# Queueing Performance Estimation for General Multifractal Traffic

Trang Dinh Dang, Sándor Molnár, István Maricza

High Speed Networks Laboratory
Department of Telecommunications and Telematics
Budapest University of Technology and Economics
H–1117, Magyar tudósok körútja 2, Budapest, Hungary
Tel: (361) 463 3889, Fax: (361) 463 3107
E-mail: {trang, molnar, maricza}@ttt-atm.ttt.bme.hu

**Abstract**

We present an approximation for the tail asymptotics in an infinite capacity single server queue serviced at a constant rate driven by general *multifractal* input process. We show that in the special and important case of the *monofractal* fractional Brownian motion (fBm) input traffic our result gives the well-known Weibullian tail. We prove that the class of Gaussian processes with scaling properties is in the class of monofractal processes and we derive the related characterization functions. Our formula in the case of Gaussian input processes also gives a queueing result which is in good agreement with the theory of Gaussian processes. Applying the approximation we provide a new practical method for queueing performance estimation of general multifractal traffic. The validation of the method based on both analysis of simulations and measured network traffic have also been presented.
**Keywords:** Gaussian processes, multifractality, network traffic, queueing analysis.

## 1 Introduction

In the last decade a number of teletraffic research papers have reported the high variability and burstiness nature of LAN/WAN traffic in a wide range of network environments. One of the remarkable characteristics is the *fractal* nature of the investigated LAN/WAN traffic [37, 33, 31, 12, 11, 15, 30, 32, 26]. This is in close relationship to traffic burstiness because bursts should be defined in terms of time scales over which clustering activities occur. The surprising *scaling phenomenon* observed in data traffic is that these clustering activities are presented over several time scales [37].

Historically first, *long-range dependence* (LRD) and *self-similarity* have been detected and a group of studies is concentrated on how to detect accurately the LRD property and how to estimate the Hurst parameter [3, 2]. The LRD property of a traffic flow is revealed in the power law decay of the autocorrelation function at large lags, i.e.,

$$r(k) \sim c|k|^{2H-2}, \qquad k \to \infty,\ H \in (0.5, 1),$$

where $c$ is a constant [3]. The degree of this slow decay is determined by the Hurst parameter ($H$).

Traffic models (fractional Brownian motion (fBm) models, FARIMA models, Cox's M/G/$\infty$ models, on/off models, etc.) to capture LRD and self-similar properties have also been developed [37, 18, 27, 26]. Among these models the fBm [28] was found to be a popular parsimonious and tractable model of traffic aggregation. It was shown that the fBm is an accurate model if the traffic is aggregated from a large number of independent users whose peak rates are small relative to link capacity and the flow control has

no significant impact, see [5, 4, 19, 34, 36] for more details. The performance implications of the fractal property are also addressed in a series of studies [10, 9].

During the fractal research of network traffic it was discovered that the LAN/WAN traffic has a more complex scaling behaviour which cannot be described by LRD and self-similarity [31, 23, 12, 11]. More precisely, it has been found that aggregate network traffic is asymptotically self-similar over time scales on the order of few hundreds of milliseconds and above but it exhibits *multifractal* scaling below this time scale. It has been also pointed out that the transition from the multifractal to self-similar scaling occurs around time scales of a typical packet round-trip time in the network [12, 11]. However, some studies showed that multifractal scaling can also be present even at large time scales [25]. Therefore the monofractal traffic models (e.g. fBm) are inadequate to characterize the network traffic and multifractal traffic models with a much more flexible rule for the scaling law seem to be needed, especially for some WAN environments. The physical explanations and engineering implications are also addressed in several papers [11, 32].

A stochastic process $X(t)$ is called *multifractal* [22, 13] if it has stationary increments and satisfies $\mathrm{E}[|X(t)|^q] = c(q)t^{\tau(q)+1}$ for some positive $q$, where $\tau(q)$ is called the *scaling function* of multifractality and $c(q)$ is independent of $t$. If the scaling function $\tau(q)$ is a linear function of $q$ the process is called *monofractal*. Multifractality is thus defined as a *global* property of the process moments. The definition is very general and it covers a very large class of processes. Multifractal processes are also called processes with *scaling property*.

From a practical point of view queueing analysis of fractal traffic is a very important issue for network dimensioning and management. Therefore the study of queueing systems with fractal traffic input is a challenge in queueing theory. In the recent years the performance of queues with LRD or self-similar input has been deeply analyzed [16, 21, 35, 37]. A collection of studies has proven that the fBm based models have a tail queue distribution that decays asymptotically like a Weibullian law [37]

$$\mathrm{P}[Q > b] \simeq \exp(-\delta b^{2-2H}) \tag{1}$$

where $\delta$ is a positive constant that depends on the service rate of the queue [28, 8]. This important result shows that queues with fBm input ($H > 1/2$) have a much slower decay than that of the exponential.

However, there is a lack of queueing results available in the cases when the input traffic has a more complex scaling behaviour. Especially, queueing systems with multifractal input are an undiscovered field and only a few results published in the literature. Véhel *et al.* [20] suggested a cascade model for TCP traffic based on the retransmission and congestion avoidance mechanisms with no performance analysis. Riedi *et al.* [30] developed a multiscale queueing analysis in the case of tree-based multiscale input models. Gao *et al.* simulated queues fed by multiplicative multifractal processes in [14] but provided no analytical results. In contrast to these results we consider *general multifractal process* without any restrictions and derive an *analytical approximation for the queue tail asymptotics*.

Our aim is to contribute to the queueing theory of multifractal queues and also to the traffic engineering implications. In this paper we present a novel analysis of multifractal queues. We give an approximation formula for the tail of the queue with multifractal traffic input. The formula results in the asymptotically Weibullian decay for the special monofractal case of fractional Brownian motion input process which is consistent with the previously mentioned results [28, 8]. In this framework we show that any Gaussian process with scaling properties is in the class of monofractal processes and derive the related characterization functions. We also consider the special case when the queue is driven by a Gaussian process with independent, stationary increments (Brownian Motion) and show that our result is in a good agreement with results known from the theory of Gaussian processes. Practical applications of the results are also included in the paper. We also present a detailed queueing analysis for some measured WAN traffic which justifies the practical use of our formula. Our hope is that these results contribute as first steps to get the full understanding of multifractal queueing behavior. We address the most important open problems of this field in our future research.

The rest of the paper is organized as follows. Section II introduces our queueing system under investigation. Section III presents our main result, the approximation for queue tails with the proofs.

Section IV considers important applications and the simulation validation of the suggested queue tail estimation on some synthetic multifractal data series. The description of measured data traffic and our queueing analysis is given in Section V. Finally, Section VI concludes the paper and presents our targets for future research directions.

## 2 Queueing model

We consider a simple queueing model: a single server queue in continuous time, the serving principle for offered work is defined to be FIFO (First In, First Out), the queue has infinite buffer and constant service rate $s$. Denote by $X(t)$ the total size of work arriving to the queue from time instant $-t$ in the past up to this moment, time instant 0. The so called *workload process* $W(t)$ is the total amount of work stored in the buffer in time interval $(-t, 0)$, i.e.,

$$W(t) = X(t) - st \tag{2}$$

Our interest, however, is the current buffer length of the queue, denoted by $Q$. This is the queue length in the equilibrium state of the queue when the system has been running for a long time and the initial queue length has no influence. If this state of the system does exist, i.e., stationarity and ergodicity of the workload process hold, and the stability condition for the system is also satisfied, i.e., $\limsup_t \mathrm{E}[X(t)]/t < s$, then:

$$Q = \sup_{t \geq 0} W(t), \tag{3}$$

where $W(0)$ is assumed to be 0. This equation is also referred to as *Lindley's equation*.

The input process $X(t)$ is considered as a generally defined multifractal process by Mandelbrot *et al.* in [22]. The definition presents multifractal processes in terms of moments which leads to a more intuitive understanding of multifractality.

**Definition 1** *A stochastic process $X(t)$ is called multifractal if it has stationary increments and satisfies*

$$\mathrm{E}\left[|X(t)|^q\right] = c(q)t^{\tau(q)+1} \tag{4}$$

*for some positive $q \in \mathrm{Q}, [0,1] \subset \mathrm{Q}$, where $\tau(q)$ is called the scaling function and the moment factor $c(q)$ is independent of $t$.*

An easy consequence of this definition is that $\tau(q)$ is a concave function [22].

An alternative approach to multifractal processes, also found in literature, is based on the study of the local erratic behaviour of the process by means of its local Hölder exponents. For details on this approach see [29] and references therein. The most obvious examples of multifractals are self-similar and multiplicative processes.

## 3 Approximation for queue tail probabilities

We now state our main proposition:

**Proposition 1** *The probabilities for the queue tail asymptotic of a single queueing model with general multifractal input is accurately approximated by:*

$$\log(\mathrm{P}[Q > b]) \approx \min_{q>0} \log \left\{ c(q) \frac{\left[\frac{b\,\tau_0(q)}{s\,(q-\tau_0(q))}\right]^{\tau_0(q)}}{\left[\frac{b\,q}{q-\tau_0(q)}\right]^{q}} \right\}, \qquad b \text{ large} \tag{5}$$

*where $\tau_0(q) := \tau(q) + 1$. The scaling function $\tau(q)$ and $c(q)$ are the functions which define the multifractal input process.*

**Proof**

Using Lindley's equation the tail probabilities of queue length can be rewritten of the form: $\mathrm{P}[Q > b] = \mathrm{P}[\sup_{t\geq 0} W(t) > b]$. First let consider the quantity $\mathrm{P}[W(t) > b]$:

Replacing $W(t)$ by Eq. (2) we have

$$\begin{aligned}
\mathrm{P}[W(t) > b] &= \mathrm{P}[X(t) - st > b] \\
&\leq \mathrm{P}[|X(t)| > b + st] \qquad (6) \\
&= \mathrm{P}[|X(t)|^q > (b + st)^q], \qquad \text{for arbitrary } q > 0 \\
&\leq \frac{E[X(t)^q]}{(b + st)^q}, \qquad \text{using Markov's inequality.} \qquad (7)
\end{aligned}$$

Since the input process is multifractal defined by Def. 1 then:

$$\begin{aligned}
\mathrm{P}[W(t) > b] &\leq \frac{c(q)t^{\tau_0(q)}}{(b + st)^q} \\
\Rightarrow \sup_{t\geq 0} \mathrm{P}[W(t) > b] &\leq \sup_{t\geq 0} \frac{c(q)t^{\tau_0(q)}}{(b + st)^q} =: \sup_{t\geq 0} f(t). \qquad (8)
\end{aligned}$$

By straightforward calculation it is easy to see that that derivative of $f(t)$ equals zero when

$$t_0 = \frac{b\tau_0(q)}{s[q - \tau_0(q)]} > 0.$$

The second derivative of $f(t)$ at this point is

$$f''(t_0) = c(q)t_0^{\tau_0(q)-1}(b + st_0)^{-q-1}s[\tau_0(q) - q] < 0,$$

thus $f(t)$ has its maximal value at $t = t_0$. Note that we have assumed that $q > \tau_0(q)$, this is justified by the fact that in the monofractal case we have $\tau_0(q) = qH < q$ and by the concavity of $\tau(q)$.

Therefore

$$\begin{aligned}
\sup_{t\geq 0} \mathrm{P}[W(t) > b] &\leq \sup_{t\geq 0} f(t) = c(q) \frac{\left[\frac{b\,\tau_0(q)}{s(q-\tau_0(q))}\right]^{\tau_0(q)}}{\left[\frac{b\,q}{q-\tau_0(q)}\right]^q} \\
\Rightarrow \log\left(\sup_{t\geq 0} \mathrm{P}[W(t) > b]\right) &\leq \log\left(c(q) \frac{\left[\frac{b\,\tau_0(q)}{s(q-\tau_0(q))}\right]^{\tau_0(q)}}{\left[\frac{b\,q}{q-\tau_0(q)}\right]^q}\right), \qquad \text{for arbitrary } q > 0 \\
\Rightarrow \log\left(\sup_{t\geq 0} \mathrm{P}[W(t) > b]\right) &\leq \min_{q>0} \log\left(c(q) \frac{\left[\frac{b\,\tau_0(q)}{s(q-\tau_0(q))}\right]^{\tau_0(q)}}{\left[\frac{b\,q}{q-\tau_0(q)}\right]^q}\right). \qquad (9)
\end{aligned}$$

For a large class of stochastic processes (including fBm) the following limit holds [17]:

$$\lim_{b\to\infty} \frac{\log(\mathrm{P}[Q > b])}{\log(\sup_{t\geq 0} \mathrm{P}[W(t) > b])} = 1. \qquad (10)$$

In addition,

$$\log(\mathrm{P}[Q > b]) \geq \log(\sup_{t\geq 0} \mathrm{P}[W(t) > b]), \qquad (11)$$

then the right-hand side of Eq. (9) is a upper bound of a lower bound on $\log(\mathrm{P}[Q > b])$. The used inequalities in Eq. (11) and Eq. (7) become tight for finite large $b$. Thus our approximation for the queue tail asymptotics is the following:

$$\log(\mathrm{P}[Q > b]) \approx \min_{q>0} \log\left(c(q) \frac{\left[\frac{b\,\tau_0(q)}{s(q-\tau_0(q))}\right]^{\tau_0(q)}}{\left[\frac{b\,q}{q-\tau_0(q)}\right]^q}\right), \qquad b \text{ large.}$$

□

For positive multifractal processes, i.e. $X(t) > 0$, Eq. (6) is an equality. In addition, the approximation in Eq. (11) and the inequality in Eq. (7) turn to be more accurate approximations as $b$ tends to infinity. Thus the presented approximation is supposed to be asymptotically tight. The tightness and accuracy of the approximation is also experimentally investigated in Section IV.4 and Section V.

Considering the formula in Eq. (5) we see that it has an implicit form and just the given form of the functions $c(q)$ and $\tau(q)$ can provide the final result. The reason behind this is that the definition for the class of multifractal processes gives no restrictions for the functions $c(q)$ and $\tau(q)$ (beyond that $\tau(q)$ is concave). *Our conjecture is that the analysis of queueing systems with general multifractal input may produce some similar general results*. It means that there is no general queueing behaviour for these systems as the Weibullian decay in the case of Gaussian self-similar processes [28]. An actual multifractal model will determine, for example, the queue length probabilities of the system.

# 4 Applications

## 4.1 Fractional Brownian motion

As a simple application first we consider a monofractal Gaussian process, called fractional Brownian motion (fBm). The fBm is self-similar which is a simple case of monofractality and it is also Gaussian. The increment process of fBm is called fractional Gaussian noise (fGn). Queueing analysis of a single queue with fBm input is first presented by Norros [28] which showed the Weibullian decay for the asymptotic tail behaviour, i.e., $\mathrm{P}[X > x] \sim \exp(-\gamma x^{\beta})$ with $\beta \leq 1$. This result is also justified by Large Deviation techniques in [8]. Applying this input process model to our formula should show its use and robustness when comparing to these available results.

First we prove that any Gaussian process with scaling property is in the class of monofractal processes. Furthermore we give the explicit forms for $\tau(q)$ and $c(q)$.

Consider the following lemma:

**Lemma 1** *A Gaussian process with scaling property is monofractal with parameters*

$$\begin{cases} \tau(q) &= \frac{q}{2}\left[\tau(2)+1\right]-1 \\ c(q) &= \frac{[2c(2)]^{q/2}}{\sqrt{\pi}}\Gamma\left(\frac{q+1}{2}\right), \end{cases}$$

*where* $\Gamma(\cdot)$ *denotes the Gamma function,* $\Gamma(z) = \int_0^{+\infty} x^{z-1}\exp^{-x}\mathrm{d}x,\ z > 0.$

The proof of this Lemma is provided in Appendix 7.1.

Turning back to our case of fBm with $c(2) = 1$ and $\tau(2) = 2H - 1$ where $H$ is referred to as the Hurst parameter, we have

$$\begin{cases} \tau(q) &= qH-1 \\ c(q) &= \frac{2^{q/2}}{\sqrt{\pi}}\Gamma\left(\frac{q+1}{2}\right). \end{cases}$$

Insert these two functions into our formula in Eq. (5) we get

$$\log(\mathrm{P}[Q > b]) \approx \log\left(\min_{q>0}\left\{\frac{2^{q/2}}{\sqrt{\pi}}\Gamma\left(\frac{q+1}{2}\right)\frac{\left(\frac{bH}{s(1-H)}\right)^{qH}}{\left(\frac{b}{1-H}\right)^{q}}\right\}\right) =: \log(\min_{q>o} g(q)).$$

The minimum value of the $g(q)$ for $q > 0$ function can be easily determined by taking its derivatives. The result is the following:

$$\log(\mathrm{P}[Q > b]) \approx \log(\min_{q>o} g(q)) = \log\left(\frac{1}{\sqrt{\pi}}\frac{\Gamma\left(\Psi^{-1}\left(\log K\right)\right)}{K^{\Psi^{-1}(\log K)-1/2}}\right) =: \log(T_{fBm}(H,s,b)), \qquad (12)$$

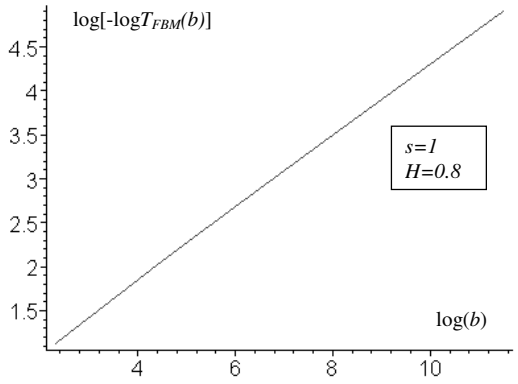


Figure 1: By setting fixed values for $H$ and $s$, the line in the log-log plot of $-\log T_{fBm}(b)$ versus $b$ clearly shows the Weibullian decay for $T_{fBm}(H,s,b)$.

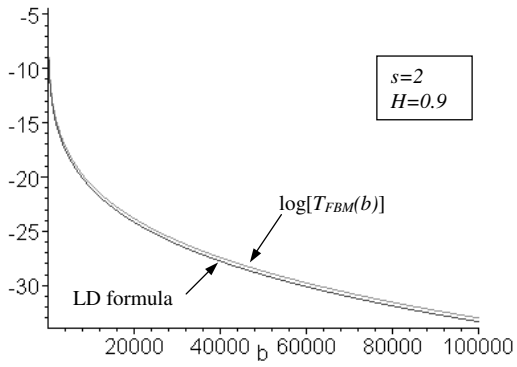


Figure 2: Our approximation compared to the Large Deviation technique result: the two plots almost coincide for all values of queue size.

where $K = K(H,s,b) = \frac{1}{2}b^{2(1-H)}s^{2H}(1-H)^{-2(1-H)}H^{-2H}$, $\Psi(\cdot)$ is the *digamma* function, $\Psi(x) = \frac{\mathrm{d}}{\mathrm{d}x}\log\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$, and $\Psi^{-1}(\cdot)$ denotes the inverse function of $\Psi(\cdot)$.

The $T_{fBm}(H,s,b)$ function is quite complex with the presence of Gamma, digamma, and its inverse function. However, we have quite a good approximation of $T_{fBm}(H,s,b)$:

**Proposition 2** *The approximation*

$$\frac{1}{\sqrt{\pi}}\frac{\Gamma\left(\Psi^{-1}\left(\log x\right)\right)}{x^{\Psi^{-1}(\log x)-1/2}} \approx \exp(-x) \tag{13}$$

*holds for large $x$, $x > 0$.*

The proof and the precise sense of this approximation can be found in Appendix 7.2.

Applying this approximation we find that the queue tail for the fBm case satisfies:

$$\log\left(T_{fBm}(H,s,b)\right) \approx -\frac{1}{2}b^{2(1-H)}s^{2H}(1-H)^{-2(1-H)}H^{-2H}, \qquad b \text{ large.} \tag{14}$$

Eq. (14) shows the Weibullian decay of this queue which was first recognized and proven by Norros [28]. Numerical evaluations of the result are presented in Fig. 1 and Fig. 2. In Fig. 1 we fix the values of $H$ and $s$ and then calculate the values of the queue tail approximation $T_{fBm}(H,s,b)$ versus the queue size $b$ and then plot it in the log-log scale. The linearity of the plot also demonstrates the Weibullian decay.

Now we compare our result to the result obtained by Duffield and O'Connell. The asymptotic formula for queue tail probabilities provided by Large Deviation technique presented in [8] is

$$\begin{aligned} &\lim_{b\to\infty} b^{-2(1-H)}\log \mathrm{P}[Q>b] = -\inf_{c>0} c^{-2(1-H)}\frac{(c+s)^2}{2} \\ \Leftrightarrow\quad &\log \mathrm{P}[Q>b] \to -\frac{1}{2}b^{2(1-H)}s^{2H}(1-H)^{-2(1-H)}H^{-2H}, \quad \text{as } b\to\infty, \end{aligned} \tag{15}$$

where $s$ also denotes the service rate. Therefore we can conclude that our approximation yields the Large Deviation result, see Eq. (14) and Eq. (15). The two results are depicted in Fig. 2 and we can see that the plots almost coincide for all calculated values of the queue size.

Our conclusions can be summarized in two main points: (i) the asymptotic tail approximation for the case of fBm has Weibullian decay; (ii) this result is also consistent with the formula presented by Norros [28] and by Duffield *et al.* with Large Deviation technique [8].

### 4.2 Gaussian input process with independent, stationary increments

In the case of a Gaussian input model $\{Z(t), t > 0\}$ with independent, stationary increments (Brownian Motion) the following equation holds for its variance: $\sigma_t^2 = \mathrm{E}[|Z(t)|^2] = \sigma^2 t$. Then the similar calculation as in the proof of Lemma 1 shows that:

$$\mathrm{E}[|Z(t)|^q] = \frac{\left(2\sigma^2\right)^{q/2}\Gamma(\frac{q+1}{2})}{\sqrt{\pi}} t^{q/2}, \tag{16}$$

which means that the concerned Gaussian input model also satisfies the general definition of multifractality with

$$\begin{cases} \tau(q) &= q/2 - 1 \\ c(q) &= \frac{\left(2\sigma^2\right)^{q/2}}{\sqrt{\pi}}\Gamma\left(\frac{q+1}{2}\right). \end{cases}$$

Thus our approximation formula can also be applied for this case. The same minimization procedure for this process shows that

$$\begin{aligned} \log(\mathrm{P}[Q > b]) &\approx \log\left(\min_{q>0}\left\{\frac{\left(\frac{2sb}{\sigma^2}\right)^{-q/2}}{\sqrt{\pi}}\Gamma\left(\frac{q+1}{2}\right)\right\}\right) \\ &= \log\left(\frac{1}{\sqrt{\pi}}\frac{\Gamma\left(\Psi^{-1}\left(\log(\frac{2sb}{\sigma^2})\right)\right)}{(\frac{2sb}{\sigma^2})^{\Psi^{-1}(\log(\frac{2sb}{\sigma^2}))-1/2}}\right) =: \log(T_{Gaussian}(\sigma^2, s, b)) \approx -\frac{2sb}{\sigma^2} \end{aligned} \tag{17}$$

The result is also in agreement with the queueing formula provided in the theory of Gaussian processes [24, 8].

### 4.3 Practical solutions

We show here the practical use of the formula. Assume that we are interested in the behaviour of the tail of the steady-state buffer occupancy (queue length) distribution at a specific multiplexer in our network. The first step should be the fine resolution measurements of the input process. We also assume that the input process exhibits multifractal scaling properties. Then the scaling function $\tau(q)$ and the function $c(q)$ can be estimated from the collected data for some available parameters $q > 0$. *We emphasize the importance of the function $c(q)$ as the quantity factor of multifractal processes which is sometimes neglected in a number of studies dealing with multiscaling properties of the high-speed network traffic. The scaling function $\tau(q)$ defines only the quality of multiscaling and it is not enough for the description of a multifractal model and therefore for the analysis of queueing models with multifractal input processes.*

Now we suggest two practical methods for the approximation of the queue tail distribution:

1. Given the service rate $s$ and the two sets $\{c(q)\}$ and $\{\tau(q)\}$, using Eq. (5) the approximation of $\log(\mathrm{P}[Q > b])$ can be computed for each value of $b$. This method is very simple but it is the more useful from network planning and capacity dimensioning point of view since we are only interested in some values of the tail probabilities. We mainly focus on the practical use of this method in this study.

2. The input process is fitted to a multifractal model. The two measured sets of $c(q)$ and $\tau(q)$ are fitted by $\tilde{c}(q)$ and $\tilde{\tau}(q)$. Then the analysis of the Eq. (5) with these functions can result in simple closed form of the queue tail probabilities. We use this method when studying the queue tail behaviour of a multifractal model.

### 4.4 Simulation evaluation of the approximation

For simulation validation of the presented approximation for queueing performance estimation we consider the queueing model with the simplest multifractal input: multiplicative cascades. The construction of the model is briefly presented below, see details in [6].

Consider a unit interval associated with a unit mass. At state $k = 1$ divide the unit interval into two equal subintervals and associate with them the mass $r$ and $1 - r$, respectively. The parameter $r$ is chosen to be a random variable governed by the probability function $\mathrm{F}_r(x)$, $0 \leq x \leq 1$. The fraction $r$ is called the multiplier. We also chose the multiplier $r$ to have a symmetric density function then $r$ and $1 - r$ have the same marginal distribution. The same rule is applied to each subinterval and its associated mass. At the state $k$ an interval of length $\triangle t = 2^{-k}$ has the mass (measure)

$$\mu(\triangle t) = r_{(1,*)} r_{(2,*)} \dots r_{(k,*)},$$

where $r_{(i,*)}$ indicates the multiplier at state number $i$. Since multipliers $r$ are i.i.d. it is easy to show that the measure $\mu$ satisfies the scaling relationship:

$$\mathrm{E}[\mu(\triangle t)^q] = \triangle t^{-\log_2 \mathrm{E}[r^q]},$$

which defines a multifractal.

We use $2^N$ synthetic data of this multiplicative multifractal for queueing evaluation of our method. After rescaling the multifractal model of unit mean and time has the following characteristic functions:

$$\begin{cases} \tau_0(q) &= -\log_2 \mathrm{E}[r^q] \\ c(q) &= 2^{N(q-\tau_0(q))} \end{cases}$$

We simulate two multifractals which are differentiated by the chosen distribution for $r$: uniform and symmetric $\beta(\alpha, \alpha)$ distribution. In the case of $\beta$-multiplier a multifractal example with $\alpha = 20$ is chosen for simulation. The queueing simulations are compared to the numerical calculation of our formula in a simple queueing system with a single server, constant service rate, infinite buffer, and FIFO serving principle. The numerical calculation method is discussed in the previous subsection (see more details of this estimation method in the next section). The results are presented in Figure 3.

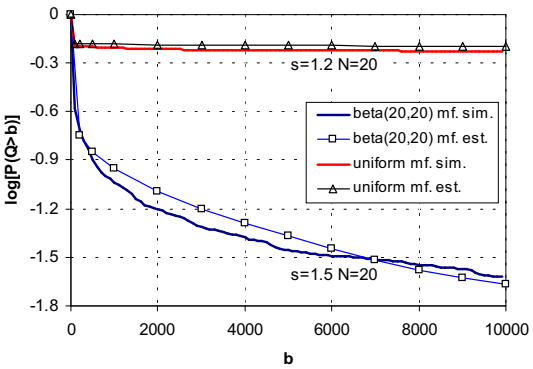


Figure 3: Comparison of simulated queue tail probabilities and their theoretical values.

Note that the scaling function $\tau_0(q)$ is $\log_2(q + 1)$ and $\log_2 \left[\frac{\Gamma(\alpha)\Gamma(2\alpha+q)}{\Gamma(\alpha+q)\Gamma(2\alpha)}\right]$ for uniform multiplier and $\beta$-multiplier, respectively. With given values of these functions the tail probabilities of the system can be numerical counted using the queueing formula. As seen in the figure the theoretical queue tail probabilities are really close to the simulation results. This shows that the suggested queueing estimation method can be effectively used in practice.

# 5 Queueing analysis

In this section we show the validation for the mentioned practical solution presented above by the queueing analysis of some real traffic traces. We also provide a simple method for estimation of multiscaling functions $c(q)$ and $\tau(q)$.

## 5.1 Data traces

Three data traces were considered in our analysis which are freely available at the Internet Traffic Archive [1]. These traces contain an hour's worth of all wide-area traffic each between Digital Equipment Corporation and the rest of the world. The traces, denoted by DEC-PKT-1, DEC-PKT-2, and DEC-PKT-3, were gathered at Digital's primary Internet access point, which is an Ethernet DMZ network operated by Digital's Palo Alto research groups. The raw traces were made using tcpdump on a DEC Alpha running Digital's OSF/1 operating system, which includes a kernel filter with capabilities comparable to those of BPF. `Tcpdump` captured all IP packet header information with millisecond precision timestamps. Each trace contains more than 3 million packet headers.

| Data set | Number of arrivals | Mean | Variance |
|---|---|---|---|
| DEC-PKT-1 | 3 027 907 | 2.5232 | 4.4153 |
| DEC-PKT-2 | 3 987 942 | 3.3234 | 5.2416 |
| DEC-PKT-3 | 4 518 090 | 3.7652 | 5.9968 |

Table 1: Summary of the investigated data sets.

We constructed the packet arrival counts traces of time sample of 3 milliseconds from the raw data. Preliminary analysis of these traces exhibits scaling properties, we therefore use them as the inputs of the queue system under investigation.

## 5.2 Simple method for multiscaling functions estimation

The full description of a multifractal model involves both $c(q)$ and the scaling function $\tau(q)$. We present here a simple method for testing of scaling properties and also for the estimation of these functions.

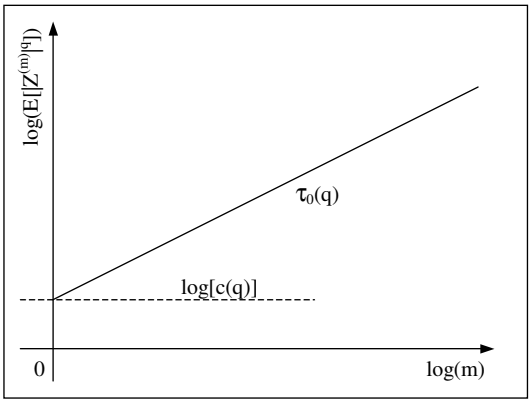


Figure 4: A simple method for scaling test and the estimation of $c(q)$ and the scaling function $\tau(q)$.

The definition of multifractal processes (Def. 1) claims the stationarity condition for the increments. Therefore it is easy to verify the following relation for the moments of the increments:

$$\mathrm{E}[|Z^{(\triangle t)}|^q] = c(q)(\triangle t)^{\tau(q)+1} = c(q)(\triangle t)^{\tau_0(q)}, \qquad q > 0, \tag{18}$$

where $Z^{(\triangle t)}$ denotes the increment process of time sample $\triangle t$. Thus this equality also holds for $m = 1, 2, \ldots$

$$\mathrm{E}[|Z^{(m\triangle t)}|^q] = c(q)(m\triangle t)^{\tau_0(q)}, \qquad q > 0. \tag{19}$$

Choose $\triangle t$ as the time unit, then

$$\log \mathrm{E}[|Z^{(m)}|^q] = \tau_0(q) \log m + \log c(q), \qquad q > 0. \tag{20}$$

Based on this property, the method is the following: Given a data series of a process increments $Z_1, Z_2, \ldots, Z_n$ and define its corresponding *real* aggregated sequence $\{Z^{(m)}\}$ of the aggregation level $m$ by

$$Z_k^{(m)} = Z_{(k-1)m+1} + Z_{(k-1)m+2} + \ldots + Z_{km}, \qquad k, m = 1, 2, \ldots \tag{21}$$

If the sequence $\{Z_k\}$ has scaling property then the plot of absolute moments $\mathrm{E}[|Z^{(m)}|^q]$ versus $m$ on a log-log plot should be a straight line due to Eq. (20). The slope of the straight line provides the estimate of $\tau_0(q)$ and the intercept is the value for $\log c(q)$. The illustration of the method can be seen in Fig. 4.

Note that we have no need to estimate $c(q)$ and $\tau_0(q)$ for all positive value of $q$, which is an impossible task. In fact, the largest value of $q$ we should considered depends on the interested finite queue length of the involved queue length probability, see below.

### 5.3 Analysis of theoretical and simulation results

We present here our analysis results of the mentioned data sets. We validate the use of our approximation in a single queue with constant service rate and general multifractal input.

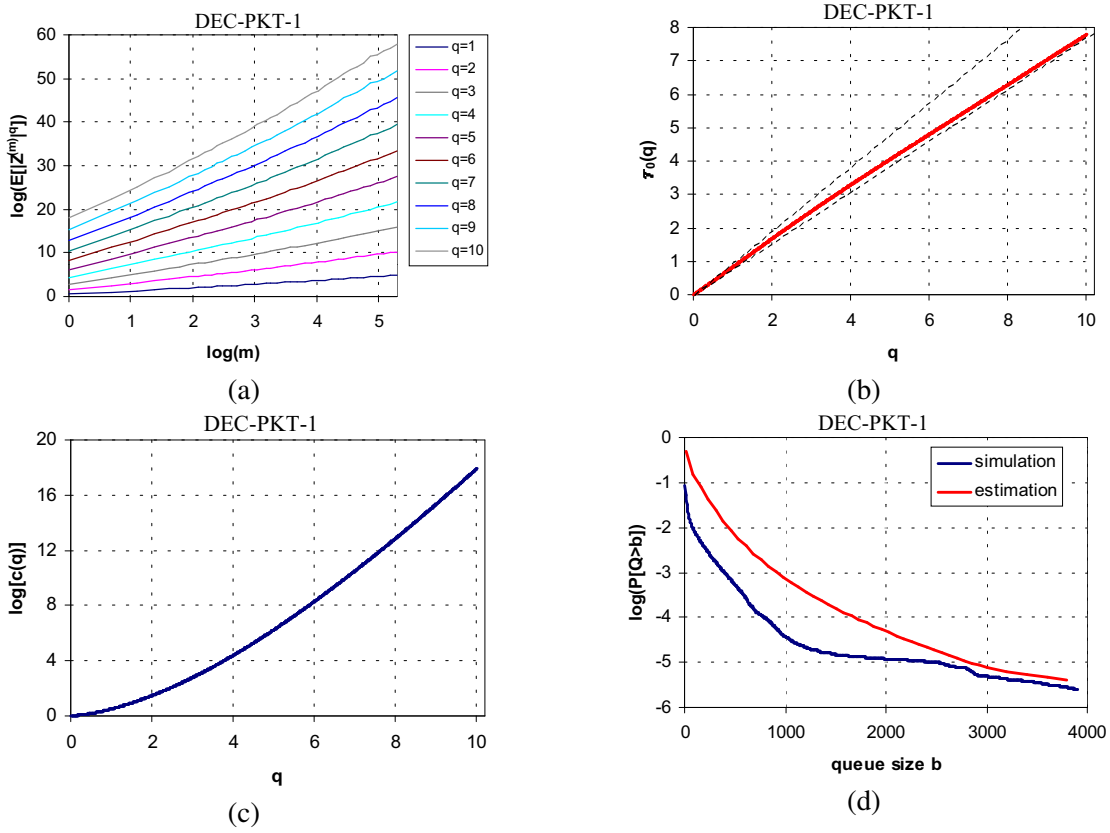


Figure 5: Analysis results of the DEC-PKT-1 data set: (a) the plots of the absolute moments, (b) and (c) estimated form of $\tau_0(q)$ and $c(q)$, respectively, (d) queueing simulation compared with the theoretical approximation.

Figure 5(a) shows the plot of absolute moments of the aggregated sets of the set DEC-PKT-1 versus the aggregation level in a log-log plot for some values of moment $q$. The linearity of the plots observed in the figure clearly indicates the scaling property of this data set. After applying the estimation method we presented in the previous subsection we get the two sets of estimated $\tau_0(q)$ and $c(q)$ which are drawn in Fig. 5(b) and Fig. 5(c) (we estimate $\log c(q)$ instead of $c(q)$). The plot of the function $\tau_0(q) = \tau(q)+1$ is a concave curve which suggests the multifractal property of DEC-PKT-1.

We then make a comparison between our approximation and the queueing simulation of real data traces to validate the use of the formula in practice. The approximation for probabilities of queue tail presented in Proposition 1 can be rewritten in the form

$$\begin{aligned} \log \mathrm{P}[Q > b] &\approx \min_{q>0} \left\{ \log c(q) + \tau_0(q) \log \frac{b\tau_0(q)}{s(q-\tau_0(q))} - q \log \frac{bq}{q-\tau_0(q)} \right\} \\ &=: \min_{q>0} \{\log T^*(s,b)\} = T(s,b). \end{aligned} \tag{22}$$

For the sake of calculation simplicity we choose the service rate such that $s = 1$. The lower curve in Fig. 5(d) shows the simulation result of the DEC-PFT-1 data set. Using Eq. (22) the value of the logarithmic tail probability at each concerned value of queue size $b$ is taken by the numerical minimization of $\log T^*(s,b)$ with the estimated sets $\{c(q)\}$ and $\{\tau_0(q)\}$. An example is shown in Fig. 6. In addition,

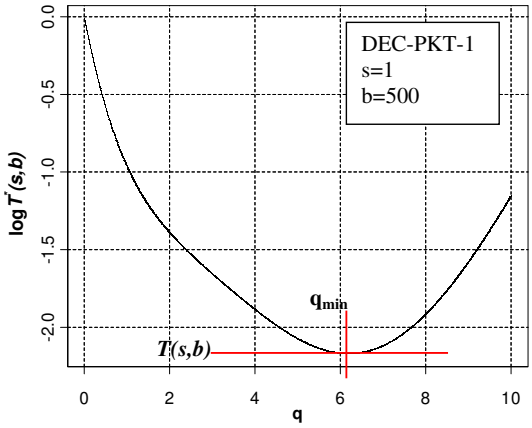


Figure 6: Theoretical queue tail probability at each value of queue size $b$ is the minimum of $\log T^*(s,b)$, for example DEC-PKT-1, $s = 1, b = 500$.

we do not need to plot $\log T^*(s,b)$ at each value of $q$ to find its minimum. A simple program routine can do it for all concerned value of $b$ at once. Our theoretical tail probabilities are on the upper curve in Fig. 5. As comparing with the simulation result which is seen in the same figure we found that it has the similar shape and becomes tight as $b$ increases. This validates our result.

We have performed the same analysis with two other data sets DEC-PFK-2 and DEC-PKT-3. The results are summarized in Fig. 7 and Fig. 8. The DEC-PKT-3 set also exhibits multifractal structure as the DEC-PKT-1 and its queueing analysis gives some similar results. The DEC-PKT-2 data set, however, has the exact monofractal structure and can be well modelled by statistical self-similarity with Hurst parameter $H = 0.8$. Our queueing model deals with general multifractal input so it also involves the case of monofractal processes. Thus it is not surprising that the analysis also provides the correct queueing results in this case.

**Some notes:**

- Our analysis experiences show that the WAN traffic can be multifractal or monofractal. The traffic does not necessarily have multiscaling property at small time scales and uniscaling property at time scales larger than round trip time as some authors argued [11, 32].

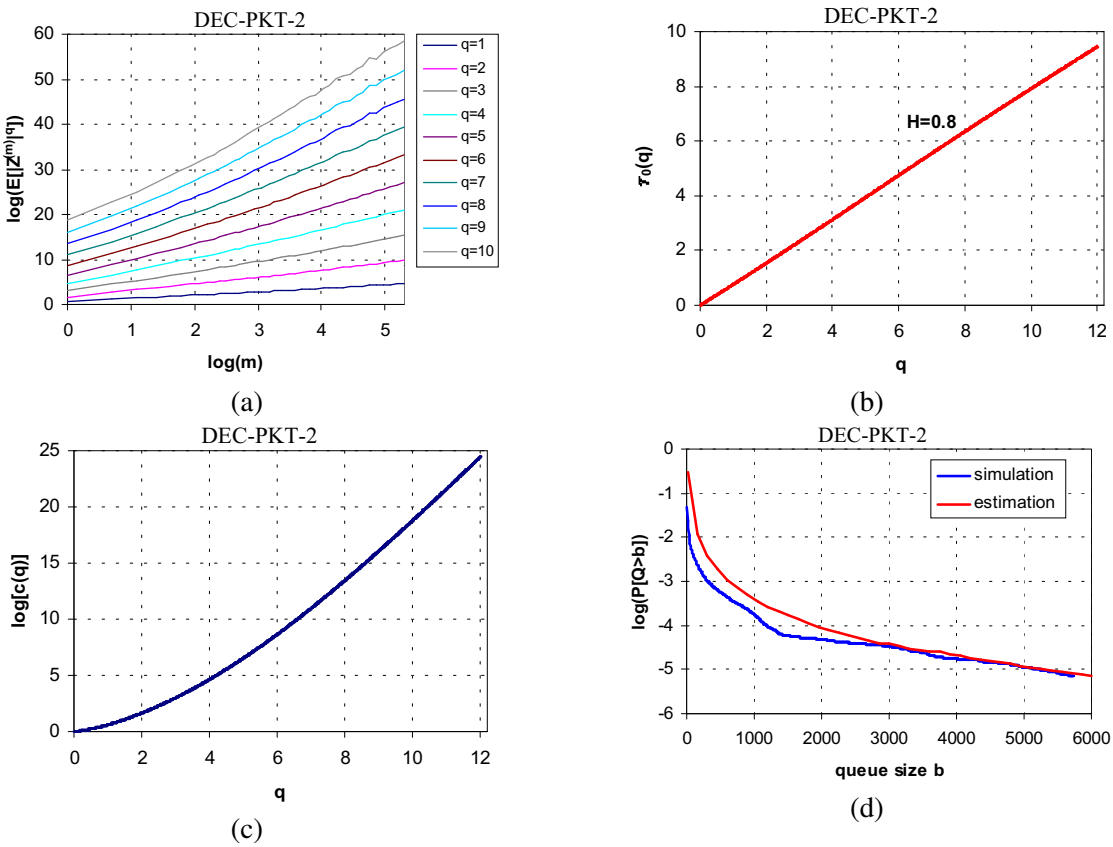


Figure 7: Analysis results of the DEC-PKT-2 data set: (a) the plots of the absolute moments, (b) estimated form of $\tau_0(q)$ which shows the data set can be modelled by statistical self-similarity with $H = 0.8$, (c) estimation of $\log[c(q)]$, (d) queueing simulation compared with the theoretical tail probabilities.

- Considering the shape of the estimated $\tau_0(q)$ functions we suggest the following model for the scaling function:

$$\tau_0(q) = (qH)^{\alpha}, \qquad 0.5 < H < 1,\ 0 < \alpha \leq 1.$$

  $\alpha = 1$ it is the case of monofractal processes. The mathematical form for the $c(q)$ function together with the suggested model for $\tau_0(q)$ would simplify the minimization task in our formula.

# 6 Conclusion

In this paper we studied the queueing performance of a single server infinite capacity queue with a constant service rate fed by general multifractal input process. We have provided the following results:

*(i) We derived an asymptotic approximation of the steady-state queue length probabilities.*

*(ii) We showed that our results gives the well-known Weibullian queue tail in case of the monofractal fractional Brownian motion input process.*

*(iii) We proved that the class of Gaussian processes with scaling properties is limited to monofractal processes and derived their characteristic functions.*

*(iv) We applied our formula for the case of Gaussian input processes and verified the results from the theory of Gaussian processes.*

*(v) We demonstrated the practical applicability of our approximation and validated the method by simulation with two multiplicative multifractal processes.*

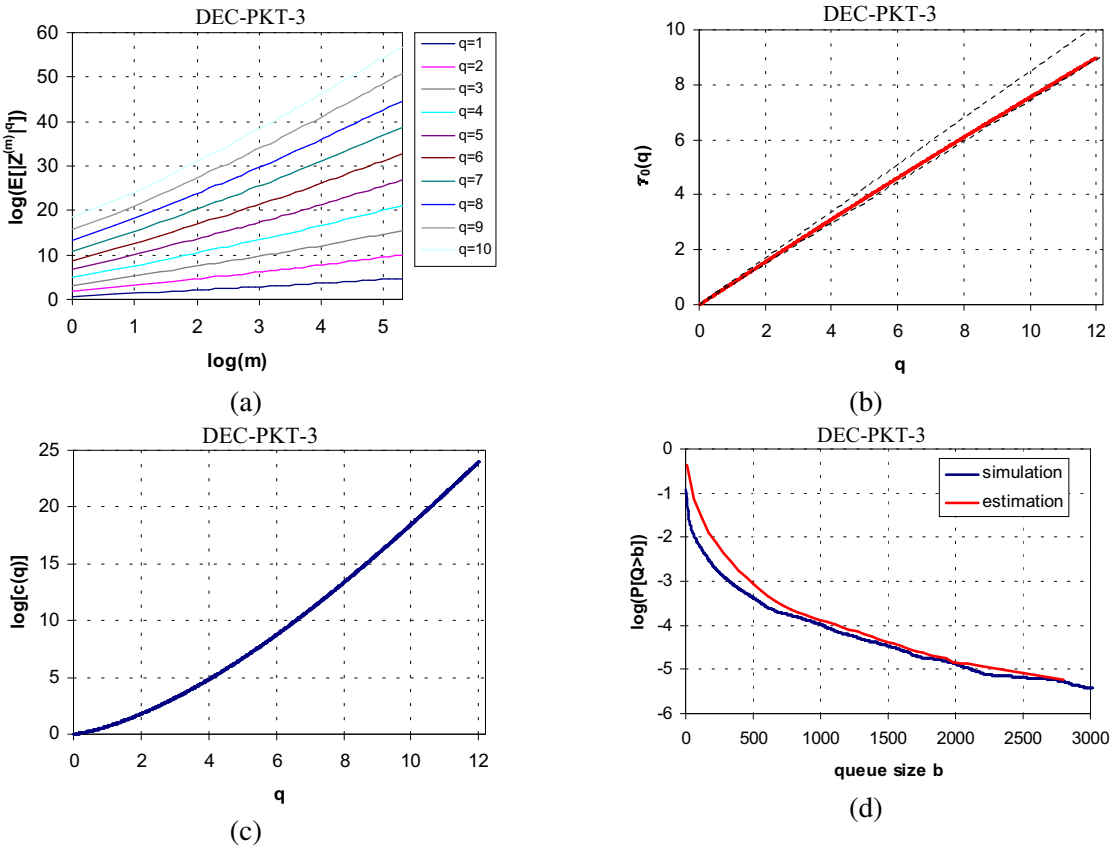


Figure 8: Analysis results of the DEC-PKT-3 data set: (a) the plots of the absolute moments, (b) and (c) estimated form of $\tau_0(q)$ and $\log[c(q)]$, respectively; the concavity of $\tau_0(q)$ shows the multiscaling nature of the data set, (d) queueing simulation compared with the theoretical tail probabilities.

*(vi) We have also shown that our formula gives the correct result in analysis of both multifractal and monofractal network traffic cases.*

There are several interesting topics for further research. Based on the multifractal process characterization one of our goal is to build a multifractal traffic model parameterized by the multifractal functions. We also intend to carry out more multifractal analyses of measured LAN/WAN traffic with corresponding performance analysis.

# 7 Appendices

## 7.1 Proof of Lemma 1

Denote by $X(t)$ the Gaussian process. Since $X(t)$ has scaling property it satisfies the general definition for multifractal process, i.e., $\mathrm{E}[|X(t)|^q] = c(q)t^{\tau(q)+1}$. Thus the variance of the $X(t)$ process should be $\sigma_t^2 = c(2)t^{\tau(2)+1}$. The Gaussian process $X(t) \sim N\left(0, c(2)t^{\tau(2)+1}\right)$ has the normal distribution and we have

$$f(x) = \frac{1}{\sqrt{2\pi c(2)t^{\tau(2)+1}}} \exp\left(-\frac{x^2}{2c(2)t^{\tau(2)+1}}\right).$$

The $q^{th}$ moment of $X(t)$ can be calculated by the definition:

$$\mathrm{E}[|X(t)|^q] = \int_{-\infty}^{+\infty} |x|^q f(x)\mathrm{d}x$$

$$= \quad 2\int_0^{+\infty} x^q \frac{1}{\sqrt{2\pi c(2) t^{\tau(2)+1}}} \exp\left(-\frac{x^2}{2c(2)t^{\tau(2)+1}}\right) \mathrm{d}x.$$

Introduce $y := \frac{x^2}{2c(2)t^{\tau(2)+1}}$. The formula above can be rewritten as follows:

$$\begin{aligned} \mathrm{E}[|X(t)|^q] &= \frac{2^{q/2}}{\sqrt{\pi}} \left[c(2)t^{\tau(2)+1}\right]^{q/2} \int_0^{+\infty} y^{\frac{q-1}{2}} \exp(-y)\mathrm{d}y \\ &= \frac{[2c(2)]^{q/2}}{\sqrt{\pi}} \Gamma\left(\frac{q+1}{2}\right) t^{\frac{q}{2}[\tau(2)+1]}, \end{aligned} \tag{23}$$

which concludes our proof. □

### 7.2 Proof of Proposition 2

The approximation in Proposition 2 states that

$$\begin{aligned} \frac{1}{\sqrt{\pi}} \frac{\Gamma\left[\Psi^{-1}(\log x)\right]}{x^{\Psi^{-1}(\log x)-1/2}} &\approx \exp(-x), \qquad x \text{ large} \\ \Leftrightarrow \quad \log\Gamma\left[\Psi^{-1}(\log x)\right] &\approx \left[\Psi^{-1}(\log x) - 1/2\right]\log x - x + \log\sqrt{\pi}, \end{aligned} \tag{24}$$

By the *Stirling*'s formula we have

$$\log\Gamma(z+1) - (z+1/2)\log z + z \approx \log\sqrt{2\pi} \qquad \text{as } z \to \infty. \tag{25}$$

Put $z + 1 = \Psi^{-1}(\log x)$, hence $x = \exp\Psi(z+1) \to \infty$ as $z \to \infty$ by the monotonicity of $\Psi(\cdot)$. We obtain

$$\begin{aligned} \log\Gamma\left[\Psi^{-1}(\log x)\right] - \left[\Psi^{-1}(\log x) - 1/2\right]\log\left[\Psi^{-1}(\log x) - 1\right] + \left[\Psi^{-1}(\log x) - 1\right] \approx \\ \approx \log\sqrt{2\pi} \qquad \text{as } x \to \infty. \end{aligned} \tag{26}$$

In [7] it is derived that

$$\lim_{z\to\infty} \frac{\Psi(z+1)}{\log z} = 1, \tag{27}$$

hence $x = \Psi^{-1}\left[\omega(x)\log x\right] - 1$ where $\lim_{x\to\infty}\omega(x) = 1$. Using the approximation $\Psi^{-1}(\log x) - 1 \approx x$ for large $x$, we get the equation

$$\log\Gamma\left[\Psi^{-1}(\log x)\right] - \left[\Psi^{-1}(\log x) - 1/2\right]\log x + x \approx \log\sqrt{2\pi}. \tag{28}$$

The approximation above is also verified by numerical computations and it is shown that the approximation can be effectively applied for large $x$.

Then the approximation in Eq. (24) follows if we change the constant $\log\sqrt{2\pi}$ by $\log\sqrt{\pi}$. □

## References


[1] The internet traffic archive. http://ita.ee.lbl.gov.

[2] P. Abry and D. Veitch. Wavelet analysis of long range dependent traffic. *IEEE Trans. Inform. Theory*, 44(1):2–15, January 1998.

[3] J. Beran. *Statistics for Long-Memory Processes*. Chapman & Hall, One Penn Plaza, New York, NY 10119, 1995.

[4] D. R. Cox. *Statistics: An Appraisal, Proc. 50th Anniversary Conference*, chapter Long-Range Dependence: A Review. Iowa State University Press, 1984.

[5] M. E. Crovella and A. Bestavros. Self-similarity in world wide web traffic: Evidence and possible causes. *IEEE/ACM Trans. on Networking*, 5(6):835–846, December 1997.

[6] T. D. Dang and S. Molnár. A new model for multifractal network traffic. Technical report, Budapest University of Technology and Economics, 2002.

[7] L. Devroye. Random variate generation for the digamma and trigamma distributions. *J. of Statistical Computation and Simulation*, 43:197–216, 1992.

[8] N. G. Duffield and N. O'Connell. Large deviations and overflow probabilities for the general single-server queue, with applications. In *Proc., Cam. Phil. Soc.*, volume 118, pages 363–374, 1994.

[9] A. Erramilli, O. Narayan, A. L. Neidhardt, and I. Saniee. Performance impacts of multi-scaling in wide-area TCP/IP traffic. In *Proc., IEEE INFOCOM 2000*, volume 1, pages 352–359, Tel Aviv, Israel, 2000.

[10] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Trans. on Networking*, 4(2):209–223, April 1996.

[11] A. Feldmann, A. C. Gilbert, and W. Willinger. Data Networks as Cascades: Investigating the Multifractal Nature of Internet WAN Traffic. *ACM Computer Communication Review*, 28:42–55, September 1998.

[12] A. Feldmann, A. C. Gilbert, W. Willinger, and T. G. Kurtz. The Changing Nature of Network Traffic: Scaling Phenomena. *ACM Computer Communication Review*, 28:5–29, April 1998.

[13] A. Fisher, L. Calvet, and B. B. Mandelbrot. *Multifractality of Deutschmark/US Dollar Exchanges Rates*. Yale University, 1997. Working Paper.

[14] J. Gao and I. Rubin. Multifractal modeling of counting processes of long-range dependent network traffic. In *Proceedings SCS Advanced Simulation Technologies Conf.*, San Diego, CA, April 1999.

[15] A. C. Gilbert, W. Willinger, and A. Feldmann. Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Trans. Inform. Theory*, 45(3):971–991, April 1999.

[16] S. Giordano, N. O'Connell, M. Pagano, and G. Procissi. A variational approach to the queueing analysis with fractional brownian motion input traffic. In *7th IFIP Workshop on Performance Modelling and Evaluation of ATM Networks*, Antwerp, Belgium, June 1999.

[17] J. Hüsler and V. Piterbarg. Extremes of a certain class of Gaussian processes. *Stochastic Process. Appl.*, 83:257–271, 1999.

[18] D. L. Jagerman, B. Melamed, and W. Willinger. Stochastic Modeling of Traffic Processes. In *Frontiers in Queueing*, pages 271–370. CRC Press, 1997.

[19] T. G. Kurtz. *Stochastic Networks: Theory and Applications*, chapter Limit Theorems for Workload Input Models. Oxford University Press, 1996.

[20] J. Lévi-Véhel and B. Sikdar. A multiplicative multifractal model for TCP traffic. In *Proc., IEEE ISCC*, Hammamet, Tunisia, July 2001.

[21] Z. Lui, P. Nain, D. Towsley, and Z. L. Zhang. Asymptotic behavior of a multiplexer fed by a long-range dependent process. *J. Appl. Prob.*, 36:105–118, 1999.

[22] B. B. Mandelbrot, A. Fisher, and L. Calvet. *A Multifractal Model of Asset Return*. Yale University, 1997. Working Paper.

[23] P. Mannersalo and I. Norros. Multifractal analysis of real ATM traffic: A first look. COST257TD(97)19, January 1997.

[24] M. B. Marcus and L. A. Shepp. Sample behaviour of Gaussian processes. In *Proceedings of the Sixth Berkeley Symposium*, 1972.

[25] S. Molnár and T. D. Dang. Scaling analysis of IP traffic components. In *ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management*, Monterey, CA, USA, 18-20 September 2000.

[26] S. Molnár and I. Maricza eds. Source Characterization in Broadband Networks. Interim report, COST 257, Vilamoura, Portugal, January 1999.

[27] S. Molnár and A. Vidács. On Modeling and Shaping Self-Similar ATM Traffic. In *15th International Teletraffic Congress*, Washington, DC, USA, June 1997.

[28] I. Norros. A storage model with self-similar input. *Queueing Systems*, 16:387–396, 1994.

[29] R. H. Riedi. Multifractal processes. In Doukhan, Oppenheim, and Taqqu, editors, *Long range dependence : theory and applications*, 2001.

[30] R. H. Riedi, M. S. Crouse, V. J. Ribeiro, and R. G. Baraniuk. A multifractal wavelet model with application to network traffic. *IEEE Trans. Inform. Theory*, 45(3):992–1018, April 1999.

[31] R. H. Riedi and J. Lévy-Véhel. Multifractal properties of TCP traffic: a numerical study. INRIA research report 3129, Rice University, February 1997.

[32] R. H. Riedi and W. Willinger. Toward an improved understanding of network traffic dynamics. In K. Park and W. Willinger, editors, *Self-Similar Network Traffic and Performance Evaluation*. Wiley, June 1999.

[33] M. S. Taqqu, V. Teverovsky, and W. Willinger. Is network traffic self-similar or multifractal? *Fractals*, 5:63–73, 1997.

[34] M. S. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *Computer Communication Review*, 27:5–23, 1997.

[35] B. Tsybakov and N. D. Georganas. On self-similar traffic in ATM queue: Definitions, overflow probability bound, and cell delay distribution. *IEEE/ACM Trans. on Networking*, 5(3):397–409, 1997.

[36] W. Willinger, V. Paxson, and M. S. Taqqu. Self-similarity and heavy tails: Structural modeling of network traffic. In R. J. Adler, R. E. Feldman, and M. S. Taqqu, editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pages 27–53. Birkhäuser, Boston, 1998.

[37] W. Willinger, M. S. Taqqu, and A. Erramilli. A Bibliographical Guide to Self-Similar Traffic and Performance Modeling for Modern High-Speed Networks. In *Stochastic Networks: Theory and Applications*, volume 4 of *Royal Statistical Society Lecture Notes Series*. Oxford University Press, 1996.