \right) \\ \bar{T}_Q &= \frac{\bar{N}_Q}{\bar{\lambda}} = \dots = T_s \left(\frac{M}{1 - p_0} - \frac{1 + a}{a} \right) \\ \bar{T}_{serv} &= T_s \end{aligned}$$

Ejemplo

- Se supone que $\alpha^{-1} = \{2 \ 4 \ 6 \ 8 \ 10\}$ (seg.) y que $T_s = 2$ seg.

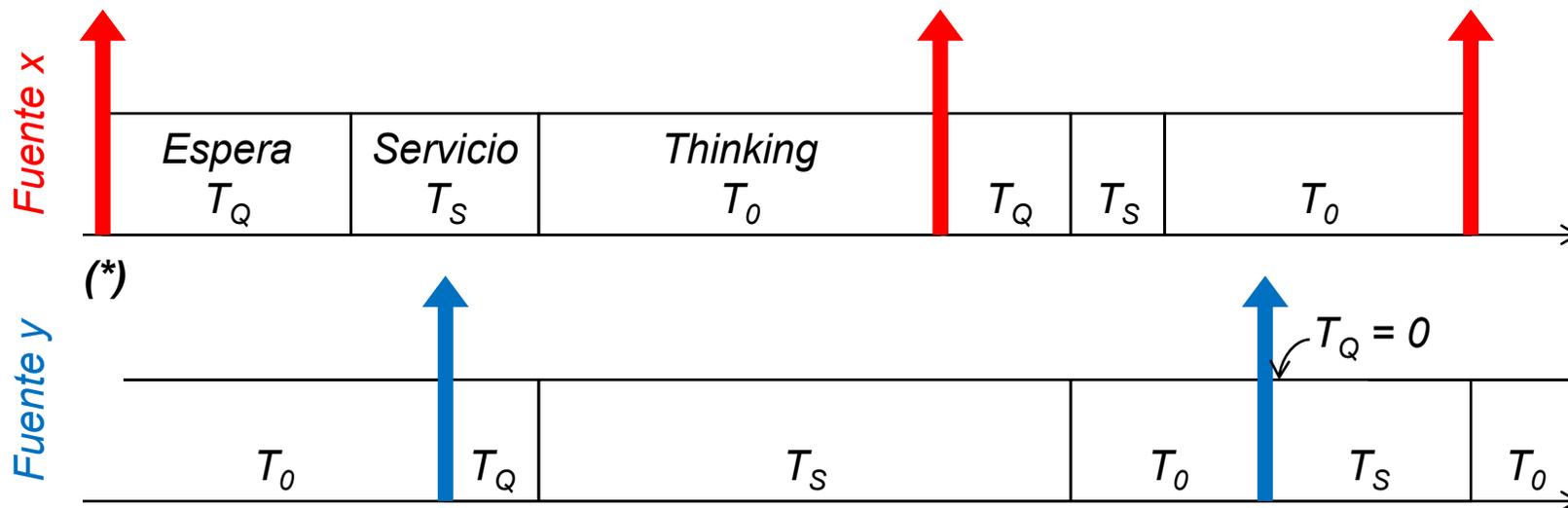


Consideración práctica

- Para poder analizar el rendimiento del sistema se ha partido de tres parámetros: M , α y μ
- M y μ se pueden conocer (estimar) de manera razonablemente sencilla
- La estimación de α es más complicada, ya que la va subyacente está condicionada a que la fuente esté libre
- De hecho, al modelo de fuente visto previamente, habría que añadir el tiempo que dicha fuente permanece en el sistema, esperando
- Es más razonable pensar que se puede estimar la tasa media de generación de paquetes, a partir de observaciones externas al sistema

Estimación de α

- A partir de observaciones externas se puede encontrar el valor de $\bar{\lambda}$
- Se asume, entonces, que $\bar{\lambda} = M \cdot \tilde{\alpha}$, con $\tilde{\alpha}$ siendo la tasa equivalente por fuente *no condicionada* a estar libre
- Asumiendo que $\alpha^{-1} = T_0$, un observador externo estimaría $\tilde{\alpha}^{-1}$ como la suma de tres tiempos: $T_0 + T_Q + T_S$, esto es: tiempo de *thinking*, tiempo de espera y tiempo de servicio



(*) Se asume que al comienzo hay una única unidad en el sistema (en el recurso)

Estimación de α

- Se puede escribir, entonces que...

$$\tilde{\alpha}^{-1} = T_Q + T_S + T_0 \rightarrow \frac{1}{\tilde{\alpha}} = T_Q + T_S + \frac{1}{\alpha} \rightarrow \alpha = \frac{\tilde{\alpha}}{1 - \tilde{\alpha}(T_S + T_Q)}$$

con $\tilde{\alpha} = \frac{\tilde{\lambda}}{M}$, siendo $\tilde{\lambda}$ la tasa estimada a partir de observaciones externas

- Se sigue sin poder calcular α a partir de la tasa estimada, ya que T_Q depende de α y no se conoce
- La solución es emplear un algoritmo iterativo, comenzando con un valor de T_Q de 0 y que vaya haciendo sucesivas aproximaciones al valor adecuado de α

Estimación de α - Algoritmo

IN: $\lambda, M, T_S, \varepsilon$

OUT: τ

$$(1) \quad \underline{\alpha} = \lambda/M, \quad T_Q = 0$$

$$(2) \quad \alpha_0 = \underline{\alpha}, \quad \alpha_1 = \frac{\underline{\alpha}}{1 - \underline{\alpha}(T_S + T_Q)}$$

(3) While (abs ($\alpha_0 - \alpha_1$) > ε)

$$(4) \quad a = \alpha_1 \cdot T_S, \quad p_0 = \frac{1}{\sum_{i=0}^M \frac{M!}{(M-i)!} a^i} \quad T_Q = T_S \left(\frac{M}{1-p_0} - \frac{1+a}{a} \right)$$

$$(5) \quad \alpha_0 = \alpha_1, \quad \alpha_1 = \frac{\underline{\alpha}}{1 - \underline{\alpha}(T_S + T_Q)}$$

(6) End while

$$(7) \quad \tau = T_Q + T_S$$

Contenido

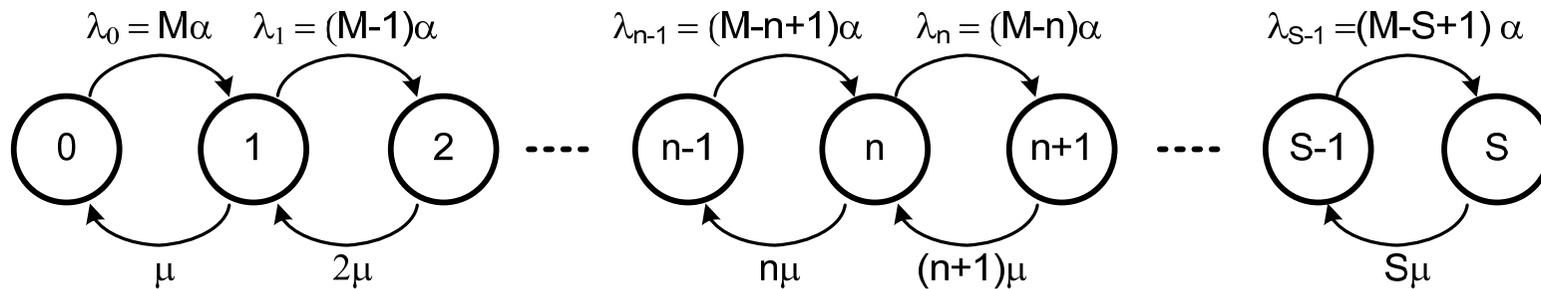
- Introducción
- Modelos de fuente
- Sistemas de espera pura (1 servidor)
- Sistemas de pérdida pura – Fórmula de Engset

Introducción

- En este caso se considera un sistema con S recursos y M fuentes ($S \leq M$), sin capacidad adicional para mantener peticiones en espera
- Como sucedía en el modelo anterior, las fuentes (libres) generan peticiones a una tasa α
 - Cuando una fuente está en el sistema no puede generar más peticiones
- La duración media del servicio sigue una distribución exponencial negativa, con media μ^{-1}
- Como el sistema no tiene capacidad suficiente para todas las posibles peticiones (y no hay espera) se producirán pérdidas: sistema de pérdida pura

Modelo

- La cadena de Markov correspondiente tendrá $S+1$ estados
- La tasa de nacimiento no es constante, ya que depende del número de fuentes libres
 - $\lambda_n = (M - n)\alpha$
- La tasa de muerte también depende del estado ($n \cdot \mu$), ya que en el estado n hay n “servicios” en curso



- Aplicando balance de flujos y resolviendo para p_n , se llega a...

$$p_n = \binom{M}{n} a^n \cdot p_0 \qquad p_0 = \frac{1}{\sum_{i=0}^S \binom{M}{i} a^i}$$

Probabilidad de bloqueo y pérdida

- La probabilidad de bloqueo se corresponde con la probabilidad de que todos los recursos estén ocupados $PB = \Pr\{\text{estado } S\}$

$$PB = p_S = \frac{\binom{M}{S} a^S}{\sum_{i=0}^S \binom{M}{i} a^i}$$

- En este caso (al contrario de lo que ocurría con el modelo M/M/S/S) la probabilidad de bloqueo NO coincide con la de pérdida
 - En este caso la probabilidad de pérdida es la probabilidad de que, estando el sistema bloqueado, llegue una petición (idea de probabilidad condicionada)
- Se calcula la probabilidad de pérdida como el cociente entre la tasa de llegadas que se pierden entre la tasa de llegadas media

$$P_L = \frac{\alpha(M - S)p_S}{\sum_{i=0}^S \alpha(M - i)p_i}$$

Prob. de pérdida – Fórmula de Engset

- Operando con la expresión anterior, se llega a la fórmula de *Engset*

$$P_L = Eng(S, M, a) = \frac{\binom{M-1}{S} a^S}{\sum_{i=0}^S \binom{M-1}{i} a^i}$$

- Destacar que $Eng(S, M, a)$ es la probabilidad de bloqueo que se tendría en un sistema con 1 fuente menos, con lo que $P_L < P_B$
- Al igual que la fórmula de ErlangB se puede resolver de manera recursiva...

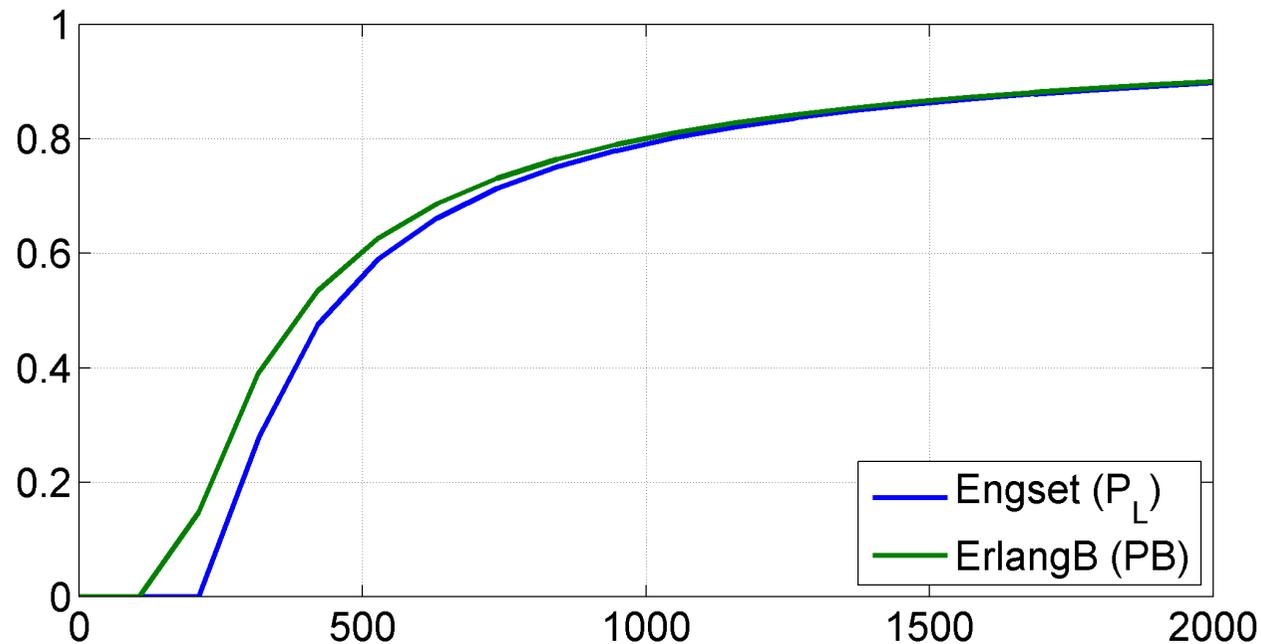
$$Eng(k, M, a) = \frac{a \cdot (M - k) \cdot Eng(k - 1, M, a)}{k + a \cdot (M - k) \cdot Eng(k - 1, M, a)}$$

- Propiedades
 - La fórmula de Engset se puede utilizar para cualquier distribución del tiempo de servicio

Engset Vs. ErlangB

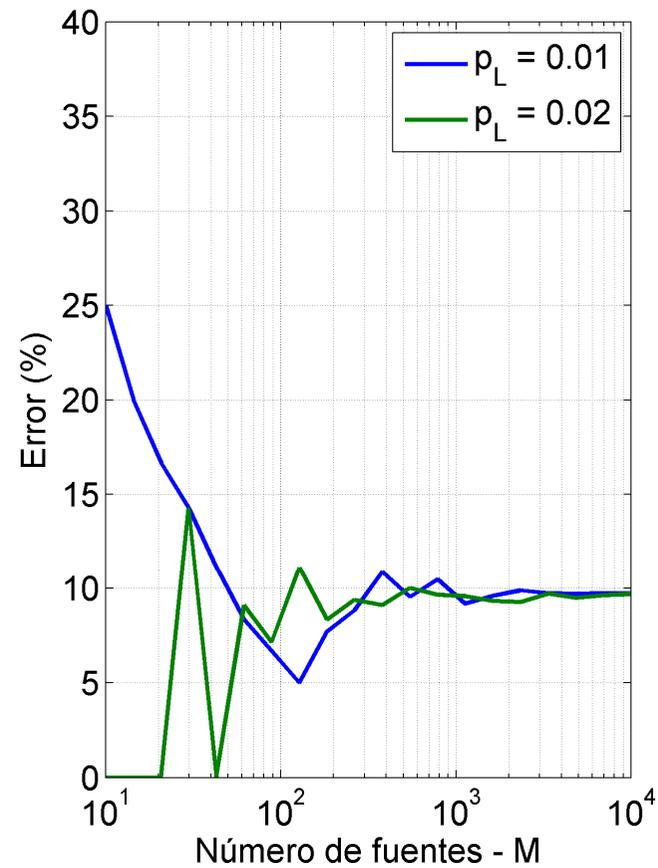
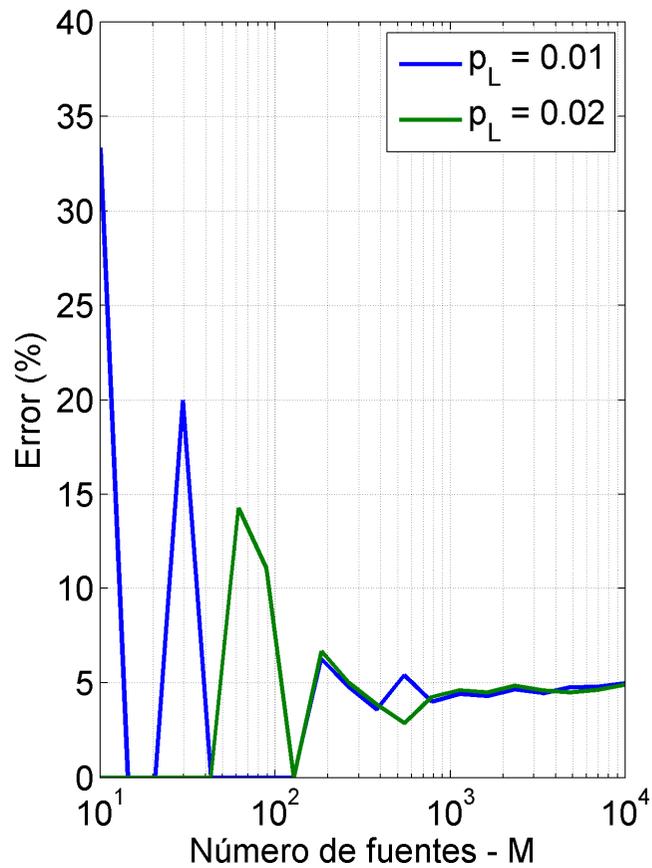
■ Parámetros

- Se asume que el tráfico por fuente (libre) $a = 0.2$
- $M = 1 \dots 2000$
- $S = 40$
- En ErlangB el tráfico ofrecido será el producto de a y M



Engset Vs. ErlangB

- Diferencia relativa entre Engset y Erlang-B para $a = 0.05$ (izquierda) y $a = 0.1$ (derecha)



Consideración práctica - Estimación de α

- Al igual que sucedía con el sistema M/M/1/K/M (con $K+1 \geq M$), la estimación del tráfico ofrecido por terminal libre (α) es complicado
- A partir de observaciones externas se puede obtener la tasa de peticiones media $E(\lambda)$, a partir de la que se puede estimar el tráfico ofrecido...

$$TO = E(\lambda) \cdot T_S$$

- Alternativamente, el TO también se puede calcular a partir la tasa de peticiones ofrecida al sistema como...

$$TO = \bar{\lambda}_0 \cdot T_S = T_S \sum_{i=0}^S \lambda_i \cdot p_i = T_S \sum_{i=0}^S (M - i)\alpha \cdot p_i = T_S \left(M \cdot \alpha - \alpha \sum_{i=0}^S i \cdot p_i \right)$$

Consideración práctica - Estimación de α

- Además, se sabe que el tráfico cursado se puede calcular como...

$$TC = \sum_{i=0}^S i \cdot p_i$$

o, alternativamente...

$$TC = TO(1 - P_L)$$

- Con lo que se llega a la siguiente expresión, a partir de la que se puede conocer el tráfico individual (por fuente libre) en función del TO global al sistema

$$\alpha = \frac{TO}{M - TO(1 - P_L)}$$

$$\alpha = \frac{\lambda_o}{M - \lambda_o \cdot T_S(1 - P_L)}$$

- A partir de las expresiones anteriores se puede “construir” un algoritmo para estimar α en función del tráfico ofrecido al sistema, y calcular así la probabilidad de pérdida

Estimación de α y P_L - Algoritmo

IN: M, T_O, S, ε

OUT: P_L

(1) $P_L = 0$

(2) Do

(3) $P_L^{aux} = P_L$

(4)
$$a = \frac{T_O}{M - T_O(1 - P_L^{aux})}$$

(5)
$$P_L = \binom{M-1}{S} \frac{a^S}{\sum_{i=0}^S \binom{M-1}{i} a^i}$$

(6) While ($\text{abs}(P_L - P_L^{aux}) > \varepsilon$)