Tráfego de Serviços Orientados a Conexões

- **Caracterização do tráfego**
- **Análise de sistemas de estados**
- **Análise de tráfego em sistemas de perda**
- **Análise de tráfego em sistemas de atraso**
- **Bloqueio em sistemas de andares múltiplos**

Intencionalmente em branco

Caracterização do tráfego

Intensidade de tráfego

A utilização de um nó ou de uma ligação de uma rede que suporta um serviço orientado a conexões é determinada por dois fatores:

– chegada de chamadas;

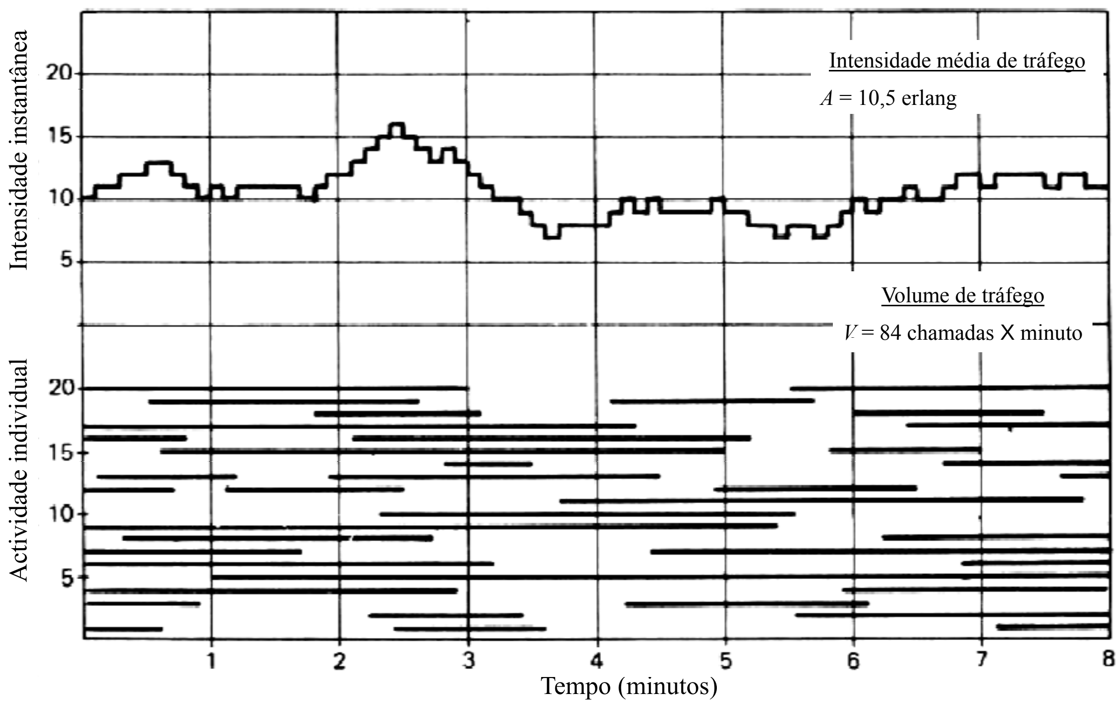
– duração de chamadas.

O tráfego toma em consideração estas componentes, podendo quantificar-se através da intensidade em unidades de erlang (segundo A.K. Erlang, pioneiro dinamarquês da teoria do teletráfego): se um sistema suporta A chamadas, diz-se que transporta A erlangs de tráfego (note-se que esta unidade é adimensional).

De um modo geral, a intensidade de tráfego num intervalo de tempo pode exprimir-se através da razão entre o volume de tráfego V - somatório das durações d_i de todas as chamadas n ocorridas no intervalo - e o período de tempo T:

$$\boxed{A = \frac{V}{T} = \frac{\sum_{i=1}^{n} d_i}{T}}$$

Caracterização do tráfego



Atividade de tráfego de um serviço orientado a conexões num nó ou numa ligação

Caracterização do tráfego

Por outro lado, a taxa média de chegada de chamadas λ e a duração média das chamadas d_m são dadas por:

$$\boxed{\lambda = \frac{n}{T}} \qquad \boxed{d_m = \frac{\sum_{i=1}^{n} d_i}{n}}$$

Logo, a intensidade de tráfego virá:

$$\boxed{A = \lambda\, d_m}$$

A intensidade de tráfego é apenas uma medida da utilização média durante um período de tempo e não reflete a relação entre chegada e duração de chamadas: muitas chamadas de curta duração podem produzir a mesma intensidade de tráfego que poucas chamadas longas.

Nos serviços orientados a conexões, é geralmente suficiente caracterizar o tráfego apenas em termos de intensidade de tráfego, podendo por vezes ser necessário considerar os padrões de chegada de chamadas ou as distribuições de durações.

Caracterização do tráfego

Variação do tráfego – exemplo do tráfego de voz

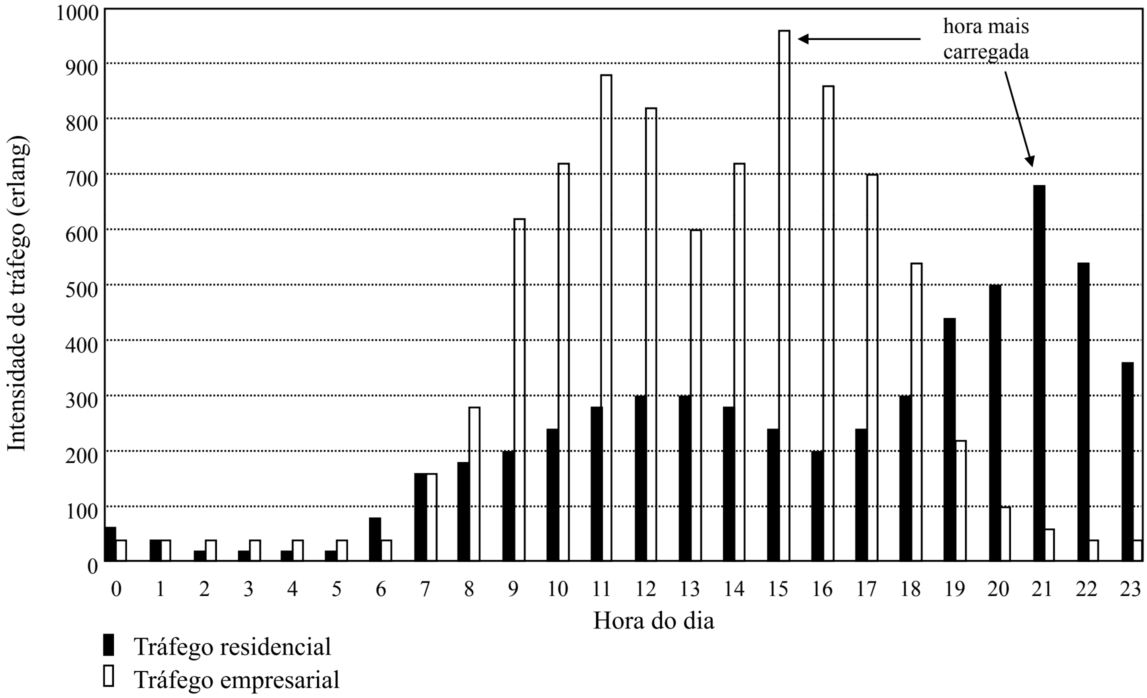
O dimensionamento do tráfego de serviços de voz é geralmente efetuado em termos da atividade durante a **hora mais carregada** do dia, procurando-se um compromisso entre dois extremos:

- projetar para a utilização média total, que inclui tempos noturnos virtualmente sem tráfego;
- projetar para picos de curta duração resultantes de acontecimentos esporádicos (fim de programas de TV, concursos por telefone, etc.).

É habitual considerar dois tipos de utilizadores no que respeita ao tráfego telefónico:

- **residenciais**, com utilização por acesso de 0,05 a 0,1 erlang na hora mais carregada, e tempos médios de duração de chamadas de 3 a 4 minutos; a hora mais carregada ocorre ao fim do dia/noite;
- **empresariais**, com utilizações superiores em horas correspondentes ao fim da manhã/meio da tarde.

Caracterização do tráfego



Intensidade de tráfego de voz em função da hora do dia

Caracterização do tráfego

Sistemas de perda e sistemas de espera

Há dois grandes tipos de sistemas usados serviços orientados a conexões:

– **sistema de perda**, em que as chamadas que excedem em qualquer instante a capacidade máxima do sistema são rejeitadas, podendo ou não regressar em novas tentativas;

– **sistemas de espera**, em que as chamadas em excesso são colocadas em filas de espera até estarem disponíveis os recursos necessários para as suportar (excepcionalmente, se a fila encher, o sistema passa a comportar-se como na situação de perda).

Nos sistemas de perda, o tráfego transportado é sempre menor ou igual ao tráfego oferecido ao sistema pelas fontes, enquanto nos sistemas de espera estas duas quantidades são iguais, desde que a média a longo prazo do tráfego oferecido seja inferior à capacidade máxima do sistema e as filas de espera tenham profundidade suficiente para absorver os picos.

Os sistemas telefónicos clássicos são sistemas de perda. Contudo, são cada vez mais comuns sistemas de espera, sendo exemplos significativos os centros de atendimento (*call centres*) e certos tipos de redes de comunicação móvel.

Caracterização do tráfego

Chegada de chamadas

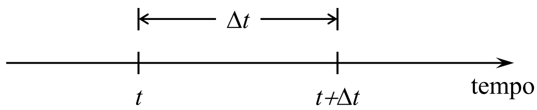
- *Processo de Poisson*

O processo de Poisson é o mais usado para modelizar o tráfego em sistemas de comunicações (sistemas telefónicos e redes de dados).

Assume o princípio básico de independência de chegada de chamadas entre fontes, podendo formular-se através da probabilidade de chegada num intervalo infinitesimal $[t, t+\Delta t]$, introduzindo uma constante de proporcionalidade λ:

$$\boxed{\begin{array}{l} \text{P (uma chegada no intervalo } [t, t+\Delta t]) = \lambda\Delta t \\ \text{P (nenhuma chegada no intervalo } [t, t+\Delta t]) = 1 - \lambda\Delta t \end{array}}$$

Note-se que um processo de Poisson não tem memória: um acontecimento num intervalo Δt é independente de acontecimentos nos intervalos anteriores ou futuros.



Intervalo temporal usado para definir um processo de Poisson

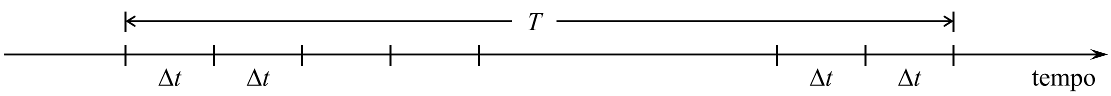
Caracterização do tráfego

- *Probabilidades de chegada de chamadas*

Para um intervalo finito T, a probabilidade $p_T(k)$ de k chegadas é dada por:

$$\boxed{\mathrm{p_T}(k) = (\lambda T)^k \,\mathrm{e}^{-\lambda T} / k!} \qquad k = 0, 1, 2, \ldots$$

DEMONSTRAÇÃO...



Derivação da distribuição de Poisson: $T = m\,\Delta t$, $\Delta t \rightarrow 0$

Esta é a chamada distribuição de Poisson, caracterizada por:

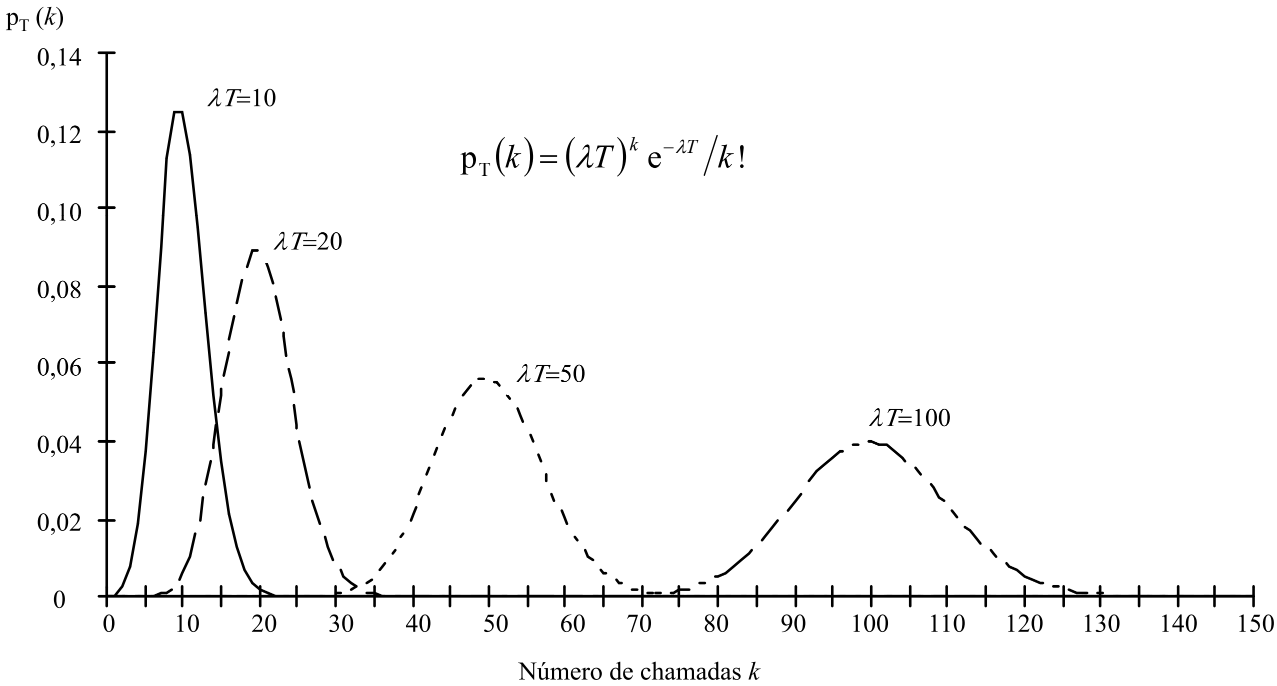
$\boxed{\text{Média: E}(k) = \lambda T}$ $\boxed{\text{Variância: } \sigma_k^2 = \lambda T}$ $\boxed{\text{Desvio padrão: } \sigma_k = \sqrt{\lambda T}}$ *DEMONSTRAÇÃO...*

Algumas conclusões

– o parâmetro λ é afinal a taxa média de chegada de chamadas ($\lambda = \mathrm{E}(k)/T$);

– como a média cresce mais rapidamente do que o desvio padrão, para valores grandes de λT a distribuição torna-se mais compacta em torno da média - logo o número n (aleatório) de chamadas chegadas num intervalo grande T conduz a uma boa estimativa de λ através de $\lambda \approx n / T$.

Caracterização do tráfego



Distribuição de Poisson

Caracterização do tráfego

- *Combinação de fontes independentes*

Na análise de tráfego em sistemas de comunicação é muitas vezes necessário combinar o tráfego gerado por fontes independentes.

Assumindo n fontes de tráfego de Poisson com taxas arbitrárias λ_i (i=1, 2, ... n), a definição de processo de Poisson conduz directamente a que o agregado das fontes é também um processo de Poisson com taxa:

$$\boxed{\lambda = \sum_{i=1}^{n} \lambda_i}$$

Exprime-se este facto dizendo que a soma de processos de Poisson é conservativa relativamente à distribuição: isto é, a soma retém a propriedade de Poisson.

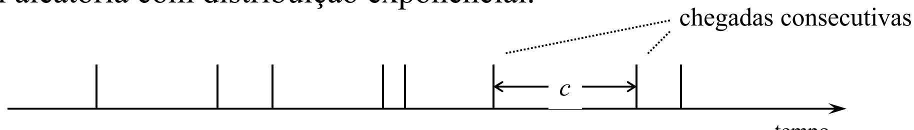
Note-se que a probabilidade de ocorrência de uma chegada no intervalo [t, t+Δt] é a soma das probabilidades de chegada de cada uma das n fontes independentes (assume-se que Δt é tão pequeno que não ocorrem múltiplas chamadas). Ou seja:

$$\boxed{\text{P (uma chegada no intervalo}\,[t, t+\Delta t]) = \sum_{i=1}^{n} \left(\lambda_i \Delta t\right) = \left(\sum_{i=1}^{n} \lambda_i\right) \Delta t = \lambda \Delta t}$$

Caracterização do tráfego

- <u>*Tempos entre chegadas de chamadas*</u>

Num processo de Poisson, o tempo c entre acontecimentos consecutivos é uma variável aleatória com distribuição exponencial.



Tempo entre chegadas de chamadas

De facto, a probabilidade de o tempo entre chamadas c ser $\leq t$ é:

$$\boxed{\mathrm{P}(c \leq t) = 1 - \mathrm{P}(c > t) = 1 - \mathrm{p_t}(0) = 1 - \mathrm{e}^{-\lambda t}}$$

A função de densidade de probabilidade virá então, por derivação:

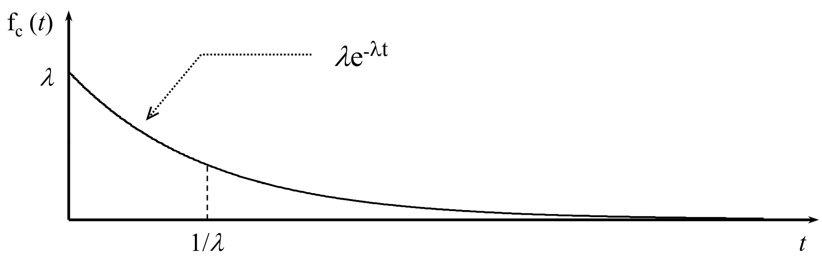
$$\boxed{\mathrm{f_c}\,(t) = \lambda\,\mathrm{e}^{-\lambda t}}$$

O valor médio e a variância desta distribuição exponencial são:

$$\boxed{\text{Média: } \mathrm{E}(c) = \int_0^\infty t\,\mathrm{f}_c(t)\,\mathrm{d}t = 1/\lambda} \qquad \boxed{\text{Variância: } \sigma_c^2 = 1/\lambda^2}$$

Caracterização do tráfego

O valor médio do tempo entre chegadas consecutivas atrás obtido é o esperado, uma vez que, sendo a taxa de chegadas do processo de Poisson λ, então o tempo entre chegadas será $1/\lambda$.



Distribuição exponencial de tempos de chegada entre chamadas

Note-se que uma variável aleatória exponencial é a única variável aleatória que tem a propriedade já referida de não memória: em termos de chegada de chamadas, a distribuição temporal dos próximos eventos é a mesma independentemente do tempo presente.

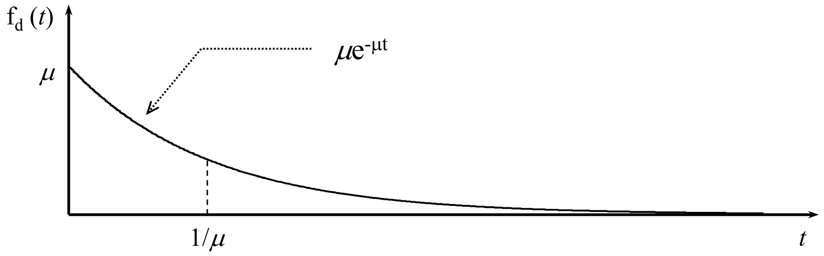
Caracterização do tráfego

Duração de chamadas

- *Distribuição de durações*

Para as chamadas entradas no sistema, é habitual assumir uma distribuição exponencial da duração d, com valor médio $E(d) = 1/\mu$:

Distribuição cumulativa de probabilidades	$P(d \leq t) = 1 - e^{-\mu t}$
Função de densidade de probabilidade	$f_d(t) = \mu\, e^{-\mu t}$



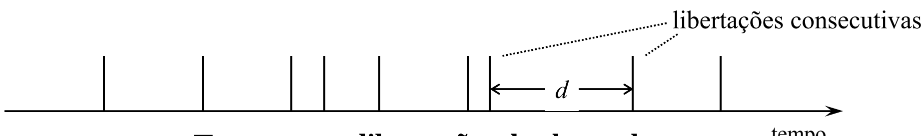
Distribuição exponencial de duração de chamadas

Esta hipótese tem sido considerada adequada para modelizar o tráfego, e tem a vantagem de apresentar a característica já referida de "sem memória": a probabilidade de uma chamada terminar num intervalo Δt é independente da sua duração anterior.

Caracterização do tráfego

- *Probabilidade de libertação de chamadas*

Considere-se um sistema constituído por uma fila de espera com um certo número de chamadas à espera de serem atendidas e uma única saída (uma única chamada de cada vez no sistema com possibilidade de ser libertada). Marquem-se à saída os pontos correspondentes à libertação de cada chamada.



Tempo entre libertações de chamadas

Uma vez que os intervalos entre libertação de chamadas têm distribuição exponencial, o processo de libertação é análogo ao de chegada de chamadas, podendo concluir-se que os tempos de libertação de chamadas representam eles próprios um processo de Poisson:

$$P(\text{uma libertação no intervalo } [t, t+\Delta t]) = \mu\Delta t$$

$$P(\text{nenhuma libertação no intervalo } [t, t+\Delta t]) = 1 - \mu\Delta t$$

Caracterização do tráfego

De igual modo, para um intervalo finito T, a probabilidade $p_T(k)$ de k libertações virá:

$$\boxed{\mathrm{p_T}(k)=(\mu T)^k\,\mathrm{e}^{-\mu T}/k!} \qquad k = 0, 1, 2, \ldots$$

Logo:

$\boxed{\text{Média: } \mathrm{E}(k) = \mu T}$ $\boxed{\text{Variância: } \sigma_k^2 = \mu T}$ $\boxed{\text{Desvio padrão: } \sigma_k = \sqrt{\mu T}}$

O parâmetro μ é, portanto:

$$\boxed{\mu = \text{taxa média de libertação de chamadas} = \frac{1}{\text{duração média de chamadas}}}$$

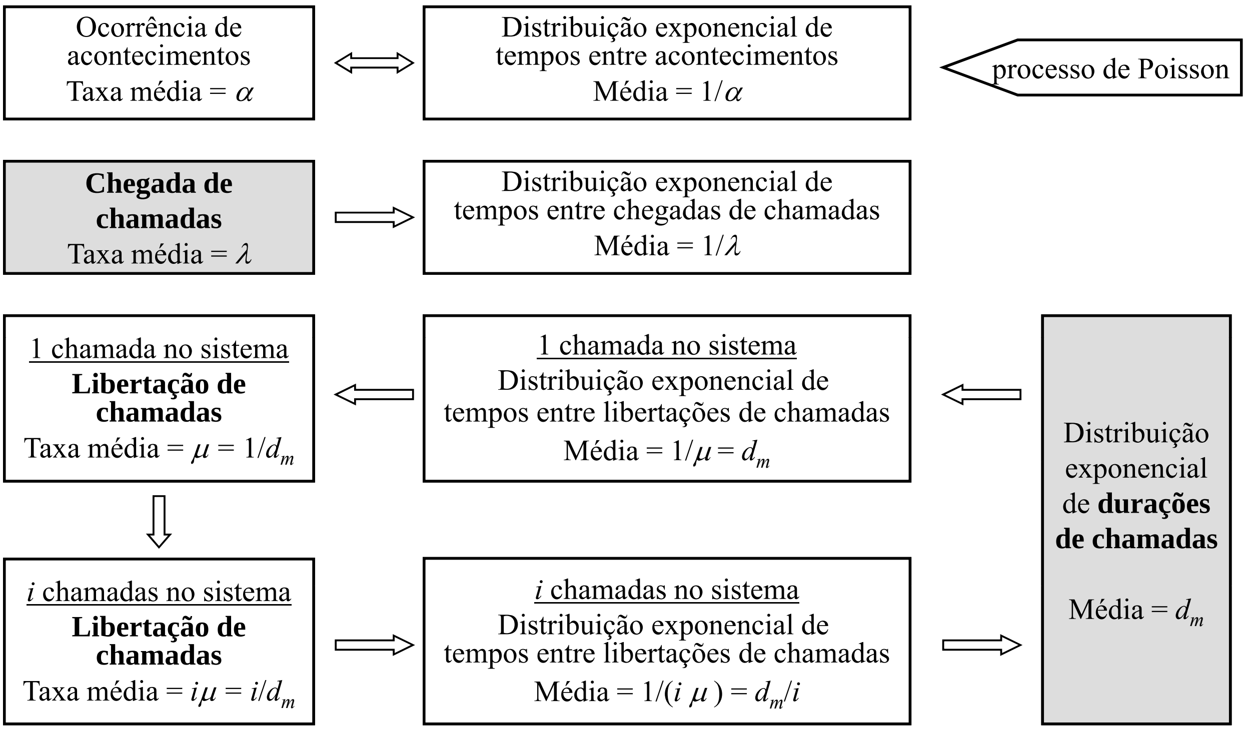
No caso mais geral de existirem no sistema i chamadas activas com possibilidade de serem libertadas de forma independente entre si, o processo de libertação resultante continuará a ser de Poisson (propriedade conservativa), sendo:

$$\boxed{\mu_i = i\mu}$$

Ou seja:

$$\boxed{\mu = \text{taxa média de libertação de chamadas} = \frac{\text{número de chamadas activas}}{\text{duração média de chamadas}}}$$

Caracterização do tráfego



Análise de sistemas de estados

Probabilidades de permanência nos estados do sistema

Considere-se um sistema de estados definidos pelo número *i* de chamadas em curso.

Para o sistema estar no estado *i* no tempo $t+\Delta t$, poderá, no tempo *t*:

- estar no estado i-1 e chegar uma chamada (probabilidade $\lambda_{i-1}\Delta t$) ou
- estar no estado i+1 e libertar-se uma chamada (probabilidade $\mu_{i+1}\Delta t$) ou
- estar no estado *i* e não ocorrer nenhuma das situações anteriores (probabilidade $1-\lambda_i\Delta t-\mu_i\Delta t$).

Os estados do sistema e as transições entre eles poderão ser modelizados através de um diagrama simplificado:

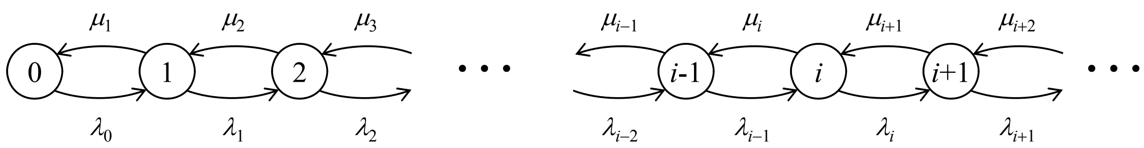


Diagrama de estados de um sistema

Análise de sistemas de estados

A probabilidade $[i]_{t+\Delta t}$ de o sistema estar no estado i no tempo $t+\Delta t$ virá então:

$$[i]_{t+\Delta t} = [i]_t(1-\lambda_i\Delta t-\mu_i\Delta t)+[i+1]_t\,\mu_{i+1}\Delta t+[i-1]_t\,\lambda_{i-1}\Delta t$$

Rearranjando os termos:

$$\frac{[i]_{t+\Delta t}-[i]_t}{\Delta t} = -[i]_t(\lambda_i+\mu_i)+[i+1]_t\,\mu_{i+1}+[i-1]_t\,\lambda_{i-1}$$

Fazendo $\Delta t \to 0$, a equação anterior poderá passar para a forma diferencial:

$$\boxed{\frac{\mathrm{d}[i]_t}{\mathrm{d}t} = -[i]_t(\lambda_i+\mu_i)+[i+1]_t\,\mu_{i+1}+[i-1]_t\,\lambda_{i-1}}$$

Note-se que $i = 0$ é um caso especial em que:

- $\mu_0 = 0$ (não há libertação de chamadas no estado 0);
- $\lambda_{-1} = 0$ (não existe o estado −1).

Logo:

$$\boxed{\frac{\mathrm{d}[0]_t}{\mathrm{d}t} = -[0]_t\,\lambda_0+[1]_t\,\mu_1}$$

Análise de sistemas de estados

Condição de equilíbrio

Se as taxas de chegada e libertação de chamadas forem constantes, o sistema atingirá ao fim de algum tempo uma situação de equilíbrio em que as probabilidades de permanência nos estados não variam com o tempo (d[i]/dt = 0). Logo:

$$[i](\lambda_i + \mu_i) = [i+1]\mu_{i+1} + [i-1]\lambda_{i-1} \qquad\qquad [0]\lambda_0 = [1]\mu_1$$

Calculando sucessivamente para i = 1, 2, ... obtém-se a expressão geral:

$$\boxed{[i-1]\lambda_{i-1} = [i]\mu_i}$$

Ou seja:

(probabilidade de estar no estado i-1) x (taxa de chegada de chamadas no estado i-1) =
(probabilidade de estar no estado i) x (taxa de libertação de chamadas no estado i)

Escrevendo a expressão anterior sucessivamente para i=1, 2, ... chega-se facilmente a:

$$\boxed{[i]/[0] = \prod_{k=0}^{i-1} \lambda_k \Big/ \prod_{k=1}^{i} \mu_k}$$

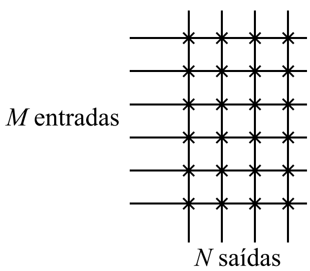
Para um sistema de N estados, a probabilidade [0] obtém-se invocando a condição de normalização:

$$\sum_{k=0}^{N} [k] = 1$$

Análise de tráfego em sistemas de perda

Sistema de acesso total

Um sistema de acesso total é um sistema de M fontes e N canais, em que qualquer fonte livre tem acesso a qualquer canal livre independentemente do estado do sistema. O exemplo mais simples é uma matriz de comutação de M x N pontos de cruzamento.



Sistema de acesso total

Em cada estado i ($0 \leq i \leq N$) o sistema é caracterizado pelos seguintes parâmetros:

- taxa de chegada de chamadas $\lambda_i = (M-i)\lambda$ em que λ é a taxa de chegada de cada fonte livre;
- taxa de libertação de chamadas $\mu_i = i\mu$, em que $1/\mu$ é a duração média das chamadas.

Análise de tráfego em sistemas de perda

Distribuição de Erlang-B (*M*»*N*)

Neste caso, a taxa de chegadas λ_i e o tráfego oferecido ao sistema *A* são constantes, isto é, independentes do número *i* ($0 \leq i \leq N$) de chamadas em curso; logo:

$$\lambda_i \cong M\lambda = \lambda_S \qquad\qquad A \cong M\lambda/\mu = \lambda_S/\mu$$

onde λ_S é a taxa total de chegada de chamadas ao sistema.

Assim a condição de equilíbrio vem

$$[i]/[0] = \prod_{k=0}^{i-1} \lambda_k \Big/ \prod_{k=1}^{i} \mu_k = (M\lambda)^i / (i!\,\mu^i) = A^i/i!$$

Efectuando a normalização $\sum_{k=0}^{N}[k]=1$, facilmente se obtém $[0]=1\Big/\sum_{k=0}^{N} A^k/k!$, pelo que

$$\boxed{[i] = \frac{A^i}{i!}\Big/\sum_{k=0}^{N}\frac{A^k}{k!}} \qquad 0 \leq i \leq N$$

Esta equação exprime a distribuição de probabilidade dos estados do sistema (distribuição de Erlang-B), sendo usada correntemente desde que a hipótese de taxa de chamadas seja aproximadamente independente do número de chamadas activas.

Análise de tráfego em sistemas de perda

No caso de o tráfego oferecido conduzir a um reduzido número de chamadas *i* em curso relativamente ao número de canais *N*, os últimos termos do somatório da expressão da distribuição de Erlang vão-se tornando desprezáveis, podendo então utilizar-se a aproximação:

$$\boxed{\mathrm{e}^A \approx \sum_{k=0}^{N}\frac{A^k}{k!}}$$

Substituindo:

$$\boxed{[i] = \frac{A^i}{i!}\mathrm{e}^{-A}}$$

Esta equação evidencia que a distribuição de Erlang-B é uma forma truncada da distribuição de Poisson.

Há uma interpretação simples para esta relação entre distribuições: sendo o tráfego oferecido $A=\lambda_S/\mu$, pode concluir-se que a probabilidade [*i*] de o sistema ter *i* chamadas activas num certo instante é igual à probabilidade de terem chegado *i* chamadas no período de tempo $1/\mu$ (duração média das chamadas) imediatamente anterior (no caso de durações constantes iguais a $1/\mu$, esta relação é óbvia).

Análise de tráfego em sistemas de perda

- *Congestão temporal*

Chama-se congestão temporal E à probabilidade de o sistema estar no estado N, isto é

$$\boxed{E = [N] = \frac{A^N}{N!}\Big/\sum_{k=0}^{N}\frac{A^k}{k!}}$$

Esta expressão é conhecida como a fórmula de Erlang de tipo B para a probabilidade de bloqueio em sistemas de perda, constituindo um dos resultados fundamentais da teoria de tráfego.

Os valores da fórmula de Erlang-B podem ser calculados por iteração, recorrendo à seguinte expressão

$$\boxed{E_N = \frac{A\,E_{N-1}}{N + A\,E_{N-1}}} \qquad E_0 = 1$$

Em alternativa, poderão ser utilizadas tabelas ou gráficos, os quais permitem obter valores com grande rapidez e com precisão suficiente para aplicações de engenharia de teletráfego.

Análise de tráfego em sistemas de perda

- *Congestão de chamadas*

A congestão de chamadas B é a probabilidade de uma chamada chegar e encontrar o sistema totalmente ocupado (bloqueado). Logo, para um intervalo T

$$\boxed{B = \frac{\text{número esperado de chamadas chegadas com o sistema no estado } N}{\text{número total esperado de chamadas chegadas}}}$$

A congestão de chamadas pode exprimir-se através de uma razão de taxas de chegadas, pesadas pelas probabilidades de ocorrência dos respectivos estados

$$\boxed{B = \lambda_N [N]\Big/\sum_{k=0}^{N}\lambda_k [k]}$$

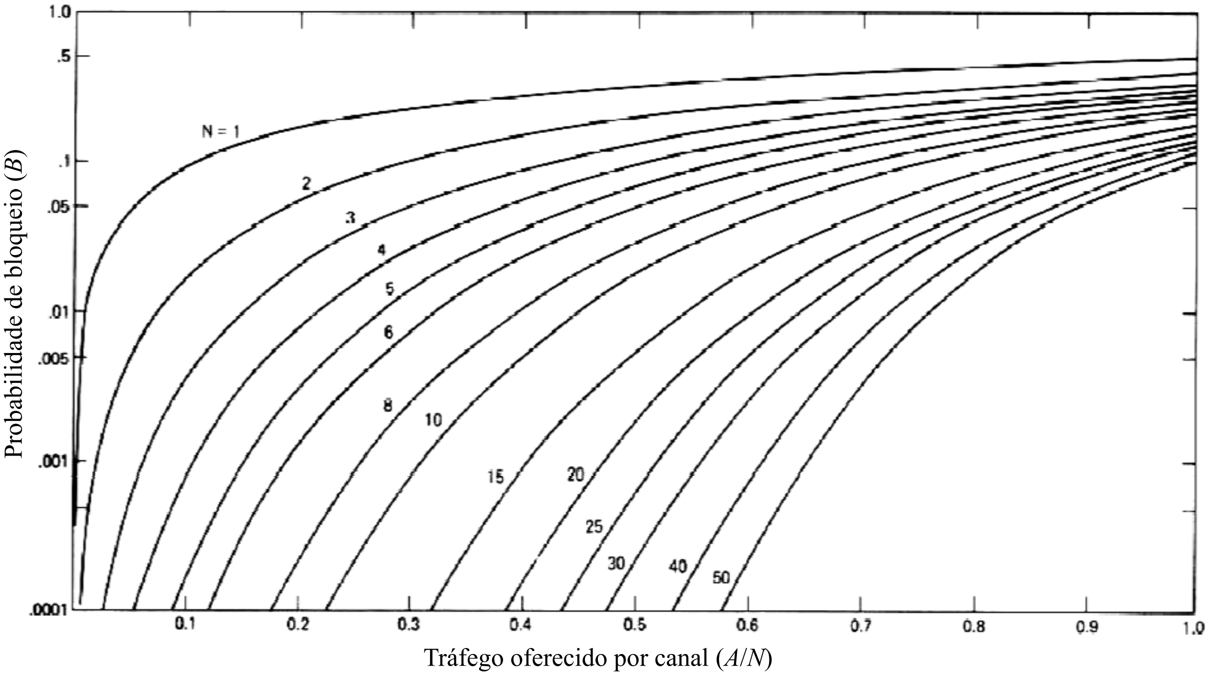
Como a taxa de chegada de chamadas é constante e sendo $\sum_{k=0}^{N}[k] = 1$ virá

$\boxed{B = [N]}$ $\boxed{B = E}$ congestão de chamadas = congestão temporal

A probabilidade de bloqueio permite relacionar directamente o tráfego oferecido A com o tráfego efectivamente transportado A_t, sendo a diferença o tráfego perdido:

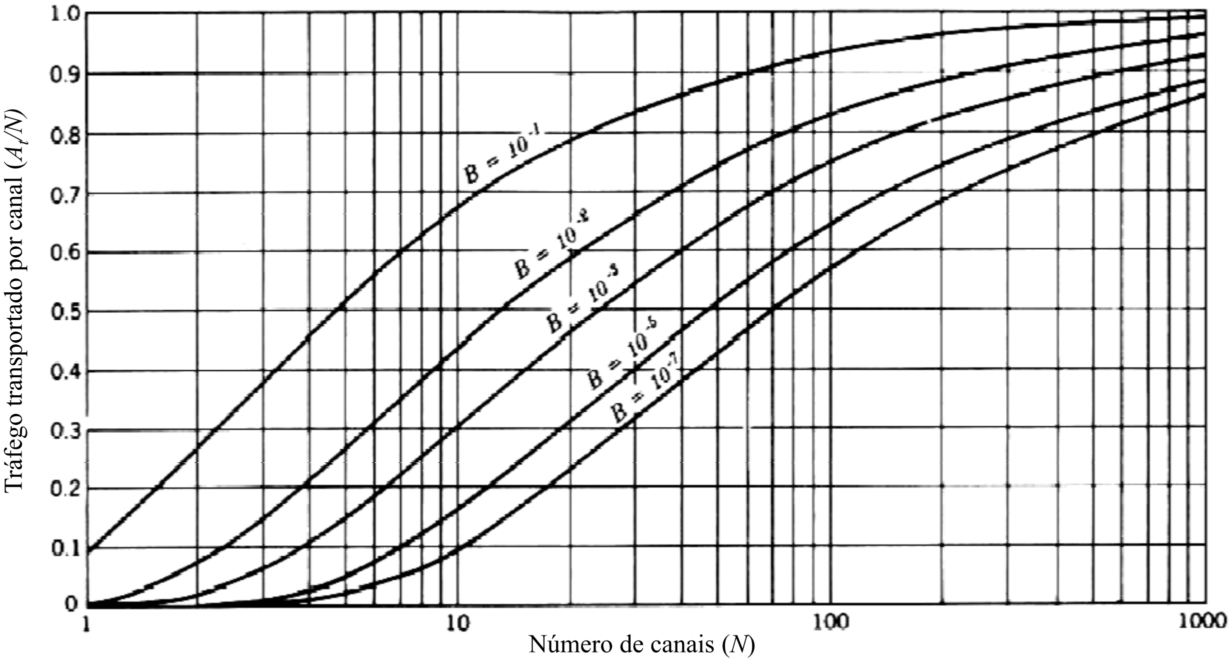
$$\boxed{A_t = A\,(1\text{-}B)}$$

Análise de tráfego em sistemas de perda



Probabilidades de bloqueio (*B*) de sistemas de perda na situação de Erlang, em função do tráfego oferecido por canal (*A*/*N*) e do número de canais (*N*)

Análise de tráfego em sistemas de perda



Tráfego transportado por canal (A_t/*N*) em sistemas de perda na situação de Erlang, em função do número de canais (*N*) e da probabilidade de bloqueio (*B*)

Análise de tráfego em sistemas de perda

Distribuição de Engset (*M*>*N*)

Corresponde ao caso geral, em que o número de fontes é superior ao número de canais, mas não tanto que se considere o tráfego oferecido ao sistema como constante.

Neste caso, a condição de equilíbrio vem

$$[i]/[0]=\prod_{k=0}^{i-1}\lambda_k \Big/ \prod_{k=1}^{i}\mu_k = \binom{M}{i}\left(\frac{\lambda}{\mu}\right)^i$$

Normalizando e calculando [0], obtém-se a distribuição de Engset para a probabilidade [*i*] do número de chamadas em curso

$$\boxed{[i]=\frac{\binom{M}{i}\left(\frac{\lambda}{\mu}\right)^i}{\sum_{k=0}^{N}\binom{M}{k}\left(\frac{\lambda}{\mu}\right)^k}} \qquad 0\le i\le N$$

Como é óbvio, a distribuição de Erlang-B é um caso particular da distribuição de Engset quando *M*»*N*.

Análise de tráfego em sistemas de perda

- *Congestão temporal*

A congestão temporal pode calcular-se através de

$$\boxed{E=[N]=\frac{\binom{M}{N}\left(\frac{\lambda}{\mu}\right)^N}{\sum_{k=0}^{N}\binom{M}{k}\left(\frac{\lambda}{\mu}\right)^k}}$$

- *Congestão de chamadas*

Pela definição

$$B=\lambda_N[N]\Big/\sum_{k=0}^{N}\lambda_k[k]=(M-N)\lambda[N]\Big/\sum_{k=0}^{N}(M-k)\lambda[k]$$

Logo

$$\boxed{B=\frac{\binom{M-1}{N}\left(\frac{\lambda}{\mu}\right)^N}{\sum_{k=0}^{N}\binom{M-1}{k}\left(\frac{\lambda}{\mu}\right)^k}}$$

Sistemas de Telecomunicações *Tráfego de Serviços Orientados a Conexões*

Análise de tráfego em sistemas de perda

A equação anterior exprime B em função da razão entre os parâmetros λ e μ, a qual deverá ser relacionados com o tráfego oferecido A, ou com a actividade média ρ de cada fonte, mais facilmente mensuráveis.

Sendo a taxa total λ_S de chegada de chamadas ao sistema igual ao produto do número médio de fontes livres pela taxa λ de chegada de cada fonte livre, o tráfego oferecido virá

$$A = \lambda_S / \mu = (\mathrm{M} - \mathrm{A_t})\lambda/\mu$$

em que A_t é o tráfego transportado (número médio de fontes ocupadas).

Introduzindo as relações $A_t = A(1-B)$ e $\rho = A/M$ obtém-se

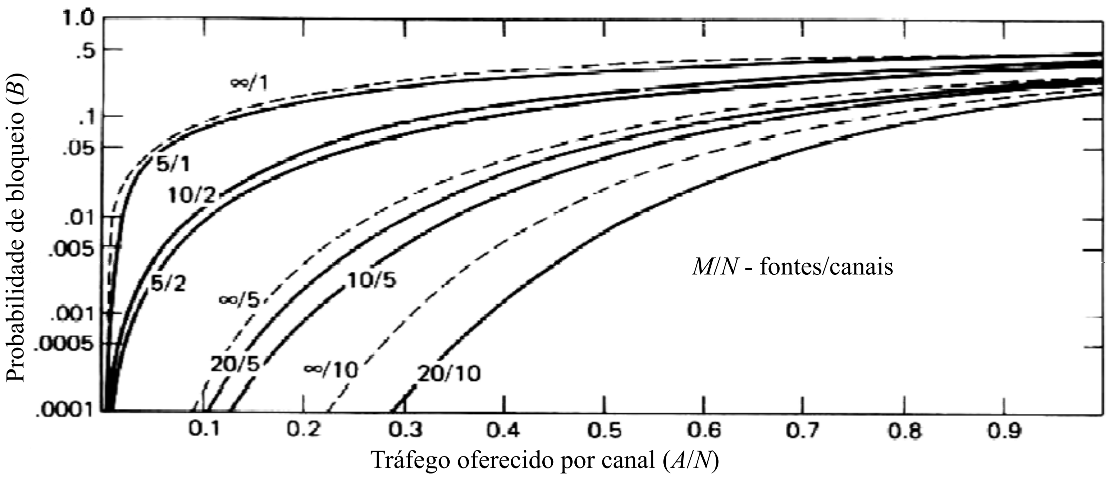
$$\boxed{\lambda/\mu = \frac{A}{M - A(1-B)} = \frac{\rho}{1-\rho(1-B)}} \qquad \boxed{A = \frac{M\,\lambda/\mu}{1+\lambda/\mu(1-B)}}$$

Embora esta expressão não estabeleça uma relação explícita entre λ/μ e A ou ρ, por incluir B, igualmente função de λ/μ, permite calcular B a partir de A ou ρ, por iteração.

Em alternativa, poderão ser utilizadas tabelas ou gráficos, como referido para a situação de Erlang.

Análise de tráfego em sistemas de perda

A figura seguinte apresenta valores da probabilidade de bloqueio para a distribuição de Engset, e também para a distribuição de Erlang, como caso particular daquela ($M=\infty$). Pode verificar-se que, quando se considera a situação de Erlang como aproximação, resultam estimativas pessimistas da probabilidade de bloqueio para um dado número de canais, ou o sobre-dimensionamento do sistema para uma dada probabilidade.



Probabilidades de bloqueio (B) de sistemas de perda na situação de Engset, em função do tráfego oferecido por canal (A/N) para diversas combinações de fontes e canais

Análise de tráfego em sistemas de perda

Distribuição binomial (Bernoulli) (*M*≤*N*)

Neste caso só poderão estar ocupadas, no máximo, *M* canais, pelo que

$$\boxed{[i]=\frac{\binom{M}{i}\left(\frac{\lambda}{\mu}\right)^i}{\sum_{k=0}^{M}\binom{M}{k}\left(\frac{\lambda}{\mu}\right)^k}} \qquad 0\le i\le M$$

O denominador é a expansão binomial de $(1+\lambda/\mu)^M$, pelo que se se substituir e fizer

$$\boxed{p=\lambda/(\lambda+\mu)}$$

obtém-se a distribuição binomial (de Bernoulli)

$$\boxed{[i]=\binom{M}{i}p^i(1-p)^{M-i}}$$

Note-se que desta expressão se pode concluir que $p=\lambda/(\lambda+\mu)$ é precisamente a probabilidade de ocupação de cada fonte.

Análise de tráfego em sistemas de perda

- *Congestão temporal*

Se $M<N$ não há congestão temporal (E=0), uma vez que o estado *N* não é atingido mesmo que as *M* fontes estejam todas activas (*M* chamadas).

Se $M=N$, obtém-se directamente da distribuição binomial a seguinte expressão para a congestão temporal:

$$\boxed{E=[N]=p^N}$$

Este resultado seria de esperar, uma vez que se *p* é a probabilidade de ocupação de cada fonte, p^N será a probabilidade de ocupação das *N* fontes/canais.

- *Congestão de chamadas*

Neste caso, não haverá congestão de chamadas (B=0), mesmo na situação de $M=N$, uma vez que não haverá novas chamadas quando o sistema está no estado *N* com todas as fontes ocupadas.

O tráfego transportado e o oferecido são iguais ($A_t=A$) e a probabilidade de ocupação de cada fonte coincide com a respectiva actividade média ($p=\rho$).

Análise de tráfego em sistemas de espera

Distribuição de Erlang-C

Neste caso, assume-se que as chamadas que não têm recursos disponíveis são colocadas numa fila de espera com profundidade infinita, pelo que

$$\lambda_i = \lambda_S \qquad \begin{array}{ll} \mu_i = i\mu, & 1 \le i \le N \\ \mu_i = N\mu, & N < i \end{array}$$

λ_S - taxa total de chegada de chamadas
$d_m = 1/\mu$ - duração média das chamadas

Aplicando a condição de equilíbrio anteriormente obtida, chega-se à fórmula de Erlang-C para a probabilidade de uma chamada sofrer atraso num sistema de espera

$$\boxed{\mathrm{P}(\Delta > 0) = \frac{\dfrac{A^N}{N!}}{\left(\displaystyle\sum_{k=0}^{N} \frac{A^k}{k!}\right)(1-\alpha) + \dfrac{A^N}{N!}\alpha} = \frac{B}{1-\alpha+B\alpha}}$$

N - número de canais
$A = \lambda_S/\mu$ - tráfego oferecido ao sistema
$\alpha = \mathrm{A}/N$ - tráfego oferecido por canal ($\alpha < 1$)

B - probabilidade de bloqueio de um sistema de perda (Erlang-B)

$$\boxed{B = \frac{A^N}{N!} \Bigg/ \sum_{k=0}^{N} \frac{A^k}{k!}}$$

Análise de tráfego em sistemas de espera

Se a fila de espera for de tipo FIFO, o atraso Δ terá uma distribuição exponencial

$$\boxed{\mathrm{P}(\Delta > t) = \mathrm{P}(\Delta > 0)\mathrm{e}^{-(N-A)\frac{t}{d_m}}}$$

Integrando para todos os valores obtém-se o atraso médio de todas as chamadas

$$\boxed{\overline{\Delta} = \mathrm{P}(\Delta > 0)\frac{d_m}{(N-A)}}$$

Considerando apenas as chamadas que sofrem atraso, resulta um atraso médio

$$\boxed{\overline{\Delta}\Big|_{\Delta>0} = \frac{d_m}{(N-A)}}$$

Estas equações são normalmente utilizadas para dimensionar o número de canais (servidores) de um sistema, de forma a satisfazer objectivos de desempenho do tipo:

– uma certa probabilidade de as chamadas que chegam ao sistema sofrerem um atraso superior a um dado valor; ou
– um certo valor do atraso médio das chamadas.

Análise de tráfego em sistemas de espera

Fila de espera finita

Neste caso o sistema trata as chegadas de três formas diferentes:

- serviço imediato se um ou mais canais estiver livre;
- serviço com atraso se todos os canais estiverem ocupados e menos de *L* pedidos na fila de espera;
- bloqueio se a fila de espera estiver preenchida com *L* pedidos pendentes.

Esta situação corresponde ao caso geral, mais realista, de um sistema com recursos finitos.

Pode demonstrar-se que, para este sistema, a probabilidade de atraso é dada por

$$\mathrm{P}(\varDelta > 0) = \frac{\dfrac{A^N}{N!}}{\left(\displaystyle\sum_{k=0}^{N-1} \frac{A^k}{k!}\right)\dfrac{1-\alpha}{1-\alpha^{L+1}} + \dfrac{A^N}{N!}}$$

Análise de tráfego em sistemas de espera

A probabilidade de bloqueio pode obter-se através de

$$B = \frac{\dfrac{A^N}{N!}\alpha^L \dfrac{1-\alpha}{1-\alpha^{L+1}}}{\left(\displaystyle\sum_{k=0}^{N-1} \frac{A^k}{k!}\right)\dfrac{1-\alpha}{1-\alpha^{L+1}} + \dfrac{A^N}{N!}} = \mathrm{P}(\varDelta > 0)\, \alpha^L \frac{1-\alpha}{1-\alpha^{L+1}}$$

O atraso médio de todas as chamadas pode ser calculado pela expressão

$$\overline{\varDelta} = \mathrm{P}(\varDelta > 0)\frac{d_m}{(N-A)} - \frac{L\,B\,d_m}{(N-A)}$$

As distribuições de Erlang-B e Erlang-C correspondem a casos extremos da fila de espera de dimensão *L:*

- se L=0 obtém-se a distribuição de Erlang-B;
- se $L \to \infty$ obtém-se a distribuição de Erlang-C.