

On Computing Bounds on Average Backlogs and Delays with Network Calculus

Florin Ciucu Oliver Hohlfeld
 Deutsche Telekom Labs / Technische Universität Berlin
 {florin, oliver}@net.t-labs.tu-berlin.de

Abstract—The stochastic network calculus is an analytical tool which was mainly developed to compute tail bounds on backlogs and delays. From these, bounds on average backlogs and delays are derived in the literature by integration. This paper improves such bounds on average backlogs by using Jensen’s inequality; furthermore, improved bounds on average delays follow immediately from Little’s Law. The gain factor can be substantial especially at high utilizations, e.g., of order $\Omega \frac{1}{1-\rho}$ when $\rho \rightarrow 1$. This gain is further numerically illustrated for Markov-modulated On-Off arrival processes. Moreover, the paper shows how to improve standard standard stochastic network calculus performance bounds by suitably using FIFO service curves.

I. INTRODUCTION

In the early 1990s, Cruz proposed an alternative approach to the classical queueing networks theory for analyzing backlog and delays in networks, which evolved in what is currently known as the deterministic network calculus [4], [3]. The novelty of this approach is that arrivals and service are characterized in terms of worst-case bounds, upper and lower, respectively [8], [9]. Although backlogs and delays are consequently derived in terms of worst-case bounds as well, the deterministic network calculus is in essence an exact analytical tool for worst-case analysis; this means that the derived worst-case backlogs and delays can be actually attained on some sample paths [3].

A valid concern, however, is whether the worst-case behavior of arrivals and service, leading to worst-case performance bounds, can simultaneously happen in scenarios with many arrival flows. In fact, by increasing the number of flows and assuming statistical independence among these, it becomes less and less unlikely for the joint worst-case sample paths to be realized. Moreover, even in single-flow scenarios, the derived bounds can be too conservative estimates for the average backlogs and delays because statistical information about moments higher than one, for instance, is usually not captured. Therefore, the deterministic network calculus can lead to conservative bounds because many of the statistical properties of the arrivals are not accounted for.

In order to improve these bounds, the deterministic network calculus has been extended in a probabilistic framework. An immediate benefit of the emerged stochastic network calculus [4], [13] is that, in addition to worst-case arrivals, it can be applied to many typical arrivals in network analysis (e.g., Markov-modulated processes, fractional Brownian motion, or heavy-tailed traffic). In scenarios with the same input,

the stochastic network calculus yields significantly tighter performance bounds than its deterministic counterpart [2]. Moreover, in M/M/1 and M/D/1 queuing scenarios where exact results are available, the stochastic network calculus bounds are reasonably accurate [5].

The fundamental difference between the bounds obtained with the deterministic and stochastic formulations of the network calculus is that the latter are expressed as probabilistic tail bounds, i.e., they can be violated with some probabilities. For instance, probabilistic bounds on the steady-state backlog B take the form for all $\sigma \geq 0$

$$Pr(B > \sigma) \leq \varepsilon(\sigma), \quad (1)$$

where $\varepsilon(\sigma) \geq 0$ is an error function describing the violation probabilities. Note that in the deterministic network calculus, $\varepsilon(\sigma_{max}) = 0$ for some positive σ_{max} , i.e., $B < \sigma_{max}$ with probability one. Moreover, σ_{max} is a bound on the average backlog, i.e., $E[B] \leq \sigma_{max}$; however, as mentioned earlier, this bound can be rather conservative.

In turn, in the stochastic network calculus, bounds on the average backlogs and delays can be obtained in a straightforward manner by integrating the corresponding tail bounds [1], [6]. For instance, a bound on the average steady-state backlog can be derived from Eq. (1) as

$$E[B] \leq \int_0^\infty \varepsilon(\sigma) d\sigma. \quad (2)$$

The problem with this integration is that the resulting average bound accumulates the errors of all the tail bounds, and consequently can be conservative.

In this paper we carry out an improved method to derive average backlog bounds by using Jensen’s inequality for convex functions¹. Average delay bounds follow then directly from the Little’s Law. This method yields bounds on averages which are qualitatively tighter than those obtained by integrating the tail bounds as in Eq. (2). We present the relative gain factor for both general and FIFO scheduling in a scenario with two aggregates of arrival flows having bounded moment generating functions (MGF). In the case of general scheduling the gain is lower bounded by $\Omega \left(\frac{1}{1-\rho} \right)$ when $\rho \rightarrow 1$, i.e., the gain can be substantial at high utilizations. This suggests that the tail bounds in such extreme scenarios are not tight, as pointed out

¹For the exponential function and the steady-state backlog B , Jensen’s inequality states that $E[B] \leq \frac{1}{\theta} \log E[e^{\theta B}]$, for all $\theta > 0$.

in [5] for some compound Poisson arrival processes. In turn, in the case of FIFO, the gain factor is smaller and depends on the fraction of traffic carried by the flow for which the bounds are computed.

Apart from [1], [6], which use integration, a recent stochastic network calculus formulation introduced a novel idea to compute bounds on averages [12]. The key idea therein is to use a statistical envelope model involving bounds on the average of the arrivals. A drawback of such statistical envelopes is that, unlike deterministic envelopes, they do not retain information about the higher moments of the arrivals, and consequently the derived bounds may be too conservative². In turn, in this paper we use a more general statistical envelope model based on bounds on the MGF of the arrivals [4], [11], and which intrinsically retains information about higher moments. Therefore, this paper provides in essence an extension of the stochastic calculus formulations with MGFs from [4], [11] in order to compute improved bounds on average backlogs and delays.

In addition to carrying out the method of using Jensen's inequality to improve the bounds on averages obtained from tail bounds, this paper presents a rather counterintuitive result in the area of the stochastic network calculus. Concretely, we show that in a single flow scenario with bounded MGF on its arrivals and traversing a constant rate server, replacing the commonly used constant rate service curve with a smaller FIFO service curve yields tighter results. An explanation of this finding is closely related to the application of Boole's inequality to evaluate sample path bounds in the stochastic network calculus (for further details see Section III). Our finding suggests that many results from the stochastic network calculus literature (e.g., related to backlog and delay bounds), and which are consistently obtained with a constant rate service curve, can be improved by using a FIFO service curve instead.

The rest of this paper is structured as follows. Section II introduces the arrival and service models used in this paper. Section III proceeds with the derivation of bounds on average backlogs and delays using two methods, and discusses on their relative tightness. Numerical illustrations of these bounds are shown in Section IV. Brief conclusions are presented in Section V.

II. ARRIVAL AND SERVICE MODELS

The time model is continuous. We consider a single network node with constant service rate and infinite-sized buffer, and which serves the arrivals in a work-conserving manner and locally FIFO. The arrivals and departures at the node are modelled with non-decreasing, left-continuous processes. For an arrival process $A(t)$ (also referred to as a flow) and the corresponding departure process $D(t)$, we assume the initial condition $A(0) = 0$ and the causal condition $D(t) \leq A(t)$.

²We point out that the bounds derived in Theorems 5, 6, and 7 from [12] are flawed. The reason is that the proofs apply Definition 5 with τ replaced by a random variable which, however, depends on $A(t)$ (a more elaborate discussion on this pitfall is given in [14]).

For convenience we introduce the bivariate process $A(s, t) = A(t) - A(s)$. The corresponding backlog and delay processes at some time $t \geq 0$ are denoted by $B(t) = A(t) - D(t)$ and $W(t) = \inf \{d : A(t - d) \leq D(t)\}$, respectively. Furthermore, the steady-state backlog and delay, assuming their existence, are denoted by $B = \lim_{t \rightarrow \infty} B(t)$ and $W = \lim_{t \rightarrow \infty} W(t)$, respectively.

To describe arrival processes we adopt the representation with exponential bounds on their moment generating functions [4], [11]. Concretely, we say that an arrival process $A(t)$ is bounded by an MGF envelope, with rate r for some choices of a parameter $\theta > 0$, if for all $0 \leq s \leq t$

$$E \left[e^{\theta A(s,t)} \right] \leq e^{\theta r(t-s)} . \quad (3)$$

The definition appears sometime more generally with an additional pre-factor of the exponential; for the purpose of this paper it suffices to take this factor one.

The rate r in Eq. (3) depends on the parameter θ whose optimal value in the expressions of performance bounds can be numerically determined. The upper limit of the range of θ is generally inversely proportional to the data unit scale such that numerical optimizations can be done over a relatively small space. On the other hand we restrict the arrivals to the case when r is invariant to time parameters. As such, the arrival model includes for instance many Markov-modulated or multiplexed-regulated processes, but excludes self-similar processes (e.g., fractional Brownian motion where r also depends on time). The model also excludes heavy-tailed processes which have infinite MGFs.

As for traffic representation, the network calculus also uses bounds for service representation. The key idea is the concept of a service curve which relates the arrival and departure processes of a traffic flow through a lower bound. Concretely, a service curve specifies a lower bound on the amount of service received by a flow either at a network node or across an entire network path.

In this paper we adopt a service curve model from [4]. A doubly-indexed random process $S(s, t)$ is a statistical service curve for an arrival process $A(t)$ if the corresponding departure process $D(t)$ satisfies for all $t \geq 0$

$$D(t) \geq A * S(t) , \quad (4)$$

where $*$ denotes the $(\min, +)$ convolution of $A(t)$ and $S(t)$, defined as $A * S(t) = \inf_{0 \leq s \leq t} \{A(s) + S(s, t)\}$. For each sample path the random process $S(s, t)$ is decreasing in s , increasing in t , and satisfies $S(s, t) = S(s, u) + S(u, t)$ for all $0 \leq s \leq u \leq t$.

III. BOUNDS ON AVERAGE BACKLOGS AND DELAYS

In this section we derive bounds on the average backlogs and delays in the single-node scenario from Figure 1, and discuss some quantitative aspects. The node is work-conserving and serves two aggregate of flows at constant rate C . The arrival processes are denoted by $A(t)$ and $A_c(t)$, and the corresponding departure processes are denoted by $D(t)$ and $D_c(t)$, respectively.

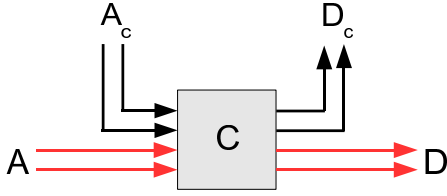


Fig. 1. A network node with cross traffic.

The next theorem provides average backlog bounds corresponding to $A(t)$, for both general scheduling and FIFO. General scheduling means that there are no assumptions on the scheduling between the two flows; this implies that the server may implement a static priority scheduling scheme, wherein $A(t)$ receives the lowest priority.

Theorem 1: (AVERAGE BACKLOG BOUNDS) Consider the node from Figure 1. The arrivals $A(t)$ and $A_c(t)$ are statistically independent and are bounded by MGF envelopes with rates r and r_c , respectively, both depending on some $\theta > 0$. Assume for stability that $r + r_c < C$. Then we have the following bounds on the average backlog according to the scheduling at the node

1) GENERAL SCHEDULING:

$$E[B] \leq \frac{1}{\theta} \log \frac{Ce}{C - (r + r_c)} \quad (5)$$

2) FIFO SCHEDULING:

$$E[B] \leq \frac{1}{\theta} \left[\log \frac{C}{C-r} + \frac{r}{C} \log \frac{C(C-r)e}{(C - (r + r_c))r} \right] \quad (6)$$

If r_A denotes the long term rate of $A(t)$ (i.e., $r_A = \lim_{t \rightarrow \infty} \frac{A(t)}{t}$), then Little's Law states that $E[W] = E[B]/r_A$. Therefore, bounds on the average steady-state delay $E[W]$ follow immediately from the theorem by using the bounds from Eqs. (5) and (6).

Let us also point out that, as hinted in the Introduction, the proof of the theorem essentially applies Jensen's inequality to the MGF of the backlog process. Bounds on such MGF's appeared also in the stochastic network calculus literature for discrete-time models [4], [13]; our proof uses the same network calculus arguments adapted to the continuous-time model.

PROOF. Fix $t \geq 0$, $\theta > 0$ such that $r + r_c < C$, and a parameter $x > 0$. For $0 < s < t - x$ let a discretization parameter τ_0 , and denote $j = \lfloor \frac{t-x-s}{\tau_0} \rfloor$ the integer part of $\frac{t-x-s}{\tau_0}$.

First we prove the FIFO case. We know from [10], [3] that the process

$$S(s, t) = [C(t-s) - A_c(s, t-x)]_+ 1_{\{t-s > x\}}$$

is a statistical service curve for $A(t)$ (this follows by extending the proof of Theorem 6.2.1 from [3] to bivariate random

processes). Then we have the following bound on the MGF of the backlog process $B(t)$

$$\begin{aligned} E[e^{\theta B(t)}] &= E[e^{\theta(A(t)-D(t))}] \\ &\leq E[e^{\theta(A(t)-A*S(t))}] \\ &\leq E\left[\sup_{0 \leq s \leq t} e^{\theta(A(s,t) - [C(t-s) - A_c(s,t-x)]_+ 1_{\{t-s > x\}})}\right] \\ &\leq E[e^{\theta A(t-x,t)}] + E\left[\sup_{j \geq 1} e^{\theta A(t-x-j\tau_0,t)}\right. \\ &\quad \left. e^{-\theta(C(x+(j-1)\tau_0) - A_c(t-x-j\tau_0,t-x))}\right] \\ &\leq e^{\theta r x} + e^{\theta C \tau_0} e^{-\theta(C-r)x} \sum_{j \geq 1} e^{-\theta(C-(r+r_c))j\tau_0} \\ &\leq e^{\theta r x} + \frac{Ce}{C - (r + r_c)} e^{-\theta(C-r)x}. \end{aligned} \quad (7)$$

In the second line we applied the definition of the statistical service curve. In the fourth line we restricted the supremum to $0 \leq s < t - x$ and applied $\sup(a, b) \leq a + b$ for positive a and b . In the sixth line we used Boole's inequality. Then we estimated the sum with $\sum_{j \geq 1} e^{-aj} \leq \int_0^\infty e^{-ax} dx = \frac{1}{a}$ for $a > 0$, and optimized $\tau_0 = \frac{1}{\theta C}$.

To continue the derivation we use the infimum

$$\inf_{x > 0} \{\alpha e^{-\beta x} + e^{\gamma x}\} = \left(\frac{\alpha\beta}{\gamma}\right)^{\frac{\gamma}{\beta+\gamma}} \frac{\beta + \gamma}{\beta},$$

which can shown by convex optimizations. We then get

$$E[e^{\theta B(t)}] \leq \frac{C}{C-r} \left(\frac{C(C-r)e}{(C - (r + r_c))r}\right)^{\frac{r}{\theta}}$$

Using Jensen's inequality and letting $t \rightarrow \infty$ completes the proof of Eq. (6).

In turn, in the case of general scheduling, we know from [11] that the process

$$S(s, t) = [C(t-s) - A_c(s, t)]_+$$

is a statistical service curve for $A(t)$. Note that this is a special case of the $S(s, t)$ obtained for FIFO scheduling with $x = 0$. A bound on the average backlog follows then directly from Eq. (7). The proof is thus complete. \square

Next we discuss on the improvement of the bounds from the theorem relative to the corresponding bounds obtained by integrating tail bounds (such bounds appeared in [6]). Assume the general case when a statistical service curve $S(t)$ for the flow $A(t)$ is known. Then, following network calculus arguments, we can write for the tail bound on the backlog for all $\sigma \geq 0$ and some $t \geq 0$

$$\begin{aligned} Pr(B(t) > \sigma) &\leq Pr\left(\sup_{0 \leq s \leq t} (A(s, t) - S(s, t)) > \sigma\right) \\ &\leq Pr\left(e^{\theta \sup_{0 \leq s \leq t} (A(s, t) - S(s, t))} > e^{\theta \sigma}\right) \\ &\leq Ke^{-\theta \sigma}, \end{aligned}$$

where $K = E \left[e^{\theta \sup_{0 \leq s \leq t} (A(s,t) - S(s,t))} \right]$, for some $\theta > 0$. The last line follows from the Chernoff bound. Note that K is a bound on the MGF of $B(t)$ (see Eq. (7)).

Therefore, we get from Eq. (2) a bound on the average backlog by integration, i.e.,

$$\begin{aligned} E[B] &\leq \int_0^\infty K e^{-\theta \sigma} d\sigma \\ &= \frac{1}{\theta} K. \end{aligned}$$

In turn, Theorem 1 gives the improved bound

$$E[B] \leq \frac{1}{\theta} \log K,$$

such that the achieved gain factor by using the method from Theorem 1 is $\frac{K}{\log K}$ (the latter bound is smaller than the former bound by this much).

In particular, in the case of general scheduling, we have from the proof of the theorem that $K = \frac{C e}{C - (r + r_c)}$. Denoting $\rho = \frac{r + r_c}{C}$, which depends on θ as well, we obtain the gain factor

$$\frac{e}{1 - \rho + \log \left(\frac{1}{1 - \rho} \right)^{1 - \rho}}.$$

Letting the actual utilization at the node go to one implies that ρ goes to one as well, because r and r_c are upper bounds on the average rates of $A(t)$ and $A_c(t)$, respectively. Using further that $\lim_{x \rightarrow 0} \left(\frac{1}{x} \right)^x = 1$, it follows that the gain factor above is of order

$$\text{Gain}_{\text{gen. sched.}} = \Omega \left(\frac{1}{1 - \rho} \right). \quad (8)$$

In turn, in the case of FIFO scheduling, the gain factor is

$$\frac{a \left(\frac{1}{1 - \rho} \right)^{\frac{r}{C}}}{\log a \left(\frac{1}{1 - \rho} \right)^{\frac{r}{C}}},$$

where $a = \frac{C}{C - r} \left(\frac{(C - r)e}{r} \right)^{\frac{r}{C}}$. Following a similar argument as for general scheduling yields a gain factor of a smaller order, i.e.,

$$\text{Gain}_{\text{FIFO}} = \Omega \left(\frac{1}{(1 - \rho)^\phi} \right), \quad (9)$$

where $\phi = \frac{rA}{C}$ is the percentage of how much the flow $A(t)$ uses out of the total rate C .

Let us also point out that in the case when $A(t)$ is alone in the system (i.e., $A_c(t) = 0$), then the bound from Eq. (6) is tighter than the bound from Eq. (5). This can be shown by replacing $r_c = 0$ and using the inequality $a \left(1 + \log \frac{1}{a} \right) \leq 1$ which holds for all $0 < a < 1$. This result indicates that at least in scenarios with a single flow traversing a server with constant rate C , using a FIFO statistical service curves yields tighter bounds than by using the statistical service curve $S(t) = Ct$ which has a purely deterministic form; note that when $A_c(t) = 0$ then the bound from Eq. (5) is obtained with such a service curve.

The intuition behind this rather counterintuitive result lies in the application of Boole's inequality in Eq. (7), adapted for $A_c(t) = 0$. Although the FIFO service curve $S'(t) = Ct 1_{\{t > x\}}$, for some $x > 0$, is smaller than the commonly used $S(t) = Ct$, and hence appears to lead to larger bounds, the application of Boole's inequality with the FIFO service curve prevents the summation of the largest terms, more exactly at the time scales less than x . Overall, by optimizing x , the service curve $S'(t)$ yields tighter bounds than $S(t)$.

IV. NUMERICAL EVALUATIONS

This section illustrates numerically the improvement of the bounds on average backlogs derived with Theorem 1 relative to the corresponding bounds derived by integrating tail bounds with Eq. (2).

We consider Markov-modulated On-Off (MMOO) processes. One such process is defined by first considering a homogenous and continuous-time Markov chain $X(t)$ with two states denoted by 'On' and 'Off', and with the transition matrix

$$Q = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix}.$$

Here, the parameters μ and λ represent the transition rates from the 'On' state to the 'Off' state, and vice-versa, respectively. In the steady-state, the average dwell time of the process $X(t)$ in the 'On' state is $\frac{1}{\mu}$, and the average dwell time in the 'Off' state is $\frac{1}{\lambda}$.

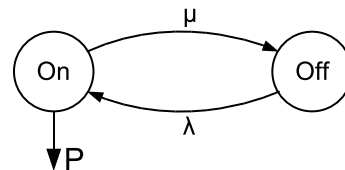


Fig. 2. A Markov-modulated On-Off traffic voice model.

A continuous-time arrival process $A(t)$ is a Markov-modulated On-Off process driven by the Markov process $X(t)$ if its instantaneous arrival rate is either P or zero, depending whether $X(t)$ is in the 'On' or 'Off' states, respectively (see Figure 2). This MMOO process has an MGF envelope with rate [7]

$$r' = \frac{1}{2\theta} \left(P\theta - \mu - \lambda + \sqrt{(P\theta - \mu + \lambda)^2 + 4\lambda\mu} \right).$$

Assuming further that $A(t)$ and $A_c(t)$ from Figure 1 consist of N and N_c , respectively, homogeneous and statistically independent MMOO processes with rate r' , it follows that the rates of the corresponding MGF envelopes are $r = Nr'$ and $r_c = N_c r'$.

For numerical illustrations we choose the following values for the parameters of an MMOO flow, typically used to model voice flows [15]: $\frac{1}{\mu} = 0.4$ s, $\frac{1}{\lambda} = 0.6$ s, peak rate $P = 64$ Kb/s, and average rate 25.6 Kb/s. Fig. 3 illustrates the bounds on the average backlog for an aggregate of N such flows, at a server with rate $C = 100$ Mbps and also serving

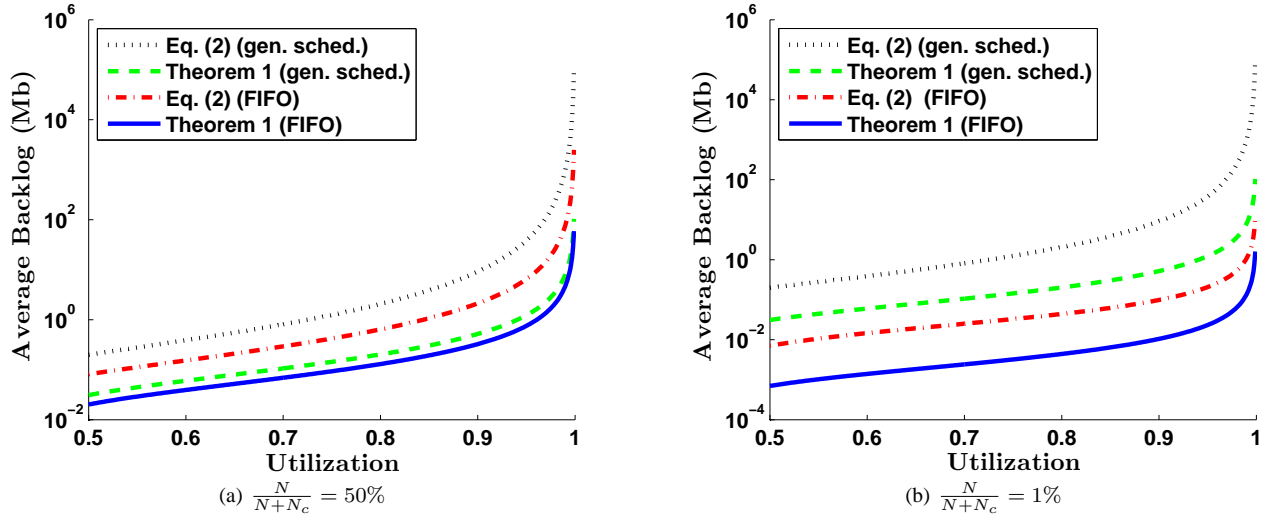


Fig. 3. Bounds on the average backlog computed with Theorem 1 and Eq. (2) (integrating tail bounds), for both general scheduling and FIFO, as a function of node utilization (constant rate $C = 100$ Mbps, MMOO flows with $\frac{1}{\mu} = 0.4$ s, $\frac{1}{\lambda} = 0.6$ s, average rate 25.6 Kb/s, and peak rate $P = 64$ Kb/s, different percentages $\frac{N}{N+N_c}$ of the aggregate of flows for which the bounds are shown)

an aggregate of N_c other flows, as a function of the utilization. Fig. 3.(a) considers a fraction $\frac{N}{N+N_c} = 50\%$ of the aggregate of flows for which the bounds are computed. Remarkably, for both general and FIFO scheduling, the gain factor of using Theorem 1 relative to using the integration of tail bounds as in Eq. (2), is substantial at all utilizations. Moreover, as it was analytically shown in Eqs. (8) and (9), the gain factor blows up at very high utilizations. The figure also indicates that the choice of scheduling makes a significant difference when using Eq. (2), and a lesser difference when using Theorem 1, mainly because the fraction $\frac{N}{N+N_c}$ is quite large.

In turn, in the case of a very small fraction $\frac{N}{N+N_c} = 1\%$, as considered in Fig. 3.(b), the choice of the scheduling algorithm makes a significant difference when using either Eq. (2) or Theorem 1. Moreover, in the case of FIFO scheduling, the figure shows that the gain factor of using Theorem 1 relative to using Eq. (2) is much smaller than illustrated in Fig. 3.(a) at very high utilizations (see the order of actual growth of the gain from Eq. (9) which decays with $\frac{N}{N+N_c}$ when the utilization $\rho \rightarrow 1$).

V. CONCLUSIONS

In this paper we have carried out the derivation of bounds on average backlogs and delays, in a stochastic network calculus framework based on moment generating functions, by using Jensen's inequality. We compared the obtained results with corresponding bounds which have been derived in the literature by integrating a-priori tail bounds, and showed that the gain of using the former method can be significant. Moreover, we have found a rather counterintuitive result which leads to the improvement of many standard bounds in the stochastic network calculus literature.

REFERENCES

- [1] S. Ayyorgun and W. Feng. A systematic approach for providing end-to-end probabilistic QoS guarantees. In *Proceedings of the International Conference On Computer Communications and Networks (ICCCN)*, pages 115–122, October 2004.
- [2] R. Boorstyn, A. Burchard, J. Liebeherr, and C. Oottamakorn. Statistical service assurances for traffic scheduling algorithms. *IEEE Journal on Selected Areas in Communications Special Issue on Internet QoS*, 18(12):2651–2664, December 2000.
- [3] J. Y. Le Boudec and P. Thiran. *Network Calculus*. Springer Verlag, Lecture Notes in Computer Science, LNCS 2050, 2001.
- [4] C.-S. Chang. *Performance Guarantees in Communication Networks*. Springer Verlag, 2000.
- [5] F. Ciucu. Network calculus delay bounds in queueing networks with exact solutions. In *20th International Teletraffic Congress (ITC)*, June 2007.
- [6] F. Ciucu. *Scaling Properties in the Stochastic Network Calculus*. PhD thesis, University of Virginia, Charlottesville, VA, 2007.
- [7] C. Courcoubetis and R. Weber. Buffer overflow asymptotics for a buffer handling many traffic sources. *Journal of Applied Probability*, 33(3):886–903, 1996.
- [8] R. Cruz. A calculus for network delay, parts I and II. *IEEE Transactions on Information Theory*, 37(1):114–141, Jan. 1991.
- [9] R. L. Cruz. Quality of service guarantees in virtual circuit switched networks. *IEEE Journal on Selected Areas in Communications*, 13(6):1048–1056, Aug. 1995.
- [10] R. L. Cruz. SCED+: Efficient management of quality of service guarantees. In *Proceedings of IEEE Infocom*, pages 625–634, Mar. 1998.
- [11] M. Fidler. An end-to-end probabilistic network calculus with moment generating functions. In *IEEE 14th International Workshop on Quality of Service (IWQoS)*, pages 261–270, June 2006.
- [12] A. Gulyás and J. Bıró. A stochastic extension of network calculus for workload loss examinations. *IEEE Communications Letters*, 10(5):399–402, 2006.
- [13] Y. Jiang and Y. Liu. *Stochastic Network Calculus*. Springer, 2008.
- [14] C. Li, A. Burchard, and J. Liebeherr. A network calculus with effective bandwidth. *IEEE/ACM Transactions on Networking*, 15(6):1442–1453, 2007.
- [15] M. Schwartz. *Broadband Integrated Networks*. Prentice Hall, 1996.