Performance Bounds Analysis on Different Scheduling Algorithms for Networks with Self-Similar Traffic

Chao Luo
National Laboratory for Optoelectronics
Huazhong University of Science and Technology
Wuhan, China
lc19851205@gmail.com

Li Yu
National Laboratory for Optoelectronics
Huazhong University of Science and Technology
Wuhan, China
hustlyu@mail.hust.edu.cn

Abstract—**There are kinds of Internet applications such as VoIP, online gaming and video Conferencing which need quality of service guarantees making the performance analysis critical, and the traffic of such applications usually exhibits the properties of self-similarity and heavy tail. In this paper, we use the generalized Stochastically Bounded Burstiness model to characterize the network traffic with self-similar property and self-similar property. And we derive the delay bounds for such kind of traffic in the system under three different scheduling algorithms: First In First Out, Static Priority and Generalized Processor Sharing. We utilize the relationships between the traffic-amount-centric stochastic arrival curves and the virtual-backlog-centric stochastic arrival curve as well as the relationships between the weak stochastic service curve and the stochastic service curve making the analysis of the performance bounds of the through traffic in the tandem systems feasible.**

Keywords-network calculus; self-similar; scheduling; performance analysis

I. INTRODUCTION

Network calculus is a theory for performance bounds analysis in computer networks. The deterministic network calculus, firstly proposed by Cruz [11], can be used to get the delay bound and backlog bound in the worst case. However, the deterministic network calculus usually results in overly pessimistic performance bounds which are rarely attained. Hence, the stochastic network calculus which can get better utilization of networks has been studied by many researchers such as in [6] and [16]. Jiang has proposed some different kinds of stochastic arrival curves such as the traffic-amount-centric stochastic arrival curve and the virtual-backlog-centric stochastic arrival curve as well as some different types of stochastic service curves such as the weak stochastic service curve and the stochastic service curve [6], [7], [17].

The network traffic related to specific sources, such as video, or specific applications such as FTP cannot be modeled by Poisson and Markov processes. It has been observed that Ethernet traffic have self-similarity [1], [2], [9]. Some stochastic processes such as Fractional Brownian motion (FBM) and α-stable process have been proved having the property of self-similarity. Since the network traffic not only exhibits the self-similar property but also extends to the heavy-tailed distribution which FBM doesn't have [15], and the α-stable process captures not only the self-similar property but also the heavy-tailed property. We use α-stable process to simulate network traffic in this paper. There are several traffic models including the Exponentially Bounded Burstiness (EBB) [11], the Stochastically Bounded Burstiness (SBB) [12] and the generalized Stochastically Bounded Burstiness (gSBB) [14] have been proposed to describe the traffic arriving into the network for the stochastic network calculus. We use the gSBB model in this paper because it can apply to α-stable process. The gSBB is one kind of virtual-backlog-centric traffic models [17].

In this paper, we consider a one node network scenario as well as a scenario of *N* nodes in tandem. In the both two network scenarios, there are two kinds of traffic arriving in the network. One is called the through traffic and the other is called the cross traffic. Both the through traffic and the cross traffic are self-similar and heavy-tailed which are simulated by α-stable process and modeled by the gSBB model. Firstly, we analyze the performance bounds for both the through traffic and the cross traffic under some different scheduling disciplines. And then we analyze the performance bounds for the through traffic in the tandem systems: We firstly derive the weak stochastic service curve for the through traffic at each node. Next, we change the weak stochastic service curve into the stochastic service curve. The reason we do this transformation is that the weak stochastic service curve does not have the concatenation property with out which the service provided by the tandem systems can not be derived, and then we get the stochastic service curve for the through traffic provided by the tandem systems by using the concatenation property. With the traffic-amount-centric stochastic arrival curve and the stochastic service curve, we readily get the delay bound and the backlog bound of the through traffic.

The rest of the paper is organized as follows. In Section II, we define notations that will be widely used in this paper. In Section III, we introduce network calculus basics and the traffic model to characterize the α-stable traffic. In Section

978-1-4673-2815-9/12/$31.00 ©2012 IEEE

IV, firstly we derive the delay bounds for the two traffic in a single node which offers a constant rate service given by the scheduling algorithms FIFO, SP, and GPS. And then we present the end-to-end performance bounds for the through traffic with single-hop persistent cross traffic in tandem systems. We give brief conclusions in Section V.

II. NOTATION

In this section some notations are defined which will be used in this paper. We consider a discrete time system, of which a process is defined as a function of time t, (t = 0, 1, 2...). We use $A(t)$ denoting the arrival process, $D(t)$ denoting the departure process and $S(t)$ denoting the service process respectively. We denote $A(s, t) = A(t) – A(s)$ for any $0 \le s \le t$, as well as $D(s, t) = D(t) - D(s)$ and $S(s, t) = S(t) - S(s)$. We assume $A(t) = D(t) = S(t) = 0$. We also assume the networks are work-conserving. At time t, the backlog $B(t)$ and the delay $W(t)$ in the system are given by $B(t) = A(t) - D(t)$ and $W(t) = \inf\{d \ge 0: A(t) \le D(t + d)\}$ respectively. We also define the queue length of a system with service rate ρ as

$$Q(t;\rho) = \sup_{0\le s\le t}\{A(s,t)-(t-s)\rho\},$$

where ρ is a positive constant.

Given two functions α(t) and β(t), we define $h(\alpha, \beta)$ and $v(\alpha, \beta)$ as follows

$$h(\alpha,\beta) = \sup_{s\ge 0}\{\inf\{\tau \ge 0 : \alpha(s) \le \beta(s+\tau)\},$$

$$v(\alpha,\beta) = \sup_{s\ge 0}\{\alpha(s)-\beta(s)\}.$$

If the arrival curve and the departure curve of a system are denoted by α(t) and β(t) respectively, and $h(\alpha, \beta)$ and $v(\alpha, \beta)$ are the maximum delay and the maximum backlog of the system respectively.

For given functions f and g, the min-plus algebra formulation [3-4] with the convolution operator $\otimes$ and the de-convolution operator $\oslash$ is as follows

$$f \otimes g(t) = \inf_{0\le\tau\le t}\{f(t\text{-}\tau) + g(\tau)\},$$

$$f \oslash g(t) = \sup_{\tau\ge 0}\{f(t+\tau) - g(\tau)\}.$$

III. NETWORK MODELS

A. Nework Calculus Basics

There are some basic definitions for deterministic network calculus.

Definition 1 ([8]). Given a wide-sense increasing function α defined for $t \ge 0$, we say that a flow R is constrained by α if and only if for all $0 \le s \le t$:

$$R(t) - R(s) \le \alpha(t-s).$$

Definition 2 ([8]). Given a wide-sense increasing function β defined for $t \ge 0$, we say that a server provides a service cure β to its arrival A if and only if for all $t \ge 0$, its departure D satisfies

$$D(t) \ge A \otimes \beta(t).$$

Several probabilistic extensions of arrival curve and service curve have been proposed for the stochastic network calculus. These arrival curves and service curves given in [6] are as follows.

Definition 3 ([6]). Given a wide-sense increasing function α and a wide-sense decreasing function f, we say a flow has a traffic-amount-centric stochastic arrival curve α to its arrival A denoted by $A \sim_{ta} \langle f, \alpha\rangle$, if and only if for all $t \ge 0$ and all $x \ge 0$, there holds

$$P\{A(s,t)-\alpha(t-s) > x\} \le f(x). \tag{1}$$

Definition 4 ([6]). Given a wide-sense increasing function α and a wide-sense decreasing function f, we say a flow has a virtual-backlog-centric stochastic arrival curve α to its arrival A denoted by $A \sim_{vb} \langle f, \alpha\rangle$, if and only if for all $t \ge 0$ and all $x \ge 0$, there holds

$$P\{\sup_{0\le s\le t}\{A(s,t)-\alpha(t-s)\} > x\} \le f(x). \tag{2}$$

The relationship between the traffic-amount-centric stochastic arrival curve and the virtual-backlog-centric stochastic arrival curve has been given as follows.

Theorem 1 ([7]). If a flow has a traffic-amount-centric stochastic arrival curve α with bounding function f, it also has a virtual-backlog-centric stochastic arrival curve α_ε with bounding function f_ε, where

$$\alpha_\varepsilon(t) = \alpha(t) + \varepsilon \cdot t,$$

$$f_\varepsilon(x) = [f(x) + \frac{1}{\varepsilon}\int_x^\infty f(u)du]_1,$$

for any $\varepsilon > 0$.

Definition 5 ([7]). Given a wide-sense increasing function α and a wide-sense decreasing function g, we say a system provides a weak stochastic service curve β with bounding function g denoted by $S \sim_{ws} \langle g, \beta\rangle$, if for all $t \ge 0$ and all $x \ge 0$ there holds

$$P\{A \otimes \beta(t) - D(t) > x\} \le g(x). \tag{3}$$

Definition 6 ([7]). Given a wide-sense increasing function α and a wide-sense decreasing function g, we say a system provides a stochastic service curve β with bounding function g denoted by $S \sim_{sc} \langle g, \beta\rangle$, if for all $t \ge 0$ and all $x \ge 0$ there holds

$$P\{\sup_{0\le s\le t}[A \otimes \beta(s) - D(s)] > x\} \le g(x). \tag{4}$$

The relationship between the weak stochastic service curve and the the stochastic service curve is also given as

Theorem 2 ([7])**.** If the server provides to the input a weak stochastic service curve β with bounding function g, it provides to the input a stochastic service curve β_λ with bounding function g_λ, where

$$\beta_\lambda(t) = \beta(t) - \lambda \cdot t,$$

$$g_\lambda(x) = [\frac{1}{\lambda}\int_{x-\lambda\cdot t}^{t} g(u)du]_1,$$

for any $\lambda > 0$.

It has been shown that stochastic service curve has the concatenation property as follows.

Theorem 3 ([6])**.** Consider a flow passing through a network of N nodes in tandem. If each node $n(= 1, 2, \ldots , N)$ provides stochastic service curve $S^n \sim_{sc} \langle g^n, \beta^n\rangle$ to its input, the network guarantees to the flow a stochastic service cure $S \sim_{sc} \langle g, \beta\rangle$ with

$$\beta(t) = \beta^1 \otimes \beta^2 \otimes \cdots \otimes \beta^N(t)$$
$$g(x) = g^1 \otimes g^2 \otimes \cdots \otimes g^N(x)$$

If the arrival curve and the service curve to the input are known, the performance bounds of the input can be obtained by

Theorem 4 ([7])**.** Consider a system with input A. If the input has a virtual-backlog-centric stochastic arrival curve α with bounding function f and the system provides to the input a stochastic service curve β with bounding function g, the backlog $B(t)$ and the delay $W(t)$ are bounded by

$$P\{B(t) > x\} \le f \otimes g(x - \alpha \oslash \beta(0)),$$
$$P\{W(t) > h(\alpha + x, \beta)\} \le f \otimes g(x).$$

B. Generalized Stochastically Bounded Bursty for α-stable traffic

Letting $\alpha(t) = \rho t$ in Definition. 3., we get the definition of gSBB as follows.

Definition 7. Given a wide-sense decreasing function f, we say a flow A is generalized stochastically bounded bursty (gSBB) with upper rate ρ and bounding function f denoted by $A \sim_{vb} \langle f, \rho\rangle$, if and only if for all $t \ge 0$ and all $x \ge 0$, there holds

$$P\{\sup_{0\le s\le t}\{A(s,t) - (t-s)\cdot\rho\} > x\} \le f(x). \qquad (5)$$

The α−stable is a traffic model is defined by four parameters: (α, H, $c1$, $c2$) [17]. It has been shown that we can use the gSBB model to represents α-stable traffic with $A_{\alpha\text{-stable}} \sim_{vb} \langle f, \rho\rangle$ where

$$f(x) = C_\alpha(\frac{\rho - m}{c_1})^{-\alpha}, \qquad (6)$$

in which C_α and m are determined from (α, H, $c1$, $c2$) [7].

Since the gSBB is one kind of virtual-backlog-centric traffic models. It has the same properties which are introduced for virtual-backlog-centric traffic models.

Theorem 5 ([7])**.** Consider 2 flows with arrival process A_1 and A_2 respectively. Let A denote the aggregate arrival process. If $A_1 \sim_{vb} \langle f_1, \rho_1\rangle$ and $A_2 \sim_{vb} \langle f_2, \rho_2\rangle$, than $A \sim_{vb} \langle f, \rho\rangle$ with $\rho = \rho_1 + \rho_2$ and $f(x) = f_1 \otimes f_2(x)$.

Theorem 6 ([7])**.** Consider a work-conserving system with constant service rate C. Let A and D be the input and output process of the system respectively. If the input $A \sim_{vb} \langle f, \rho\rangle$ and $\rho \le C$, the output $D \sim_{vb} \langle f, \rho\rangle$.

Theorem 7 ([7])**.** Consider a system with input A. If the input A is gSBB as $A \sim_{vb} \langle f, \rho\rangle$ and the system provides the input a weak stochastic service curve as $S \sim_{ws} \langle g, \beta\rangle$, the output has a traffic-amount-centric stochastic arrival curve as $D \sim_{ta} \langle f \otimes g, \rho \oslash \beta\rangle$.

IV. PERFORMANCE BOUNDS ANALYSIS.

A. Single Node Delay Analysis

We consider a single node network as shown in Fig. 1. There are two traffic flows, the through traffic and the cross traffic, flowing into the node. Assuming both the through traffic and the cross traffic are α-stable traffic with parameters (α_1, H_1, c_1, c_2) and (α_2, H_2, c_3, c_4) respectively. They can be modeled as $A_t \sim vb \langle f_t, \rho_t\rangle$ and $A_c \sim vb \langle f_c, \rho_c\rangle$. We have f_t and f_c from (6) as

$$f_t(x) = C_t\left(\frac{\rho_t - m_t}{c_1}\right)^{-\alpha_1}, \qquad (7)$$

$$f_c(x) = C_c\left(\frac{\rho_c - m_c}{c_3}\right)^{-\alpha_2}, \qquad (8)$$

in which C_t , m_t determined from (α_1, H_1, c_1, c_2) and C_c , m_c determined from (α_2, H_2, c_3, c_4). We assume the network provides a service with constant rate C and analyze the delay bounds under three scheduling algorithms: FIFO, SP, and GPS.

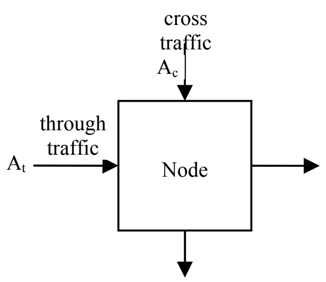

Figure 1. A single node network with cross traffic.

1) FIFO: FIFO is also known as First Come, First Served (FCFS), is the simplest scheduling algorithm, simply queues processes in the order that they arrive in the ready queue.

Theorem 8. Assume that the cross traffic $A_c \sim_{vb} \langle f_c, \rho_c \rangle$ and the through traffic $A_t \sim_{vb} \langle f_t, \rho_t \rangle$ arrive into the single node network under FIFO scheduling algorithm. For $k \geq 1$ and $C > \rho_c + \rho_t$, the through traffic and cross traffic have the same delay bound

$$P\{W(t) \geq k\} \leq f(kC - \rho), \tag{9}$$

where $f(x) = f_c \otimes f_t(x)$ and $\rho = \rho_c + \rho_t$.

Proof: Obviously the through traffic and cross traffic have the same delay bound because they both arrive in the system at time t, which also means there are both through traffic and cross traffic at the queue tail.

By the superposition property of the gSBB from Theorem 1., we can aggregate the through traffic and the cross traffic as a new traffic which can be modeled as $A \sim \langle f, \rho \rangle$ with $f(x) = f_c \otimes f_t(x)$ and $\rho = \rho_c + \rho_t$.

We can easily get $W(t) = \left\lceil \dfrac{Q(t)}{C} \right\rceil$ by the definition of FIFO.

Hence, for all $k \geq 1$,

$$\begin{aligned}
&P\{W(t) \geq k\} = P(Q(t) > (k-1)C\} \\
&= P\{\sup_{0 \leq s < t}\{A(s,t) - C(t-s)\} > (k-1)C\} \\
&\leq P\{\sup_{0 \leq s < t}\{A(s,t) - \rho(t-s) + \rho - C\} > (k-1)C\} \\
&= P\{\sup_{0 \leq s < t}\{A(s,t) - \rho(t-s) > kC - \rho\} \leq f(kC - \rho).
\end{aligned}$$

2) SP: In an SP scheduler, every flow is assigned a priority index, where a lower priority index indicates a higher priority. An SP scheduler selects for transmission the earliest arrival from the highest priority class with a nonzero backlog [16].

Theorem 9. Assume that the cross traffic $A_c \sim_{vb} \langle f_c, \rho_c \rangle$ and the through traffic $A_t \sim_{vb} \langle f_t, \rho_t \rangle$ arrive into the single node network under SP scheduling algorithm. We also assume the cross traffic has the higher priority, and we have, for $k \geq 1$ and $C > \rho_c + \rho_t$,

$$P\{W_c(t) \geq k\} \leq f_c(kC - \rho_c), \tag{10}$$

and

$$P\{W_t(t) \geq k\} \leq f_c \otimes f_t(k(C - \rho_c - \rho_t)). \tag{11}$$

Proof: By the definition of SP, we know that if there is cross traffic in the queue the through traffic won't be served. We can get (10) directly in the same way as in the FIFO system mentioned above.

For the through traffic, we can aggregate the through traffic and cross traffic as a new traffic which can be modeled as $A \sim \langle f, \rho \rangle$ with $f(x) = f_c \otimes f_t(x)$ and $\rho = \rho_c + \rho_t$. It is obvious that A_t and A has the same delay bound since the through traffic has the lower priority. We can get the delay bound from [14],

$$\begin{aligned}
&P\{W_t(t) \geq k\} = P\{W(t) \geq k\} \\
&\leq f(kC - k\rho).
\end{aligned}$$

3) GPS: In a GPS scheduler, the through traffic and the cross traffic are assigned with the weight index ϕ_t and ϕ_c respectively. If the through traffic is continuously backlogged in the time interval $(s, t]$, its departures D_t satisfy the defining weighting [5]

$$\frac{D_t(s,t)}{D_c(s,t)} \geq \frac{\phi_t}{\phi_c}.$$

Theorem 10. Assume that the cross traffic $A_c \sim_{vb} \langle f_c, \rho_c \rangle$ and the through traffic $A_t \sim_{vb} \langle f_t, \rho_t \rangle$ arrive into the single node network under GPS scheduling algorithm. The through traffic and the cross traffic are assigned with the weight index ϕ_t and ϕ_c respectively and $\phi_t + \phi_c = 1$. For $k \geq 1$ and $C > \rho_c + \rho_t$, if $\phi_t C > \rho_t$, we have,

$$P\{W_t(t) \geq k\} \leq f_t(k\phi_t C - \rho_t).$$

Otherwise, we get the same result as (11)

$$P\{W_t(t) \geq k\} \leq f_c \otimes f_t(k(C - \rho_c - \rho_t)).$$

The proofing process is omitted since the results are obvious. The delay bound for the cross traffic can be derived in the same way.

We can get the delay bounds for the α-stable traffic by using (7) and (8) in the results derived above, instead of f_t and f_c.

B. Tandem Systems Performance Analysis

We now consider tandem systems as shown in Fig. 2. The through traffic traverses N nodes in series. At each node, there is cross traffic passing through it. To easy the expression, we assume the through traffic has the same parameters (α_1, H_1, c_1, c_2) as we describe in the single node scenario above and the cross traffic at each node has the same parameters (α_2, H_2, c_3, c_4). We also assume every node provides the same constant service rate C. We will present the derivation process of the delay bound and the backlog bound for the through traffic in this network scenario.

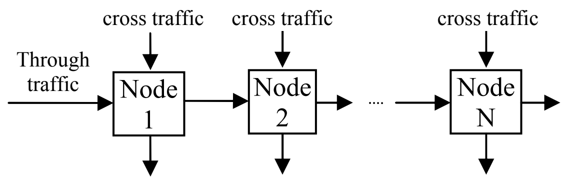

Figure 2. A tandem nodes network scenario.

We assume every node adopts FIFO scheduling. The following weak stochastic service curve for the through traffic at node 1 can be derived based on the theorems in [7]:

$$S_t^1 \sim_{ws} < f_c \otimes f_t, Ct - \rho_c t > .$$

By Theorem 7., we can get the departure traffic of the through traffic at node 1 as

$$D_t^1 \sim_{ta} < f_c \otimes f_t, (\rho_t t) \oslash (Ct - \rho_c t) > .$$

Since $C - \rho_c > \rho_t$, we have

$$\begin{aligned}
&(\rho_t t) \oslash (Ct - \rho_c t)(t) \\
&= \sup_{s \geq 0} \{\rho_t (t+s) - (C - \rho_c) s\} \\
&= \rho_t t.
\end{aligned}$$

The departure traffic of the through traffic at node 1 can be now represented as

$$D_t^1 \sim_{ta} < f_c \otimes f_t, \rho_t t > .$$

According to the relationship between the traffic-amount-centric stochastic arrival curve and the virtual-backlog-centric stochastic arrival curve from Theorem 1., we get the input at node 2 as

$$A_t^2 = D_t^1 \sim_{vb} < f_t^2, \rho_t t + \varepsilon_1 t > ,$$

where

$$f_t^2(x) = [f_c \otimes f_t(x) + \frac{1}{\varepsilon_1} \int_x^\infty f_c \otimes f_t(u) du]_1 ,$$

for any $\varepsilon_1 > 0$.

We can get the weak stochastic service curve for the through traffic at node 2 in the same way as we derived at node 1 as

$$S_t^2 \sim_{ws} < f_c \otimes f_t^2, Ct - \rho_c t > .$$

Hence, we can get the weak stochastic service curve for the through traffic at node i in the same way as

$$S_t^i \sim_{ws} < f_c \otimes f_t^i, Ct - \rho_c t > ,$$

where

$$f_t^i(x) = [f_c \otimes f_t^{i-1}(x) + \frac{1}{\varepsilon_{i-1}} \int_x^\infty f_c \otimes f_t^{i-1}(u) du]_1$$

for $2 \leq i \leq N$, any $\varepsilon_{i-1} > 0$.

Since the combination of the virtual-backlog-centric stochastic arrival curve and the weak stochastic service curve does not have the concatenation property, we change the weak stochastic service curve into the stochastic service curve. According to the relationship between the weak stochastic service curve and the stochastic service curve by Theorem 2., we have

$$S_t^i \sim_{sc} < g_t^i, \beta_t^i > ,$$

where

$$g_t^i(x) = [\frac{1}{\lambda_i} \int_{x - \lambda_i t}^t f_c \otimes f_t^i(y) dy]_1$$

and

$$\beta_t^i = Ct - \rho_c t - \lambda_i t$$

for any $\lambda_i > 0$.

By Theorem 3., we have the stochastic service curve for the through traffic

$$S_t \sim_{sc} < g_t, \beta_t > ,$$

where

$$\begin{aligned}
g_t(x) &= g_t^1 \otimes g_t^2 \otimes \ldots \otimes g_t^N(x), \\
\beta_t(t) &= \beta_t^1 \otimes \beta_t^2 \otimes \ldots \otimes \beta_t^N(t).
\end{aligned}$$

We can readily get the performance bounds for the through traffic from Theorem 4.

$$\begin{aligned}
&P\{B_t(t) > x\} \leq f_t \otimes g_t(x - v(\rho_c t, \beta_t(t))), \\
&P\{W(t) > h(\rho_t t + x, \beta_t(t))\} \leq f_t \otimes g_t(x).
\end{aligned}$$

The performance bounds under SP and GPS can be derived in a similar way.

V. CONCLUSION

This paper introduced a way to analyze the performance bounds for the networks with heavy-tailed self-similar traffic. We used the gSBB model to represent the α-stable traffic which can be used to simulate the network traffic with heavy-tailed self-similar characterization. Firstly, we explored the delay bounds of two flows for the through traffic in a single node network under several scheduling algorithms (FIFO, SP, GPS), and then we derived the leftover service for the through traffic at each node in tandem systems. Finally, we got the delay bound and backlog bound of the through traffic by using some properties of stochastic network calculus.

ACKNOWLEDGMENT

This work is supported by National 863 Projects of China (2009AA01Z205), Fund of National Laboratory (P080010), Natural Science Foundation of China (60972016) and Program for New Century Excellent Talents in University (NCET070339).

REFERENCES

[1] W. Leland, M. Taqqu, W. Willinger, and D. Willson, "On the self-similar nature of Ethernet traffic (extended version)," IEEE/ACM Trans. Networking, vol. 2, no. 1, pp. 1-15, Feb. 1994.

[2] M. Grossglauser and J.-C. Bolot, “On the relevance of long-range dependence in network traffic,” IEEE/ACM Trans. Networking, vol. 7, pp. 629-640, 1999.

[3] R. Agrawal, R. L. Cruz, C. Okino, and R. Rajan, "Performance bounds for flow control protocols", IEEE/ACM Trans. Networking, vol. 7, pp. 310-323 , 1999.

[4] J. Y. Le Boudec, “Application of network calculus to guaranteed service networks,” IEEE Trans. Inform. Theory, vol. 44, pp.1087 - 1096 , 1998.

[5] M. Fidler, “A survey of deterministic and stochastic service curve models in the network calculus,” IEEE Communications Surveys and Tutorials, vol. 12, no. 1, 2010.

[6] Y. Jiang, “A basic stochastic network calculus,” in Proc. ACM SIGCOMM, Oct. 2006, pp. 123-134.

[7] Y. Jiang and Y. Liu, Stochastic Network Calculus. Springer, 2008.

[8] J.-Y. LeBoudec and P. Thiran, Network Calculus - A Theory of Deterministic Queuing System for the Internet. Springer, 2004.

[9] M. Crovella and A. Bestavros. “Self-similarity in World Wide Web traffic: evidence and possible causes. IEEE/ACM Transactions on Networking, 5(6):835-846, Dec. 1997.

[10] R. Cruz, “A calculus for network delay, parts I and II,” IEEE Trans. Inf. Theory, vol. 37, no.1, pp. 114-141, Jan. 1991.

[11] O. Yaron and M. Sidi, “Generalized processor sharing networks with exponentially bounded Burstiness arrivals,” J. High Speed Networks, vol. 3, pp. 375-387, 1994.

[12] D. Starobinski and M. Sidi, “Stochastically bounded burstiness for communication networks,” IEEE Trans. Inf. Theory, vol. 46, no. 1, pp. 206-212, Jan. 2000.

[13] Q. Yin, Y. Jiang, S. Jiang, and P. Y. Kong, “Analysis of generalized stochastically bounded bursty traffic for communication networks,” in Proc. IEEE LCN, Nov. 2002, pp. 141-149.

[14] M. Fidler, “An End-to-End Probabilistic Network Calculus with Moment Generating Functions,” in Proc. IEEE IWQOS, Jun. 2006, pp. 261-270.

[15] J. Liebeherr, A. Burchard and F. Ciucu, “Non-asymptotic Delay Bounds for Networks with Heavy-Tailed Traffic,” in Proc. IEEE INFCOM, Mar. 2010, pp. 1-9.

[16] C. Li, A. Burchard and J. Liebeherr, “A Network Calculus with Effective Bandwith,” IEEE/ACM Trans. Networking, vol. 15, no. 6, pp. 1442-1543, Dec. 2007.

[17] Y. Jiang, Q. Yin, Y. Liu and S. Jiang, “Fundamental Calculus on Generalized Stochastically Bounded Bursty Traffic for Communication Networks,” J. Computer Networks, vol. 53, pp. 2011-2021, Aug. 2009.