

# Computação de Banda Passante Equivalente de Fluxos Multifractais

César Augusto V. Melo\*, Nelson L. S. Fonseca†

<sup>1</sup>Instituto de Computação – UNICAMP  
Av. Albert Einstein, 1251  
Cx. 6176  
13083-970 Campinas, SP

**Resumo.** Neste artigo, apresenta-se uma solução para sistemas de fila alimentados por fluxos que apresentam natureza multifractal. Este modelo leva em consideração os períodos em que o sistema permanece ocupado e ocioso. Expressões para calcular o instante em que o sistema alcança sua ocupação máxima são derivadas. Além disso, expressões para determinar a banda passante equivalente de um agregado de fluxos multifractais são calculadas.

**Abstract.** This paper presents a new model for a queuing system fed by multifractal flows. This model considers not only busy periods but also idle periods. We derived an equation to calculate the time instant at which a queue reaches its maximum occupancy. Moreover, equations to estimate the effective bandwidth of multifractal flows are derived.

## 1. Introdução

O fenômeno do cascadeamento (do inglês *scaling*) foi identificado pela primeira vez em traços de tráfego de redes locais [Leland et al., 1994]. Desde então, tem-se empreendido esforços no sentido de melhor compreender esse fenômeno e o seu impacto nos mecanismos de controle de redes. O cascadeamento caracteriza-se pela ausência de uma escala de tempo específica na qual a explosividade de um fluxo de tráfego possa ser caracterizada, assim faz-se necessário uma descrição do tráfego em diferentes escalas de tempo. O tráfego de redes que apresenta cascadeamento é denominado tráfego fractal.

O cascadeamento em tráfego fractal é definido por um valor constante denominado de parâmetro de Hurst. O processo mais utilizado na modelagem do tráfego fractal é o Movimento Browniano Fractal (fBm -*fractal Brownian Motion*) que, com um reduzido número de parâmetros, caracteriza com precisão este tipo de tráfego. De fato, o processo fBm é um modelo preciso quando: i) o tráfego modelado é resultado da agregação do tráfego gerado por fontes que operam em baixas taxas quando comparadas com a banda passante do canal, ii) o impacto dos mecanismos de controle sobre as estatísticas do tráfego é irrelevante, e iii) a escala de tempo de interesse do tráfego encontra-se em regiões em que a sua natureza fractal se manifesta. O processo Movimento Browniano multifractal (mBm-*multifractal Brownian motion*) é a generalização do processo fBm e suas realizações em pequenas escalas de tempo são definidas por fBm's.

Verificou-se também em traços de tráfego IP o fenômeno do cascadeamento em pequenas escalas de tempo, bem como a existência de dependências de longa duração

---

\* Este trabalho foi financiado pela FAPESP processo # 00/09772 – 6

† Este trabalho foi financiado pelo CNPq processo # 305076/2203 – 5

[Riedi and Levy-Vehel, 1997] [A. Erramilli et al., 2001]. Em pequenas escalas de tempo, o tráfego IP apresenta alta variabilidade e o cascadeamento não é definido por um único parâmetro. Para esses tráfegos, a distribuição marginal é não-gaussiana sugerindo a necessidade de uma modelagem envolvendo estatísticas superiores à de segunda ordem.

Se por um lado, no interior da rede, as dependências de longa duração são mais importantes que a variabilidade verificada em pequenas escalas de tempo devido a agregação do tráfego (Propriedade aditiva) [Abry et al., 2002], por outro lado, nas bordas da rede, onde o controle de admissão é executado, a variabilidade do tráfego em pequenas escalas de tempo (propriedade multiplicativa) é uma das características mais importantes a ser considerada [Gilbert et al., 1999]. Traços de rede que apresentam dependência de longa duração e alta variabilidade em pequenas escalas de tempo, podem ser modelados por processos multifractais.

Em redes baseadas em multiplexação estatística, fluxos de tráfego são multiplexados nos comutadores e transferidos para o canal de saída. Resolver um sistema de fila com multiplexação estatística que possui como entrada um fluxo (multi/mono)fractal é de suma importância para o controle de admissão. Soluções baseadas na Teoria dos Desvios Largos têm sido propostas recentemente [Kelly, 1996]. Entretanto, nestas soluções supõe-se que o buffer no multiplexador tenha tamanho infinito, o que é uma suposição não-realista.

Em [Melo and Fonseca, 2003], introduziu-se um processo envelope para modelagem de tráfego multifractal. Esse processo envelope é um limitante superior para o acumulado de tráfego que chega até um certo instante de tempo em um processo movimento Browniano multifractal. Validou-se extensivamente esse processo envelope com traços de tráfego real e com traços sintéticos. Mostrou-se que, embora o mBm seja um processo Gaussiano, o seu processo envelope é um limitante superior justo para a quantidade de tráfego verificada em fluxos IP.

A principal contribuição do presente artigo é a definição de um método para computar o instante em que a fila alcança o seu máximo quando ela é alimentada por fluxos multifractais. Essa computação pode ser usada para determinar a banda passante equivalente do agregado de vários fluxos multifractais. Tal computação é útil na formulação de mecanismos de controle, como por exemplo em mecanismo de controle de admissão na entrada de redes de domínio DiffServ. Nessas redes, um novo fluxo é admitido se a quantidade de recursos disponíveis (banda passante e buffer) forem suficientes para atender a Qualidade de Serviço requerida pelo novo fluxo sem qualquer prejuízo ao serviço sendo prestado aos fluxos já estabelecidos. Além disso, as expressões introduzidas neste artigo podem ser usadas em abordagens baseadas em medições.

A diferença entre a contribuição do presente artigo e a contribuição de [Melo and Fonseca, 2003] é explicada a seguir. Em [Melo and Fonseca, 2003], derivou-se expressões para calcular o instante em que o trabalho não-realizado num sistema alcança o seu máximo, considerando-se apenas um único período de ocupação (do inglês *busy period*). No presente artigo, as expressões derivadas levam em consideração períodos em que o sistema fica ocioso (do inglês *idle period*), sendo portanto, uma descrição mais realista do sistema. Em outras palavras, neste artigo as expressões derivadas levam em consideração todos os períodos de ocupação e de ociosidade do sistema. Assim sendo, resultados mais precisos são obtidos como demonstrado nas Seções 4 e 5.

## 2. O processo movimento Browniano multifractal

O movimento Browniano multifractal (mBm) é um processo Gaussiano contínuo com incrementos não-estacionários e definido no intervalo  $(0, \infty)$  da seguinte forma [Peltier and J.L. Vehel, 1995]:

$$W_{H(t)} = \frac{1}{\Gamma(H(t) + 1/2)} \left\{ \int_{-\infty}^0 [(t-s)^{H(t)-1/2} - (-s)^{H(t)-1/2}] dB(s) + \int_0^t (t-s)^{H(t)-1/2} dB(s) \right\}$$

onde  $B(s)$  é o movimento Browniano.

O processo movimento Browniano multifractal (mBm) é uma generalização do processo movimento Browniano fractal (fBm - *fractional Brownian Motion*) e é assintoticamente auto-similar, ou seja :

$$\lim_{\rho \rightarrow 0^+} \left\{ \frac{W(t + \rho u) - W(t)}{\rho^{H(t)}} \right\}_{u \in \mathbb{R}^+} = \{B_{H(t)}(u)\}_{u \in \mathbb{R}^+}$$

onde  $W(\cdot)$  é um mBm e  $B_{H(t)}(u)$  é um processo fBm com parâmetro Hurst  $H = H(t)$ .

A regularidade local de um processo é expressa pelo expoente Holder que mede o cascadeamento (do inglês *scaling*), nesse processo. O expoente Holder define o cascadeamento em pequenas escalas de tempo, através de uma potência cujo expoente  $h(t)$  inclui tanto dependências temporais quanto espaciais. O expoente Holder é o maior valor que  $h$  pode assumir,  $0 \leq h \leq 1$ , tal que :

$$|X(t + \gamma) - X(t)| \leq k|\gamma|^h \text{ para } \gamma \rightarrow 0$$

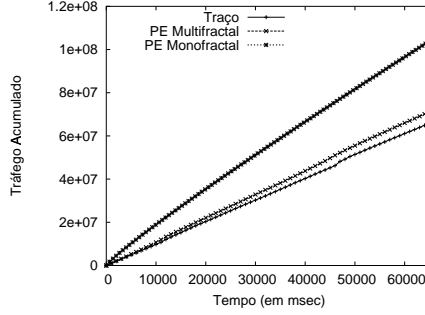
Para processos monofractais o expoente Holder é um valor constante (parâmetro Hurst), enquanto que para processos multifractais este expoente é uma função aleatória chamada de função Holder.

### 3. Um processo envelope para o processo movimento Browniano multifractal

A solução de um sistema de filas requer o conhecimento da quantidade de trabalho que chega ao sistema, bem como a taxa de serviço com a qual o trabalho é realizado. Processos envelope (PE) são limitantes superiores para o acumulado do trabalho que chega a um sistema de filas. Processos envelope permitem soluções menos complexas que as obtidas quando se considera processos que descrevem com exatidão o trabalho acumulado que chega ao sistema. Processos envelope podem ser probabilísticos ou determinísticos, a diferença fundamental entre estas abordagens é a caracterização obtida para o processo real. Numa representação determinística, o trabalho acumulado do processo real é sempre menor que o limite estabelecido pelo processo envelope, enquanto em uma representação probabilística, o trabalho acumulado do processo real pode ser maior que o limite estabelecido pelo processo envelope dentro de um certo limite estabelecido pelo valor da probabilidade de violação.

O dimensionamento feito utilizando processos envelope determinísticos pode gerar desperdício de recursos, uma vez que sempre se considera o volume máximo de trabalho que chega em qualquer intervalo de tempo. Por outro lado, podem ocorrer perdas quando processos envelope probabilísticos são usados.

Um limitante superior para o acumulado do trabalho que chega em um sistema de filas pode ser computado pela média da quantidade de trabalho mais um limite para o acumulado dos



**Figura 1: Modelagem baseada em processos envelope monofractal e multifractal**

incrementos. Um limite superior para os incrementos do mBm é dado pelo limite superior das realizações de um processo fBm, uma vez que os incrementos do processo mBm na vizinhança do instante  $t$  são realizações de um processo fBm  $Z_H(\cdot)$  com parâmetro Hurst  $H$  dado pelo valor da função Holder  $H(\cdot)$  no instante  $t$  [N.L.Fonseca et al., 2000]; assim:

$$Z_H(t) \leq \kappa H t^{H-1} \quad (1)$$

onde  $Z_H(t)$  é a realização do processo fBm  $Z_H(\cdot)$  no instante  $t$ .

Assim sendo, o processo envelope  $\hat{A}(\cdot)$  de um mBm com média  $\bar{a}$ , desvio padrão  $\sigma$  e função Holder  $H(\cdot)$  é dado por [Melo and Fonseca, 2003]:

$$\hat{A}(t) = \int_0^t \bar{a} + \kappa \sigma H(x) x^{H(x)-1} dx \quad (2)$$

Um caso particular, desse processo envelope, é o processo envelope de um fBm derivado em [N.L.Fonseca et al., 2000] quando  $H(\cdot)$  é um valor constante, ou seja,

$$\hat{A}(t) = \bar{a}t + \kappa \sigma t^H \quad (3)$$

A precisão do processo envelope do mBm foi avaliada utilizando-se tráfego de rede real e tráfego sintético. Verificou-se que o processo envelope do mBm é capaz de caracterizar com precisão processos multifractais [Melo and Fonseca, 2003]. A Figura 1 mostra dois modelos, PE multifractal e PE monofractal, caracterizando um traço de tráfego real coletado em uma rede operacional da Universidade de Memphis em 24 de Maio de 2003 (ver Tabela 1). A média  $\bar{a}$  e a variância  $\sigma^2$  desse traço é dada por  $\bar{a} = 1013,83$ ,  $\sigma^2 = 3245708,75$ , e o valor estimado para o parâmetro Hurst é  $H = 0,78$ . O método apresentado em [J.E.Cavanaugh, 2003] foi usado para estimar a função Holder  $H(\cdot)$  usada na derivação do PE multifractal. Observa-se que o processo envelope multifractal é um limite superior justo enquanto o processo envelope monofractal superestima a quantidade de tráfego registrada no traço. Portanto, o dimensionamento da rede com fluxos multifractais baseado em modelos monofractais resulta em sub-utilização dos recursos. Resultados similares foram obtidos para os demais traços apresentados na Tabela 1.

#### 4. A análise de um sistema de filas alimentado por um fluxo multifractal

Nesta seção, apresenta-se um modelo baseado no processo envelope do mBm capaz de caracterizar um sistema alimentado por fluxos multifractais. Seja  $A(\cdot)$  trabalho acumulado que chega a um

sistema operando com taxa de serviço constante  $C$ . O trabalho não-realizado  $Q(\cdot)$  nesse sistema, num intervalo  $[0, t]$ , é dado por [Benes, 1962]:

$$Q(t) = A(t) - Ct - \min\{0, \inf_{t>0}\{A(t) - Ct\}\} \quad (4)$$

onde  $\min\{0, \inf_{t>0}\{A(t) - Ct\}\}$  contabiliza os períodos ociosos do sistema.

A função que descreve o volume de trabalho servido pelo sistema  $S(\cdot)$  é dada por:

$$S(t) = Ct + \min\{0, \inf_{t>0}\{A(t) - Ct\}\} \quad (5)$$

ou seja

$$S(t) = Ct + A(t^\bullet) - A(t^\diamond) - Ct^\bullet + Ct^\diamond \quad (6)$$

onde  $t^\bullet = \arg \inf_{t>0}\{A(t) - Ct\}$  e  $t^\diamond = \max\{t | A(t) - Ct = 0 \text{ e } t \leq t^\bullet\}$ .

Seja  $r^\bullet$  e  $r^\diamond$  tal que  $r^\bullet = \frac{t}{t^\bullet}$  e  $r^\diamond = \frac{t}{t^\diamond}$ . A Equação 6 pode ser reescrita da seguinte forma:

$$S(t) = Ct + A(t/r^\bullet) - A(t/r^\diamond) - Ct/r^\bullet + Ct/r^\diamond. \quad (7)$$

Substituindo  $S(\cdot)$  na Equação 4 tem-se:

$$\begin{aligned} Q(t) &= A(t) - S(t) \\ &= A(t) - A(t/r^\bullet) + A(t/r^\diamond) - Ct(1 - 1/r^\bullet + 1/r^\diamond). \end{aligned} \quad (8)$$

Um limitante superior para o volume de trabalho servido pelo sistema  $S(\cdot)$  é dado por:

$$\hat{S}(t) = Ct + \hat{A}(t/\hat{r}^\bullet) - \hat{A}(t/\hat{r}^\diamond) - Ct/\hat{r}^\bullet + Ct/\hat{r}^\diamond \quad (9)$$

onde  $\hat{A}(\cdot)$  é processo envelope para o trabalho acumulado que chega ao sistema.

Um limitante superior,  $\hat{Q}(\cdot)$ , para  $Q(\cdot)$  pode ser obtido substituindo  $S(\cdot)$  na Equação 8 por  $\hat{S}(\cdot)$ , ou seja

$$\begin{aligned} \hat{Q}(t) &= \hat{A}(t) - \hat{S}(t) \\ &= \hat{A}(t) - \hat{A}(t/\hat{r}^\bullet) + \hat{A}(t/\hat{r}^\diamond) - Ct(1 - 1/\hat{r}^\bullet + 1/\hat{r}^\diamond) \end{aligned} \quad (10)$$

A efetividade do limite derivado na Equação 10,  $\hat{Q}(\cdot)$ , foi verificado comparando-se a probabilidade de perda medida (PPM) e a probabilidade de perda calculada (PPC), derivadas, respectivamente, a partir das Equações 8 e 10.

A probabilidade de perda medida é definida por:

$$P_{loss} = 1 - \frac{S(t^*) + q_{max}}{A(t^*)} \quad (11)$$

Traço	Data	# de pacotes	ponto de agregação
ANL-1050127417	11/04/03	121998	Argonne NL to STARTAP
ANL-1050225668	13/04/03	105641	Argonne NL to STARTAP
MEM-1053844177	24/05/03	220904	University of Memphis
MEM-1054459191	01/06/03	266708	University of Memphis
COS-1057970154	12/07/03	1247518	Colorado State University
BWY-1058086940	13/07/03	1168143	Columbia University
20000125-143640-1	25/01/00	7000000	University of Auckland

**Tabela 1: Traços reais com tráfego coletado na Internet 2 ABILENE**

onde  $q_{max}$  é um limiar para o volume de trabalho não-realizado no sistema,  $S(\cdot)$  é a função que define o volume de trabalho realmente servido pelo sistema no intervalo  $[0, t]$ , e  $t^* = \arg \sup_{t>0} \left\{ \frac{S(t) + q_{max}}{A(t)} \right\}$ . Enquanto que, a probabilidade de perda calculada é definida por:

$$\hat{P}_{loss} = 1 - \frac{\hat{S}(t^*) + q_{max}}{\hat{A}(t^*)} \quad (12)$$

onde  $\hat{S}(\cdot)$  é um limite para  $S(\cdot)$  e  $t^* = \arg \sup_{t>0} \left\{ \frac{\hat{S}(t) + q_{max}}{\hat{A}(t)} \right\}$ .

Para validar a Equação 10, experimentos de simulação foram realizados. Coletou-se a PPM e a PPC quando  $A(\cdot)$  e  $\hat{A}(\cdot)$  são, respectivamente, um traço de tráfego IP que apresenta natureza multifractal e o processo envelope derivado para este traço. Na Tabela 1, mostra-se os traços de tráfego utilizados neste experimento. Estes traços encontram-se disponíveis no sítio [www.nlanr.net](http://www.nlanr.net). No presente artigo, os resultados apresentados correspondem as avaliações realizadas com os traços 20000125-143640-1 (Figura 2.a), MEM-1054459191 (Figura 2.b) e MEM-1053844177 (Figura 2.c).

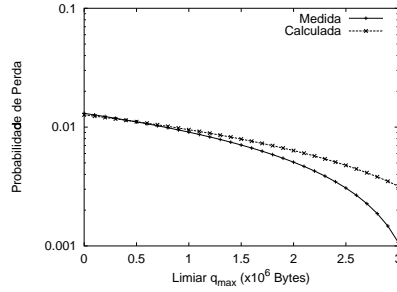
As Figuras 2.a-2.c mostram as probabilidades de perdas medida (Equação 11) e calculada (Equação 12) para diferentes valores do limiar  $q_{max}$  e diferentes níveis de utilização do canal. Os resultados mostram que PPC é um limitante superior para PPM.

A Figura 2.a mostra os resultados coletados para um canal operando com utilização média de 80%. Neste cenário verifica-se que a PPC é um limite superior justo quando os valores do limiar  $q_{max}$  são pequenos e médios. Para valores altos do limiar verifica-se uma diferença maior entre as probabilidades medida e calculada. As Figuras 2.b e 2.c mostram os resultados coletados em um cenário com tráfego intenso, ( $\rho \approx 1$ ). Nesse cenário, o comportamento verificado anteriormente se repete, mas, como o servidor passa a maior parte do tempo ocupado,  $\hat{S}(t) \approx Ct$ , o erro gerado pelo fato de  $\hat{A}(\cdot)$  ser um limitante superior é amenizado.

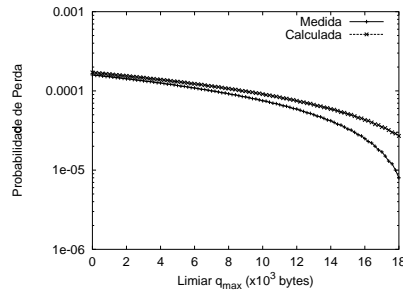
Estes resultados reforçam a viabilidade do modelo de fila proposto. Nota-se ainda que, neste modelo, o custo computacional envolvido na derivação dos limites para o sistema é baixo, sendo necessário somente a determinação de raízes e máximos de funções determinísticas. Resultados similares foram obtidos para os demais traços apresentados na Tabela 1.

## 5. A escala de tempo de interesse de uma fila alimentada por um fluxo multifractal

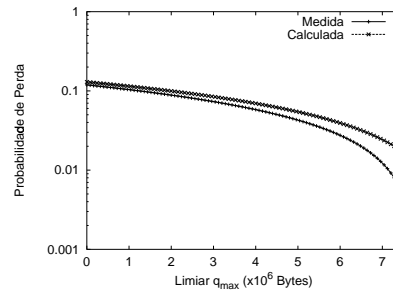
Nesta seção, expressões para calcular o instante em que o volume de trabalho não-realizado em um sistema de filas alimentado por um fluxo multifractal alcança o seu máximo são derivadas. Nos



(a) Auckland ( $\rho = 0, 80$ )



(b) MEM-1054459191 ( $\rho = 0, 95$ )



(c) MEM-1053844177 ( $\rho = 1, 15$ )

**Figura 2: Probabilidade de perda medida e calculada em diferentes cenários**

mecanismos de controle de admissão, o valor da escala de tempo de interesse é usada para calcular banda passante necessária para atender um agregado de fluxos que requer uma certa Qualidade de Serviço. Um novo fluxo é aceito se a banda passante equivalente calculada for menor ou igual ao volume de recursos disponíveis.

A solução de sistema de filas alimentado por processos estocásticos é complexa, tornando difícil a definição de expressões que descrevam o comportamento deste sistema. Uma abordagem alternativa consiste em limitar o volume de trabalho que chega ao sistema, utilizando processos envelope. Isto significa transformar o problema de se obter limites para um sistema estocástico (Equação 8) em um problema de menor complexidade que é o de encontrar limites para um sistema determinístico (Equação 10).

Seja um sistema de filas operando em tempo contínuo, servido por um canal de capacidade  $C$  e alimentado por um fluxo multifractal  $A(\cdot)$ . Seja  $\hat{A}(\cdot)$  o processo envelope de  $A(\cdot)$ , um limite para a ocupação máxima do sistema, considerando-se uma política de serviço FIFO (*First In First Out*), é dado por:

$$\hat{q}_{max} = \max_{t \geq 0} \{\hat{Q}(t)\}. \quad (13)$$

Para determinar este limite toma-se  $\hat{Q}(\cdot)$  conforme definido pela Equação 10, ou seja,

$$\hat{Q}(t) = \hat{A}(t) - \hat{A}(t/\hat{r}^\bullet) + \hat{A}(t/\hat{r}^\circ) - Ct(1 - 1/\hat{r}^\bullet + 1/\hat{r}^\circ) \quad (14)$$

$$= \int_0^{t/\hat{r}^\diamond} \bar{a} + \kappa\sigma H(x)x^{H(x)-1} dx + \int_{t/\hat{r}^\bullet}^t \bar{a} + \kappa\sigma H(x)x^{H(x)-1} dx - Ct(1 + 1/\hat{r}^\diamond - 1/\hat{r}^\bullet),$$

onde  $\hat{A}(\cdot)$  é o processo envelope multifractal (Equação 2). Diferenciando-se esta equação em relação a  $t$ , obtém-se

$$\hat{t}^* = \left[ \frac{\kappa\sigma(H(\hat{t}^*)(\hat{t}^*)^{H(\hat{t}^*)} + \hat{r}^\diamond H(\hat{t}^*/\hat{r}^\diamond)(\hat{t}^*/\hat{r}^\diamond)^{H(\hat{t}^*/\hat{r}^\diamond)} - \hat{r}^\bullet H(\hat{t}^*/\hat{r}^\bullet)(\hat{t}^*/\hat{r}^\bullet)^{H(\hat{t}^*/\hat{r}^\bullet)})}{(C - \bar{a})(1 + 1/\hat{r}^\diamond - 1/\hat{r}^\bullet)} \right] \quad (15)$$

$\hat{t}^*$  é o instante de tempo no qual  $\hat{Q}(\cdot)$  atinge o seu valor máximo. Este instante de tempo é chamado de Escala de Tempo de Interesse Máxima (MaxTS). Note que a Equação 15 pode ser avaliada diretamente do traço de tráfego através da obtenção da média  $\bar{a}$ , variância  $\sigma^2$  e dos valores da função Holder  $H(\cdot)$  nos pontos desejados. Assim sendo, a Equação 15 pode ser usada no escopo das abordagens baseadas em medições.

Para validar a Equação 15 experimentos de simulação foram realizados. Registrou-se a dinâmica do trabalho não-realizado em um sistema de filas alimentado por fluxos multifractais. Para um sistema com capacidade  $C$  mediu-se a cada instante  $t$  o volume de trabalho não-realizado. Ao final do experimento, coletou-se o instante  $t$  em que este volume atingiu o seu valor máximo. Esse instante é chamado de MaxTS medido. Numa segunda etapa do experimento, calculou-se o valor MaxTS utilizando-se a Equação 15. Esse valor é chamado de MaxTS calculado.

A Figura 3 mostra os resultados do experimento realizado com uma subsequência de  $2^{16}$  pontos do traço MEM-1053844177 (Tabela 1) agregado numa escala de  $1msec$ . Esta subsequência foi caracterizada por um processo envelope com probabilidade de violação de  $10^{-3}$ , média  $\bar{a} = 1013,83$ , variância  $\sigma^2 = 3245708,75$ , e função Holder  $H(\cdot)$  estimada conforme procedimento descrito em [J.E.Cavanaugh, 2003].

As Figuras 3.a e 3.b mostram os resultados obtidos quando a utilização média do canal é igual respectivamente a  $\rho = 0,7$  e  $\rho = 0,9$ . Para  $\rho = 0,7$  o MaxTS calculado é  $45698msec$ . Este valor apresenta um erro de 3% em relação ao MaxTS medido. Quando  $\rho = 0,9$ , Figura 3.b, o MaxTS calculado é  $48130msec$  apresentando um erro de 0,65% ao MaxTS medido. Esses resultados ilustram a precisão da expressão derivada para calcular a MaxTS (Equação 15), especialmente em sistemas que operam sob altas taxas de utilização. Além disso, esses resultados reforçam o argumento de que limites justos para sistemas estocásticos podem ser derivados. Resultados qualitativamente similares foram obtidos para os demais traços na Tabela 1.

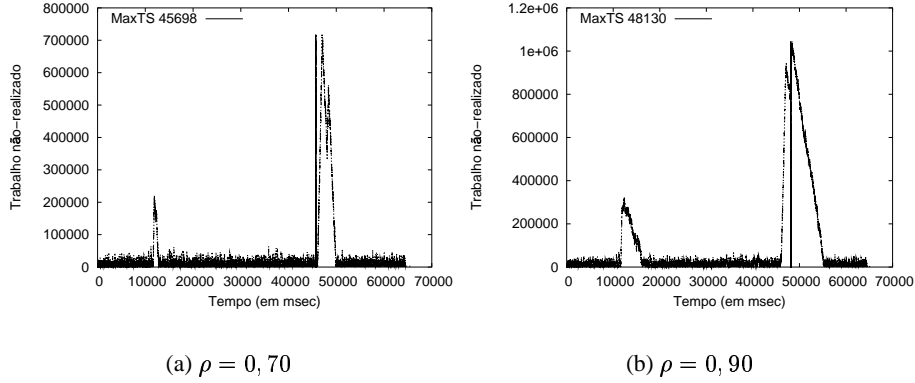
## 6. A banda passante equivalente para um agregado de $N$ fluxos multifractais

Nesta seção, deriva-se expressões para o cálculo da banda passante equivalente de um agregado de fluxos multifractais utilizando-se o valor de MaxTS, ou seja, computa-se a banda passante capaz de garantir um nível específico de Qualidade de Serviço a um agregado de fluxos multifractais. O problema fundamental estudado nesta seção pode ser definido da seguinte forma:

*Dado um conjunto de fluxos com média  $\bar{a}_i$ , desvio padrão  $\sigma_i$  e função Holder  $H_i(\cdot)$ , qual é a capacidade necessária para que o tamanho máximo da fila não ultrapasse um limiar  $q_{max}^N$  com probabilidade  $\epsilon$  ?*

A resposta a esta questão requer a definição de uma expressão para o processo envelope do agregado de  $N$  fluxos. Para computar o volume do tráfego agregado emprega-se a propriedade





**Figura 3: O MaxTS estimado em sistemas operando sob diferentes níveis de utilização**

de que em uma vizinhança  $(\Delta t)$  do instante  $t$  de um processo mBm, este comporta-se como um fBm. Em [Norros, 1994], mostrou-se que o agregado de  $N$  processos fBm com parâmetro Hurst  $H_i$ , média  $\bar{a}_i$  e variância  $\sigma_i^2 t^{2H_i}$  é um processo fBm com média  $\bar{a}$  e variância  $\sigma^2 t^{2H}$ , dadas respectivamente por  $\sum_{i=1}^N \bar{a}_i$  e por  $\sum_{i=1}^N \sigma_i^2 t^{2H_i}$ . Assim, um processo mBm gerado pela agregação de processos mBm pode ser representado localmente por um fBm resultante da agregação dos processos fBms.

Sejam  $N$  fluxos independentes definidos pelos seguintes parâmetros: média  $\bar{a}_i$ , variância  $\sigma_i^2$  e função Holder  $H_i(\cdot)$ . Seja  $W(\cdot)$  o processo agregado e  $\hat{A}_i(\cdot)$  o processo envelope de cada fluxo. O processo envelope  $\hat{A}^N(\cdot)$  para o trabalho acumulado em  $W(\cdot)$  no intervalo  $[0, t]$  é dado por:

$$\begin{aligned} \hat{A}^N(t) &= \sum_{i=1}^N \hat{A}_i(t) \\ &= \int_0^t \sum_{i=1}^N \bar{a}_i + \kappa \left( \sum_{i=1}^N \sigma_i^2 H_i(x) x^{2H_i(x)-1} \right) \left( \sum_{i=1}^N \sigma_i^2 x^{2H_i(x)} \right)^{-\frac{1}{2}} dx \end{aligned} \quad (16)$$

Pela definição de  $\hat{S}(\cdot)$  (Equação 9), um limitante  $\hat{S}^N(\cdot)$  para o volume de trabalho servido em um sistema é dado por:

$$\hat{S}^N(t) = Ct + \hat{A}^N(t/\hat{r}^\bullet) - \hat{A}^N(t/\hat{r}^\diamond) - Ct/\hat{r}^\bullet + Ct/\hat{r}^\diamond \quad (17)$$

onde  $\hat{A}^N(\cdot)$  é processo envelope para o trabalho acumulado que chega ao sistema gerado pelo agregado de  $N$  fluxos,  $C$  é a taxa de serviço do sistema,  $\hat{r}^\bullet = \frac{t}{\hat{t}^\bullet}$ ,  $\hat{t}^\bullet = \arg \inf_{t>0} \{\hat{A}^N(t) - Ct\}$ ,  $\hat{r}^\diamond = \frac{t}{\hat{t}^\diamond}$  e  $\hat{t}^\diamond = \max\{t | \hat{A}^N(t) - Ct = 0 \text{ e } t \leq \hat{t}^\bullet\}$ .

Substituindo-se  $\hat{A}(\cdot)$  e  $\hat{S}(\cdot)$  na Equação 10 por  $\hat{A}^N(\cdot)$  e  $\hat{S}^N(\cdot)$ , a questão postulada no início desta seção pode ser respondida resolvendo-se o seguinte sistema de equações não-lineares:

$$\begin{cases} \frac{d}{dt} \hat{A}^N(t) - \frac{d}{dt} \hat{S}^N(t) = 0 \\ \hat{A}^N(t/\hat{r}^\bullet) - \hat{A}^N(t/\hat{r}^\diamond) - Ct(1/\hat{r}^\bullet - 1/\hat{r}^\diamond) = 0 \end{cases} \quad (18)$$

ou seja

$$\left\{ \begin{array}{l}
\kappa \left[ \left( \sum_{i=1}^N \sigma_i^2 H_i(t) t^{2H_i(t)-1} \right) \left( \sum_{i=1}^N \sigma_i^2 t^{2H_i(t)} \right)^{\frac{-1}{2}} + \right. \\
\left. \left( \sum_{i=1}^N \sigma_i^2 H_i(t/\hat{r}^\diamond) (t/\hat{r}^\diamond)^{2H_i(t/\hat{r}^\diamond)-1} \right) \left( \sum_{i=1}^N \sigma_i^2 (t/\hat{r}^\diamond)^{2H_i(t/\hat{r}^\diamond)} \right)^{\frac{-1}{2}} - \right. \\
\left. \left( \sum_{i=1}^N \sigma_i^2 H_i(t/\hat{r}^\bullet) (t/\hat{r}^\bullet)^{2H_i(t/\hat{r}^\bullet)-1} \right) \left( \sum_{i=1}^N \sigma_i^2 (t/\hat{r}^\bullet)^{2H_i(t/\hat{r}^\bullet)} \right)^{\frac{-1}{2}} \right] - \\
(C - \sum_{i=1}^N \bar{a}_i) (1 + 1/\hat{r}^\diamond - 1/\hat{r}^\bullet) \\
\int_0^{t/\hat{r}^\bullet} \kappa \left( \sum_{i=1}^N \sigma_i^2 H_i(x) x^{2H_i(x)-1} \right) \left( \sum_{i=1}^N \sigma_i^2 x^{2H_i(x)} \right)^{\frac{-1}{2}} dx - \\
\int_0^{t/\hat{r}^\diamond} \kappa \left( \sum_{i=1}^N \sigma_i^2 H_i(x) x^{2H_i(x)-1} \right) \left( \sum_{i=1}^N \sigma_i^2 x^{2H_i(x)} \right)^{\frac{-1}{2}} dx - \\
(C - \sum_{i=1}^N \bar{a}_i) (t/\hat{r}^\diamond + t/\hat{r}^\bullet)
\end{array} \right. = 0 \quad (19)$$

A solução desse sistema implica na determinação de  $\hat{C}$  e  $\hat{t}^*$ , respectivamente, a banda passante equivalente e a escala de tempo de interesse. A solução da primeira equação do sistema implica na determinação de uma  $\hat{C}$  capaz de garantir a Qualidade de Serviço requerida, enquanto a solução da segunda equação garante a utilização racional dos recursos do sistema, reduzindo ao máximo os períodos em que o sistema permanece ocioso. Este sistema pode ser resolvido utilizando-se o método Newton descrito em [Press et al., 1993] que apresenta convergência quadrática em regiões depositárias das soluções do sistema e garante convergência global. Note que as variáveis envolvidas na solução deste sistema podem ser derivadas diretamente dos fluxos que compõem o agregado avaliado.

A grande vantagem da multiplexação estatística é o uso eficiente dos recursos alcançado com a transmissão de pacotes de diferentes fluxos, permitindo suportar um número muito maior de usuários quando comparado com a comutação de circuitos. Para avaliar os benefícios de se usar as expressões para a banda passante equivalente de um agregado de fluxos multifractais (Equação 19), definiu-se uma medida de ganho,  $G(\cdot)$ , dada pela razão entre  $n$  vezes a banda passante equivalente de um fluxo e a banda passante equivalente calculada para o agregado de  $n$  fluxos com estatísticas iguais.  $G(n)$  é dado por:

$$\begin{aligned}
G(n) &= \frac{nEB(1)}{EB(n)} \\
&= \frac{\left[ \int_0^{\hat{t}^*/\hat{r}_1^\diamond} \bar{a} + \kappa \sigma H(x) x^{H(x)-1} dx + \int_{\hat{t}^*/\hat{r}_1^\bullet}^{\hat{t}^*} \bar{a} + \kappa \sigma H(x) x^{H(x)-1} dx - K \right]}{\left[ \hat{t}^* (1 + 1/\hat{r}_1^\diamond - 1/\hat{r}_1^\bullet) \right]} \\
&= \frac{\left[ \int_0^{\hat{t}^{**}/\hat{r}_n^\diamond} \bar{a} + n \frac{-1}{2} \kappa \sigma H(x) x^{H(x)-1} dx + \int_{\hat{t}^{**}/\hat{r}_n^\bullet}^{\hat{t}^{**}} \bar{a} + n \frac{-1}{2} \kappa \sigma H(x) x^{H(x)-1} dx - K' \right]}{\left[ \hat{t}^{**} (1 + 1/\hat{r}_n^\diamond - 1/\hat{r}_n^\bullet) \right]}
\end{aligned} \quad (20)$$

onde  $EB(1)$  é a banda passante equivalente para um único fluxo e  $EB(n)$  é a banda passante para o agregado de  $n$  fluxos.  $\hat{t}^*$  e  $\hat{t}^{**}$  são, respectivamente, as escalas de tempo de interesse obtidas pela resolução do sistema definido pela Equação 19 para um único fluxo,  $EB(1)$ , e para  $n$  fluxos,  $EB(n)$ .  $K$  é o tamanho do buffer no multiplexador e  $K' = \frac{K}{n}$ .

Experimentos de simulação foram realizados para determinar o ganho  $G(\cdot)$  em cenários com um número variável de fluxos e que, individualmente, apresentam diferentes médias, variância e função Holder  $H(\cdot)$ . A Figura 4 mostra o ganho para traços com diferentes funções Holder  $H(\cdot)$ .

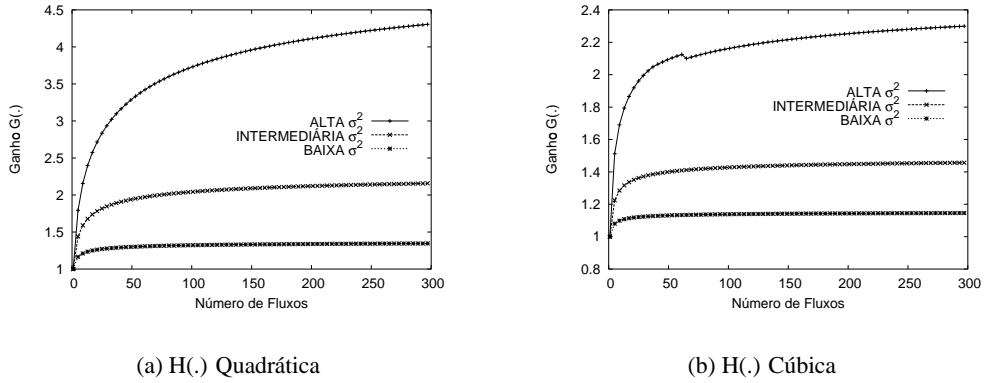


Figura 4: Ganho com multiplexação de fluxos sintético de tráfego

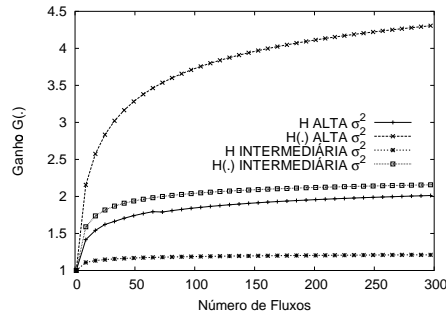
Grau	Função Holder $H(\cdot)$
Quadrático	$\frac{19}{10}t^2 - \frac{19}{10}t + 0,985$
Cúbico	$\frac{49}{10}t^3 - \frac{79}{10}t^2 + \frac{33}{10}t + 0,51$

Tabela 2: Funções  $H(\cdot)$  usadas para gerar traços sintéticos

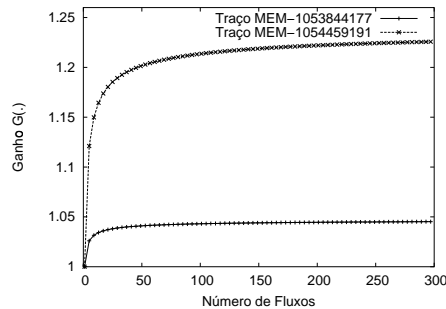
e valores de variância. Todas as funções  $H(\cdot)$  dos traços sintéticos usados na avaliação de  $G(\cdot)$  são polinômios (Tabela 2). Os fluxos chamados de variância “baixa” apresentam média  $\bar{a} = 125,09$  e variância  $\sigma^2 = 290,00$ . Os fluxos chamados de variância “intermediária” e “alta” apresentam variâncias  $10\sigma^2$  e  $100\sigma^2$ , respectivamente.

Observa-se na Figura 4 que o ganho aumenta com a variância. Por exemplo, quando a função  $H(\cdot)$  é definida pelo polinômio quadrático, o ganho máximo obtido é de 1,35 para fluxos com baixa variância, enquanto que para fluxos com alta variância, o ganho verificado é 4,0. O ganho também é influenciado pelo valor da função Holder. De fato, o valor médio da função Holder até o instante definido pelo valor da MaxTS,  $\frac{\int_0^{t^*} H(x)dx}{t^*}$ , é de grande relevância para a determinação do ganho, como pode ser visto ao se comparar as Figuras 4.a e 4.b. Considerando-se os mesmos valores de média e variância, o ganho obtido é maior em traços em que a função  $H(\cdot)$  é definida pelo polinômio quadrático em relação ao polinômio cúbico. Em fluxos de variância intermediária, o ganho máximo obtido é de 2,2 para traços em que a função  $H(\cdot)$  é definida pelo polinômio quadrático, enquanto em traços onde a função  $H(\cdot)$  é definida pelo polinômio cúbico o ganho obtido é de apenas 1,45.

Outros experimentos de simulação foram realizados considerando-se uma modelagem monofractal dos fluxos baseada no processo envelope monofractal (Equação 3), com parâmetro Hurst  $H$  igual ao valor médio da função  $H(\cdot)$ . O objetivo é comparar os ganhos obtidos com a modelagem multifractal e com a modelagem monofractal e assim, responder a questão de que modelos monofractais podem ser utilizados para modelagem de tráfego multifractal usando-se um valor de parâmetro de Hurst  $H$  igual ao valor médio da função  $H(\cdot)$ . A Figura 6 mostra o ganho obtido considerando a modelagem monofractal  $H$  e multifractal  $H(\cdot)$ . A função  $H(\cdot)$  é definida pelo polinômio quadrático (Tabela 2). Em cenários onde os fluxos apresentam valores de variância intermediário, a modelagem multifractal gera um ganho duas vezes maior que o obtido na modelagem monofractal derivada. Quando os fluxos apresentam valores de variância alto, o ganho verificado com a modelagem multifractal é ainda maior. Este comportamento da modelagem monofractal é devido a regularidade local ser superestimada pelo parâmetro Hurst, causando a subutilização dos



**Figura 5: Ganho com Multiplexação de fluxos caracterizado por modelos monofractal e multifractal**



**Figura 6: Ganho com a multiplexação de fluxos de tráfego real**

recursos e, conseqüentemente, ganhos menores.

Avaliações do ganho usando-se traços de tráfego IP (Tabela 1) foram realizadas. A Figura 6 mostra os resultados coletados com os traços MEM-1053844177 e MEM-1054459191, que possuem médias  $\bar{a} = 1014,88$  e  $\bar{a} = 1810,98$ , e variâncias  $\sigma^2 = 3,2 \times 10^6$  e  $\sigma^2 = 8,6 \times 10^6$ , respectivamente. O ganho verificado nas avaliações usando o traço MEM-1054459191 é maior que o verificado com o traço MEM-1053844177. Novamente isto acontece porque o traço MEM-1054459191 apresenta alta variância e um valor médio de  $H(\cdot)$  maior até o instante MaxTS.

## 7. Trabalhos Relacionados

Em [A.Erramilli et al., 2001], considera-se a existência de uma escala de tempo  $\theta$  que delimita a manifestação da natureza monofractal e multifractal do tráfego. Assim, o tráfego é modelado por dois processos: cascata multiplicativa, para escalas de tempo menores que  $\theta$ , e o fBm, para escala de tempo maiores que  $\theta$ . Mostra-se que em fluxos IP a escala  $\theta$  é da ordem de um RTT, enquanto que para fluxos de vídeo essa escala corresponde a duração de um frame. Os autores mostraram que uma maior utilização do sistema é alcançada quando essa abordagem é empregada em substituição a uma abordagem puramente monofractal.

Outros modelos baseados em cascata multiplicativa foram propostos. Esses modelos mapeiam uma amostra do tráfego em uma árvore binária multiescalar. Cada nó na árvore representa o tráfego agregado numa escala de tempo, e que é descrito pelos seus descendentes. Dessa forma, os nós nos níveis mais altos da árvore correspondem as escalas de tempo maiores, enquanto os nós nos níveis mais baixos correspondem às escalas de tempo menores. Os pesos que são atribuídos para cada descendente do nó determinam a distribuição marginal e o cascadeamento. No modelo WIG (Wavelet-Domain Independent Gaussian) [Ma and Ji, 2001] os pesos são inovações aditivas e

representam coeficientes da ondaleta Haar (do inglês *Wavelet Haar*), representado por uma árvore binária multiescalar. Quando a profundidade da árvore tende para o infinito, a distribuição marginal do tráfego aproxima-se de uma distribuição Gaussiana. No MWM, (*Multifractal Wavelet Model*) [Riedi et al., 1999] os pesos são fatores que geram uma distribuição log-normal. Nesses dois modelos precisa-se ajustar  $2 + \log(N)$  parâmetros onde  $N$  é o tamanho da amostra. As principais restrições desses modelos são o número de parâmetros que devem ser ajustados e a necessidade de construir a árvore binária, inviabilizando sua aplicação em modelagens feitas em tempo real.

Investigações recentes identificaram monofractalidade no tráfego Internet observado em pequenas escalas de tempo [Zhang et al., 2003]. Neste estudo, as correlações são apontadas como as principais responsáveis pelo tráfego de pacotes em rajadas, e não o mecanismo de confirmação do TCP. Entretanto, outras investigações conduzidas com traços de tráfego Internet, também em pequenas escalas de tempo, mostram claramente a natureza multifractal desse tráfego [Melo and Fonseca, 2003].

Modelagem de tráfego baseada em medições é outro tipo de abordagem relacionada a esse artigo. Em [Eun and Shroff, 2003], a escala de tempo em que perdas provavelmente ocorrem é computada. Essa escala de tempo é chamada de escala de tempo de interesse dominante. Usa-se uma abordagem híbrida (medição/analítica) para computar essa escala. A escala de tempo é determinada verificando-se uma fila virtual com capacidade menor que uma de fato estudada, na qual perdas são muito mais frequentes. Em [Qiu and Knightly, 2001], um processo envelope baseado numa taxa máxima é usado para caracterizar as chegadas sendo a taxa de serviço para esses envelopes derivadas via medições. A escala de tempo máxima é derivada usando essas duas medidas.

## 8. Conclusões

O oferecimento de Qualidade de Serviço para aplicações na Internet é um dos problemas mais desafiadores atualmente na área de modelagem de tráfego. As soluções propostas concentram-se na definição de mecanismos de controle capazes de quantificar os recursos da rede de tal forma que a Qualidade de Serviço requerida seja atendida. No caso de mecanismos de controle de admissão, espera-se que estes possuam a habilidade de quantificar a banda passante demandada por fluxos que requerem Qualidade de Serviço sem prejuízo aos fluxos em andamento. Além disso, o tráfego Internet apresenta cascadeamento em escalas de tempo pequenas que é modelado com precisão por processos multifractais.

Este artigo introduziu um modelo para caracterização de um sistema de fila alimentado por fluxos multifractais. O modelo é realista uma vez que considera tanto períodos de ocupação quanto períodos de ociosidade do sistema. Expressões para calcular a escala de tempo de interesse, o instante em que a quantidade de trabalho não-realizado no sistema é máxima, foram derivadas. Expressões para o cálculo de banda passante equivalente de um agregado de  $N$  fluxos multifractais foram também derivadas. Essas expressões podem ser usadas na definição de mecanismos de controle no escopo de abordagens baseadas em medições, dado que o valor de todas as variáveis do modelo pode ser obtido diretamente do fluxo modelado. Sugere-se como trabalho futuro um estudo comparativo com abordagens baseadas em medições.

## Referências

Abry, P., Baraniuk, R., Flandrin, P., Riedi, R., and Veitch, D. (2002). The multiscale nature of network traffic: Discovery, analysis and modelling. *IEEE Signal Processing Magazine*, 19(3):28–46.

- A. Erramilli, O. Narayan, A. Neidhart, and I. Saniee (2001). Multi-scaling models of TCP/IP and sub-frame VBR video traffic. *Journal of Communications and Networks*, 3(4):383–395.
- Benes, V. E. (1962). *General Stochastic Process in the Theory of Queues*. ADDISON-WESLEY.
- Eun, D. Y. and Shroff, N. B. (2003). A measurement-analytic approach for qos estimation in a network based on the dominant time scale. *IEEE/ACM Transaction on Networking*, 11(2):222–235.
- Gilbert, A., Willing, W., and Feldmann, A. (1999). Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Transactions on Information Theory*, 45(3):971–992.
- J.E. Cavanaugh, Y. Wang, J. (2003). *Handbook of Statistics 21: Stochastic Processes: Modeling and Simulation*, chapter Self-similar processes and their wavelet analysis. Elsevier Science.
- Kelly, F. P. (1996). *Stochastic Networks: Theory and Applications*, chapter Notes on effective bandwidths, pages 141–168. Oxford University Press.
- Leland, W., Taqqu, M., Willinger, W., and Wilson, D. (1994). On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transaction on Networking*, 2(1):1–15.
- Ma, S. and Ji, C. (2001). Modeling Heterogeneous Network Traffic in Wavelet Domain. *IEEE/ACM Transactions on Networking*, 9(5):635–649.
- Melo, C. A. and Fonseca, N. (2003). Um processo envelope para modelagem de tráfego multifractal. In *XXI Simpósio Brasileiro de Redes de Computadores*.
- N.L. Fonseca, Mayor, G., and Neto, C. (2000). On the equivalent bandwidth of self-similar source. *ACM Transactions on Modeling and Computer Simulation*, 10(2):104–124.
- Norros, I. (1994). A storage model with self-similar input. *Queueing Systems*, 18:387–396.
- Peltier, R. and J.L. Vehl (1995). Multifractional brownian motion: definition and preliminary results. Technical report, INRIA.
- Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (1993). *Numerical Recipes in C : The Art of Scientific Computing*. Cambridge University Press.
- Qiu, J. and Knightly, E. (2001). Measurement-based admission control using aggregate traffic envelopes. *IEEE/ACM Transaction on Networking*, 9(2):199–210.
- Riedi, R., Crouse, M., Ribeiro, V., and Baraniuk, R. (1999). A Multifractal Wavelet Model with Application to TCP Network Traffic. *IEEE Transactions on Information Theory*, 45:992–1018.
- Riedi, R. and Levy-Vehl, J. (1997). Tcp traffic is multifractal: A numerical study. Technical Report 3129, INRIA Rocquencourt.
- Zhang, Z., Ribeiro, V., Moon, S., and Diot, C. (2003). Small-time scaling behaviors of internet backbone traffic: An empirical study. In *IEEE Infocom*, San Francisco.