

Performance Bounds for a Cascade based Multifractal Traffic Model with Generalized Multiplier Distributions

Flávio Henrique Teles Vieira and Luan Ling Lee, *Member, IEEE*

Abstract—In this paper we propose a multifractal traffic model that is based on a multiplicative cascade with generalized multiplier distributions (CGMD). The multipliers are determined through their probability densities estimated from real network traffic flows by using Kernel and Acceptance/Rejection methods. Statistical analysis and queueing behavior studies were carried out for the model validation in comparison to other multiplicative cascade based models. In order to build an efficient estimation method of performance bounds for network traffic flows that takes account of multifractal characteristics, we derive the effective bandwidth for the CGMD model as well as its Hurst parameter. The proposed performance bounds are computed by relating the CGMD based effective bandwidth to statistical network calculus concepts. Our performance bound estimation approach is evaluated through simulations with Internet and Ethernet traffic traces, verifying its efficiency in describing the byte loss probability and mean buffer occupation.

Index Terms—Performance bounds, Multifractal traffic, Effective bandwidth, Network calculus, Hurst parameter.

I. INTRODUCTION

ANALYSIS of measured traffic streams taken from a wide range of sources has indicated that many traffic sources present correlation structure that decays much more slowly than exponentially. These slowly decaying correlation structures have been shown in VBR video [1], Ethernet LAN traffic [2], MAN traffic [3] and general Internet WAN traffic [4]. Such a slow decay property in autocorrelation has been related to the concept of self-similarity and long-range dependence (LRD). While the short-range dependent traffic models are significant for their tractability, they cannot capture the long-range dependencies present in real packet data traffic. Many studies have revealed high variability of today's Internet traffic, i.e., traffic is bursty over a wide range of time scales in contrast to the assumption that traffic burstiness exists only at short time scales while traffic is smooth at large time scales [2] [4]. This multiscale burstiness has been shown to make a significant impact on network performance [2] [4] [5].

Most LRD traffic models are based on self-similar processes. Precisely, asymptotically second order self-similarity implies LRD and vice versa. In traffic modeling, the term self-similarity is usually used to refer to the asymptotically

second order self-similar or monofractal processes [6]. The Hurst parameter H is used to measure the degree of 'self-similarity'. For random processes suitable for network traffic modeling, the Hurst parameter is basically a measure of the tail decaying rate of the autocorrelation function. However, it is known that distinct long-range dependent processes with the same measured Hurst parameter value can produce vastly different queueing behavior [7]. In addition, as argued in [5], the queueing performance mostly depends on the traffic's variability over certain time scales rather than on the value of H .

For many real network traffic processes, the wavelet energy-scale or variance-time plots usually do not tend to straight lines. Instead, these processes have piecewise fractal behavior with varying Hurst parameter over some small ranges of time scales [8]. Processes with such fractal behavior are usually referred to as multifractal processes. It has been observed that queueing performance greatly depends on the degree of traffic irregularities at small time scales which are believed to be due to the complex dynamics of data networks [7]. Feldmann, et al. claimed that this combination of scaling behaviours is best represented via a multifractal process [9]. In [10], Erramilli et al. confirmed this finding of multifractal behaviour on fine time scales, and indicated that this behaviour can have a significant impact on queueing performance when utilizations are low. In [11], Lévy Véhel and Riedi showed that a multifractal version of fractional Brownian motion (fBm) may better reflect the properties of measured network traffic than the fBm itself. In [12], the authors presented a multifractal wavelet model (MWM) and argued that this model better predicted the behavior of both a TCP trace [4] and an Ethernet trace [2] when compared to the fractional Brownian motion. The MWM can closely match the properties of the original traffic stream in both marginal distribution and correlation structure. In [13], Gao and Rubin showed that both packet interarrival times and packet length sequences could be modeled by multifractal processes. In [14], for packet interarrival time modeling, Krishna et al. proposed a multiplicative cascade with the multiplier distribution being assumed Gaussian (VVGM- Variable Variance Gaussian Model). The VVGM model assumes that the multiplier distribution has a constant mean value equal to 0.5 in every stage of the cascade. In this work, we demonstrate that traffic modeling can be significantly enhanced by a more precise estimation of the multiplier distributions.

Besides traffic modeling, several studies have aimed at characterizing the mean queue length and the packet arrival

Manuscript received July 17, 2006; revised July 7, 2007. This work was supported by Capes and Fapesp.

Flávio Henrique Teles Vieira and Luan Ling Lee are with the Department of Communications, School of Electrical and Computer Engineering, University of Campinas (DECOM-FEEC-UNICAMP), PO Box 6101, 13083-852 Campinas, SP, Brazil Phone: +55 (19) 3521-3811, (email:flavio.lee@decom.fee.unicamp.br).

distribution in buffers in order to establish bounds for performance measures, such as loss, delay and backlog. The knowledge of these performance bounds explicitly allows us to guarantee the QoS parameters requested by the traffic flows.

Using the notion of the linear envelope process (LEP), Cruz in [15][16] derived some performance bounds in a deterministic framework. Based on the minimum LEP, Chang established performance bounds for both deterministic and stochastic systems [17]. Later, Liyi et al. suggested tighter bounds asymptotically achievable that resulted in a better performance for traffic management [18]. Interesting enough, in these two latter works, effective bandwidth equations are explicitly involved in the network calculus theory [19][20][21]. We revisit this connection by incorporating into network calculus, the proposed multifractal model properties such as its effective bandwidth and Hurst parameter.

In this paper, we firstly develop a new multifractal traffic modeling framework by extending the VVGM model in terms of the cascade multiplier distributions. The multiplier distributions are estimated through the Kernel method and are used to generate multiplier samples by applying the Acceptance-Rejection method. We also propose a novel approach to estimate network performance bounds, such as byte loss probability and mean queue length, by relating the effective bandwidth for the CGMD model to network calculus. We show that the proposed performance bounds are tighter than those given by the Large Deviations theory.

The paper is organized in the following way: in Section II, we describe cascade based multifractal processes and review some multifractal modeling concepts. In Section III, we present the proposed multifractal model (CGMD) and in subsection III.D, its Hurst parameter. In Section IV, we show the results of the tests to validate the CGMD multifractal model. In Section V, we derive the CGMD based effective bandwidth. In Section VI, we show how the effective bandwidth notion is related to the statistical network calculus and in the subsection VI.B, we validate our performance bound proposal by simulation tests. Finally, in Section VII, we conclude.

II. MULTIFRACTAL PROCESSES

Multifractal processes are defined by a scaling law for moments of the processes' increments over finite time intervals. This means the traffic has complex and strong dependence structures inherently, appearing very bursty and the burstiness looks similar over a wide range of scales [6]. Traffic flows with such properties make the network performance much worse than those of Gaussian and short-range dependent traffic types [22]. It is well-known that the increment process of the cumulative traffic load is generally strongly Gaussian at large scales while non-Gaussian at small scales. Now, let us formally introduce the concept of "multifractality".

Definition 1 A stochastic process $X(t)$ is called multifractal if it satisfies [23]:

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

for $t \in T$ and $q \in Q$, where T and Q are intervals on the real line, and $\tau(q)$ and $c(q)$ are functions with domain Q ,

known as the scaling factor and the moment factor of the multifractal process, respectively. Moreover, we assume that T and Q have positive lengths, and that $0 \in T$, $[0, 1] \subseteq Q$. If $\tau(q)$ is linear in q , the process is called monofractal; otherwise, it is multifractal. For self-similar processes with Hurst parameter H , it can be shown that $\tau(q) = qH - 1$ and $c(q) = E(|X(1)|^q)$ [24].

A. Binomial Measure

The simplest multifractal, the binomial measure can be generated through an iterative procedure, called multiplicative cascade, on the compact interval $[0,1]$. Let m_0 and m_1 be two positive numbers adding up to 1. At stage $k = 0$, we start the construction with the uniform probability measure μ_0 on $[0,1]$. In step $k = 1$, the measure μ_1 uniformly spreads mass equal to m_0 on the subinterval $[0,1/2]$ and mass m_1 on $[1/2, 1]$. In step $k = 2$, the set $[0,1/2]$ is split into subintervals $[0,1/4]$ and $[1/4,1/2]$ and interval $[1/2,1]$ into $[1/2,3/4]$ and $[3/4,1]$, with [24]:

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

A dyadic interval at iteration stage k is denoted by $[t, t + 2^{-k}]$ where $t = 0.\eta_1 \dots \eta_k = \sum_{i=1}^k \eta_i 2^{-i}$. Let φ_0 and φ_1 be the relative frequencies of 0's and 1's in the binary development of t . Then, the measure on the dyadic interval is given by:

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

The process preserves at each stage the mass of split dyadic intervals, so it's called conservative or microcanonical. If the multipliers used have the same fixed value m_0 then the multiplicative cascade is deterministic [24]. The binomial measure is a deterministic cascade with scaling function $\tau(q) = -\log_2(m_0^q + m_1^q) + 1$. At each stage of the cascade the intervals can be split into $b > 2$ intervals of equal size; this defines the class of multinomial measures.

B. Multiplicative Cascade

Allowing the cascade multipliers to be random variables, we get a stochastic multiplicative cascade. Denoted by R , these multipliers are chosen to be independent and identically distributed random variables concentrated on $[0,1]$ with probability distribution function $f_R(x)$. At stage k , a dyadic interval $[t, t + 2^{-k}]$ of length $\Delta t_k = 2^{-k}$ ($t = 0.\eta_1 \dots \eta_k = \sum_{i=1}^k \eta_i 2^{-i}$), has the measure μ given by:

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

where $R(\eta_1, \dots, \eta_i)$ represents the multiplier at stage i . Since the multipliers are assumed i.i.d, it can be shown that the measure μ in (3) satisfies the following scaling relationship:

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

which defines a multifractal process with the scaling function given by:

$$\tau(q) = -1 - \log_2 E(R^q). \quad (5)$$

C. Multifractal Wavelet Model

R. Riedi et al. proposed an important multifractal model, namely the Multifractal Wavelet Model (MWM). The MWM is based on the Haar wavelet to characterize network traffic [12]. This model can be viewed as a multiplicative cascade in the Haar wavelet domain that aims to capture the decay of wavelet energy on scale [25].

In the MWM synthesis process, one needs to apply the Haar DWT (Discrete Wavelet Transform) to network traffic, to calculate the second moments of the wavelet coefficients at each scale, and to estimate the mean and the variance of the scaling coefficient at the coarsest scale [12]. The MWM model efficiently approximates the network traffic properties in terms of marginal distribution (it produces approximately lognormal distribution) and correlation structure [26].

D. Estimation of the Cascade Multiplier Distributions

Let X_i^N represent the data at stage N of the cascade. The data at stage $(N-1)$ is obtained by aggregating the consecutive values at stage N over non overlapping blocks of size two. More explicitly, given the data at a coarser scale $(N-j)$, X_i^{N-j} ($i = 1, \dots, 2^{N-j}$), we obtain the data at stage $(N-j-1)$ by the following ordinary sum operation:

$$X_i^{N-j-1} = X_{2i-1}^{N-j} + X_{2i}^{N-j} \tag{6}$$

for $i = 1, \dots, 2^{N-j-1}$. The procedure ends when the aggregates form one point at the coarsest scale. An estimate of the multiplier values for proceeding from stage j to stage $j+1$ can be obtained as [9]:

$$r_j^{(i)} = \frac{X_i^{N-j}}{X_{2i-1}^{N-j-1}} \tag{7}$$

for $i = 1, \dots, 2^{N-j-1}$. The estimate $r_j^{(i)}$ in (7) can be viewed as samples of the multiplier distribution $f_{R_j}(r)$ at stage j . The multiplier distribution at scale j can be obtained from histograms of $r_j^{(i)}$'s. The VVGM model, for example, consists of a multiplicative cascade that assumes that the multiplier distribution at scale j is Gaussian with fixed mean equal to 0.5 and varying variances from stage to stage. In the next sections, we will develop a new multiplicative cascade using more general multiplier distributions in order to achieve better traffic process representation.

III. CASCADE WITH GENERALIZED MULTIPLIER DISTRIBUTIONS (CGMD)

In this section we present the methods that compose the CGMD traffic model, and then in subsection C, we show how to synthesize corresponding process samples.

A. Non-parametric Probability Density Estimation: The Kernel Estimator

The probability density function is a fundamental concept in statistics. Its estimation can be either parametric or non-parametric, assuming or not that the data are drawn from a known distribution, respectively. In this work, we will consider

a non-parametric approach to estimate the cascade multipliers based on observed data.

The Kernel estimator has wide applicability, particularly in the univariate case. The Kernel estimator apart from the histograms is probably the most commonly used non-parametric density estimator and certainly the most studied mathematically.

Let X_1, \dots, X_n be i.i.d samples observed from a continuous univariate distribution with unknown probability density function $f(x)$. The Kernel estimator of $f(x)$ with Kernel K is defined as:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) \tag{8}$$

where h is the window width, also called smoothing parameter.

The Kernel function K must satisfy the following condition:

$$\int_{-\infty}^{\infty} K(x)dx = 1 \tag{9}$$

The Kernel estimator is like a sum of ‘bumps’ placed at the observation points. The Kernel K determines the shape and the widths h 's of the bumps. The following two properties (mean and variance) of the Kernel estimator will be used later into some equation derivations:

$$E(\hat{f}(x)) = \left\{ \int_0^1 \frac{1}{h} K\left(\frac{x-y}{h}\right) f(y)dy \right\} \tag{10}$$

$$var(\hat{f}(x)) = \frac{1}{n} \left(\int_0^1 \frac{1}{(h)^2} K\left(\frac{x-y}{(h)^2}\right)^2 f(y)dy \right) - \left\{ \int_0^1 \frac{1}{h} K\left(\frac{x-y}{h}\right) f(y)dy \right\}^2 \tag{11}$$

When considering estimation of the discrepancy between the density estimation $\hat{f}(x)$ and the true density $f(x)$, a commonly used measure is the mean integrated square error (MISE) defined as:

$$EQM(\hat{f}) = \int_{-\infty}^{\infty} E\{\hat{f}(x) - f(x)\}^2 dx + \int_{-\infty}^{\infty} var \hat{f}(x) dx \tag{12}$$

For a generic Kernel K , the optimal value of h from the viewpoint of minimizing the MISE is given by [27]:

$$h_{ot} = K_2^{-\frac{2}{5}} \left\{ \int_{-\infty}^{\infty} K(t)^2 dx \right\}^{\frac{1}{5}} \left\{ \int_{-\infty}^{\infty} f''(x)^2 dx \right\}^{-\frac{1}{5}} n^{-\frac{1}{5}} \tag{13}$$

For a Gaussian Kernel the optimum window width h_{ot} can be explicitly estimated by [27]:

$$h_{ot} = \frac{4}{3} \sigma n^{-\frac{1}{5}} = 1.06 \sigma n^{-\frac{1}{5}} \tag{14}$$

where σ is the standard deviation of the process X .

Fig. 1 shows the histogram of the multipliers between stages 1 and 2 obtained using the method of Section II.D and their density function through the Kernel method for the dec-pkt-1 multifractal traffic trace¹. It can be noticed that the assumed Gaussian multiplier with mean 0.5 by the VVGM model cannot sometimes be realistic. In fact, in this case, the distribution tends to be lognormal.

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

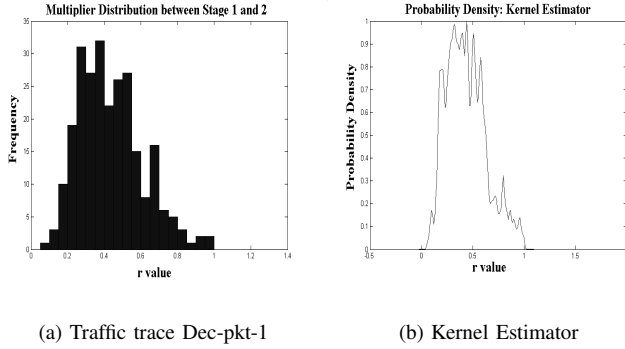


Fig. 1. Multiplier Distribution between Stages 1 and 2

B. Acceptance/Rejection Method

The Acceptance/Rejection method presented is used to generate random samples of a process X in accordance with the multiplier distribution $f(x)$. The basic Acceptance/Rejection algorithm has the following formulation: Let $g(x)$ be another probability density defined in the support of f such that:

$$f(x) \leq cg(x) \quad (15)$$

where $c > 1$ is a known constant that holds for all x in the support. The function $q(x) = cg(x)$ can be viewed as an envelope function of $f(x)$. Random samples of X can be generated by running the following algorithm [28]:

Acceptance/Rejection Algorithm:

Repeat

1. Generate sample y from $g(x)$;
 2. Generate sample u from the uniform distribution $U(0,1)$;
- Until
3. $u \leq f(y)/cg(y)$;
- Return $X = y$.

It can be shown that the process X obtained by the Acceptance-Rejection method has the required distribution as stated by the following theorem.

Theorem 1: Let U be a uniform random variable and Y another random variable whose density probability is $g(x)$. The random variable X generated by the Acceptance-Rejection method has probability density $f(x)$.

The proof of this theorem can be easily done by using the fact that $P(X \leq x) = P\{X \leq x | U \leq f(x)/(cg(x))\}$ with U and X being independent.

The closer the enveloping function to $f(\cdot)$, the better the acceptance rate. For this method to be efficient, an appropriate constant c must be selected in order to get low rejection rate. One way to choose an optimum c is by setting [28]:

$$c = \sup_x f(x)/g(x) \quad (16)$$

From the Gaussian sum obtained by the Kernel method, we can determine the constant c by using the following equation [28]:

$$c = 1/(\min(\sigma_g^2) \cdot n_g \cdot \sqrt{2\pi}) \quad (17)$$

where n_g is the number of gaussian functions and σ_g^2 ($g = 1, \dots, n_g$) represents the variance of the gaussian function. In

equation (17), $g(x)$ is assumed to be a uniform distribution, i.e., $g(x) = 1$ for all x , and $\hat{f}(x)$ corresponds to the gaussian mixture given by the Kernel method.

C. CGMD Synthesis Procedure

In this section we present the CGMD synthesis method that consists of a multiplicative cascade procedure assuming different multiplier distributions in each cascade. The CGMD synthesis procedure involves the aggregated process value obtained in the coarsest scale of the cascade, the probability density of the multipliers estimated through the Kernel method and the multipliers obtained in each stage by the Acceptance-Rejection method. Explicitly, we compute the multiplier densities $\hat{f}_i^*(x)$ in each stage $i = 1, 2, \dots, N$ through the Kernel estimator by the following equation:

$$\hat{f}_i^*(x) = \frac{1}{n\delta^i} \sum_{p=1}^n K\left(\frac{x - z_i}{\delta^i}\right) \quad (18)$$

where $n = 2^N$ and the mean z_i and the variance δ^i of the cascade multipliers in each stage are respectively given by:

$$z_i = E\{\hat{f}_i^*(x)\} \quad i = 1, 2, \dots, N \quad (19)$$

$$\delta^i = \left(\int_0^1 \frac{1}{(h_i)^2} K\left(\frac{x-y}{(h_i)^2}\right)^2 \hat{f}_i(y) dy \right) - \left\{ \int_0^1 \frac{1}{h_i} K\left(\frac{x-y}{h_i}\right) \hat{f}_i(y) dy \right\}^2 \quad (20)$$

In fact, we are using a gaussian mixture to model the real probability density of the multipliers. In summary, the synthesis algorithm of the CGMD process basically consists of:

- 1) Calculate the aggregated value obtained at the coarsest scale;
- 2) Estimate the multipliers using (7);
- 3) Apply the Kernel method to estimate the multiplier distribution;
- 4) Apply the Acceptance-Rejection method in the distribution given by (18) to generate the multipliers at each stage i of the cascade;
- 5) By using the obtained multipliers and the aggregated value of the traffic process, the same procedure described previously for generation of a multiplicative binomial cascade (Sections II.A) is applied.

D. Hurst Parameter for the CGMD Multifractal Model

This section is dedicated to the derivation of a global scaling law (Hurst parameter [14]) associated with the CGMD multifractal process and the multifractal analysis. The obtained analytical expression for the Hurst parameter will be used for the estimation of the statistical performance bounds.

In the multifractal analysis, the scaling of moments are mainly evaluated through the scaling function $\tau(q)$ (also known as partition function). For a multiplicative cascade process, the scaling function $\tau(q)$ is represented by (5). For $q = 2$, we have the multifractal scaling exponent $\tau(2)$ which is closely related to the LRD parameter H since both parameters measure the power-law behavior of second-order statistics.

More precisely, $\tau(2)$ captures the scaling behavior of the second sample moments of the process while H , the decay of the covariance. Actually, one can demonstrate that [12]:

$$H = \frac{\tau(2) + 1}{2} \tag{21}$$

An important characteristic of the CGMD model is that its cascade multiplier distributions are expressed as a weighted sum of Gaussian functions by (18). Since the partition function is related to the multiplier distribution through (5) and by using (10), we can express the Hurst parameter of the CGMD model as (see [29] for details on $\tau(q)$ estimation):

$$H = \frac{-\sum_{i=1}^N \log_2 E[\hat{f}_i^*(x)^2]}{2} \tag{22}$$

and by applying the properties of the Kernel estimator (See eq. (11)), we have:

$$H = \frac{-\sum_{i=1}^N \log_2 \left(\int_0^1 \frac{1}{(\delta^i)^2} K \left(\frac{x-z}{(\delta^i)^2} \right)^2 \hat{f}_i^*(z) dz \right)}{2} \tag{23}$$

where N is the number of cascade stages, δ^i and \hat{f}_i^* are respectively given by (20) and (18). Equation (23) permit us to estimate the Hurst parameter by the knowledge of the multiplier probability density of the CGMD model, i.e., through the corresponding CGMD modeling parameters. Table I compares for two well-known network traffic traces [2][10] the Hurst parameter values given by (23) to those based on the wavelet method described in [30].

TABLE I

MEAN, VARIANCE, HURST PARAMETER AND CGMD MODEL BASED HURST PARAMETER

Trace	Mean	Variance	Hurst P.	CGMD H
lbl-tcp-5	2.6146.10 ³	1.0033.10 ⁷	0.7811	0.8050
Bc-Aug	1.3819.10 ⁴	1.7111 ⁸	0.8617	0.8789

IV. CGMD MODEL EVALUATION

In order to compare the statistical characteristics of the CGMD model to those of real network traffic, as well as to those of other traffic models, some tests were conducted to evaluate some statistical parameters such as: autocorrelation function, high order moments of aggregated data and multifractal spectrum. Besides that, to verify the modeling accuracy in representing real world traffic, simulations were carried out to analyze the queue size and byte loss in a single server system fed by CGMD model based synthetic traffic. TCP/IP traffic traces (dec-pkt-1.tcp and dec-pkt-2.tcp) from the Digital Equipment Corporation² and Ethernet traffic traces collected by Bellcore³ were used in the simulations. We considered samples of traffic in a scale aggregation of 512ms and 100ms for the TCP/IP and Ethernet traffic traces, respectively. In these scale aggregations the considered traces present self-similar and multifractal characteristics [2] [10].

²http://ita.ee.lbl.gov/html/contrib/DEC-PKT.html

³http://www.acm.org/sigcomm/ITA

A. Autocorrelation Function

The autocorrelation function reflects the second order statistics of a time series, giving an idea regarding the long-range dependence in data.

Let $y(t)$ be a process with mean μ_t and variance σ_t^2 , and $y(t+k)$ be the same process shifted in time with mean μ_{t+k} and variance σ_{t+k}^2 . The correlation coefficients for the process $y(t)$ is given by:

$$\rho(k) = \frac{E[(y(t+k) - \mu_{t+k})(y(t) - \mu_t)]}{\sigma_{t+k}\sigma_t} \tag{24}$$

Fig. 2 shows the correlation coefficients for the Dec-pkt-2 traffic traces, as well as for the CGMD and MWM based synthetic traces. Notice that the correlation coefficients of the CGMD model for the majority of the analyzed k values are closer to those of the real traffic traces.

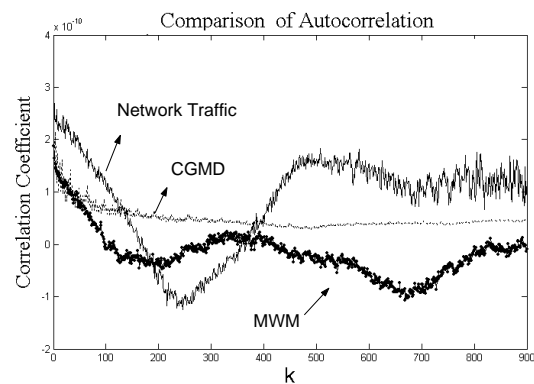


Fig. 2. Correlation Coefficient: Dec-pkt-2 Traffic Trace

B. High Order Moments of Aggregated Data

Aggregated traffic greatly influences queuing behavior in multiplexers and routers. Therefore, it is important to investigate the statistics of the aggregated traffic. In this subsection, we compare the higher order moments of the aggregated synthetic traffic traces generated by different cascade models. Let X_k^m be given by:

$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i^N \tag{25}$$

where $i = 1, 2, \dots, 2^N$ and $k = 1, 2, \dots, L$ with $L = 2^N/m$ denoting the total number of aggregations made for a fixed value of m . The q^{th} moment of the aggregated traffic is estimated as follows:

$$\hat{\mu}^{(m)}(q) = \frac{1}{L} \sum_{k=1}^L |X_k^{(m)}|^q \tag{26}$$

Fig. 3 presents the variation of $\log_{10}(\hat{\mu}^{(m)}(q))$ versus $\log_{10}(m)$ for $q = 2$ and $q = 4$. The CGMD model provides the tightest moment estimates which are also larger than those derived from the real traffic traces. This fact indicates that the CGMD model can faithfully reproduce bursty behavior indeed occurring in real traffic traces.

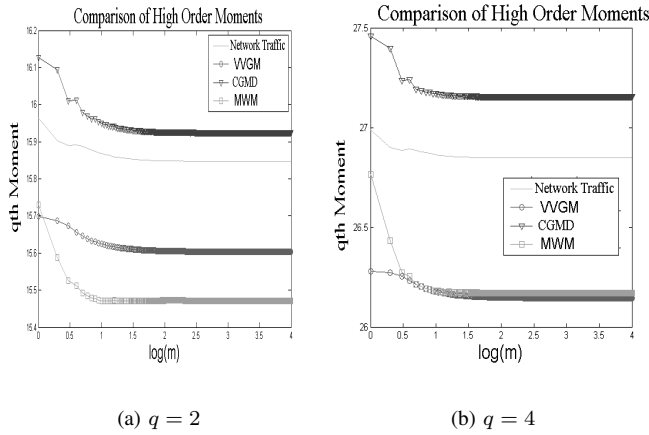


Fig. 3. High Order Moments of Aggregated Data (Bc-Aug)

C. Multifractal Spectrum

In contrast to Markovian and monofractal traffic types, multifractal traffic processes present a multiplicity of local Hölder exponents within any finite interval. The Hölder exponents describe the local scaling properties of a realized path at any point in time. Their distribution can be represented by a renormalized density called the multifractal spectrum. In an alternative interpretation, the multifractal spectrum describes the fractal dimension of the set of time instants in the process having a given local exponent [12]. While the statistical self-similarity of the sets accounts for long memory in the process, the concept of Holder exponent is related with the local singularity of a process, i.e, it characterizes the burstiness of a process at a given time instant [31]. The multifractal spectrum $f(\alpha)$ of a process $X(t)$ can be obtained via the Legendre Transform of $\tau(q)$ (partition function) defined as:

$$f(\alpha) = \min_q \{q\alpha - \tau(q)\} \tag{27}$$

Fig. 4 compares the multifractal spectra generated by different multifractal traffic models with respect to that computed directly from the Internet traffic trace dec-pkt-1. Notice that the multifractal spectrum of the Internet traffic traces (dec-pkt-1 and dec-pkt-2) presents a great amount of $\alpha < 1$ values. This property characterizes high incidence of multiscale bursts. Also the CGMD model captures efficiently the multifractal spectrum and for consequence the “multifractality” of the real traffic traces.

D. Performance and Queueing Tests

We considered a single server system with a finite buffer fed by CGMD synthetic traffic samples in order to compare its accuracy in representing real data traffic. We mainly analyzed the mean buffer occupation (queue length distribution) and the byte loss rate associated to different buffer utilizations. We defined the buffer utilization factor λ as being the ratio of the total service time to the total time for which the buffer was in use. A higher value of λ indicates that the buffer is constantly being used and consequently there could be a larger probability

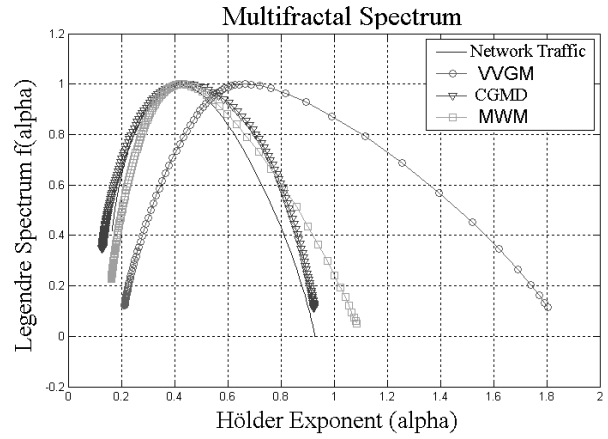


Fig. 4. Multifractal Spectrum: Dec-pkt-1 Traffic Trace

of data being discarded. A value of λ smaller than 0.4 means that the buffer is being under utilized, a value between 0.6 and 0.7 is considered ideal in real networks [32].

Fig. 5.(a) shows how the byte loss varies in function of the buffer utilization for a finite buffer size of 64Kbytes. Fig.5.(b) expresses the mean buffer occupation in bytes (queue length distribution) versus the buffer utilization. It can be observed a larger loss rate for the CGMD traffic than for the MWM model one, probably due to the higher moments of the CGMD trace.

Fig. 6 displays the relationship between the byte loss rate and the buffer size for the traffic trace dec-pkt-1 for a fixed buffer utilization of 0.5. Among the analyzed traffic models, a slightly larger loss rate was observed for the CGMD synthetic traffic trace (but closer to the real ones). It’s worth to remind that the CGMD model is based on the cascade multiplier distribution estimation while the MWM relies on the decay of wavelet energy on scale. This fact can lead to some differences in the results of these models.

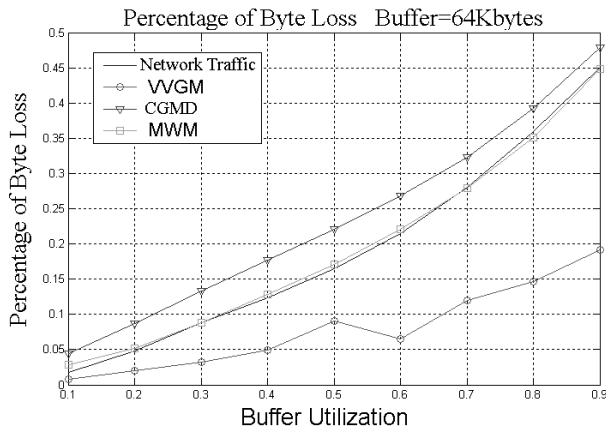
V. EFFECTIVE BANDWIDTH FOR THE CGMD MODEL

The concept of effective bandwidth for high speed networks was first independently presented in [33][34][35] where i.i.d and On-Off sources were considered. The methods of effective bandwidth estimation for Markov and other processes were carried out in [36][37][38]. Further theoretic development for admission control, traffic regulation, and other applications can be found in [39][40] and many others. The effective bandwidth of a traffic flow is a rate greater than the mean flow rate but less than its peak rate, that can be used to estimate the necessary link capacity that attains the required QoS, given the available buffer capacity.

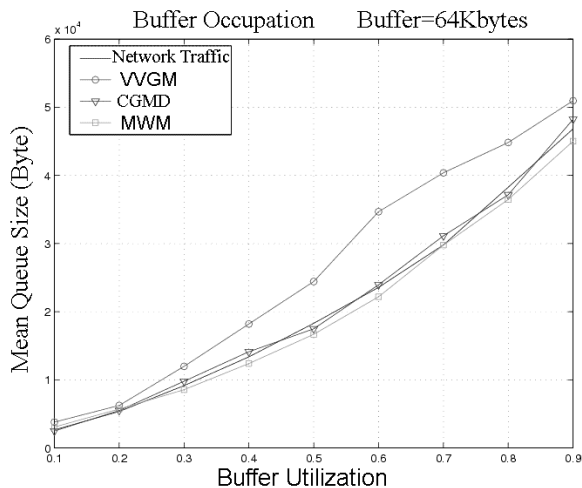
Formally, the effective bandwidth (EB) of a traffic stream is defined as [41][34]:

$$\alpha(\theta, t, N) = \frac{1}{\theta t} \log E_N[e^{\theta X(0,t)}] \quad 0 < \theta; 0 < t < N \tag{28}$$

where $X(0, t)$ indicates the total number of arrived bytes in the interval of length t , θ is the space parameter and $E_N[e^{\theta X(0,t)}$ is the measured moment generating function over a trace consisting of N samples [34]. In the following proposition, we state the effective bandwidth formula for the CGMD model.



(a) Bc-Aug Traffic Trace



(b) Dec-pkt-1 Traffic Trace

Fig. 5. Byte Loss Probability and Mean Buffer Occupation

Proposition 1 The effective bandwidth of the CGMD Multifractal Modeled process X_t can be expressed as:

$$\alpha(\theta, t) = \frac{1}{\theta t} \log \left(\sum_q \frac{\theta^q}{q!} \left(\sum_{t=1}^{2^N} X_t \right)^q \prod_{j=0}^{N-1} \left\{ \int_0^1 \frac{1}{(\delta^j)^q} K \left(\frac{x-z}{(\delta^j)^q} \right)^q f_j^*(z) dz \right\} \right) \quad (29)$$

where N is the number of cascade stages, δ^j and f_j^* are respectively given by (20) and (18).

Proof: See Appendix.

Fig. 7 presents the effective bandwidth given by (29) for the lbl-tcp-5 traffic trace in comparison to the Norros' monofractal one [42]. In this case, the loss probability was set to 10^{-7} and the buffer size to 60Kbytes. From Fig. 7, it can be seen that the proposed effective bandwidth is close to that obtained by Norros' method [42]. Notice that the effective bandwidth proposed by Norros is based on the monofractal fBm model [42], while, to contrast, we make use of multifractal model

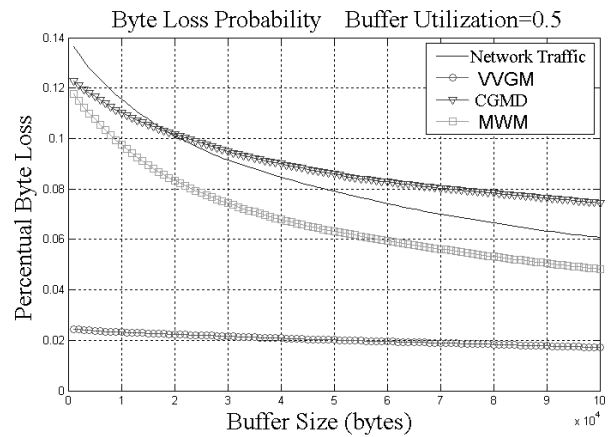


Fig. 6. Byte Loss Probability Versus Buffer Size: Dec-pkt-1 traffic trace

parameters in our approach. In fact, since the CGMD is a multifractal model, our result is more general, because the CGMD model based effective bandwidth includes the monofractal case.

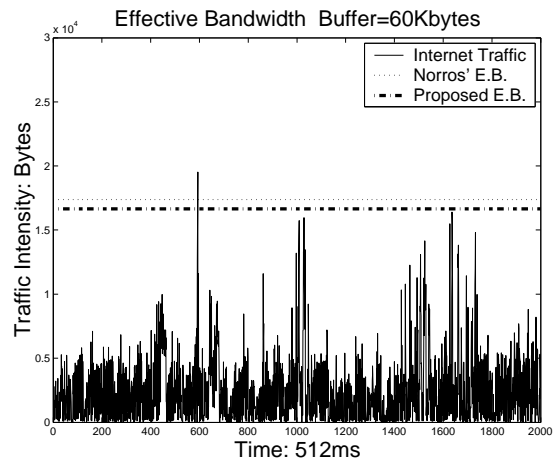


Fig. 7. Effective Bandwidth Estimation for the lbl-tcp-5 Traffic Trace

VI. PERFORMANCE BOUNDS THROUGH NETWORK CALCULUS AND THE CGMD EFFECTIVE BANDWIDTH

Research on performance bounds has opened a new direction for effective analysis and design of high-speed communication systems [43]. Performance bounds such as mean backlog and mean delay can be obtained in terms of effective bandwidth [17].

The envelope process bounds the original process for an arbitrary shift of time [15][19]. We denote the smallest envelope process as the minimum envelope process (MEP). The MEP is subadditive and has an average rate, denoted as minimum envelope rate (MER). Based on the MER, a set of rules for network operations such as superposition and input-output relation of a queue can be developed.

Before describing the stochastic framework for performance bounds, let's introduce some elements of deterministic network calculus. Consider a nonnegative sequence $\{a(t), t = 0, 1, 2, \dots\}$ corresponding to the arrival process.

Let $A(t_1, t_2) = \sum_{t=t_1}^{t_2-1} a(t)$. A bounding process $\hat{A}(t)$ has the following property:

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

where $\hat{A}(t_1, t_2) = \sum_{t=t_1}^{t_2-1} \hat{A}(t)$.

$\hat{A}(t)$ is called the envelope process of $a(t)$. A process $\hat{A}(t)$ is subadditive if $\hat{A}(t_1, t_2) \leq \hat{A}(t_1) + \hat{A}(t_2)$ for all t_1 and t_2 . $\hat{A}(t)$ is “stationary” in the sense it only depends on the difference t_1 and t_2 . We may assume that $\hat{A}(t)$ is increasing and subadditive, then:

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

where \hat{a} is the envelope rate of the process $\hat{A}(t)$. The minimum envelope process (MEP) $A^*(t)$ is also increasing and subadditive given by:

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

One can view the MEP $A^*(t)$ by the family of linear envelope process proposed by Cruz [15][16]: $A^*(t) \leq \hat{a}t + \hat{\sigma}$ for some nonnegative constant $\hat{\sigma}$. The minimum envelope rate (MER) a^* of $a(t)$ is:

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

Now, we turn our attention to the statistical network calculus framework. In a stochastic scenario for a server with capacity c and operating under a work-conserving policy if the MER of the input process is smaller than c , then we can state that [17]: i) the queue length is bounded exponentially; ii) there exists a linear envelope process of the departure process which can be represented as a function of the linear EP of the input process. Using these results, bounds for the tail distribution of queue length can be easily computed from the linear EP of the input process.

Let $\hat{A}(\theta, t)$ be the envelope process of $a(t)$ with respect to a parameter θ as defined in [17]:

$$\frac{1}{\theta} \log E e^{\theta A(t_1, t_2)} \leq \hat{A}(\theta, t_2 - t_1) \quad \forall \quad t_1 \leq t_2 \quad (34)$$

Then, the MEP with respect to θ is:

$$A^*(t) = \sup_{s \geq 0} \frac{1}{\theta} \log E e^{\theta A(s, s+t)} \quad (35)$$

Unlike the MEP in deterministic framework, this MEP is not subadditive in general. Thus, the minimum envelope rate of $a(t)$ with respect to θ is given by [17]:

$$a^*(\theta) = \limsup_{t \rightarrow \infty} \frac{A^*(\theta, t)}{t} \quad (36)$$

An arrival process $\{a(t)\}$ is said to be bounded by a linear envelope process $a(\theta)t + \sigma(\theta)$ where $a(\theta) \geq 0$, $\sigma(\theta) \geq 0$ with respect to θ if:

$$\frac{1}{\theta} \log E e^{\theta A(t_1, t_2)} \leq \alpha(\theta)(t_2 - t_1) + \sigma(\theta) \quad (37)$$

Therefore, $\alpha(\theta)$ is bounded on the “equivalent”, stationary byte arrival rate while $\sigma(\theta)$ is usually interpreted as a bound on the burstiness of the arrival process.

The definition of MER is connected to the Large Deviation theory through the Gärtner-Ellis theorem [44]. We have to consider the following conditions to $\{a(t), t \geq 0\}$:

- i) $\{a(t), t \geq 0\}$ is stationary and ergodic;
- ii) $a^*(\theta) = \lim_{t \rightarrow \infty} \frac{A^*(\theta, t)}{t}$ for all $0 < \theta < \infty$;
- iii) $\theta a^*(\theta)$ is strictly convex and differentiable for all $0 < \theta < \infty$.

Under these conditions, the sequence $\{A(0, t), t \geq 1\}$ obeys the large deviation principle with the rate function defined as [44]:

$$I(v) = \sup_{\theta} \{\theta v - \theta a^*(\theta)\} \quad (38)$$

If the arrival process $a(t)$ consist of i.i.d random variables, the MER $a^*(\theta)$ is referred to as effective bandwidth [45].

The pair $(\alpha^*(\theta), \sigma^*(\theta))$ defines a minimum LEP $\alpha^*(\theta)t + \sigma^*(\theta)$ with respect to θ :

$$\alpha^*(\theta) = \lim_{t \rightarrow \infty} \sup \frac{1}{t} \sup_{\theta \geq 0} \frac{1}{\theta} \log E e^{\theta A(t_1, t_2)} \quad (39)$$

and

$$\sigma^*(\theta) = \inf \left\{ \sigma(\theta) \mid \frac{1}{\theta} \log E e^{\theta A(t_1, t_2)} \leq \alpha^*(\theta)(t_2 - t_1) + \sigma(\theta) \right\} \quad (40)$$

where $t_2 \geq t_1 \geq 0$. The expression of the minimum LEP in (39) is valid for general traffic. Note that $a(t_1)$ and $a(t_2)$ are not necessarily independent. If the following limit exists:

$$h(\theta) = \lim_{t \rightarrow \infty} \frac{1}{t} \log E e^{\theta A(0, t)} \quad (41)$$

then $h(\theta)/\theta$ is the effective bandwidth of $a(t)$ with respect to θ . Thus, $\alpha^*(\theta)$ is exactly the effective bandwidth of the process $a(t)$.

Let $\{W(t)\}$ be the backlog process (queue length process). Assume that the arrival process $a(t)$ is independent of $W(0)$ and $\alpha^*(\theta) < c$, then the moment generating function of the backlog process is bounded by [19]:

$$E[e^{\theta W(t)}] \leq e^{\theta(\alpha^*(\theta)-c)t} e^{\theta \sigma^*(\theta)} E[e^{\theta W(0)}] + B(\theta) \quad (42)$$

where

$$B(\theta) = \frac{(1 - e^{-c\theta})e^{\theta \sigma^*(\theta)}}{1 - e^{\theta(\alpha^*(\theta)-c)}} \quad (43)$$

When $t \rightarrow \infty$, the bound $B(\theta)$ is tighter than that obtained in [17] by a factor of $(1 - e^{-c\theta})$. The bound on the moment generating function can be used to derive bounds on the backlog process, the minimum linear envelope process of the output process and the tail distribution of the backlog process. Under the condition of $a(t)$ being independent of $W(0)$ and $\alpha^*(\theta) < c$, we have the following results for a single-server connection [19]:

- i) The tail distribution of the backlog process is exponentially bounded as:

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

ii) The mean queue length is bounded by:

$$E[W(t)] \leq \frac{\{e^{\theta(\alpha^*(\theta)-c)t} e^{\theta\sigma^*(\theta)} E[e^{\theta W(0)}] + B(\theta)\}}{(1 - e^{-\theta})} \quad (45)$$

Substituting $\alpha^*(\theta)$ by the proposed effective bandwidth given by (29) and $\sigma^*(\theta)$ by (40), we can compute the performance bounds for a queue fed by a CGMD multifractal process. The parameter θ in (44) and (45), can be determined through the operating point (t^*, θ^*) , i.e., the values of time t and space parameter θ for which the effective bandwidth is related to the asymptotic overflow probability [46]. Determination of θ as the operating point θ^* guarantees a tighter performance bound. For a process with mean μ , variance σ^2 and Hurst parameter H , we have the following equations for the operating point pair (t^*, θ^*) [47]:

$$t^* = \frac{HB}{(c - \mu)(1 - H)} \quad (46)$$

and

$$\theta^* = \frac{B(t^*)^{-2H}}{\sigma^2(1 - H)} \quad (47)$$

where B is the buffer size and c is the link capacity.

A. Loss Probability for Long-Range Dependent Processes

The works of Norros [42] and Duffield and O’Connell [48] are concerned with queuing systems with self-similar input such as the fBm process. Duffield and O’Connell introduced an asymptotic expression for the loss probability $P(Q > b)$ at steady-state by applying a Large Deviation Principle (LDP). The lower-bound for $P(Q > b)$ decays asymptotically (for large buffer b) in a Weibullian fashion. The tail distribution of the queue occupancy in this case is much heavier than the exponential distribution predicted by traditional traffic models. The result presented by Duffield and O’Connell in [48] shows that $P(Q > b)$, when scaled properly, will satisfy the large deviation principle and for $H \in (0.5, 1)$ the following relation holds:

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

where $a = c/H - c$.

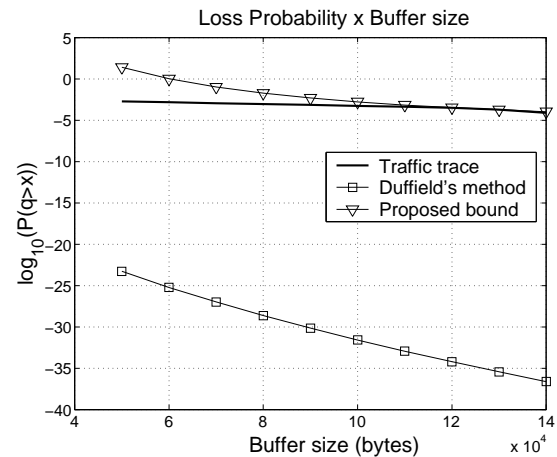
B. Experimental Results for the Performance Bounds

In this section, we validate our performance bound proposals by simulations with TCP/IP and Ethernet traffic traces and traces collected between the years 2000 and 2002 at the Petrobrás computer network [49]. We aggregated the traffic traces (lbl-pkt-5 and Bc-Aug) at the 100ms time scale because the considered traces present multifractal characteristics at that time scale [9][10]. We verified the same for the Petrobrás traffic trace 10-7-S-1 by applying multifractal analysis through its multifractal spectrum and scaling of wavelet moments [31]. The numbers of traffic samples N_t for traces lbl-pkt-5, Bc-Aug and 10-7-S-1 are 2^{15} , 2^{14} and 2^{12} , respectively.

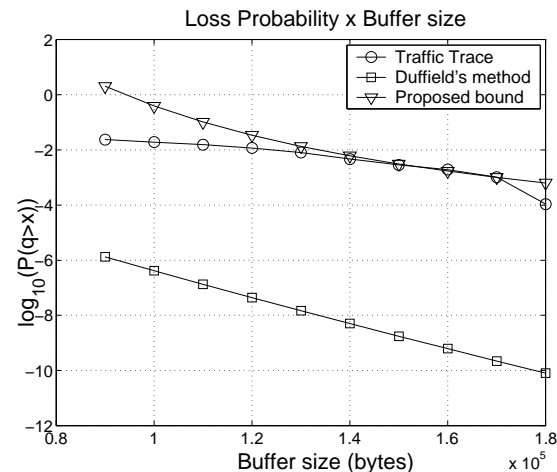
The estimation of the proposed loss probability and mean queue length bounds is accomplished through the effective bandwidth computing. In the evaluation of the loss probability

bound proposal represented by (44) in conjunction with the CGMD based effective bandwidth (29), we considered a single server connection with finite buffer and capacity c where for lbl-tcp-5, 10-7-S-1 and Bc-Aug traffic traces, c is set to 5.6μ , 1.7μ and 3.1μ , respectively, and μ is the mean traffic rate. We present in Fig. 8, the loss probability at steady state, i.e., for $t \rightarrow \infty$, the loss probability obtained for the real traffic series and that given by large deviations theory (Duffield’s method). As it can be seen from Fig. 8, the Duffield and O’Connell’s approach underestimates the loss probability rate in comparison with that given by the real traffic traces. Similarly to Duffield’s, our method provides better results for larger buffers.

Fig. 9 shows the mean queue length given by (45) and the simulated mean buffer occupation. It can be noticed that as the buffer size is increased, the proposed mean queue length bound becomes closer to the mean buffer occupation of the link fed by the lbl-tcp-5 traffic trace.



(a) Traffic trace: lbl-tcp-5



(b) Traffic trace: 10-7-S-1

Fig. 8. Loss probability versus buffer size

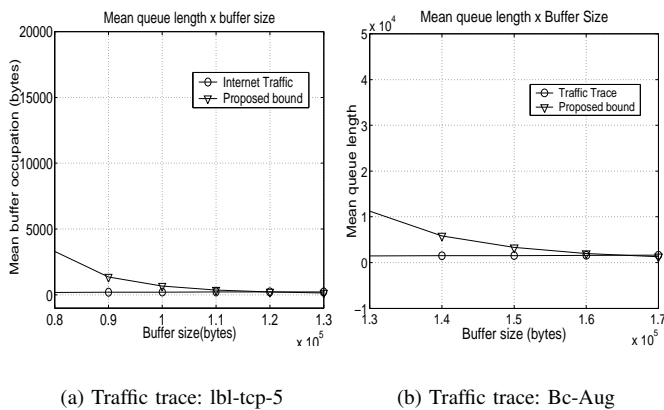


Fig. 9. Mean buffer occupation

VII. CONCLUSION

We can conclude that the proposed multifractal model efficiently captures the network traffic characteristics. It is also noticed that a better estimation of the cascade multiplier density can lead to modeling improvements. The simulations reveal that the performance of the CGMD modeling is comparable or even superior to that of the MWM and VVGM modeling.

One of the challenges of network traffic characterization is to accurately model queueing behavior. In this paper, we demonstrated that the queueing behavior characterization can be enhanced by connecting the CGMD properties, such as effective bandwidth and Hurst parameter, to statistical network calculus. This result is obtained due to both the CGMD modeling potential and the network calculus flexibility. We have contributed with a new QoS estimation tool, namely the CGMD based effective bandwidth estimator, capable of attaining the required loss probabilities for real world network flows.

The statistical network calculus allows us to achieve accurate expressions for the byte loss probability and mean queue length estimation by the envelope process concept [50]. The simulation results demonstrated that the performance bounds, based on the union of network calculus and the CGMD based effective bandwidth, are tighter than those offered by the large deviations theory for long-range dependent processes.

Some works have stated that the Internet backbone is monofractal in small time scales [31] [51]. In fact, our proposal is also suitable in this case, because the CGMD model, as a multifractal model, also includes the monofractal case in its structure. Indeed, we have verified that the proposed loss probability and mean queue length estimates are adequate for a large variety of traffic traces such as Internet and Ethernet ones.

Finally, from our analysis, it can be concluded that the proposed performance bounds can be promptly inserted into admission control and other control schemes to guarantee required QoS. Besides, more than one network performance parameter can be analyzed, providing a more complete description of the network resources.

APPENDIX I

PROOF OF PROPOSITION 1

The effective bandwidth of the traffic process X_t is defined as:

$$\alpha(\theta, t) = \frac{1}{\theta t} (\log E\{e^{\theta X_t}\}) \quad (49)$$

Using the Taylor expansion of the exponential function ($q \rightarrow \infty$), we can write (49) as:

$$\alpha(\theta, t) = \frac{1}{\theta t} \log E \left\{ \sum_q \frac{(\theta X_t)^q}{q!} \right\} \quad (50)$$

$$\alpha(\theta, t) = \frac{1}{\theta t} \log \left(\sum_q \frac{(\theta)^q}{q!} E\{X_t^q\} \right) \quad (51)$$

Let the multipliers $r_j^{(i)}$ be i.i.d random variables with probability density function f_{R_j} in stage j , then the moments for a multifractal cascade with N stages is given by [12]:

$$E\{X_t^q\} = \left(\sum_{t=1}^{2^N} X_t \right)^q \prod_{j=0}^{N-1} E\{R_j^q\} \quad (52)$$

Therefore, we can compute the moments of the CGMD model as:

$$E\{X_t^q\} = \left(\sum_{t=1}^{2^N} X_t \right)^q \prod_{j=0}^{N-1} \left\{ \left(\int_0^1 \frac{1}{(\delta^j)^q} K \left(\frac{x-z}{(\delta^j)^q} \right)^q \hat{f}_j^*(z) dz \right) \right\} \quad (53)$$

Substituting (53) into (51), we obtain the following equation for the CGMD model effective bandwidth:

$$\alpha(\theta, t) = \frac{1}{\theta t} \log \left(\sum_q \frac{\theta^q}{q!} \left(\sum_{t=1}^{2^N} X_t \right)^q \prod_{j=0}^{N-1} \left\{ \int_0^1 \frac{1}{(\delta^j)^q} K \left(\frac{x-z}{(\delta^j)^q} \right)^q \hat{f}_j^*(z) dz \right\} \right) \quad (54)$$

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their comments and suggestions.

REFERENCES

- [1] F. H. P. Fitzek and M. Reisslein. "MPEG-4 and H.263 video traces for network performance evaluation". IEEE Network, Vol.15, No.6, pp.40-54, December 2001.
- [2] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. "On the self-similar nature of Ethernet traffic (extended version)". IEEE/ACM Transactions on Networking, Vol.2, No.1, pp.1-15, 1994.
- [3] R. G. Addie, M. Zukerman, and T. D. Neame. "Fractal traffic: Measurements, modeling and performance evaluation". Proceedings of IEEE Infocom '95, pp. 977-984, April 1995.
- [4] V. Paxson and S. Floyd. "Wide-area traffic: The failure of Poisson modeling". IEEE/ACM Transactions on Networking, Vol.3, No.3, pp.226-244, 1995.
- [5] A. Erramilli, O. Narayan and W. Willinger. "Experimental Queueing Analysis with Long-Range Dependent Packet Traffic", IEEE/ACM Trans. on Net., Vol.4, No.2, April 1996.
- [6] K. Park and W. Willinger. "Self-similar Network Traffic and Performance Evaluation", John Wiley & Sons, New York, 2000.
- [7] M. Grossglauser and J.-C. Bolot. "On the relevance of long-range dependence in network traffic". IEEE/ACM Transactions on Networking, Vol.7, No.5, pp.629-640, October 1999.
- [8] Q. Li and D. Mills. "Investigating the Scaling Behavior, Crossover and Anti-persistence of Internet Packet Delay Dynamics", Globecom, pp.1843-1852, 1999.

- [9] A. Feldmann, A. C. Gilbert, and W. Willinger. "Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic". Proc. of the ACM/SIGCOMM'98, Vancouver, B.C., pp. 25-38, 1998.
- [10] A. Erramilli, O. Narayan, A. Neidhardt, and I. Sanjeev. "Performance impacts of multi-scaling in wide area TCP/IP traffic". Proceedings of IEEE Infocom 2000, March, 2000.
- [11] J. Lévy Véhel and R. Riedi. "Fractional Brownian motion and data traffic modelling: The other end of the spectrum". In J. Lvy Vehel, E. Lutton, and C. Tricot, editors, *Fractals in Engineering*. Springer, 1997.
- [12] R.H. Riedi, M.S. Crouse, V.J. Ribeiro, and R.G. Baraniuk. "A Multifractal Wavelet Model with Application to Network Traffic", *IEEE Trans. on Information Theory*, Vol. 45, No.3, April 1999.
- [13] J. Gao and I. Rubin. "Multifractal analysis and modeling of long range-dependent traffic". Proceedings of ICC'99, June 1999.
- [14] M. Krishna et al. "Global scaling exponent for variable variance gaussian multiplicative (VVGGM) multifractal cascades". Proc. SPCOM'01, pp.19-25, 2001.
- [15] R. L. Cruz, "A calculus for network delay Part I: Network elements in isolation", *IEEE Trans. Information Theory*, Vol.37, pp.114-131, 1991.
- [16] R. L. Cruz, "A calculus for network delay Part II: Network Analysis", *IEEE Trans. Information Theory*, Vol.37, pp.132-141, 1991.
- [17] C. S. Chang, "Stability, queue length and delay of deterministic and stochastic queueing networks", *IEEE Trans. Automat. Contr.*, Vol.39, No.5, pp. 913-931, 1994.
- [18] L. Dai, "Effective bandwidths and performance bounds in high-speed communication systems", *Decision and Control*, Proceedings of the 36th IEEE Conference, Vol.5, pp. 4580-4585, 10-12 December 1997.
- [19] J. Y. LeBoudec and P. Thiran, "Network Calculus", *Lecture Notes in Computer Science*. New York: Springer-Verlag, 2001.
- [20] J. Liebeherr; S.D. Patek; A. Burchard, "Statistical per-flow service bounds in a network with aggregate provisioning", *IEEE INFOCOM*. Vol.3, pp.1680-1690, 30 March to 3 April 2003.
- [21] R. R. Boorstyn; A. Burchard; J. Liebeherr; C. Oottamakorn, "Statistical service assurances for traffic scheduling algorithms", *IEEE Transactions on Selected Areas in Communications*, Vol.18, No.12, pp.2651-2664, December 2000.
- [22] Q. Li; D. Mills, "Investigating the Scaling Behavior, Crossover and Anti-persistence of Internet Packet Delay Dynamics", *Globecom 3*, pp. 1843-1852, 1999.
- [23] T. D. Dang, S. Molnr, I. Maricza, Performance analysis of multifractal network traffic, *European Transactions on Telecommunications*, Vol. 15. Issue 2, March-April 2004, pp. 63-78.
- [24] B. B. Mandelbrot, A. Fisher, and L. Calvet. "A Multifractal Model of Asset Return". Yale University, 1997.
- [25] C. K. Chui. "An Introduction to Wavelets". San Diego: Academic, 1992.
- [26] V. J. Ribeiro, R.H.Riedi, M.S.Crouse and R.G.Baraniuk, "Multiscale queueing analysis of long-range dependent traffic", Proc. IEEE Infocom, March, 2000.
- [27] B. Silverman. "Density Estimation for Statistics and Data Analysis". Chapman and Hall, London, 1986.
- [28] K.Monahan. "Numerical Methods of Statistics". Cambridge University Press, NY, 2001.
- [29] A.C.Gilbert, W.Willinger and A.Feldmann. "Scaling analysis of conservative cascades with applications to network traffic", *IEEE Trans. on Information Theory*, Vol. 45, No.3, pp.971-991, 1999.
- [30] D. Veitch; D. Abry, "A wavelet based joint estimator for the parameters of LRD", *IEEE Trans. Info. Theory.*, Vol.45, No.3, April 1999.
- [31] Z. Zhang, V. Ribeiro, S. Moon, and C. Diot. "Small-Time Scaling behaviors of internet backbone traffic: An Empirical Study," *IEEE Infocom*, San Francisco, March, 2003.
- [32] M. Krishna, V. M. Gadre and U. B. Dessay. "Multifractal based network traffic modelling". Kluwer Academic Publishers, 2003.
- [33] R. Guerin, H. Ahmadi and M.Naghshineh, "Equivalent capacity and its application to bandwidth allocation in high-speed networks", *IEEE J. Select. Areas Commun.*, Vol. 9, pp.968-981, 1991.
- [34] F. P. Kelly, "Effective bandwidth at multi-class queues", *Queue Syst.*, Vol. 9, pp.5-16, 1991.
- [35] R. J. Gibbens and P.J.Hunt, "Effective bandwidths for the multi-type UAS channel", *Queue Syst.*, Vol. 9, pp.17-28, 1991.
- [36] C. S. Chang, "Stability, queue length and delay of deterministic and stochastic queueing networks," *IEEE Trans. Automat. Contr.*, Vol. 39, No. 5, pp. 913-931, May 1994.
- [37] A. Elwalid and D. Mitra, "Effective bandwidth of general markovian traffic sources and admission control of high speed networks," *IEEE/ACM Trans. Networking*, Vol. 1, No. 3, pp.329-343, June 1993.
- [38] G. de Veciana, G. Kesidis, and C. S. Chang, "Effective bandwidths of general Markovian traffic sources and admission control of high speed networks," *IEEE/ACM Trans. Networking*, Vol. 1, pp. 329-343, 1993.
- [39] G. de Veciana, G. Kesidis, and J. Walrand, "Resource management in wide-area ATM networks using effective bandwidths," *IEEE J. Select. Areas Commun.*, Vol. 13, pp. 1081-1090, Aug. 1995.
- [40] G. de Veciana, C. Courcoubetis, and J. Walrand, "Decoupling bandwidths for networks, a decomposition approach for networks," in Proc. IEEE INFOCOM'94, vol. 2, Toronto, Ont., Canada, June 12-16, 1994.
- [41] S. Tartarelli, M. Falkner, M. Devetsikiotis, I. Lambadaris and S. Giordano, "Empirical effective bandwidths", *IEEE Globecom'00*, San Francisco, 2000.
- [42] I. Norros, "A storage model with self-similar input", *Queueing Systems*, Vol. 16, pp.387-396, 1994.
- [43] J. Kurose, "On computing per-session performance bounds in high-speed multi-hop computer networks", *Proc. ACM Sigmetrics*, pp. 128-139, 1992.
- [44] J. A. Bucklew, "Large deviation techniques in decision, simulation and estimation", J.Wiley, New York, 1990.
- [45] F. P. Kelly, "Notes on effective bandwidths," in *Stochastic Networks*, Oxford, U.K.: Clarendon, pp. 141-168, 1996.
- [46] L. Aspirot, P. Belzarena, P. Bermolen, A. Ferragut, G. Perera and M. Simon, "Quality of service parameters and link operating point estimation based on effective bandwidths", *Performance Evaluation*, Vol. 59, Issues 2-3, pp. 103-120, February 2005.
- [47] C. Courcoubetis, V. A. Siris, and G. D. Stamoulis, "Application of the many sources asymptotic and effective bandwidths to traffic engineering," *Telecommunication Systems*, Vol. 12, pp. 167-191, 1999.
- [48] N. G. Duffield and N. O'Connell, "Large deviations and overflow probabilities for the general single-server queue, with applications," *Dublin Institute for Advanced Studies-Applied Probability Group*, Tech. 1 Rep. DIAS-STP-93-30, 1993.
- [49] F. Perlingeiro and L.L Lee, "Estudo de Estimação de Banda Efetiva para Tráfego Auto-similar com Variância Infinita". SBRT'05, Campinas, pp. 326-331, September 2005.
- [50] F. H. T. Vieira and L. L. Lee. "Queueing Analysis for Multifractal Traffic through Network Calculus and Global Scaling Parameter". *International Telecommunications Symposium - ITS*. Fortaleza-CE, Brazil, September, 2006.
- [51] D. A. Rolls, G. Michailidis and F. Hernandez-Campos. "Queueing analysis of network traffic: methodology and visualization tools". *Computer Networks*, Vol. 48, pp.447-473, 2005.



Flávio Henrique Teles Vieira received his B.Sc. degree in Electrical Engineering from the Federal University of Goiás (UFG) in 2000, the M.Sc. degree in Electrical and Computer Engineering from UFG in 2002 and the doctorate degree in Electrical and Computer Engineering at State University of Campinas (FEEC-UNICAMP) in 2006. He acts in the following research areas: Network Traffic Modeling, Prediction and Control and Communication Networks. E-mails: flavio@decom.fee.unicamp.br, flaviotel@hotmail.com



Luan Ling Lee obtained the B.Sc. degree in electrical engineering from USP in 1980, the master degree in Electrical Engineering from UNICAMP in 1984 and PhD in Electrical Engineering from University of Cornell in 1991. It is the founder and it has been the coordinator of the Pattern Recognition and Communications Networks Laboratory (LRPRC) of FEEC-UNICAMP since 1994. He was the Department of Communications chief in 2000. Since 2002, he has been a Professor at FEEC-UNICAMP. He acts intensely in two research areas: Pattern Recognition and Communication Networks. E-mail: lee@decom.fee.unicamp.br