

# Análise de Fila para Tráfego Multifractal utilizando Cálculo de Rede e Parâmetro de Escala Global

Flávio Henrique Teles Vieira, Lee Luan Ling

<sup>1</sup>Departamento de Comunicações, Faculdade de Engenharia Elétrica e de Computação  
Universidade Estadual de Campinas  
Albert Einstein, 400 - PO Box 6101 - 13.083-970, Campinas, SP, Brasil)

{flavio, lee}@decom.fee.unicamp.br

**Abstract.** *In this article, we derive a global scaling parameter for network multifractal traffic. Through this parameter, we use the effective bandwidth concept and statistical network calculus to propose performance bounds for high-speed networks. Our performance bound approach was evaluated by simulations with Internet and Ethernet traffic traces, verifying its efficiency in capturing the byte loss probability and the mean buffer occupation. The bounds obtained for the byte loss probability was shown to be tighter than that given by the large deviations theory.*

**Resumo.** *Neste artigo, derivamos um parâmetro de escala global para tráfego multifractal de redes. Por meio deste parâmetro, calculamos banda efetiva para um modelo multifractal geral e a inserimos no cálculo de rede estatístico, com o intuito de propor limitantes de desempenho de fila para redes. A abordagem proposta foi avaliada por simulações com tráfego real Internet e Ethernet, constatando sua eficiência em capturar a probabilidade de perda e a ocupação média de bytes no buffer. O limitante obtido para a probabilidade de perda de bytes se mostrou mais rígido do que o esperado pela teoria dos grandes desvios.*

## 1. Introdução

Probabilidade de perda e atraso de pacotes são medidas de desempenho fundamentais associadas a qualidade de serviço (QoS) em redes de computadores, como as redes TCP-IP e ATM. Assim, vários estudos tem sido realizados com o intuito de caracterizar o tamanho médio da fila e distribuição do número de pacotes no *buffer*. Para que, dessa forma, se consiga estabelecer limitantes para essas medidas de desempenho, tais como perda, atraso e 'backlog'. O conhecimento destes limitantes permite garantir a qualidade de serviço requerida pelos fluxos de tráfego.

R.L.Cruz obteve limitantes determinísticos utilizando o conceito de processo envelope linear (PEL), oriundo do cálculo de rede [1][2]. C.S.Chang derivou limitantes de desempenho tanto determinísticos quanto estatísticos usando o conceito de mínimo PEL e relacionando o mesmo com a banda efetiva do tráfego [3]. Com base no trabalho de Chang, Liyi Dai obteve limitantes de desempenho mais rígidos [4]. Nestes dois últimos trabalhos citados, podem ser inseridas equações de banda efetiva derivadas para alguns modelos de tráfego, ao invés de somente se trabalhar com processo envelope. A vantagem disso é que pode-se fazer estimativas de banda efetiva sem a consideração de todas as amostras de uma série ao se assumir um modelo para o tráfego. Outros trabalhos se

seguiram neste sentido, principalmente usando cálculo de rede, mas sem o uso de banda efetiva [5][6][7].

A teoria de banda efetiva teve início à uma década [8]. A banda efetiva de um fluxo de tráfego está relacionada com a taxa necessária fornecida a este fluxo para que se atenda determinado critério de QoS. O conceito de banda efetiva tem sido aplicado largamente e é visto como um método apropriado para controle de admissão e alocação de recursos em redes [9].

Algumas características do tráfego Internet, não consideradas por alguns modelos de tráfego, dificultam a estimativa de desempenho de rede utilizando modelos mais simplificados, como modelos Markovianos. Há muitos estudos que revelam a alta variabilidade do tráfego Internet, ou seja, o tráfego é 'bursty' em uma gama de escalas de tempo, em contraste da suposição de que rajadas de tráfego só existem em escalas curtas de tempo [10][11]. Foi mostrado que estas incidências de rajadas multiescalas tem um impacto significativo no desempenho das redes [10][11][12]. Modelos mais realistas foram propostos, inicialmente caracterizando a auto-similaridade do tráfego Internet e Ethernet [6]. Auto-similaridade assintótica de segunda ordem implica em longa-dependência (LRD) e vice-versa. O termo auto-similaridade se refere normalmente a processos assintoticamente auto-similares de segunda ordem ou monofractais [13]. O parâmetro de Hurst mede o grau de auto-similaridade do processo.

No entanto, para muitos processos de tráfego de rede, a energia em escala dos coeficientes wavelet ou os gráficos variância-tempo normalmente não demonstram comportamento linear. Muitos destes processos têm comportamento fractal com parâmetro de Hurst variado em diferentes escalas de tempo pequenas [14], ou seja, são multifractais. De fato, o desempenho de fila depende grandemente das irregularidades do tráfego em escalas de tempo pequenas devido à dinâmica complexa das redes de dados [15][16].

Neste trabalho, derivamos um fator de escala global para tráfego multifractal, equivalente ao parâmetro de Hurst. Através deste parâmetro, calculamos sua banda efetiva e a incorporamos no cálculo de rede [3]. Nosso objetivo é a obtenção de limitantes de desempenho de fila, neste caso, para probabilidade de perda e tamanho médio da fila. Mostramos que os limitantes de desempenho obtidos são mais precisos do que a teoria dos grandes desvios prega para tráfego monofractal.

O artigo está organizado da seguinte forma: na seção 2, discursamos sobre processos multifractais baseados em cascata multiplicativa. Na seção 3, derivamos o parâmetro de escala global para tráfego multifractal. Na seção 4, descrevemos o cálculo de banda usando o parâmetro de escala obtido. Na seção 5, mostramos como o conceito de banda efetiva se relaciona com o cálculo de rede estatístico. Na seção 6, são mostrados os testes realizados para validar a proposta de cálculo de limitantes de desempenho de fila. Por fim, na seção 7, apresentamos as conclusões obtidas.

## **2. Processos Multifractais**

O conceito de multifractal foi introduzido por Mandelbrot no contexto de turbulência nos anos 70. Desde então a teoria multifractal é usada em vários campos tais como processamento de imagem, geofísica, etc. O tráfego de redes ao ser considerado multifractal, significa que possui uma estrutura de forte dependência inerente, com incidência de rajadas em várias escalas [17] [13]. Estas características fazem com que o desempenho da

rede seja pior do que o considerado por modelos Gaussianos e de curta-dependencia. O conceito de multifractalidade se estende facilmente a processos estocásticos:

**Definição 1:** Um processo estocástico  $X(t)$  é multifractal se satisfaz a equação:

$$E(|X(t)|^q) = c(q)t^{\tau(q)+1} \quad (1)$$

onde  $T$  e  $Q$  são intervalos da reta real,  $\tau(q)$  e  $c(q)$  são funções no domínio  $Q$ . Além disso, supõe-se que  $T$  e  $Q$  possuam comprimentos positivos, e que  $0 \in T$ ,  $[0, 1] \subseteq Q$ .

Assim, descrevemos multifractalidade em termos de momentos em que  $\tau(q)$  é a função de escala e  $c(q)$  é o fator de momento de um processo multifractal. Se  $\tau(q)$  é linear em  $q$ , o processo é chamado monofractal; caso contrário, é multifractal. Para processos auto-similares com parâmetro de Hurst  $H$ , pode-se mostrar que  $\tau(q) = qH - 1$  e  $c(q) = E(|X(1)|^q)$ .

## 2.1. Cascatas Multiplicativas

A cascata binomial é um método para se obter um processo multifractal, que consiste de um procedimento iterativo no intervalo compacto  $[0, 1]$ . Sejam  $m_0$  e  $m_1$  (multiplicadores da cascata) dois números positivos cuja soma é 1. No estágio  $k = 0$  da cascata, obtemos a medida inicial  $\mu_0$  do processo com distribuição de probabilidade uniforme em  $[0, 1]$ . No estágio  $k = 1$ , a medida  $\mu_1$  distribui massa utilizando a distribuição uniforme, sendo,  $m_0$  no subintervalo  $[0, 1/2]$  e massa igual a  $m_1$  em  $[1/2, 1]$ . Em  $k = 2$ , o intervalo  $[0, 1/2]$  é subdividido em  $[0, 1/4]$  e  $[1/4, 1/2]$  e o mesmo acontece com intervalo  $[1/2, 1]$ , obtendo [18]:

$$\begin{aligned} \mu_2[0, 1/4] &= m_0 m_0 & \mu_2[1/4, 1/2] &= m_0 m_1 \\ \mu_2[1/2, 3/4] &= m_1 m_0 & \mu_2[3/4, 1] &= m_1 m_1 \end{aligned}$$

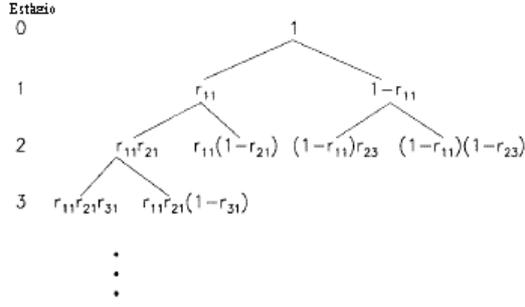
Considere o intervalo diádico  $[t, t + 2^{-k}]$  em que  $t = 0.\eta_1 \dots \eta_k = \sum_{i=1}^k \eta_i 2^{-i}$ . Sejam  $\varphi_0$  e  $\varphi_1$  as frequências relativas de 0's e 1's no desenvolvimento da cascata. A medida  $\mu$  no intervalo diádico é dada por:

$$\mu[t, t + 2^{-k}] = m_0^{k\varphi_0} m_1^{k\varphi_1} \quad (2)$$

Este processo preserva a massa dos intervalos diádicos em cada estágio, por isso é chamado de cascata conservativa ou microcanônica. Em cada estágio da cascata os intervalos podem ser divididos em intervalos de  $b > 2$  de tamanho igual, este processo é definido como cascata multinomial. Se os multiplicadores usados têm um valor fixo para  $m_0$  e  $b = 2$ , então a cascata multiplicativa é binomial determinística com função de escala:  $\tau(q) = -\log_2(m_0^q + m_1^q) + 1$  [18].

Ao se permitir que os multiplicadores da cascata  $r$  sejam variáveis aleatórias independentes em  $[0, 1]$ , com densidade de probabilidade  $f_R(x)$ , obtém-se uma estrutura mais geral do que a determinística em que os multiplicadores são valores fixos (Figura 1). Desta forma, o processo multifractal obtido terá no estágio  $k$  da cascata no intervalo diádico de comprimento  $\Delta t_k = 2^{-k}$ , que começa em  $t = 0.\eta_1 \dots \eta_k = \sum_{i=1}^k \eta_i 2^{-i}$  a seguinte medida  $\mu$ :

$$\mu(\Delta t_k) = R(\eta_1).R(\eta_1, \eta_2), \dots, R(\eta_1, \dots, \eta_k) \quad (3)$$



**Figura 1. Processo de Construção da Cascata**

onde  $R(\eta_1, \dots, \eta_i)$  é o multiplicador no estágio  $i$  da cascata. Uma vez que os multiplicadores  $r$  são independentes e identicamente distribuídos (i.i.d), pode-se demonstrar que a medida  $\mu$  satisfaz a seguinte relação de escala:

$$E(\mu(\Delta t_k)^q) = (E(R)^q)^k = \Delta t_k^{-\log_2 E(R^q)} \quad (4)$$

Esta relação define um processo multifractal com função de escala  $\tau(q) = -\log_2 E(R^q)$ .

## 2.2. Modelo Multifractal Geral

Os dados de tráfego reais apresentam suas propriedades multifractais caracterizadas pela função de escala  $\tau(q)$  e o fator de momento  $c(q)$ . Assim, um modelo multifractal geral deve capturar estas duas propriedades multifractais. Isto pode ser obtido pelo produto de uma cascata e uma variável aleatória i.i.d positiva  $Y$ . A variável  $Y$  é independente da medida da cascata  $\mu(\Delta t_k)$ , então a série obtida denotada por  $X(\Delta t_N)$  satisfaz a seguinte equação:

$$E(X(\Delta t_N)^q) = E(Y^q)E(\mu(\Delta t_N)^q) = E(Y^q)\Delta t_N^{\tau_0(q)} \quad (5)$$

Comparando-se (5) e (4), pode-se notar que as variáveis  $R$  e  $Y$  são tais que:

$$-\log_2(E(R^q)) = \tau_0(q) \quad (6)$$

$$E(Y^q) = c(q) \quad (7)$$

Este modelo multifractal, descrito em [19], pode ser visto como o produto da taxa de pico do fluxo  $Y$ , pela medida de rajada  $\mu(\Delta t_N)$  na escala de tempo aplicada  $\Delta t_N$ .

A medida  $\mu(\Delta t_N)$  possui valor pequeno, devido ao fato de ser o produto de  $N$  multiplicadores  $0 < r < 1$ . Assim, multiplicamos a medida da cascata por  $2^N$ . Desde que  $E(\mu(\Delta t_N)) = 2^{-N}$ , isto normaliza o processo, fazendo com que obtenhamos média 1. Para a unidade de intervalo de tempo também consideramos unitária no estágio  $N$  da cascata ao invés de  $\Delta t_N = 2^{-N}$ . Assim, podemos obter [19]:

$$E(X(\Delta t_N)^q) = E(Y^q)2^{N(q+\log_2 E(R^q))} \Delta t_0^{-\log_2 E(R^q)} \quad (8)$$

A partir disto, as variáveis  $R$  e  $Y$ , devem ser escolhidas de forma a atender as seguintes equações:

$$-\log_2(E(R^q)) = \tau_0(q) \quad (9)$$

$$\log E(Y^q) = \log c(q) - (q + \log_2(E(R^q)))N \log 2 \quad (10)$$

A função de escala pode ser precisamente modelada ao assumirmos que  $R$  é uma variável aleatória em  $[0,1]$  com distribuição beta simétrica  $\text{Beta}(\alpha,\alpha)$  com  $\alpha > 0$ :

$$\tau_0(q) = \log_2 \frac{\Gamma(\alpha)\Gamma(2\alpha + q)}{\Gamma(2\alpha)\Gamma(\alpha + q)} \quad (11)$$

onde  $\Gamma(\cdot)$  corresponde a função Gama.

Consideramos que a variável aleatória  $Y$  possui uma distribuição lognormal definida pelos parâmetros  $\rho$  e  $\gamma$  e momento  $E(Y^q) = e^{\rho q + \gamma^2 q^2/2}$ . Assim, este modelo multifractal tem 3 parâmetros  $(\alpha, \rho, \gamma)$ , com a seguinte função que descreve o fator de momento  $c(q)$ :

$$E(Y^q) = e^{\rho q + \gamma^2 q^2/2} 2^{N(q - \log_2 \frac{\Gamma(\alpha)\Gamma(2\alpha+q)}{\Gamma(2\alpha)\Gamma(\alpha+q)})} \quad (12)$$

Utilizando propriedades estatísticas das cascatas multiplicativas, podemos verificar para este modelo multifractal geral os seguintes resultados [19]:

(i) A média do processo é dada pela equação:

$$E(X(\Delta t_0)) = e^{\rho + \gamma^2/2} \quad (13)$$

(ii) A variância do processo é dada por:

$$\text{var}(X(\Delta t_0)) = e^{2\rho + 2\gamma^2} \left( \frac{\alpha + 1}{\alpha + 1/2} \right)^N - e^{2\rho + \gamma^2} \quad (14)$$

$X(\Delta t_0)$  tem distribuição lognormal para  $N \gg 1$ . Podemos deduzir esta propriedade pelo teorema do limite central, uma vez que  $X(\Delta t_0) = 2^N Y R(\eta_1) R(\eta_1 \eta_2) \dots R(\eta_1, \dots, \eta_N)$  [20].

### 3. Parâmetro de Escala Global para Tráfego Multifractal

Nesta seção, nós derivamos um fator de escala global associado ao modelo multifractal geral apresentado anteriormente. A vantagem de se ter um parâmetro de escala global deve-se a possibilidade de se chegar a um modelo de fila aproximado para processos multifractais [21]. Isto proporciona por exemplo, cálculo de probabilidade de perda e banda efetiva. No desenvolvimento desta teoria, usamos alguns resultados oriundos do estudo de processos auto-similares [22].

Seja  $X(t)$  um processo auto-similar com parâmetro de Hurst  $H$ , com média zero e variância  $\sigma^2$ , tal que a seguinte relação é válida:

$$X \stackrel{d}{=} m^{1-H} X^m \quad (15)$$

em que  $m$  é o parâmetro de agregação. Define-se um processo  $Y(t) = X(t) - X(t-1)$ , que é o processo de incrementos de  $X(t)$ . Uma vez que o processo agregado de  $Y(t)$  tem características auto-similares, podemos afirmar que:

$$\bar{Y} \stackrel{d}{=} m^{H-1} Y(1) \quad (16)$$

A média do processo agregado  $\bar{Y}$  é zero e a variância do mesmo, em função do parâmetro de agregação  $m$ , é dada em logarítmo por:

$$\log_2\{var[\bar{Y}]\} = (2H - 2) \log_2 m + \log_2 \sigma^2 \quad (17)$$

Agora, pretendemos estabelecer uma relação similar a anterior, assim como um parâmetro de escala global  $H_g$  para o modelo cascata de multiplicativa geral.

**Proposição 1** Seja o processo multifractal  $X(t)$  com parâmetros  $\alpha$ ,  $\rho$  e  $\gamma$ . O parâmetro de escala global para este processo é dado por:

$$H_g = 1 - \frac{\log_2\left(\frac{\alpha+1}{\alpha+1/2}\right)}{2}$$

Prova: Sejam  $m = 2^k$  e  $x_{(N-k)} = \sum_{(k-1)m+1}^{km} X_k$ . Assim, o processo agregado  $X^{(m)}$  pode ser escrito como:

$$X^m = 2^{-k} x_{(N-k)} \quad (18)$$

A variância do processo agregado  $X^m$  da cascata multifractal geral pode ser escrita como:

$$\begin{aligned} var[X^{(m)}] &= var(2^{-k} x_{N-k}) \\ &= E\{[(2^{-k} x_{(N-k)}) - E(2^{-k} x_{(N-k)})]^2\} \\ var[X^{(m)}] &= E[\{2^{-k} x_{(N-k)}\}^2] - E^2[2^{-k} x_{(N-k)}] \end{aligned} \quad (19)$$

Devido ao fato de o processo  $X$  ser o produto de uma variável aleatória lognormal  $Y$  e um cascata multiplicativa  $\mu$ , podemos obter a seguinte relação:

$$x_{(N-k)} = \sum_{(k-1)m+1}^{km} Y_k \mu(\Delta t_k) \quad (20)$$

Usando a equação (20) e os momentos do processo agregado da cascata multiplicativa dados por:

$$M_\varepsilon(q) = \frac{1}{m} \sum_{i=1}^m X(i)^q = \frac{1}{m} \sum_{i=1}^m (Y_i \mu(\Delta t_i))^q \quad (21)$$

temos:

$$\begin{aligned} var[X^m] &= E[Y^2] E[R^2]^{N-k} - E^2[Y] \left(\frac{1}{2}\right)^{2(N-k)} \\ &= E \left[ \left\{ 2^{-k} \frac{1}{m} \sum_{(k-1)m+1}^{km} Y_k \mu(\Delta t_k) \right\}^2 \right] - E^2 \left[ 2^{-k} \frac{1}{m} \sum_{(k-1)m+1}^{km} Y_k \mu(\Delta t_k) \right] \end{aligned}$$

$$\text{var}[X^m] = e^{2\rho+2\sigma^2} \left( \frac{\alpha+1}{\alpha+1/2} \right)^{N-k} - E^2[Y]2^{-2N}2^k \quad (22)$$

Quando o número de estágios  $N$  na geração da cascata é grande, o termo  $E^2[Y]2^{-2N}2^k$  pode ser seguramente ignorado. Aplicando logaritmo em (22), obtemos:

$$\begin{aligned} \log_2 \text{var}[X^m] &= \\ &= \log_2 e^{2\rho+2\sigma^2} + \log_2 \left( \frac{\alpha+1}{\alpha+1/2} \right)^N + \log_2 \left( \frac{\alpha+1}{\alpha+1/2} \right)^{-\log_2 m} \end{aligned} \quad (23)$$

Desta forma, podemos definir um parâmetro de escala global  $H_g$  para o tráfego multifractal, análogo ao parâmetro de Hurst  $H$  no caso monofractal. Comparando-se (17) e (23), pode-se verificar a correspondência entre os termos envolvidos. Assim, o parâmetro  $H_g$  é dado pela seguinte equação:

$$H_g = 1 - \frac{\log_2 \left( \frac{\alpha+1}{\alpha+1/2} \right)}{2} \quad (24)$$

Para o cálculo de  $H_g$  devemos estimar o valor de  $\alpha$  através da função escala  $t(q)$ . Podemos estimar esta função utilizando o método descrito a seguir: Dado o processo de incrementos  $\{X_1, X_2, \dots, X_n\}$ , definimos a série agregada  $\{X^m\}$  no nível  $m$  por:

$$X_k^m = X_{(k-1)m+1} + X_{(k-1)m+2} + \dots + X_{(k)m} \quad k, m = 1, 2, \dots \quad (25)$$

Se a série  $\{X_k\}$  possui propriedades em escala, então o gráfico dos momentos absolutos  $E(|x(t)|^q)$  versus  $m$  em um escala logarítmica em ambos eixos, deve resultar em uma reta descrita pela seguinte equação:

$$\log E(|x(t)|^q) = \tau_0(q) \log m + \log c(q) \quad (26)$$

A inclinação desta reta provê uma estimativa de  $\tau_0(q)$  e sua intersecção corresponde ao valor de  $\log c(q)$ . Aplicando o método de mínimos quadrados, obtemos a curva  $\hat{\tau}_0(q)$ . A partir da mesma, calculamos o parâmetro  $\alpha$  da equação (11) por meio do algoritmo de Levenberg-Marquardt [23].

#### 4. Banda Efetiva usando Parâmetro de Escala Global

O conceito de banda efetiva provê um modo de caracterizar os requisitos de recursos de uma conexão. A banda efetiva de um fluxo de tráfego é uma taxa maior que a taxa média, mas menor que a taxa de pico deste fluxo de dados. Ela corresponde a capacidade que pode ser usada para atender QoS exigido por um fluxo. Além disso, se vários fluxos de tráfego são simultaneamente servidos à uma taxa equivalente à sua banda efetiva, então as demandas de QoS não serão violadas [24]. A banda efetiva pode ser modelada de forma paramétrica. Algumas equações de banda efetiva analítica são conhecidas, por exemplo para processos de Poisson, On-Off e ruído gaussiano fracionário (fGn) [8]. Por outro

lado, temos os cálculos de 'banda efetiva medida', ou seja, na qual não se assume um modelo, mas sim, se obtém a banda efetiva diretamente pela medição da fonte. Entre estes métodos têm-se: estimador direto, estimador em bloco, estimador baseado na distância de Kullback-Leibler, baseado em regressão linear e banda efetiva empírica [25].

A banda efetiva empírica de um fluxo de tráfego é definida como [25]:

$$eb_{emp}(\theta, t, N) = \frac{1}{\theta t} \log \hat{E}_{N_t}[e^\theta X(0, t)] \quad 0 < \theta; 0 < t < N_t \quad (27)$$

em que  $X(0, t)$  indica o número agregado de chegadas de bytes dentro um intervalo de comprimento  $t$  e  $\hat{E}_{N_t}[e^\theta X(0, t)]$  é a função geradora de momento medida para a série de tráfego com  $N_t$  amostras. Para processos de Poisson e On-Off, a banda efetiva empírica é muito próxima de suas respectivas bandas efetivas analíticas [25].

Seja  $X(t)$  um processo multifractal geral com parâmetro de escala global  $H_g$ . Note que, em escalas de tempo grandes, a função geradora de momento do fBm (movimento Browniano fracionário) pode ser considerada [26]. Desta forma, a função geradora de momento deste processo multifractal com média  $\mu$  e variância  $\sigma^2$  pode ser dada por:

$$\phi(\theta) = \exp\left(\mu\tau\theta + \frac{\tau^{2H_g}\sigma^2\theta^2}{2}\right) \quad (28)$$

Assim, usando a média e a variância do modelo multifractal geral (equações (13),(14)), a banda efetiva do processo usando o parâmetro de escala global  $H_g$  pode ser dada por:

$$eb(\theta, \tau) = \frac{1}{\theta\tau} \log \phi(\theta) = \mu + \frac{\tau^{2H_g}\sigma^2\theta}{2}$$

$$eb(\theta, \tau, \alpha, \rho, \gamma, N) = e^{\rho+\gamma^2/2} + \frac{\theta}{2} \left( e^{2\rho+2\gamma^2} \left( \frac{\alpha+1}{\alpha+1/2} \right)^N - e^{2\rho+\gamma^2} \right) \tau \left[ 2 \left( 1 - \frac{\log_2\left(\frac{\alpha+1}{\alpha+1/2}\right)}{2} \right) - 1 \right] \quad (29)$$

onde  $s$  é o parâmetro de espaço e  $\tau$  é a escala de tempo [8]. Usamos a banda efetiva empírica como exemplo de banda efetiva 'medida' e a de Norros [26], que se baseia no modelo monofractal fBm, para efeito de comparação nas simulações. Diferentemente da de Norros, a banda efetiva proposta utiliza parâmetros de um modelo multifractal e um parâmetro de escala global para esse modelo multifractal. Mesmo o tráfego sendo multifractal possui um parâmetro de escala global  $H_g$ , como o parâmetro de Hurst para tráfego monofractal. Se o tráfego for monofractal, os resultados da nossa proposta de banda efetiva e a de Norros serão semelhantes.

## 5. Limitantes de Desempenho de Redes

As pesquisas em limitantes de desempenho de redes têm aberto novos rumos à análise e projeto de redes de alta velocidade [27]. Limitantes de desempenho para atraso e tamanho de fila podem ser obtidos em termos de banda efetiva de tráfego.

O processo envelope limita o processo original mesmo sob um deslocamento de tempo arbitrário (semelhante ao conceito de estacionaridade) [1] [5]. Chamamos de PEM, o processo envelope mínimo. O PEM é subaditivo e sua média é denominada como taxa de envoltória mínima (TEM).

Antes de descrevermos o cálculo de rede em sua forma estatística [3], mencionaremos alguns elementos do cálculo de rede determinístico [5]. Considere uma sequência não-negativa  $\{a(t), t = 0, 1, 2, \dots\}$  correspondente ao processo de chegada de tráfego e seja  $A(t_1, t_2) = \sum_{t=t_1}^{t_2-1} a(t)$ . Um processo envelope tem a seguinte propriedade:

$$A(t_1, t_2) \leq \hat{A}(t_1, t_2) \quad \text{qq } t_1 \leq t_2 \quad (30)$$

$\hat{A}(t)$  é um processo envelope de  $a(t)$ . Um processo  $\hat{A}(t)$  é subaditivo se  $\hat{A}(t_1, t_2) \leq \hat{A}(t_1) + \hat{A}(t_2)$  para todo  $t_1$  e  $t_2$ .  $\hat{A}(t)$  é "estacionário" no sentido de que depende apenas da diferença de  $t_1$  e  $t_2$ . Supondo que  $\hat{A}(t)$  é crescente e subaditivo [3], então:

$$\lim_{t \rightarrow \infty} \frac{\hat{A}(t)}{t} = \inf_{t \geq 1} \frac{\hat{A}(t)}{t} = \hat{a} \quad (31)$$

onde  $\hat{a}$  é a taxa do processo envelope  $\hat{A}(t)$ . O processo envelope mínimo (PEM)  $A^*(t)$  também é crescente e subaditivo, dado por:

$$A^*(t) = \sup_{s \geq 0} A(s, s + t) \quad (32)$$

O processo envelope mínimo pode ser obtido através dos processos envelopes lineares propostos por Cruz [1][2]:  $A^*(t) \leq \hat{a}t + \hat{\sigma}$  para algum valor de  $\hat{\sigma}$  constante e não-negativa. A taxa de envelope mínima (TEM)  $a^*$  de  $a(t)$  é:

$$\lim_{t \rightarrow \infty} \frac{A^*(t)}{t} = a^* \quad (33)$$

Considerando-se um sistema com um servidor com capacidade  $c$ , *buffer* infinito e uma disciplina de serviço conservativa, se  $a^* < c$ , existe uma constante  $d < \infty$ , tal que o retardo máximo não é maior do que  $d$ .

Podemos dizer que uma variável aleatória  $X$  é limitada exponencialmente em relação a  $\theta$  ( $0 < \theta < \infty$ ) se existe uma constante  $g < \infty$ , tal que:

$$(Ee^{\theta X})^{\frac{1}{\theta}} \leq g \quad (34)$$

Assim, utilizando o limite de Chernoff, temos:

$$P(X \geq x) \leq g^{\theta} e^{-\theta x} \quad \text{for all } x \quad (35)$$

que provê um limite para a distribuição de cauda de  $X$ .

Em um cenário estocástico (cálculo de rede estatístico), para um servidor com capacidade  $c$  operando com uma disciplina de serviço conservativa, se a TEM do processo

de entrada é menor do que  $c$ , então podemos afirmar que [3]: i) O comprimento de fila é limitado exponencialmente com relação a  $\theta$ ; ii) O retardo virtual é limitado exponencialmente com relação a  $\theta c$  se a política de escalonamento for FIFO (First-In First-Out). Usando estes resultados, limitantes para a distribuição de cauda do comprimento de fila podem ser estimados através do processo envelope linear do tráfego de entrada.

Seja  $\hat{A}(\theta, t)$  um processo envelope de  $a(t)$  com relação a  $\theta$ :

$$\frac{1}{\theta} \log E e^{\theta A(t_1, t_2)} \leq \hat{A}(\theta, t_2 - t_1) \quad \text{qq} \quad t_1 \leq t_2 \quad (36)$$

Então, o PEM em relação a  $\theta$  é:

$$A^*(t) = \sup_{s \geq 0} \frac{1}{\theta} \log E e^{\theta A(s, s+t)} \quad (37)$$

Ao contrário do PEM no cálculo determinístico, este PEM estatístico não é em geral subaditivo. Portanto, a taxa mínima de  $a(t)$  em relação a  $\theta$  é definida como [3]:

$$\alpha^*(\theta) = \limsup_{t \rightarrow \infty} \frac{A^*(\theta, t)}{t} \quad (38)$$

Um processo de chegada  $a(t)$  é dito limitado por um processo envelope linear  $a(\theta)t + \sigma(\theta)$  onde  $a(\theta) \geq 0$ ,  $\sigma(\theta) \geq 0$  com relação a  $\theta$ , se:

$$\frac{1}{\theta} \log E e^{\theta A(t_1, t_2)} \leq \alpha(\theta)(t_2 - t_1) + \sigma(\theta) \quad (39)$$

Portanto,  $a(\theta)$  é um limite de taxa de chegada estacionário, enquanto  $\sigma(\theta)$  pode ser interpretado como um limite do grau de rajada presente no processo de chegada.

A definição de TEM está conectada à Teoria dos Grandes Desvios através do teorema de Gärtner-Ellis [20]. É necessário considerar as seguintes condições para  $a(t)$ ,  $t \geq 0$ : i)  $\{a(t), t \geq 0\}$  deve ser estacionário e ergódico; ii)  $a^*(\theta) = \lim_{t \rightarrow \infty} \frac{A^*(\theta, t)}{t}$  para todo  $0 < \theta < \infty$ ; iii)  $\theta a^*(\theta)$  deve ser estritamente convexo e diferenciável para todo  $0 < \theta < \infty$ . Sob estas condições, a sequência  $\{A(0, t), t \geq 1\}$  obedece ao princípio dos grandes desvios com função de taxa  $I(v)$  dada por [28]:

$$I(v) = \sup_{\theta} \{\theta v - \theta a^*(\theta)\} \quad (40)$$

A TEM  $a^*(\theta)$  para variáveis aleatórias i.i.d é citada como banda efetiva por Kelly [8]. O par  $(\alpha^*(\theta), \sigma^*(\theta))$  define um processo envelope linear (PEL) mínimo  $\alpha^*(\theta)t + \sigma^*(\theta)$  com relação a  $\theta$ :

$$\alpha^*(\theta) = \limsup_{t \rightarrow \infty} \sup_{\theta \geq 0} \frac{1}{t} \log E e^{\theta A(t_1, t_2)} \quad (41)$$

e

$$\sigma^*(\theta) = \inf \left\{ \sigma(\theta) \mid \frac{1}{\theta} \log E e^{\theta A(t_1, t_2)} \leq \alpha^*(\theta)(t_2 - t_1) + \sigma(\theta) \right\} \quad \text{qq} \quad t_2 \geq t_1 \geq 0 \quad (42)$$

O PEL mínimo é definido para qualquer tipo de processo. Os valores de  $a(t_1)$  e  $a(t_2)$  não são necessariamente independentes. Caso o seguinte limite exista:

$$h(\theta) = \lim_{t \rightarrow \infty} \frac{1}{t} \log E e^{\theta A(0,t)} \quad (43)$$

então  $h(\theta)/\theta$  é a banda efetiva de  $a(t)$  com relação a  $\theta$ . Assim,  $\alpha^*(\theta)$  é exatamente a banda efetiva com relação a  $\theta$ .

No cálculo de rede estatístico, vamos considerar limitantes para a função geradora de momento em vez de se ter limitantes determinísticos para as variáveis aleatórias. Seja  $\{W(t)\}$  o processo correspondente ao tamanho da fila no *buffer* (*backlog*). Considerando-se que  $a(t)$  é independente de  $W(0)$  e  $\alpha^*(\theta) < c$ , então a função geradora de momento para o processo de *backlog* é limitado por [4]:

$$E[e^{\theta W(t)}] \leq e^{\theta(\alpha^*(\theta)-c)t} e^{\theta\sigma^*(\theta)} E[e^{\theta W(0)}] + B(\theta) \quad (44)$$

onde

$$B(\theta) = \frac{(1 - e^{-c\theta})e^{\theta\sigma^*(\theta)}}{1 - e^{\theta(\alpha^*(\theta)-c)}} \quad (45)$$

Quando  $t \rightarrow \infty$ , verifica-se que o termo  $B(\theta)$  é mais rígido do que o obtido em [3] por um fator igual a  $(1 - e^{-c\theta})$ . O valor limite para a função geradora de momento pode ser usado para se derivar limitantes para o *backlog*, atraso médio, o processo envelope mínimo e a distribuição de cauda do tamanho da fila no *buffer*. Considerando-se que  $a(t)$  é independente de  $W(0)$  e  $\alpha^*(\theta) < c$ , para um nó com um único servidor podemos afirmar que [4]:

i) A probabilidade de perda de bytes é limitada exponencialmente por:

$$P[W(t) \geq w] \leq e^{-\theta w} \{e^{\theta(\alpha^*(\theta)-c)t} e^{\theta\sigma^*(\theta)} E[e^{\theta W(0)}] + B(\theta)\} \quad (46)$$

ii) O tamanho médio da fila é limitado em:

$$E[W(t)] \leq \frac{\{e^{\theta(\alpha^*(\theta)-c)t} e^{\theta\sigma^*(\theta)} E[e^{\theta W(0)}] + B(\theta)\}}{(1 - e^{-\theta})} \quad (47)$$

O parâmetro  $\theta$  pode ser ajustado para a obtenção de limites mais justos. Dessa forma, usando a equação de banda efetiva (29), calculamos limitantes de desempenho para uma fila alimentada com processo multifractal. Verificamos neste artigo, limitantes para a probabilidade de perda de bytes e tamanho médio da fila.

### 5.1. Probabilidade de Perda para Processos com Longa-dependência

Várias questões de engenharia de tráfego, como dimensionamento de *buffer* e controle de fluxo, estão relacionadas ao comportamento de fila do tráfego. A característica de longa-dependência do tráfego tem um significativo impacto em seu comportamento de fila [15].

Norros [26] e Duffield e O'Connell [24] apresentaram limitantes inferiores de probabilidade de perda para processos auto-similares. Entretanto, em muitos casos, esta

aproximação subestima  $P(Q > b)$ . O limitante inferior para  $P(Q > b)$  decai assintoticamente (para *buffer* muito grande) de acordo com uma função de Weibull. A probabilidade de cauda de ocupação da fila é muito mais 'pesada' que a distribuição exponencial predita por modelos de tráfego tradicionais de curta-dependência. A distribuição do tamanho da fila ou a probabilidade de perda para processos que têm um parâmetro de escala global  $H \in (0.5, 1)$  pode ser dada, segundo [24], por:

$$\lim_{b \rightarrow \infty} b^{-2(1-H)} \ln P(Q > b) = -a^{-2(1-H)}(a+c)^2/2 \quad (48)$$

onde  $a = c/H - c$ .

## 6. Simulações e Resultados

Utilizamos nas simulações, traços de tráfego TCP/IP (lbl-pkt-5) obtidos da Digital Equipment Corporation <sup>1</sup>, traços Ethernet (Bc-Aug89) obtidos da Bellcore<sup>2</sup> e traços capturados entre os anos de 2000 e 2002 na rede Petrobras através de um analisador de dados DA350 da Acterna<sup>TM</sup>, com uma resolução de 32 microsegundos [29]. Consideramos amostras de tráfego em uma escala de agregação de 100ms, devido ao fato de os traços apresentarem característica multifractal nesta escala [30]. Apresentamos neste artigo, os resultados obtidos com a séries: lbl-pkt-5 com  $N_t = 2^{15}$  amostras, Bc-Aug89 com  $N_t = 2^{14}$  amostras e 10-7-S-1 com  $N_t = 2^{12}$ .

A Tabela 1 apresenta algumas estatísticas para duas das séries utilizadas na análise de desempenho de fila. O parâmetro de Hurst é calculado segundo o método descrito em [31]. A Tabela 1 mostra que o parâmetro de escala global proposto  $H_g$  é próximo de  $H$ .

**Tabela 1. Média, Variância, Parâmetro de Hurst e  $H_g$ .**

Série de Tráfego	Média	Variância	P.Hurst	$H_g$
lbl-tcp-5	$2,6146.10^3$	$1,0033.10^7$	0,7811	0,8062
Bc-Aug89	$1,3819.10^4$	$1,7111^8$	0,8617	0,8797

Para realizarmos as estimativas de probabilidade de perda e tamanho médio da fila precisamos, segundo a nossa abordagem, da banda efetiva das séries de tráfego. A Figura 2 apresenta a banda efetiva obtida usando a equação (29) em comparação à banda efetiva empírica e a de Norros, para uma probabilidade de perda de  $10^{-7}$  e tamanho do *buffer* igual a 60Kbytes para a série lbl-tcp-5. Como era de se esperar, a banda efetiva dada pela equação (29) é bastante próxima à proposta por Norros. O valor da banda efetiva empírica ficou abaixo das outras duas. Sabe-se que, realmente a banda efetiva proposta por Norros é mais conservadora [26].

Na avaliação da proposta de cálculo de probabilidade de perda utilizando a equação (46), consideramos um servidor com *buffer* finito com capacidade igual a  $c=(5.6)$ x média para a série lbl-tcp-5,  $c=(1.7)$ x média para série 10-7-S-1 e  $c=(3.1)$ x média para série Bc-Aug89. Apresentamos na Figura 3, a probabilidade de perda em regime permanente, ou seja, para  $t \rightarrow \infty$  e a probabilidade de perda obtida para a série real de

<sup>1</sup><http://ita.ee.lbl.gov/html/contrib/DEC-PKT.html>

<sup>2</sup><http://ita.ee.lbl.gov/html/contrib/BC.html>

tráfego. A abordagem de Duffield e O'Connell subestima a porcentagem de perda para o tráfego real, como pode ser visto pela Figura 3. Assim, como a proposta de Duffield, a nossa provê melhores resultados para *buffers* maiores.

A Figura 4 mostra o tamanho médio da fila utilizando a equação (47). Pode-se notar que a medida que se aumenta o tamanho do *buffer*, o limitante para a ocupação média do *buffer* se torna mais próximo da ocupação obtida com o traço de tráfego real.

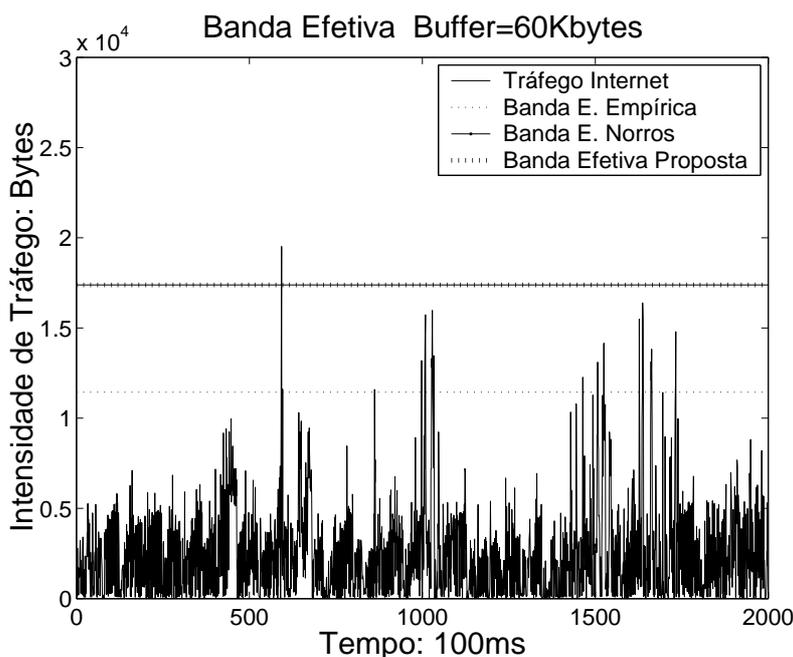


Figura 2. Banda efetiva para série lbi-tcp-5

## 7. Conclusões

O tráfego de redes, mesmo ao possuir características multifractais, possui uma lei de escala global que tem grande influência no comportamento de fila nos buffers. Demonstramos neste artigo, que este comportamento de fila pode ser melhor caracterizado utilizando-se um parâmetro de escala global e cálculo de rede estatístico.

Derivamos inicialmente, o parâmetro de escala global para tráfego multifractal e associamos o cálculo de rede à teoria da banda efetiva. Os valores para o parâmetro de escala global  $H_g$  se mostraram próximos ao parâmetro de Hurst. No entanto, o parâmetro  $H_g$  é obtido a partir de um modelo multifractal.

O cálculo de rede nos permitiu obter expressões para a probabilidade de perda de bytes e tamanho médio da fila. Mostramos que a inserção do cálculo de rede torna mais preciso o limitante para a probabilidade de perda de bytes em comparação, por exemplo, ao obtido pela teoria dos grandes desvios para processos com longa-dependência, que não considera um processo envelope para o tráfego de entrada.

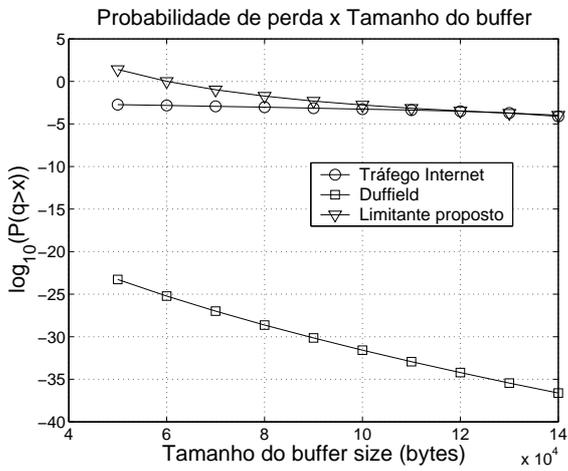
Os limitantes para probabilidade de perda e tamanho médio da fila se mostraram adequados tanto para tráfego Internet quanto Ethernet. Mesmo o tráfego de 'backbone' Internet sendo monofractal em escalas de tempo pequenas (1 a 100ms) como afirmam

alguns trabalhos [32], nossa proposta também é aplicável a esses casos, como pôde ser observado nas simulações com tráfego monofractal. Isso possibilita a inclusão desta abordagem em esquemas de controle de admissão e outros tipos de controle de tráfego para garantir QoS. Além do mais, vários parâmetros de QoS podem ser analisados em conjunto e então, oferecidos aos fluxos das redes atuais.

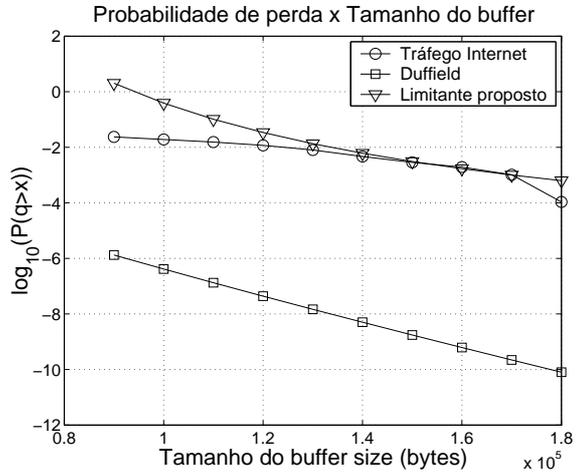
## Referências

- [1] Cruz, R.L. (1991). A calculus for network delay Part I: Network elements in isolation, IEEE Trans. Information Theory, v. 37, pp.114-131.
- [2] Cruz, R.L. (1991). A calculus for network delay Part II: Network Analysis”, IEEE Trans. Information Theory, v. 37, pp.114-131.
- [3] Chang, C.S. (1994). Stability, queue length and delay of deterministic and stochastic queueing networks, IEEE Trans. Automat. Contr., v. 39, n. 5, pp. 913-931.
- [4] Dai, L. (1997). Effective bandwidths and performance bounds in high-speed communication systems, Decision and Control, Proceedings of the 36th IEEE Conference, v. 5, pp.4580-4585.
- [5] LeBoudec, J. Y. e Thiran P. (2001). Network Calculus, Lecture Notes in Computer Science. New York: Springer-Verlag.
- [6] Liebeherr, J.; Patek, S.D.; Burchard, A. (2003). Statistical per-flow service bounds in a network with aggregate provisioning, IEEE INFOCOM., v. 3, 30 Março a 3 Abril, pp.1680-1690.
- [7] Boorstyn, R.R.; Burchard, A.; Liebeherr, J.; Ottamakorn, C. (2000). Statistical service assurances for traffic scheduling algorithms, IEEE Transactions on Selected Areas in Communications, v. 18, n. 12, Dezembro, pp.2651-2664.
- [8] Kelly, F. P. (1996). Notes on effective bandwidths, Stochastic Networks. Oxford, U.K.: Clarendon, pp. 141-168.
- [9] Gibbens, R.J. e Kelly, F.P. (1991). Measurement-based connection admission control, Proc. of the 15th International Teletraffic Symposium, pp. 879-888.
- [10] Leland, W. E.; Taqqu, M. S.; Willinger, W. e Wilson, D. V. (1994). On the self-similar nature of Ethernet traffic (extended version). IEEE/ACM Transactions on Networking, v. 2, n. 1, pp.1-15.
- [11] Paxson, V. e Floyd, S. (1995). Wide-area traffic: The failure of Poisson modeling. IEEE/ACM Transactions on Networking, v.3, n.3, pp.226-244.
- [12] Erramilli, A.; Narayan, O.; Willinger, W. (1996). Experimental Queueing Analysis with Long-Range Dependent Packet Traffic, IEEE/ACM Trans. on Net., v. 4, n.2, Abril.
- [13] Park, K. e Willinger, W. (2000). Self-similar Network Traffic and Performance Evaluation”, John Wiley and Sons, New York.
- [14] Li, Q.; Mills, D. (1999). Investigating the Scaling Behavior, Crossover and Anti-persistence of Internet Packet Delay Dynamics, Globecom 3, pp. 1843-1852.
- [15] Grossglauser, M. e Bolot, J.-C. (1999). On the relevance of long-range dependence in network traffic. IEEE/ACM Transactions on Networking, v. 7, n.5, pp.629-640, Outubro.

- [16] Kant,K.(1999).On aggregate Traffic Generation with Multifractal Properties, Globecom 2, pp.801-804.
- [17] Riedi,R.H.; Crouse,M.S.;Ribeiro, V.J. e Baraniuk,R.G.(1999). A Multifractal Wavelet Model with Application to Network Traffic, IEEE Trans. on Information Theory, v. 45, n.3, Abril.
- [18] Mandelbrot,B. B.; Fisher,A. e Calvet,L.(1997). A Multifractal Model of Asset Return. Yale University, Yale University.
- [19] Dang,T. D.; Molnár,S. e Maricza,I.(2002). Capturing the complete characteristics of multifractal network traffic,"GLOBECOM 2002, Taipei, Taiwan, Novembro.
- [20] Papoulis,A.(1991). Probability, Random Variables, and Stochastic Processes,3rd ed. New York: McGraw-Hill.
- [21] Murali, K.P.; Gadre,V.M.(2001). Global scaling exponent for variable variance gaussian multiplicative (VVGm) multifractal cascades, Proc. SPCOM'01, pp. 19-25.
- [22] Beran,J.(1995). Statistics for Long-Memory Processes. Chapman and Hall, One Penn Plaza, New York, NY 10119.
- [23] Marquardt, D.W.(1963). An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics, v.11, pp.431-441.
- [24] Duffield,N.G. e O'Connell,N.(1993). Large deviations and overflow probabilities for the general single-server queue, with applications,"Dublin Institute for Advanced Studies-Applied Probability Group, Tech. l Rep. DIAS-STP-93-30.
- [25] Tartarelli,S.; Falkner,M.; Devetsikiotis,M. e Lambadaris,I. e Giordano,S.(2000). Empirical effective bandwidths, IEEE Globecom'00, v.33, San Francisco.
- [26] Norros,I.(1994).A storage model with self-similar input,Queueing Systems, v. 16, pp.387-396.
- [27] Kurose,J.(1992).On computing per-session performance bounds in high-speed multi-hop computer networks, Proc. ACM Sigmetrics, pp. 128-139.
- [28] Bucklew,J.A.(1990).Large deviation techniques in decision, simulation and estimation, J.Wiley, New York.
- [29] Perlingeiro, R. F. e Lee, L. L. Estudo de Estimaco de Banda Efetiva para Tráfego Auto-similar com Varncia Infinita. SBRT'05, Campinas, Setembro, 2005, pp. 326-331.
- [30] Erramilli,A.; Narayan,O.; Neidhardt,A. e Sanjee,I.(2000).Performance impacts of multi-scaling in wide area TCP/IP traffic. Proc.IEEE Infocom 2000, Maro.
- [31] Veitch,D.; Abry,D.(1999). A wavelet based joint estimator for the parameters of LRD, IEEE Trans. Info. Theory., v. 45, n.3, Abril.
- [32] Zhang, Z.-L; Ribeiro, V.; Moon, S. e Diot, C. (2003). Small-time scaling behaviors of Internet backbone traffic: an empirical study, IEEE Infocom, San Francisco, Maro.

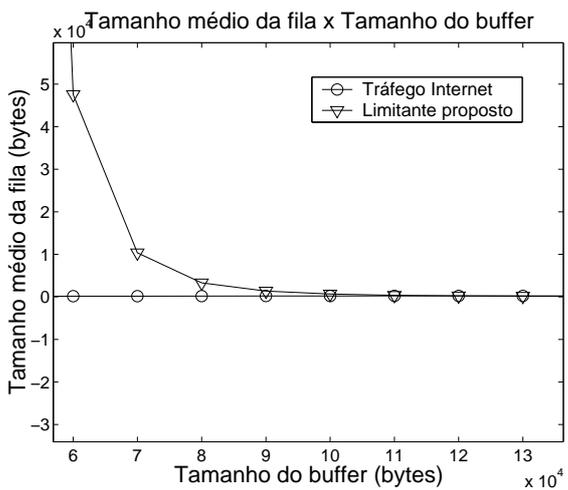


(a) Traço de tráfego lbl-tcp-5

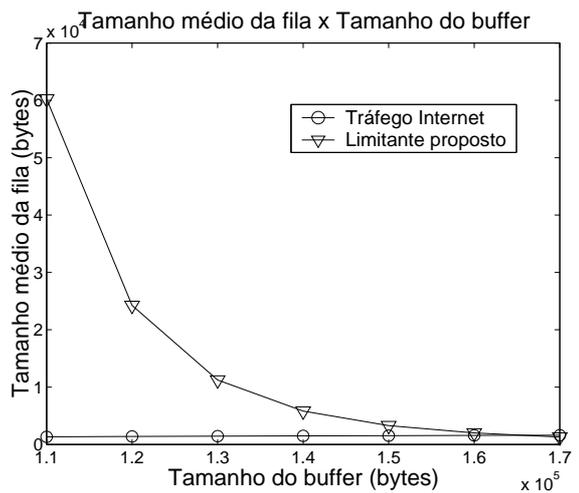


(b) Traço de tráfego 10-7-S-1

**Figura 3. Probabilidade de perda de bytes versus tamanho do buffer**



(a) Traço de tráfego lbl-tcp-5



(b) Traço de tráfego Bc-Aug89

**Figura 4. Ocupação média do buffer**