Tractable Effective Bandwidths for End-to-end Evaluation and Fractional Brownian Motion Traffic

Kazutomo Kobayashi
Graduate School of Engineering
Nagasaki University
Nagasaki 852-8521, Japan
kobayashi@cis.nagasaki-u.ac.jp

Yukio Takahashi
Tokyo Institute of Technology
Tokyo 152-8552, Japan

ABSTRACT

Effective bandwidth is a concept that has been developed for admission control as an indicator of the traffic load given to a network. However, that concept has been discussed mainly under a single node. That's because, while there are many studies on evaluation formulas for backlog at a single node, studies on end-to-end evaluation have been almost limited to ones by network calculus.

In this paper, we develop an end-to-end backlog evaluation formula for a heterogeneous tandem network with cross traffic, as an extension of the one obtained in previous papers by the authors. The formula evaluates the asymptotic tail probability of the end-to-end backlog by the total traffic load at a bottleneck node, and here the traffic load of a flow is evaluated by a kind of effective bandwidth named tractable effective bandwidths (tEBW). The tEBW has a special property that makes the formula simple and tractable, and it is applicable to various types of input flows frequently used in performance analyses.

Unfortunately, however, fractional Brownian motion (fBm) flows with long range dependency do not have any tEBW. For fBm flows, we show that another new tighter end-to-end backlog evaluation formula is effective, and using it we can show that, in a homogeneous tandem network with fBm cross traffic, the asymptotic tail probability of the end-to-end backlog is the same as that in the single node.

Categories and Subject Descriptors

G.3 [Probability and statistics]: Queueing theory; C.2 [Computer communication networks]: Miscellaneous; B.8.2 [Performance and reliability]: Performance analysis and design aids

General Terms

Theory, Performance

Keywords

Effective bandwidths, End-to-end performance evaluation, Stochastic network calculus, Fractional Brownian motion

1. INTRODUCTION

Effective bandwidth is a concept that has been developed for admission control as an index of the traffic load given to a network, which is more sophisticated than the average rate or the peak rate. In [9], for a flow with stationary increments, the effective bandwidth of the arrivals $A(t)$ during $(0, t]$ is defined as

$$a(\theta, t) = \frac{1}{\theta t} \log E[e^{\theta A(t)}], \quad \theta, t \in (0, \infty), \tag{1}$$

and for a single node with J types of arrival flows and a constant service rate C, a simple evaluation formula for admission control is given as follows: If

$$\sum_{j=1}^{J} N_j \alpha_j(\theta^*, t^*) \le C + \frac{B}{t^*} - \frac{\gamma}{\theta^* t^*},$$

then

$$P(Q(t) > B) \lesssim e^{-\gamma B},$$

where N_j, $\alpha_j(\theta, t)$, B, γ and $Q(t)$ are the number of flows of type j, the effective bandwidth of flows of type j, the buffer threshold, the decay rate and the backlog at time t, respectively, and the pair (θ^*, t^*) is determined through some procedure.

However, it is difficult to apply this formula to a network with many nodes since the effective bandwidth of the departure process of a flow from a node, which becomes the arrival process to the next node, is generally unknown, or very difficult to be known even when that of the arrival process to the node is known. We need to discuss effective bandwidths and evaluation formulas for admission control through studies on end-to-end evaluation for networks.

In [10], the authors proposed a stochastic network calculus for many flows using an approach similar to the large deviations techniques (see for the proper large deviations techniques, e.g., [5, 6]), and provided asymptotic end-to-end evaluation formulas. In [12], the authors applied this stochastic network calculus to a heterogeneous tandem network with many forwarding flows and cross traffic flows constrained by leaky buckets, and provided a simple end-to-end evaluation formula for admission control. This formula has a convenient nature that the tail probability of the end-to-end backlog can be evaluated from the stochastic characteristics of total traffic load at a bottleneck node only, and hence the evaluation is independent of the number of nodes.

VALUETOOLS 2013, December 10-12, Torino, Italy

DOI 10.4108/icst.valuetools.2013.254393

In this paper, then, we discuss whether such a simple and convenient end-to-end evaluation formula can be obtained or not for various traffic flows including fractional Brownian motion (fBm) flows.

To treat various traffic flows including fBm flows, we provide two end-to-end backlog evaluation formulas in this paper, using a slightly different approach from in [10], one of which is a little tighter than that in [10, 12] and the other is the same as in them.

Using the latter evaluation formula, we extend the simple end-to-end evaluation formula for constrained flows with leaky buckets in [12] to that for generalized traffic flows, and show that the tail probability of the end-to-end backlog can be evaluated by the total amount of effective bandwidths of a limited type at a bottleneck node. Here, the effective bandwidth of a limited type means that it is a continuous and strictly increasing function of θ but independent of time t (see (37)). We call it as a tractable effective bandwidths, shortly tEBW, since it is so tractable that the evaluation formula using tEBW is very simple and convenient.

If an effective bandwidth is bounded above by a tEBW, we say that it has a tEBW. We check whether each of the effective bandwidths discussed in [9] has a tEBW or not. The results in Section 3 show that fBm flows with long-range dependency don't have any tEBW while all other flows have. Hence, for most of flow types frequently used in performance evaluation, the asymptotic tail probability of the end-to-end backlog can be evaluated from the stochastic characteristics of total traffic load at a bottleneck node.

On the other hand, using the tighter evaluation formula, we evaluate the end-to-end backlog in a homogeneous tandem network with fBm cross traffic. Our result indicates that the asymptotic tail probability of the end-to-end backlog in the homogeneous tandem network is the same as that in a single node. So, the authors conjecture that, even in a heterogeneous tandem network, a similar property holds.

There are a number of studies on performance evaluation of fBm traffic [13, 6, 17], but most of them are just for a single node. Recently there appear some studies on end-to-end evaluation for a tandem network with fBm cross traffic, using stochastic network calculus [15, 18]. However, their performance bounds grow as the number of nodes increases, and they don't match the authors' conjecture stated above.

The remainder of the paper is constructed as follows. In Section 2, we discuss the stochastic network calculus with many flows using a slightly different approach. In Section 3, we obtain a simple evaluation formula for admission control. Then we define tractable effective bandwidths and check whether or not frequently used effective bandwidths have tEBW. In Section 4 we discuss a backlog evaluation for a network with fBm cross traffic, and in Section 5 we give some numerical results and discussions on tEBW.

2. STOCHASTIC NETWORK CALCULUS WITH MANY FLOWS

2.1 Preliminaries

We consider a discrete-time tandem network with m nodes and L forwarding flows, as illustrated in Figure 1. Here, we don't consider cross traffic flows. Time t takes discrete values $0, 1, 2, \ldots$. For time $t \geq 0$, let $A^L(t)$ and $S_i^L(t)$, $i = 1, 2, \ldots, m$, be random variables representing the total arrivals to the network and the total offered services at node i, respectively, during time interval $(0, t]$ for L flows, with a convention $A^L(0) = S_i^L(0) = 0$. Further, let $Q^L(t)$ be the total backlog of L flows in the whole network at time t. In this paper, an arrival means an arrival of one packet, and packets are assumed to be of the same size. So, all flows seem like discrete-time fluid flows.

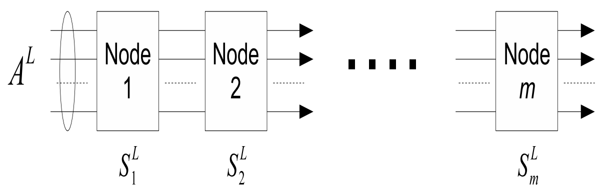

Figure 1: Tandem network with m nodes and L flows

For a pair of times s and t such that $0 \leq s \leq t$, we let $\overline{A}^L(s,t) = A^L(t) - A^L(s)$ and $\overline{S}_i^L(s,t) = S_i^L(t) - S_i^L(s)$. The bar indicates the operation of deriving a bi-variate function from a single-variable function.

We define a convolution operator $*$ and a deconvolution operator $\oslash$ for functions $f(s,t)$ and $g(s,t)$ of two variables s, t such that $0 \leq s \leq t$ as

$$f * g(s,t) = \min_{s \leq \tau \leq t} \{f(s,\tau) + g(\tau,t)\} \quad \text{and} \tag{2}$$

$$f \oslash g(s,t) = \max_{0 \leq \tau \leq s} \{f(\tau,t) - g(\tau,s)\}. \tag{3}$$

Then, as shown in the previous works [2, 8, 10], the backlog $Q^L(t)$ is written as

$$Q^L(t) = \overline{A}^L \oslash S^L(t,t) \tag{4}$$

with probability one, where

$$S^L(s,t) = \overline{S}_1^L * \overline{S}_2^L * \cdots * \overline{S}_m^L(s,t). \tag{5}$$

From the definitions (2) and (3), $Q^L(t)$ may be rewritten as

$$Q^L(t) = \max_{0 \leq s_0 \leq \cdots \leq s_m = t} U^L(s_0, \cdots, s_m), \tag{6}$$

where

$$U^L(s_0, \cdots, s_m) = \overline{A}^L(s_0, s_m) - \overline{S}^L(s_0, s_1) - \cdots - \overline{S}^L(s_{m-1}, s_m). \tag{7}$$

In the subsequent sections we discuss asymptotic tail probability of end-to-end backlog as $L \to \infty$. We have defined L as a parameter indicating the number of flows in the traffic. But, in our analysis, flows are treated as a whole collectively, and not treated individually. Hence, L is not necessary to be the number of flows. It is a mere parameter used for asymptotic analysis in our discussion.

2.2 Asymptotic tail probability of end-to-end backlog

Before discussing the tail probability of $Q^L(t)$, we will evaluate the tail probability of $U^L(s_0, \cdots, s_m)$ by using cumulant generating functions.

The cumulant generating function (cgf) of a random variable X^L is the logarithm of the moment generating function of X^L, and defined by

$$\mathcal{X}^L(\theta) = \log \mathrm{E}\left[e^{\theta X^L}\right] \tag{8}$$

for $\theta \in \mathbb{R}$. Further, for a sequence of random variables $\{X^L\}_{L=1,2,\dots}$, we let

$$\mathcal{X}(\theta) = \limsup_{L\to\infty} L^{-1}\, \mathcal{X}^L(\theta), \tag{9}$$

and call it the asymptotic cumulant generating function (asymptotic cgf) of $\{X^L\}$. We allow $\pm\infty$ for the values of the cgf and the asymptotic cgf [1]. On the treatment of $\pm\infty$ in mathematical operations we follow the usual convention. We note that, if X^L is nonnegative with probability one, $\mathcal{X}^L(\theta)$ and $\mathcal{X}(\theta)$ are nonnegative for $\theta > 0$ and nonpositive for $\theta < 0$.

For given $0 \le s_0 \le \cdots \le s_m$, we let $\mathcal{U}^L(\theta; s_0, \cdots, s_m)$ be the cgf of $U^L(s_0, \cdots, s_m)$ and $\mathcal{U}(\theta; s_0, \cdots, s_m)$ be the asymptotic cgf of $\{U^L(s_0, \cdots, s_m)\}$. We denote the abscissa of convergence of $\mathcal{U}(\theta; s_0, \cdots, s_m)$ by

$$\delta_U(s_0, \cdots, s_m) = \sup\left\{\theta \,|\, \mathcal{U}(\theta; s_0, \cdots, s_m) < +\infty\right\}, \tag{10}$$

and assume that $\delta_U(s_0, \cdots, s_m) > 0$ including the case $\delta_U(s_0, \cdots, s_m) = +\infty$. Then the asymptotic tail probability of $U^L(s_0, \cdots, s_m)$ is evaluated as follows.

Lemma 1. For given $0 \le s_0 \le \cdots \le s_m$ and any $b \ge 0$, we have

$$\limsup_{L\to\infty} L^{-1} \log \mathrm{P}(U^L(s_0, \cdots, s_m) > Lb) \le \inf_{\theta \in (0, \delta_U(s_0,\cdots,s_m))} \{-\theta b + \mathcal{U}(\theta; s_0, \cdots, s_m)\}. \tag{11}$$

Proof. We apply the Chernoff's bound (or the Markov inequality, see, e.g., p.240 of [2]) to the tail probability $\mathrm{P}(U^L(s_0, \cdots, s_m) > Lb)$. For any $\theta > 0$, we have

$$\begin{aligned}\mathrm{P}(U^L(s_0, \cdots, s_m) > Lb) &= \mathrm{P}(e^{\theta U^L(s_0,\cdots,s_m)} > e^{\theta L b}) \\ &\le e^{-\theta L b}\, \mathrm{E}[e^{\theta U^L(s_0,\cdots,s_m)}].\end{aligned}$$

Taking the logarithm, dividing by L, and then taking the limit superior on both sides, we have

$$\limsup_{L\to\infty} L^{-1} \log \mathrm{P}(U^L(s_0, \cdots, s_m) > Lb) \le -\theta b + \mathcal{U}(\theta; s_0, \cdots, s_m).$$

Since the parameter θ does not appear in the left-hand side, we may take the infimum on θ in the right-hand side. If $\theta > \delta_U(s_0, \cdots, s_m)$, $\mathcal{U}(\theta; s_0, \cdots, s_m) = +\infty$. Hence it is natural to restrict the range on which the infimum is taken to the interval $(0, \delta_U(s_0, \cdots, s_m))$ as in (11). □

Using Lemma 1, we have the following lemma in terms of the end-to-end backlog.

[1] The terminology "asymptotic cgf" is usually used when the sequence $\{L^{-1}\,\mathcal{X}^L(\theta)\}$ converges to a finite limit function. However, in this paper, for brevity of exposition, we use it for the function defined above even when the sequence does not converge.

Lemma 2. Let $\delta_Q(t) = \min_{0\le s_0 \le \cdots \le s_m = t} \delta_U(s_0, \cdots, s_m)$. Then we have

$$\limsup_{L\to\infty} L^{-1} \log \mathrm{P}(Q^L(t) > Lb) \le \max_{0\le s_0\le\cdots\le s_m = t} \inf_{\theta\in(0,\delta_Q(t))} \{-\theta b + \mathcal{U}(\theta; s_0, \cdots, s_m)\}. \tag{12}$$

Proof. Using the relation $\mathrm{P}(\max\{X, Y\} > x) \le \mathrm{P}(X > x) + \mathrm{P}(Y > x)$ for any x and random variables X and Y, and the inequality $\sum_{0\le s_0\le\cdots\le s_m = t} 1 \le \binom{t+m}{m}$, we have from(6)

$$\begin{aligned}&\mathrm{P}(Q^L(t) > Lb) \\ &\quad = \mathrm{P}\left(\max_{0\le s_0\le\cdots\le s_m=t} U^L(s_0, \cdots, s_m) > Lb\right) \\ &\quad \le \sum_{0\le s_0\le\cdots\le s_m=t} \mathrm{P}(U^L(s_0, \cdots, s_m) > Lb) \\ &\quad \le \binom{t+m}{m} \max_{0\le s_0\le\cdots\le s_m=t} \mathrm{P}(U^L(s_0, \cdots, s_m) > Lb).\end{aligned}$$

Taking the logarithm, dividing by L and then taking the limit superior on both sides, we have

$$\limsup_{L\to\infty} L^{-1}\log \mathrm{P}(Q^L(t) > Lb) \le \max_{0\le s_0\le\cdots\le s_m=t} \limsup_{L\to\infty} L^{-1}\log \mathrm{P}(U^L(s_0, \cdots, s_m) > Lb).$$

Here we use the monotonicity of the logarithmic function and the relations of the limit superiors

$$\lim_{n\to\infty} n^{-1}x = 0 \quad \text{for} \quad |x| < \infty, \qquad \text{and}$$

$$\limsup_{n\to\infty} \max\{a_n, b_n\} = \max\{\limsup_{n\to\infty} a_n, \limsup_{n\to\infty} b_n\}.$$

Applying Lemma 1 to the right hand of the above inequality and using the relation $(0, \delta_Q(t)) \subseteq (0, \delta_U(s_0, \cdots, s_m))$ for any $0 \le s_0 \le \cdots \le s_m = t$, we obtain (12). □

Now we shall express $\mathcal{U}(\theta; s_0, \cdots, s_m)$ in terms of the asymptotic cgf's of the arrivals and the offered services. We denote by $\overline{\mathcal{A}}(\theta; s, t)$ the asymptotic cgf of $\{\overline{A}^L(s, t)\}_{L=1,2,\dots}$ and by $\overline{\mathcal{S}}_i(\theta; s, t)$ the asymptotic cgf of $\{\overline{S}_i^L(s, t)\}_{L=1,2,\dots}$, for $i = 1, 2, \dots, n$.

For an arbitrarily fixed $t > 0$, we make the following assumptions.

A1. For any $0 \le s_0 \le s_1 \le \cdots \le s_m = t$, random variables $\overline{A}^L(s_0, t)$, $\overline{S}_1^L(s_0, s_1)$, $\cdots$, $\overline{S}_{m-1}^L(s_{m-2}, s_{m-1})$ and $\overline{S}_m^L(s_{m-1}, s_m)$ are mutually independent.

A2. There exists $\delta_{\mathrm{A}}(t) > 0$, allowing $\delta_{\mathrm{A}}(t) = +\infty$, such that $\overline{\mathcal{A}}(\theta; 0, t)$ is finite for $\theta < \delta_{\mathrm{A}}(t)$.

Under the above assumptions, we get the following lemma.

Lemma 3. For a given $t > 0$ and $\theta \in (0, \delta_A(t))$, under A1 and A2 we have

$$\begin{aligned}\mathcal{U}(\theta; s_0, \cdots, s_m) &= \overline{\mathcal{A}}(\theta; s_0, s_m) \\ &+ \overline{\mathcal{S}}_1(-\theta; s_0, s_1) + \cdots + \overline{\mathcal{S}}_m(-\theta; s_{m-1}, s_m),\end{aligned} \tag{13}$$

and $\delta_A(t) \le \delta_Q(t)$.

Proof. Since the random function $U^L(s_0,\cdots,s_m)$ is written as in (7), its cgf is given as

$$\begin{aligned}&\mathrm{E}[e^{\theta U^L(s_0,\cdots,s_m)}]\\&\quad = \mathrm{E}[e^{\theta(\overline{A}^L(s_0,s_m)-\overline{S}^L(s_0,s_1)-\cdots-\overline{S}^L(s_{m-1},s_m))}].\end{aligned}$$

From A1, the right-hand side above can be written in the product form

$$\mathrm{E}[e^{\theta\overline{A}^L(s_0,t)}]\,\mathrm{E}[e^{-\theta\overline{S}_1^L(s_0,s_1)}]\cdots\mathrm{E}[e^{-\theta\overline{S}_m^L(s_{m-1},s_m))}].$$

Taking the logarithm, dividing by L, and taking the limit superior on both sides, we have (13).

Since $\overline{S}_i^L(s_{i-1},s_i)$ is positive, $\overline{\mathcal{S}}_i(-\theta;s_{i-1},s_i)$ is nonpositive. Thus the right-hand side of (13) is less than or equal to $\overline{\mathcal{A}}(\theta;0,t)$ and it is finite or equal to $-\infty$ for $\theta\in(0,\delta_{\mathrm{A}}(t))$ from A2. Hence $\delta_{\mathrm{A}}(t)\le\delta_U(s_0,\cdots,s_m)$ for any $0\le s_0\le\cdots\le s_m=t$ and then $\delta_{\mathrm{A}}(t)\le\delta_Q(t)$. □

Applying Lemmas 2 and 3, we have the following theorem.

Theorem 1. For given $t>0$ and $\theta\in(0,\delta_A(t))$, under A1 and A2 we have

$$\begin{aligned}&\limsup_{L\to\infty}L^{-1}\log\mathrm{P}(Q^L(t)>Lb)\\&\le\max_{0\le s_0\le\cdots\le s_m=t}\inf_{\theta\in(0,\delta_A(t))}\{-\theta b+\overline{\mathcal{A}}(\theta;s_0,s_m)\\&\qquad+\overline{\mathcal{S}}_1(-\theta;s_0,s_1)+\cdots+\overline{\mathcal{S}}_m(-\theta;s_{m-1},s_m)\}.\quad(14)\end{aligned}$$

In the previous paper [10], a slightly different evaluation formula was presented as in the corollary below. In these formulas (14) and (15), the order of "max" and "inf" is reverse. Generally $\sup_x\inf_y f(x,y)\le\inf_y\sup_x f(x,y)$, and so the formula (14) is tighter than (15). From the same reason, the formula (15) can be considered as a direct consequence of (14).

Corollary 1. For given $t>0$ and $\theta\in(0,\delta_A(t))$, under A1 and A2 we have

$$\begin{aligned}&\limsup_{L\to\infty}L^{-1}\log\mathrm{P}(Q^L(t)>Lb)\\&\le\inf_{\theta\in(0,\delta_A(t))}\{-\theta b+\max_{0\le s_0\le\cdots\le s_m=t}\{\overline{\mathcal{A}}(\theta;s_0,s_m)\\&\qquad+\overline{\mathcal{S}}_1(-\theta;s_0,s_1)+\cdots+\overline{\mathcal{S}}_m(-\theta;s_{m-1},s_m)\}\}.\end{aligned}\quad(15)$$

If we use the convolution and deconvolution operators, the right hand side of (15) is more compactly rewritten as

$$\inf_{\theta\in(0,\delta_A(t))}\{-\theta b+\overline{\mathcal{A}}^\theta\oslash(-\mathcal{S}^{-\theta})(t,t)\}\quad(16)$$

with

$$-\mathcal{S}^{-\theta}(s,t)=(-\overline{\mathcal{S}}_1^{-\theta})*(-\overline{\mathcal{S}}_2^{-\theta})*\cdots*(-\overline{\mathcal{S}}_m^{-\theta}(s,t)).\quad(17)$$

Here, in order to apply the convolution and/or deconvolution operators to asymptotic cgfs, we write as $\overline{\mathcal{A}}^\theta(s,t)$ instead of $\overline{\mathcal{A}}(\theta;s,t)$ and as $-\overline{\mathcal{S}}_i^{-\theta}(s,t)$ instead of $-\overline{\mathcal{S}}_i(-\theta;s,t)$.

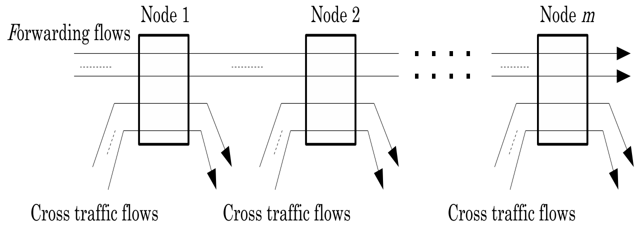

Figure 2: Tandem network with cross traffic

2.3 Tandem network with cross traffic

Now we consider a tandem network with cross traffic depicted in Figure 2. We let the link capacity at node i be constant in time and equal to Lc_i, that is, c_i is the link capacity per forwarding flow at node i, and the arrivals of the cross traffic at node i during $(s,t]$ be $A_i^{L,\mathrm{cross}}(s,t)$. We assume[2]that the offered services $S_i^L(s,t)$ for the forwarding traffic at node i during $(s,t]$ is given by the "leftover" bandwidth

$$S_i^L(s,t)=[Lc_i(t-s)-A_i^{L,\mathrm{cross}}(s,t)]^+\quad(18)$$

where $[X]^+=\max\{0,X\}$. We also assume that the cross traffic at one node is independent of the others as well as of the forward traffic.

Since $\mathrm{E}[\min\{X,Y\}]\le\min\{\mathrm{E}[X],\mathrm{E}[Y]\}$, we have

$$\begin{aligned}\mathrm{E}\left[e^{-\theta S_i^L(s,t)}\right]&=\mathrm{E}\left[\min\left\{1,e^{-\theta\left\{Lc_i(t-s)-A_i^{L,\mathrm{cross}}(s,t)\right\}}\right\}\right]\\&\le\min\left\{1,\mathrm{E}\left[e^{-\theta\left\{Lc_i(t-s)-A_i^{L,\mathrm{cross}}(s,t)\right\}}\right]\right\}.\end{aligned}$$

Taking the logarithm, dividing by L, and then taking the limit superior on both sides, we have

$$\overline{\mathcal{S}}_i(-\theta;s,t)\le-\left[c_i\theta(t-s)-\overline{\mathcal{A}}_1^{\mathrm{cross}}(\theta;s,t)\right]^+,\quad(19)$$

where $\overline{\mathcal{A}}_i^{\mathrm{cross}}(\theta;s,t)$ is the asymptotic cgf of $A_i^{L,\mathrm{cross}}(s,t)$. Using the inequality (19), the right hand side of (14) and that of (15) can be evaluated as

$$\begin{aligned}&\max_{0\le s_0\le\cdots\le s_m=t}\inf_{\theta\in(0,\delta_A(t))}\Big\{-\theta b+\overline{\mathcal{A}}(\theta;s_0,s_m)\\&\quad-\left[c_1\theta(s_1-s_0)-\overline{\mathcal{A}}_1^{\mathrm{cross}}(\theta;s_0,s_1)\right]^+-\cdots\\&\quad-\left[c_m\theta(s_m-s_{m-1})-\overline{\mathcal{A}}_m^{\mathrm{cross}}(\theta;s_{m-1},s_m)\right]^+\Big\},\end{aligned}\quad(20)$$

and

$$\begin{aligned}&\inf_{\theta\in(0,\delta_A(t))}\Big\{-\theta b+\max_{0\le s_0\le\cdots\le s_m=t}\Big\{\overline{\mathcal{A}}(\theta;s_0,s_m)\\&\quad-\left[c_1\theta(s_1-s_0)-\overline{\mathcal{A}}_1^{\mathrm{cross}}(\theta;s_0,s_1)\right]^+-\cdots\\&\quad-\left[c_m\theta(s_m-s_{m-1})-\overline{\mathcal{A}}_m^{\mathrm{cross}}(\theta;s_{m-1},s_m)\right]^+\Big\}\Big\},\end{aligned}\quad(21)$$

[2]In the previous paper [10], it was shown that a tandem network with cross traffic, in which the cross traffic is served with higher priority than the forwarding traffic, can be evaluated by a corresponding tandem network without cross traffic dealt with in the preceding subsections, in which the offered service for the forwarding traffic is given by the leftover bandwidth above.

respectively. We have obtained two evaluation formulas (20) and (21), and they are just different in the operation order of "max" and "inf". The formula (20) is tighter than (21), but we can handle (21) more easily. In the next section, using (21) we will obtain a simple evaluation formula for admission control and discuss tractable effective bandwidths.

3. TRACTABLE EFFECTIVE BANDWIDTHS

3.1 Simple evaluation formula

In this subsection, we extend the simple end-to-end evaluation formula for constrained flows with leaky buckets in [12] to that for generalized traffic flows. This extension is not difficult since basic ideas are the same as in the previous paper.

We consider a tandem network with cross traffic as discussed in the previous section. In addition, we assume that the forwarding traffic and the cross traffic consist of some of J types of flows. The types of flows are labeled as $j = 1, 2, \ldots, J$. Let the number of forwarding flows of type j be $L\alpha_j$ and the number of cross traffic flows of type j at node i be $L\beta_{ij}$. If there exist no forwarding flows of type j or no cross traffic flows of type j at node i, then we consider $\alpha_j = 0$ or $\beta_{ij} = 0$. However, for brevity of discussion, without loss of generality, we assume $\alpha_j + \sum_{i=1}^{m} \beta_{ij} > 0$ for every j. The total number of forwarding flows is equal to $L\sum_{j=1}^{J} \alpha_j$, and hence $\sum_{j=1}^{J} \alpha_j = 1$. The total number of cross traffic flows at node i is given by $L\sum_{j=1}^{J} \beta_{ij}$. The link capacity at node i is constant in time and equal to Lc_i, as stated before. When we move L later, α_j's, β_{ij}'s, and c_i's are kept constant.

For individual flows, we make the following assumptions:

C1. All flows (both forwarding flows and cross traffic flows) are mutually independent.

C2. Each flow has stationary increments.

C3. Flows of type j are subjecting to a common probabilistic law. Let $A_j(t)$ be the random variable representing the total arrivals to the network during time interval $(0, t]$ of a typical flow of type j, with a convention $A_j(0) = 0$.

C4. The cgf $\mathcal{A}_j(\theta; t) = \log E[e^{\theta A_j(t)}]$ of $A_j(t)$ is bounded from above as

$$\mathcal{A}_j(\theta; t) \le \theta t \cdot \phi_j(\theta) \quad \text{for } \theta \in (0, \infty), \tag{22}$$

where $\phi_j(\theta)$ is a function such that it is continuous and strictly increasing on the interval $(0, \delta_j)$ for some $\delta_j > 0$ (including the case $\delta_j = \infty$) and that $\phi_j(\theta) = \infty$ for $\theta \ge \delta_j$ if $\delta_j < \infty$. We write $\phi_j(0) = \lim_{\theta \downarrow 0} \phi_j(\theta)$. (This function $\phi_j(\theta)$ will be called a tractable effective bandwidth, later.)

We let $\overline{A}_j(s, t) = A_j(t) - A_j(s)$. Under the assumptions C2 and C4, the cgf $\overline{\mathcal{A}}_j(\theta; s, t) = \log E[e^{\theta \overline{A}_j(s,t)}]$ of $\overline{A}_j(s, t)$ can be evaluated as

$$\overline{\mathcal{A}}_j(\theta; s, t) = \overline{\mathcal{A}}_j(\theta; 0, t - s) \le \theta(t - s) \cdot \phi_j(\theta). \tag{23}$$

Denoting the kth flow of type j of the forwarding traffic as $A_{j,k}(t)$, the increment of arrivals in the forwarding traffic during $(s, t]$ is given by

$$\overline{A}^L(s, t) = \sum_{j=1}^{J} \sum_{k=1}^{L\alpha_j} \{A_{j,k}(t) - A_{j,k}(s)\},$$

and under the assumptions C1 and C3, its cgf is given by

$$\sum_{j=1}^{J} L\alpha_j \overline{\mathcal{A}}_j(\theta; s, t).$$

Since $\overline{\mathcal{A}}_j(\theta; s, t)$ is evaluated as in (23), the asymptotic cgf of the forwarding flows is evaluated as

$$\overline{\mathcal{A}}(\theta; s, t) \le \theta(t - s) \sum_{j=1}^{J} \alpha_j \phi_j(\theta). \tag{24}$$

Similarly, denoting the kth flow of type j of the cross traffic at node i as $A_{i,j,k}^{\text{cross}}(t)$, the increment of arrivals in the cross traffic at node i during $(s, t]$ is given by

$$\overline{A}_i^{L,\text{cross}}(s, t) = \sum_{j=1}^{J} \sum_{k=1}^{L\beta_{ij}} \{A_{i,j,k}^{\text{cross}}(t) - A_{i,j,k}^{\text{cross}}(s)\},$$

and under the assumptions C1 and C3, its cgf is given by

$$\sum_{j=1}^{J} L\beta_{ij} \overline{\mathcal{A}}_j(\theta; s, t).$$

Then the asymptotic cgf's of the cross traffic at node i is evaluated as

$$\overline{\mathcal{A}}_i^{\text{cross}}(\theta; s, t) \le \theta(t - s) \sum_{j=1}^{J} \beta_{ij} \phi_j(\theta). \tag{25}$$

Now, we put

$$\xi(\theta) = \sum_{j=1}^{J} \alpha_j \phi_j(\theta) \quad \text{and} \tag{26}$$

$$\psi_i(\theta) = \sum_{j=1}^{J} \beta_{ij} \phi_j(\theta), \quad i = 1, 2, \ldots, m. \tag{27}$$

From (25) and (27), we have

$$\left[c_i\theta(s_i - s_{i-1}) - \overline{\mathcal{A}}_i^{\text{cross}}(\theta; s_{i-1}, s_i)\right]^+ \ge \left[c_i\theta(s_i - s_{i-1}) - \theta(s_i - s_{i-1})\psi_i(\theta)\right]^+,$$

and (21) can be reevaluated from above by using $\xi(\theta)$ and $\psi_i(\theta)$'s as

$$\inf_{\theta \in (0,\delta)} \Big\{ -\theta b + \max_{0 \le s_0 \le \cdots \le s_m = t} \Big\{\theta(t - s_0) \cdot \xi(\theta) - \theta(s_1 - s_0) \cdot [c_1 - \psi_1(\theta)]^+ - \cdots - \theta(s_m - s_{m-1}) \cdot [c_m - \psi_m(\theta)]^+\Big\} \Big\}, \tag{28}$$

where $\delta = \min_{1 \le j \le J}\{\delta_j\}$. If all δ_j's are infinity, then δ is infinity.

Then we have the following theorem.

Theorem 2. Assume that, for $i = 1, 2, \ldots, m$,

$$\xi(0) + \psi_i(0) = \sum_{j=1}^{J} (\alpha_j + \beta_{ij})\phi_j(0) < c_i. \quad (29)$$

Then, for any $b > 0$, we have

$$\limsup_{L\to\infty} L^{-1} \log P(Q^L(t) > Lb) \le -\theta^* b, \quad (30)$$

where

$$\theta^* = \sup\left\{\theta \in (0,\delta) \;\middle|\; \xi(\theta) + \psi_i(\theta) \le c_i,\ i = 1, 2, \ldots, m\right\}. \quad (31)$$

Proof. From C4, the functions $\phi_j(\theta)$, $\xi(\theta)$ and $\psi_i(\theta)$ are all continuous and strictly increasing. So, under the condition (29), there exists some $\theta \in (0, \delta)$ that satisfies the inequality in the conditional part of (31). Hence θ^* is well defined and $\theta^* > 0$.

We denote the quantity in the outermost braces of (28) as $W(\theta)$ and rewrite it as

$$W(\theta) = -\theta b + \max_{0\le s_0\le\cdots\le s_m=t} \Big\{ \theta(s_1 - s_0) \cdot \min\{\xi(\theta) + \psi_1(\theta) - c_1,\ \xi(\theta)\} + \cdots + \theta(s_m - s_{m-1}) \cdot \min\{\xi(\theta) + \psi_m(\theta) - c_m,\ \xi(\theta)\} \Big\}. \quad (32)$$

Consider the case $0 < \theta < \theta^*$. Then for any i, $\xi(\theta) + \psi_i(\theta) - c_i$ is less than or equal to 0, and hence

$$\min\{\xi(\theta) + \psi_i(\theta) - c_i,\ \xi(\theta)\} \le 0.$$

So the maximum in (32) is attained at $s_0 = s_1 = \cdots = s_m = t$, and the maximum value is equal to zero. Hence in this case $W(\theta) = -\theta b$. Then (28) is evaluated from above as

$$\inf_{\theta\in(0,\delta)} W(\theta) \le \inf_{\theta\in(0,\delta^*)} W(\theta) = \inf_{\theta\in(0,\delta^*)} \{-\theta b\} = -\theta^* b.$$

This proves (30). □

Suppose that θ^* defined in the theorem satisfies $\xi(\theta^*) + \psi_i(\theta^*) = c_i$ for $i = i_0$. Then the node i_0 is a bottleneck node of the network. The theorem says that the asymptotic tail probability of the end-to-end backlog is evaluated by the total traffic load of both the forwarding flows and the cross ones at the bottleneck node.

3.2 Tractable effective bandwidth and its examples

Theorem 1 in the preceding section implies that, if L is large enough, $L^{-1} \log P(Q^L(t) > Lb)$ is less than or equal to $-\theta b$ for any positive $\theta \le \theta^*$. We will write this as

$$P(Q^L(t) > Lb) \lesssim e^{-L\theta b}. \quad (33)$$

For an application to an admission control, it is convenient to rewrite this result without using L. Let us denote the backlog threshold as $B = Lb$, the link capacity at node i as $C_i = Lc_i$, the number of forwarding flows of type j as $N_j = L\alpha_j$ and the number of cross traffic flows of type j at node i as $N_{ij}^{\text{cross}} = L\beta_{ij}$. Then, for a given tail probability bound ϵ, we obtain the following backlog evaluation formula for admission control.

A formula for admission control: If N_j's and N_{ij}^{cross}'s are large enough (except for flows such that $N_j = 0$ and $N_{ij}^{\text{cross}} = 0$) and if

$$\hat{\theta} = -\frac{\log \epsilon}{B} \quad (34)$$

satisfies the inequality

$$\sum_{j=1}^{J} (N_j + N_{ij}^{\text{cross}})\phi_j(\hat{\theta}) \le C_i \quad (35)$$

for any $i = 1, 2, \ldots, m$, then

$$P(Q^L(t) > B) \lesssim \epsilon. \quad (36)$$

From (35), we see that $\phi_j(\theta)$ is a kind of effective bandwidth of a flow of type j. Comparing $\phi_j(\theta)$ with the effective bandwidth (1) discussed in [9], the assumption C4 in the previous section requests that $a(\theta, t)$ is bounded from above by a function $\phi(\theta)$ as

$$a(\theta, t) \le \phi(\theta) \quad \text{for } \theta, t \in (0, \infty), \quad (37)$$

where $\phi(\theta)$ is continuous and strictly increasing in the interval $(0, \delta)$ such that $\phi(\theta) = \infty$ for $\theta \ge \delta$.

To distinguish $\phi(\theta)$ from $a(\theta, t)$, we call it a tractable effective bandwidth (tEBW), and we say $a(\theta, t)$ has a tEBW $\phi(\theta)$ if there exist a function $\phi(\theta)$ satisfying (37).

In the subsequent subsubsections, we check whether the effective bandwidths discussed in [9] have tEBW's or not. From this section, we consider time t is continuous as in [9], though we have assumed it is discrete up to the previous section.

3.2.1 *Periodic flows*

Consider a flow with arrivals σ at times $\{Ud + nd, n = 0, 1, \ldots\}$, where U is a random variable distributed uniformly on $[0, 1]$. The effective bandwidth of the flow is given by

$$a(\theta, t) = \frac{\sigma}{t}\left\lfloor \frac{t}{d} \right\rfloor + \frac{1}{\theta t} \log\left[1 + \left(\frac{t}{d} - \left\lfloor \frac{t}{d} \right\rfloor\right)(e^{\sigma\theta} - 1)\right]. \quad (38)$$

For an arbitrarily fixed $\theta > 0$, the function $a(\theta, t)$ has the limit

$$\lim_{t\to 0} a(\theta, t) = \frac{e^{\sigma\theta} - 1}{\theta d}, \quad (39)$$

as $t \downarrow 0$, and it takes a common value $a(\theta, d) = a(\theta, 2d) = \cdots = \sigma/d \le (e^{\sigma\theta} - 1)/\theta d$ at $t = d, 2d, \cdots$. In the interval $[nd, (n+1)d)$, $n = 0, 1, \ldots$, it is monotonically decreasing.

Since $(e^{\sigma\theta} - 1)/\theta d$ is continuous and strictly increasing in the interval $(0, \infty)$, we see that $a(\theta, t)$ has tEBW $\phi(\theta) = (e^{\sigma\theta} - 1)/\theta d$.

3.2.2 *Fluid flows with two-state Markov chains*

Consider a stationary fluid flow with an on-off type two-state Markov chain. The transition rate from off-state to on-state is λ and the opposite transition rate is μ. The arrival rate is h in the on-state. Then, the effective bandwidth is given by

$$a(\theta, t) = \frac{1}{\theta t} \times \log\left\{\begin{pmatrix} \frac{\lambda}{\lambda+\mu} & \frac{\mu}{\lambda+\mu} \end{pmatrix} \exp\left\{\begin{pmatrix} -\mu + h\theta & \mu \\ \lambda & -\lambda \end{pmatrix} t\right\} \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\right\}.$$

For an arbitrarily fixed $\theta > 0$, $a(\theta, t)$ is an increasing function of t and

$$\lim_{t\to\infty} a(\theta,t) = \frac{1}{2\theta}\left(h\theta - \mu - \lambda + \sqrt{(h\theta-\mu+\lambda)^2 + 4\lambda\mu}\right). \tag{40}$$

The right hand side of the above equation is a continuous and strictly increasing function on the interval $(0, \infty)$. Therefore, $a(\theta, t)$ has tEBW $\phi(\theta)$ given by the right hand side of (40).

3.2.3 Brownian motion flows

Consider an arrival process defined by

$$A(t) = \lambda t + Z(t) \tag{41}$$

where $Z(t)$ is a Brownian motion with variance σ^2. Then the effective bandwidth is given by

$$a(\theta,t) = \lambda + \frac{\sigma^2\theta}{2}. \tag{42}$$

In this case, $a(\theta, t)$ itself is a tEBW.

3.2.4 Compound Poisson flows

Consider an arrival process defined by

$$A(t) = \sum_{n=1}^{N(t)} Y_n$$

where $Y_1, Y_2, \ldots$ are mutually independent and identically distributed random variables subjecting to a common distribution F, and $N(t)$ is an independent Poisson process of rate ν. Then the effective bandwidth is given by

$$a(\theta,t) = \frac{1}{\theta}\int (e^{\theta x} - 1)\nu dF(x). \tag{43}$$

Thus $a(\theta, t)$ itself is a tEBW.

If $Y_1, Y_2, \ldots$ are exponentially distributed with parameter μ, then

$$a(\theta,t) = \frac{\nu}{\mu-\theta} \quad \text{for } \theta \in (0,\mu). \tag{44}$$

If $Y_1, Y_2, \ldots$ are deterministic random variables with the constant value σ, then

$$a(\theta,t) = \frac{\nu}{\theta}(e^{\theta\sigma} - 1) \quad \text{for } \theta \in (0,\infty). \tag{45}$$

3.2.5 Flows shaped with leaky buckets

Consider an arrival process $A(t)$ which is limited by a leaky bucket with token rate ρ and token bucket size σ. Namely, $A(t) - A(s) \le \rho(t - s) + \sigma$ holds with probability one for any $0 \le s \le t$. Then the effective bandwidth is evaluated as

$$a(\theta,t) \le \frac{1}{\theta t}\log\left[1 + \frac{\lambda t}{\rho t+\sigma}\left(e^{\theta(\rho t+\sigma)} - 1\right)\right] \tag{46}$$

where λ is the mean rate of $A(t)$, namely, $\lambda = a(0, t) = E[A(t)]/t$, and $\lambda \le \rho$. Let $\hat{a}_\lambda(\theta, t)$ be the right hand side of (46) and $\hat{a}_\rho(\theta, t)$ be that in which λ is replaced with ρ. Then $a(\theta, t) \le \hat{a}_\lambda(\theta, t) \le \hat{a}_\rho(\theta, t)$ and $\hat{a}_\rho(\theta, t)$ decreases as t increases. Since

$$\lim_{t\to 0} \hat{a}_\rho(\theta,t) = \frac{\rho}{\theta\sigma}\left(e^{\theta\sigma} - 1\right), \tag{47}$$

$a(\theta, t)$ has the tEBW $\phi(\theta)$ given by (47), which is the same as (45) with $\nu = \rho/\sigma$.

3.2.6 Fractional Brownian motion flows with long range dependency

Consider an arrival process given by (41) with a fractional Brownian motion $Z(t)$ having Hurst parameter $H \in (0.5, 1)$, which is a long range dependent process. The effective bandwidth is given by

$$a(\theta,t) = \lambda + \frac{\sigma^2\theta}{2}t^{2H-1}. \tag{48}$$

Since t^{2H-1} is an increasing function and $\lim_{t\to\infty} t^{2H-1} = \infty$, $a(\theta, t)$ doesn't have any tEBW.

4. EVALUATION FOR FBM TRAFFIC

We know that fractional Brownian motion flows with long range dependency have no tEBW. In this section, then, we consider a network with arrival processes of fractional Brownian motion.

4.1 Single node

To check which formula between (20) and (21) can be applied to a network with fractional Brownian motion arrivals, we first consider a single node without cross traffic. We let the arrivals to the node have the mean rate λ and variance σ^2 per a flow with Hurst parameter $H \in (0.5, 1)$, and the single node have the constant bit rate services with c per a flow. We assume that $\lambda - c < 0$ as a stable condition. We also assume that time t is enough large and s can be treated as non-negative continuous variables, though it was originally non-negative discrete time.

The asymptotic cgf of the arrivals is given by

$$\overline{\mathcal{A}}(\theta; s,t) = \lambda\theta(t-s) + \frac{\sigma^2\theta^2}{2}(t-s)^{2H}. \tag{49}$$

Then, (21) becomes to

$$\begin{aligned}
&\inf_{\theta\in(0,\infty)} \max_{0\le s\le t}\left\{ -\theta b + (\lambda - c)\theta(t-s) + \frac{\sigma^2\theta^2}{2}(t-s)^{2H}\right\} \\
&= \inf_{\theta\in(0,\infty)}\left\{ -\theta b + (\lambda-c)\theta t + \frac{\sigma^2\theta^2}{2}t^{2H}\right\} \\
&= -\frac{(b+(c-\lambda)t))^2}{2\sigma^2 t^{2H}}.
\end{aligned}$$

We see that this infimum increases as t increases and converges to 0 as $t \to \infty$. So this evaluation doesn't seem useful.

The other formula (20) is given by

$$\begin{aligned}
&\max_{0\le s\le t}\inf_{\theta\in(0,\infty)} \left\{ -\theta b + (\lambda - c)\theta(t-s) + \frac{\sigma^2\theta^2}{2}(t-s)^{2H}\right\} \\
&= \max_{0\le s\le t}\left\{ -\frac{(b+(c-\lambda)(t-s))^2}{2\sigma^2(t-s)^{2H}}\right\} \\
&= -\frac{1}{2\sigma^2}\left(\frac{c-\lambda}{H}\right)^{2H}\left(\frac{b}{1-H}\right)^{2(1-H)}.
\end{aligned} \tag{50}$$

This is a well-known result.

From these results, we see that we have to use the formula (20) for networks with fractional Brownian motion flows.

4.2 Homogeneous tandem network

Here we consider a homogeneous tandem network with cross traffic flows of fractional Brownian motion. Let the number of forwarding flows be $L\alpha$, the number of cross traffic flows $L\beta$ and the link capacity Lc. To make analyses easier, we let all flows of the cross traffic be subjecting to a common probabilistic law of fractional Brownian motion with Hurst parameter $H \in (0.5, 1)$ and variance σ^2, and let the forwarding flows be subjecting to constant bit rate. All flows of both forwarding and cross traffic flows have the same mean rate λ.

The cumulant generating function of a forwarding flow is given by

$$\mathcal{A}_f(\theta; t) = \lambda\theta t, \tag{51}$$

and that of a cross traffic flow is given by

$$\mathcal{A}_c(\theta; t) = \lambda\theta t + \frac{\sigma^2\theta^2}{2} t^{2H}. \tag{52}$$

Then (20) is evaluated as

$$\max_{0\le s_0\le\cdots\le s_m=t} \inf_{\theta\in(0,\infty)} G(\theta; s_0, \cdots, s_m) \left(= G^{\max\inf} \right) \tag{53}$$

where

$$\begin{aligned} G(\theta; s_0, \cdots, s_m) &= -\theta b + \alpha\lambda\theta(s_m - s_0) \\ &- \left[(c-\beta\lambda)\theta(s_1 - s_0) - \frac{\beta\sigma^2\theta^2}{2}(s_1 - s_0)^{2H} \right]^+ - \cdots \\ &- \left[(c-\beta\lambda)\theta(s_m - s_{m-1}) - \frac{\beta\sigma^2\theta^2}{2}(s_m - s_{m-1})^{2H} \right]^+ . \end{aligned} \tag{54}$$

As indicated in the parenthes of (53), we denote the max inf of $G(\theta; s_0, \cdots, s_m)$ as $G^{\max\inf}$.

We assume that $(\alpha+\beta)\lambda - c < 0$ for the stability of the system. We also assume that time t is enough large and $s_0, s_1, \cdots, s_m$ can be treated as non-negative continuous variables. Then we have the following theorem.

Theorem 3. $G^{\max\inf}$ of (53) is given as follows:

(i) For $H \ge \dfrac{c-(\alpha+\beta)\lambda}{c-\beta\lambda}$,

$$G^{\max\ \inf} = -\frac{1}{2\beta\sigma^2}\left(\frac{c-(\alpha+\beta)\lambda}{H}\right)^{2H}\left(\frac{b}{1-H}\right)^{2(1-H)}. \tag{55}$$

(ii) For $H \le \dfrac{c-(\alpha+\beta)\lambda}{c-\beta\lambda}$,

$$G^{\max\ \inf} = -\frac{(c-\beta\lambda)^2}{2\beta\sigma^2}\left(\frac{b}{\alpha\lambda}\right)^{2(1-H)}. \tag{56}$$

This quantity is smaller than or equal to the right hand side of (55).

The proof of the theorem is so complicated that we need a lot of space to write down it. The authors will present it elsewhere.

Notice that the results (55) and (56) are independent of m, the number of nodes, and that $\inf_{\theta\in(0,\infty)} G(\theta; s_0, \cdots, s_m)$ can attain the maximum at some of the possible combinations $(s_0, \cdots, s_m)$. One example is that $s_0 = t - Hb/(1-H)(c-(\alpha+\beta)\lambda)$, $s_2 = \cdots = s_m = t$ for (i) and $s_0 = t - b/\alpha\lambda$, $s_2 = \cdots = s_m = t$ for (ii). This implies that $G^{\max\ \inf}$ is attained by the case where one of the nodes is dominant and the others can be ignored in the sense that $G(\theta; s_0, \cdots, s_m)$ is reduced to

$$-\theta b + \alpha\lambda\theta(t-s_0) - \left[(c-\beta\lambda)\theta(t-s_0) - \frac{\beta\sigma^2\theta^2}{2}(t-s_0)^{2H} \right]^+ .$$

This is the G function of a single node. Thus Theorem 3 asserts that our homogeneous tandem network with m nodes has the same asymptotic tail probability of the end-to-end backlog as that the single node has. The authors conjecture that, even in a heterogeneous tandem network, the asymptotic tail probability of the end-to-end backlog is determined by the total traffic load at a bottleneck node.

Finally, for reference purpose, in stead of (54), we consider the case that

$$\begin{aligned} \hat{G}(\theta; s_0, \cdots, s_m) &= -\theta b + \alpha\lambda\theta(s_m - s_0) \\ &- \left\{ (c-\beta\lambda)\theta(s_1 - s_0) - \frac{\beta\sigma^2\theta^2}{2}(s_1 - s_0)^{2H} \right\} - \cdots \\ &- \left\{ (c-\beta\lambda)\theta(s_m - s_{m-1}) - \frac{\beta\sigma^2\theta^2}{2}(s_m - s_{m-1})^{2H} \right\}. \end{aligned}$$

Here, $[x]^+$ in (54) is replaced with x. Then the equation is rewritten as

$$\begin{aligned} \hat{G}(\theta; s_0, \cdots, s_m) &= -\theta b + ((\alpha+\beta)\lambda - c)\,\theta(s_m - s_0) \\ &\quad + \frac{\beta\sigma^2\theta^2}{2}\sum_{i=1}^{m}(s_i - s_{i-1})^{2H}. \end{aligned}$$

This is a parabola on θ and $\hat{G}$ attains the infimum at θ^* such that $\partial\hat{G}/\partial\theta = 0$. Then we have

$$\begin{aligned} &\inf_{\theta\in(0,\infty)} \hat{G}(\theta; s_0, \cdots, s_m) \\ &\quad = -\frac{\{b + (c-(\alpha+\beta)\lambda)(s_m - s_0)\}^2}{2\beta\sigma^2\sum_{i=1}^m (s_i - s_{i-1})^{2H}}. \end{aligned}$$

As is easily known from the convexity of x^{2H}, this function attains the maximum at a combination $(s_0, s_1, \cdots .s_m)$ such that $s_0 < s_1 = \cdots = s_m = t$. Thus, we have

$$\begin{aligned} &\max_{0\le s_0\le\cdots\le s_m=t} \inf_{\theta\in(0,\infty)} \hat{G}(\theta; s_0, \cdots, s_m) \\ &= \max_{0\le s_0\le\cdots\le s_m=t} -\frac{\{b + (c-(\alpha+\beta)\lambda)(s_m - s_0)\}^2}{2\beta\sigma^2\sum_{i=1}^m (s_i - s_{i-1})^{2H}} \\ &= \max_{0\le s_0\le t} -\frac{\{b + (c-(\alpha+\beta)\lambda)(t - s_0)\}^2}{2\beta\sigma^2 (t - s_0)^{2H}} \\ &= -\frac{1}{2\beta\sigma^2}\left(\frac{c-(\alpha+\beta)\lambda}{H}\right)^{2H}\left(\frac{b}{1-H}\right)^{2(1-H)}. \end{aligned}$$

This is the same as the one in (55). The authors think that the influence of the existence of function $[x]^+$ appears in the case (ii) of the theorem.

5. DISCUSSIONS ON TRACTABLE EBW

Here we will discuss tEBW for flows shaped with leaky buckets, showing numerical examples.

5.1 Flows shaped with leaky buckets and their tEBW

We consider a homogeneous tandem network with cross traffic. Every node has the same traffic load, and all flows, both forwarding flows and cross traffic flows, are subjecting to a common probabilistic law. We assume that the arrival process of a typical flow is a greedy process which is limited by a leaky bucket with token rate ρ and token bucket size σ. Let

$$\eta(\theta;t) = \log\left[1 + \frac{\rho t}{\rho t + \sigma}(e^{\theta(\rho t+\sigma)} - 1)\right], \qquad (57)$$

which corresponds to the right hand side of the inequality (46) with $\lambda = \rho$. Then the asymptotic cgf of the arrivals of the typical flow is given as

$$\overline{\mathcal{A}}(\theta;s,t) = \eta(\theta;t-s) \text{ for } \theta \in (0,\infty), \qquad (58)$$

and the tEBW is given by (47).

Now we compare the evaluation formula (21) with the simple formula (30) in Theorem 2 for the particular flows stated above. From (21), the backlog can be evaluated as

$$\log P(Q^L(t) > Lb) \lesssim \inf_{\theta\in(0,\infty)} \Big\{ -\theta Lb + \max_{0\le s_0\le\cdots\le s_m=t} \Big\{ L\eta(\theta;t-s_0) - [Lc\theta(s_1-s_0) - L\beta\eta(\theta;s_1-s_0)]^+ - \cdots - \Big[Lc\theta(s_m - s_{m-1}) - L\beta\eta(\theta;s_m-s_{m-1})\Big]^+\Big\}\Big\}, \qquad (59)$$

where L, $L\beta$ and Lc are the number of forwarding flows, that of cross traffic flow and the link capacity, respectively. On the other hand, the simple formula (30) is reduced as

$$\log P(Q^L(t) > Lb) \lesssim -\theta^* Lb, \qquad (60)$$

where θ^* is the value of θ such that $(1+\beta)\dfrac{\rho}{\theta\sigma}(e^{\theta\sigma}-1) = c$.

Figure 3 shows some numerical results. We set parameters as $L = 10$, $L\beta = 50$, $Lc = 2.5$Gbps, $\rho = 40$Mbps and $\sigma = 4$Mbits. Then the link utilization is 96%. The ordinate is $\log_{10} P(Q^L(t) > x)$ and the abscissa is the buffer threshold $x = Lb$. The four curves indicate the results of (59) with one node, two nodes, three nodes and ten nodes. The straight line shows the result of the simple formula (60). The curves seem to converge to the straight line as the number of nodes increases.

We will check (59) as the number of nodes m increases to infinity. In [10, 11], it was shown that, from the convexity of $c\theta - \beta\eta(\theta;t)$ and $[c\theta-\beta\eta(\theta;t)]^+$ on t for a fixed $\theta > 0$, the maximum in (59) is attained by $s_0, \cdots, s_m(=t)$ such that

$$s_1 - s_0 = s_2 - s_1 = \cdots = s_m - s_{m-1} = \frac{t-s_0}{m},$$

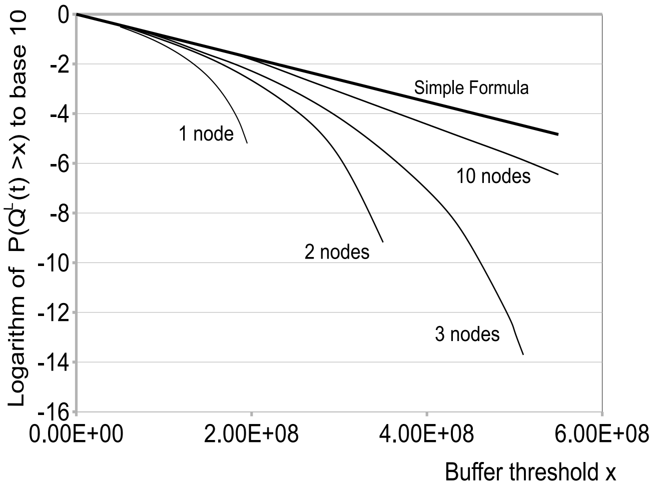

Figure 3: A tandem network with $L = 10$, $L\beta = 50$, $Lc = 2.5$Gbps, $\rho = 40$Mbps and $\sigma = 4$Mbits

and then the right hand side of (59) is reduced to

$$\inf_{\theta\in(0,\infty)} \Big\{ -\theta Lb + \max_{0\le s_0\le t} \Big\{ L\eta(\theta;t-s_0) - \Big[Lc\theta(t-s_0) - L\beta m\,\eta\Big(\theta;\frac{t-s_0}{m}\Big)\Big]^+\Big\}\Big\}. \qquad (61)$$

Giving attention to $m\,\eta(\theta;\frac{t-s_0}{m})$ in (61), we can easily check that

$$m\,\eta\left(\theta;\frac{t}{m}\right) \to \theta t\frac{\rho}{\theta\sigma}\left(e^{\theta\sigma}-1\right) \quad \text{as } m\to\infty. \qquad (62)$$

Then we see that the tEBW (47) appears in the evaluation formula. The worst case end-to-end backlog increases linearly as the number of nodes increases and it doesn't converge. However, the tail probability of the end-to-end backlog does converge. This is the reason why the numerical results of (59) approach ones of the simple formula (60) as the number of nodes increases.

5.2 Admissible region for two types of flows

Theorem 2 suggests that the asymptotic tail probability of the end-to-end backlog is determined by the total amount of the tEBWs at a bottleneck node and does not depend on the number of nodes m. So, one might think this simple evaluation formula provides too overestimated approximations for the use of admission control.

We shall look at a numerical example of an admissible region for two types of flows, quoting a result in [12]. We set the token rate and the token bucket size of type 1 flows as $\rho_1 = 40$Mbps and $\sigma_1 = 4$Mbits, and those of type 2 flows as $\rho_2 = 20$Mbps and $\sigma_2 = 10$Mbits. We calculate effective bandwidths of type 1 and type 2 using (47) and obtain an admissible region from (35). Figure 4 shows the admissible region with the link capacity $C = 10$Gbps for the buffer threshold $B = 100$Mbits or $B = 500$Mbits and the tail probability bound $\epsilon = 10^{-3}$ or $\epsilon = 10^{-6}$.

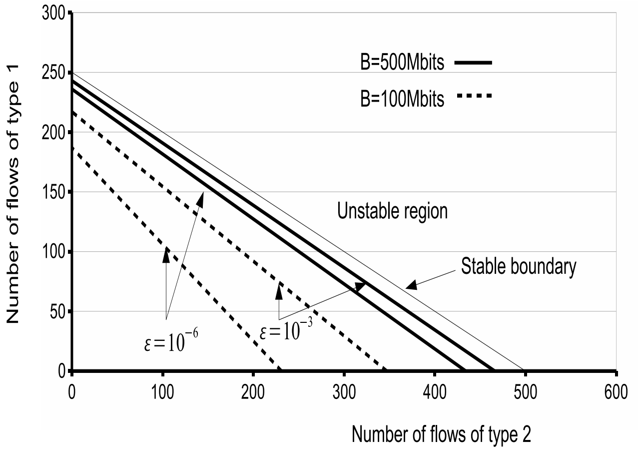

Figure 4: Admissible region with $C = 10$Gbps

We see that in the cases with $B = 500$Mbits, which is a reasonable buffer size, the boundary of the admission region comes close to the stable one. This means that, even if there exists an ideal (or a better) addmissible region, it will almost match our admissible region, so that we may consider this evaluation with tEBWs is not too overestimated in the example.

6. CONCLUSION

We provided a simple evaluation formula in which the tail probability of the end-to-end backlog can be determined by the total amount of the tEBWs at a bottleneck node. We checked whether several well-known effective bandwidths have tEBWs or not, and showed that they have tEBWs except for fractional Brownian motion flows with long range dependency. We also discussed numerical examples for the flows shaped with leaky buckets, and showed that the tEBW might be useful for admission control, though it is a rougher indicator of traffic load than the proper effective bandwidth.

Further, we evaluated the end-to-end backlog in a homogeneous tandem network with fBm cross traffic. Our results indicate that the asymptotic tail probability of the end-to-end backlog in the tandem network is the same as that in a single node. This suggests that an evaluation formula with fBm cross traffic flows might have a similar convenient nature as the evaluation formula for flows with tEBWs.

7. REFERENCES

[1] J. Y. Le Boudec and P. Thiran. Network Calculus: A Theory of Deterministic Queueing Systems for the Internet. On line Version of the Book, Springer Verlag, LNCS 2050, 2004.

[2] C. S. Chang. Performance Guarantees in Communication Networks. Springer Verlag, 2000.

[3] F. Ciucu, A. Burchard and J. Liebeherr. A network service curve approach for the stochastic analysis of networks. In Proceedings of ACM Sigmetrics '05, 2005.

[4] R. Cruz. A calculus for network delay, parts I and II. IEEE Transactions on Information Theory, 37(1):114-141, 1991.

[5] A. Dembo and O. Zeitouni. Large Deviations Techniques and Applications. Jones and Bartlett, 1993.

[6] A. Ganesh, N. O'Connell and D. Wischik. Big Queues. Springer, 2004.

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

[8] Y. Jiang, Network calculus and queueing theory: two sides of one coin, In Proceedings of Valuetools 2009, 2009.

[9] F. Kelly. Notes on effective bandwidths. In Stochastic Networks: Theory and Applications, Oxford University Press, 1996.

[10] K. Kobayashi, Y. Takahashi and H. Takada. A Stochastic network calculus for many flow. In Proceedings of ITC21, 2009.

[11] K. Kobayashi, Y. Takahashi and H. Takada. Asymptotic end-to-end stochastic evaluation for tandem networks with many flows. Proceedings of Valuetools 2009, 2009.

[12] K. Kobayashi, Y. Takahashi and H. Takada. A backlog evaluation formula for admission control based on the stochastic network calculus with many flows. IEICE Transactions on Communincations, E94-B(6):1288-1294, 2011.

[13] W. Leland, M. Taqqu, W. Willinger and D. Wilson. On the self-similar nature of Ethernet traffic (extended version), IEEE/ACM Transactions on networking, 2(1);1-15, 1994.

[14] C. Li, A. Burchard and J. Liebeherr. A network calculus with effective bandwidth. Technical Report CS-2003-20, University of Virginia, Computer Science Department, 2003.

[15] J. Lieveherr, A. Burchard and F. Ciucu. Delay bounds in communication networks with heavy-tailed and self-similar traffic. IEEE Transactions on Information Theory, 58(2);1010-1024, 2012.

[16]] Y. Liu, C.-K. Tham and Y. Jiang. A calculus for stochastic QoS analysis. Performance Evaluation, 64(6):547-572, 2007.

[17] I. Norros. A storage model with self-similar input. Queueing Systems, 16(3);387-396, 1994.

[18] A. Rizk and M. Fidler. Statistical end-to-end performance bounds for networks under long memory FBM cross traffic. In Proceedings of 18th IWQoS, 2010.