

The Burstiness Behavior of Regulated Flows in Networks

Yu Ying¹, Ravi Mazumdar², Catherine Rosenberg² and Fabrice Guillemin³ *

¹ Dept. of ECE, Purdue University, West Lafayette, IN, 47906, U.S.A.
yingy@ecn.purdue.edu

² Dept. of ECE, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada
mazum,cath@ece.uwaterloo.ca

³ France Telecom R&D DAC/CPN, 2, Ave. Pierre Marzin, 22300, Lannion, France
fabrice.guillemin@francetelecom.com

Abstract. In this paper we study the impact of statistical multiplexing on leaky-bucket regulated traffic streams as they pass through the network. In particular we show that the burstiness of a flow is randomized as it transits through the nodes with mean equal to its initial burstiness value at the ingress. We then show that the random burstiness for a single flow converges to a constant equal to the initial value at the ingress when the flow is multiplexed with a large number of sources. The results do not depend on independence or homogeneity between flows. We conclude by providing some simulation results that confirm the theory.

1 Introduction

Within the past decade, there has been a lot of work on characterizing the behavior of traffic in large networks (see [1–5, 16] and the references therein). In particular, it is well known that Internet traffic is bursty although there is not a rigid definition for burstiness. To manage the traffic in networks, a convenient approach is to shape traffic offered by users according to the parameters of an envelope at the network access point. One of the simplest regulation mechanisms is the so-called leaky bucket mechanism, which has been studied in great depth in the context of providing guaranteed quality of service. Traffic offered by a user is in general regulated by a dual leaky bucket, one leaky bucket enforcing the peak rate π and another the mean rate ρ by using a bucket size σ . This regulator is referred to as a (σ, ρ, π) regulator and a formalism to study worst-case delay bounds called *network calculus* has been developed for leaky bucket regulated inputs. Network calculus (see for instance [4–9]) provides a framework to compute end-to-end delays based on algebraic properties of the $(\min, +)$ algebra.

Network calculus is essentially a deterministic worst-case approach. This analysis yields conservative results for network resources provisioning because it neglects the statistical properties of traffic flows as well as their possible statistical independence. Moreover, the statistics of flows are altered as they interact

* This work has been supported by NSF grants ANI #0073359 and #0099137

with other flows through the scheduling and multiplexing mechanisms in routers. Thus, even though we might precisely regulate flows at the ingress, the internal behavior is more difficult to characterize. In this paper, from the viewpoint of computing statistical performance metrics, such as the mean delay through a network and overflow probability at an internal node, we aim to characterize the statistical properties of regulated flows as they progress through the network. In particular, this paper focuses on how statistical multiplexing alters the behavior of single regulated flows. We investigate how the parameters used at the network access point are modified when the single stream traverses several multiplexing nodes, such that performance characteristics at each node can be computed. This leads us to characterize the traffic parameters of flows at the output of a multiplexing stage. Insights from the single flow case will also help subsequent study on the behavior of aggregated flow when a portion of the output flow from a queue goes to the same node.

Recently, there has been a lot of interest in analyzing the statistical properties of multiplexed regulated streams. In [16] it is shown that when a large number of homogeneous sources are multiplexed, a single output flow of a switch essentially keeps its initial statistical characteristics from a large deviations point of view. The authors in [17] provide a network decomposition framework based on the same idea. In [10], the authors consider many independently multiplexed regulated streams as one flow and obtain an effective burst parameter for the aggregated input in terms of a given overflow probability. In [7, 13] the authors obtain bounds for the tail distribution of the buffer occupancy but do not directly address the burstiness characterization; and the analysis is restricted to one queue. Authors in [14] and [15] define a notion of exponentially bounded burstiness (EBB) which yields bounds on the tail distributions of the buffer occupancies. This EBB notion is mainly studied for multiplexing a few flows. In [11, 12], the authors obtained tight bounds on the mean delay and overflow probabilities for multiplexed regulated streams using an extremal traffic characterization but in single queues. None of these earlier results provide insights on how the statistical nature of the burstiness of a data flow behaves as it passes through *multiple* nodes and is multiplexed with *many* other flows in each node.

In fact, when a flow passes through several multiplexing nodes, the burstiness parameter is altered at each node and can no more be described by a constant parameter. This leads us to introduce the concept of a stochastic burst size which manages to characterize the burstiness of each sample path of the traffic flow. Using the idea of stochastic burstiness, we study the alteration of the traffic characteristics of a single flow as it passes through several multiplexing nodes.

In particular, we derive the moments of the burstiness parameter and provide some information on its distribution. We show the mean of the burstiness parameter for a single source remains a constant equal to the initial value of the burst size. The distribution spreads as the number of multiplexing nodes increases, but this distribution converges to its mean if the contribution of this flow is negligibly small at each node. *Our results hold for heterogeneous network and does not require independence between flows.*

The efficacy of this main result is then studied by simulating the mean delay and overflow probability in a tandem queue scenario. Theoretical results are validated via simulations. It turns out that when the number of sources in the network is large and the contribution of individual sources is small, then taking the burst size as unchanged is a reasonable approximation as in the large deviations setting considered in [16]. Therefore it can be used to facilitate the designs of shapers as well as admission control schemes for IntServ/DiffServ networks.

The outline of the paper is as follows: Section 2 gives a general definition of the stochastic burstiness. Section 3 studies in detail the burst size model $\tilde{\sigma}(s, t)$ for single flows. Statistical properties of this model are given. In section 4, we show the effectiveness of our burstiness model via numerical results. Finally, section 5 concludes the paper.

2 Problem Formulation

In this paper, all the traffic processes are defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ and are leaky-bucket regulated at the access point of the network. We say an input flow $I(t)$ is leaky-bucket regulated, if $\forall s \leq t \in \mathcal{R}^+$,

$$I(s, t) \leq \min\{\pi(t - s), \sigma + \rho(t - s)\}, \quad (1)$$

where π stands for peak rate, ρ for mean rate and burst size is denoted by σ .

Equation (1) actually defines a worst case envelope for a regulated flow. If we let $I(\omega, (s, t))$ denote the amount of data produced by one sample path $\omega \in \Omega$ of flow $I(t)$ in the time period of $[s, t]$, (1) then indicates that for all the sample paths in Ω and for any time period $[s, t]$, $I(\omega, (s, t))$ is confined by a deterministic linear sub-additive function. Therefore, when the burstiness of a traffic flow is measured by the burst size σ , this definition implies that the burstiness of this flow is the maximum value achieved over all possible sample paths of flow $I(t)$ and all time interval $[s, t]$, i.e.,

$$\sigma = \max_{s \leq t, \omega \in \Omega} \{I(\omega, (s, t)) - \rho(t - s)\}. \quad (2)$$

We believe this is the fundamental reason why the performance bounds given by the deterministic network calculus approach are conservative for dimensioning the network with statistical performance metrics. Furthermore, the stochastic properties of a flow may change after multiplexing inside the network; while this deterministic approach fails to characterize the multiplexing effect.

To emphasize the stochastic properties of $I(t)$, in particular, of its burstiness inside the network, we relax the above worst case condition and consider each sample path of $I(t)$ in each time interval $[s, t]$. We define our stochastic burstiness of a regulated flow by

$$\sigma_r(\omega, (s, t)) = (I(\omega, (s, t)) - \rho(t - s), 0)^+. \quad (3)$$

This original stochastic burst model keeps track of each sample path of traffic flow $I(t)$; hence it dynamically inherits the stochastic information of $I(t)$ as the

flow passes through a network of queues. Meanwhile, it satisfies the fundamental burstiness property that for each sample path, σ_r confines the maximum time period during which a regulated flow can produce data at its peak rate ([4]). It is easy to see that the peak rate of an output flow is bounded by the capacity of the node and the average rate of the flow is unchanged in a stable network. Hence to compute the performance, it is necessary to characterize the probabilistic properties of this stochastic process $\sigma_r(s, t)$.

In this paper, we study the burstiness characterization of a single output flow $\tilde{I}_i(t)$ which is the i th output stream from a previous node with traffic descriptor denoted by $(\tilde{\sigma}_i, \tilde{\rho}_i, \tilde{\pi}_i)$. We aim to understand how the burstiness parameter $\tilde{\sigma}_i$ is changed from that of the i th input stream after statistical multiplexing in the queues of previous nodes, such that the single output traffic flow can be stochastically regulated as

$$\tilde{I}(s, t) \leq \min\{\tilde{\pi}(t-s), \tilde{\sigma}(s, t) + \tilde{\rho}(t-s)\}.$$

We focus on the burst level phenomena that occurs in a time scale typical of an on-off source activity period rather than the inter-arrival periods of packets, we thus ignore the discrete nature of the packets arrivals and regard the input flows as continuous stationary fluid processes. We restrict our analysis to a stable network in which all the nodes have infinite buffer with work conserving FIFO scheduling and each flow is assigned a route without loop. Traffic flows are called “fresh” if they are regulated with fixed parameters $(\sigma_j, \rho_j, \pi_j)$. Such flows are usually inputs to the edge nodes of the network. Their burst sizes are known at the edge node via subscription or a signaling procedure. Note that independence between flows are not assumed in our analysis.

3 Burstiness of a Single Flow

In this section, we tag one specific flow, $I^i(t)$, at the entrance and characterize its burstiness behavior after it passes through several internal nodes.

3.1 Single Output Flow with Fresh Inputs

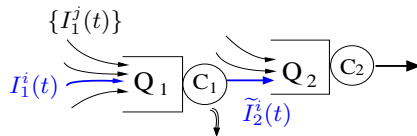


Fig. 1. Scenario S-I

We will begin our discussion with the simplest case as shown in Fig. 1. In this scenario, N_1 number of fresh flows $\{I_1^j(t)\}_{j=1}^{N_1}$ enters Q_1 where flow j has

parameters $(\sigma_1^j, \rho_1^j, \pi_1^j)$. We are interested in the burstiness process $\tilde{\sigma}_2^i(s, t)$ of the i th output flow $\tilde{I}_2^i(t)$.

To relate the output flow \tilde{I}_2^i with its input $I_1^i(t)$, we use the definition of Virtual Arrival Time Process (VATP) $v_1^i(t)$ ([9, 20]) of fresh input $I_1^i(t)$ as the following:

$$v_1^i(t) = \sup_{s \leq t} \{s : I_1^i(s) \leq \tilde{I}_2^i(t)\}. \quad (4)$$

Hence $v_1^i(t)$ denotes the arrival time of a bit from source i to the queue which departs at time t . Since we deal with FIFO queueing, $v_1^i(t)$ can equivalently be defined as:

$$v_1^i(t) = \sup_{s \leq t} \left\{ s : s + \frac{Q_1(s)}{C_1} \leq t \right\}. \quad (5)$$

where $Q_1(s)$ is the stationary workload of the first queue.

It readily follows that $v_1^i(t)$ is a nondecreasing right-continuous process and that $\tilde{I}_2^i(s, t) = I_1^i(v_1^i(s), v_1^i(t))$. Then with $v_1^i(t) = t - \frac{Q_1(v_1^i(t))}{C_1}$, and the fact that $I_1^i(s, t)$ is regulated as

$$I_1^i(s, t) \leq \min\{\pi_1^i(t - s), \sigma_1^i + \rho_1^i(t - s)\},$$

we obtain

$$\begin{aligned} \tilde{I}_2^i(s, t) &\leq \min\{C_1(t - s), \sigma_1^i + \rho_1^i(v_1^i(t) - v_1^i(s))\} \\ &= \min\{C_1(t - s), \rho_1^i(t - s) + \sigma_1^i + \frac{\rho_1^i}{C_1}[Q_1(v_1^i(s)) - Q_1(v_1^i(t))]\}. \end{aligned}$$

Here, we used the fact that the peak rate for the output flow is confined by the capacity C_1 of the first queue, $\tilde{\pi}_2^i = C_1$. Hence $\tilde{I}_2^i(s, t)$ is also regulated, and its burstiness parameter can be characterized as a stochastic process $\{\tilde{\sigma}_2^i(s, t)\}$ defined by

$$\tilde{\sigma}_2^i(s, t) = \sigma_1^i + \frac{\rho_1^i}{C_1}[Q_1(v_1^i(s)) - Q_1(v_1^i(t))]. \quad (6)$$

As seen from (6), the stochastic properties of $\tilde{\sigma}_2^i(s, t)$ are determined by another random process $Q_1(v_1^i(t))$ which is a random-time shifting of stationary workload process $Q_1(t)$ of the first queue. The following lemma shows that, with FIFO queueing, this process $Q_1(v_1^i(t))$ is also a stationary process and it has the same moment distributions as $Q_1(t)$.

Lemma 1. *Consider several input streams $\{I^j(s, t)\}$ entering a FIFO queue with capacity C , with $\rho^j = \mathbb{E}[I^j(0, 1)]$ and $\sum \rho^j < C$. Let $Q^*(t) = Q(v^i(t))$, where $v^i(t)$ is the Virtual Arrival Time Process. Then $Q^*(t)$ is a stationary process and for $n \geq 1$, $\mathbb{E}[(Q^*)^n(t)] = \mathbb{E}[Q^n]$.*

Proof. The proof follows mutatis mutandis from [20, Theorem 4].

Therefore, we can obtain the following statistical properties of $\tilde{\sigma}_2^i(s, t)$:

Proposition 1 (Properties of $\tilde{\sigma}_2^i(s, t)$). For any sample path, $\tilde{I}_2^i(s, t)$ is regulated with $(\tilde{\sigma}_2^i(s, t), \rho_1^i, C_1)$:

$$\tilde{I}_2^i(s, t) \leq \min\{C_1(t - s), \rho_1^i(t - s) + \tilde{\sigma}_2^i(s, t)\}. \quad (7)$$

The mean value of $\tilde{\sigma}_2^i(s, t)$ is given by

$$\mathbb{E}[\tilde{\sigma}_2^i(s, t)] = \sigma_1^i, \quad (8)$$

and its second moment is upper bounded.

Finally, we have $\tilde{\sigma}_2^i(s, t) \rightarrow \sigma_1^i$ in probability as $\frac{\rho_1^i}{C_1} \rightarrow 0$.

Proof. Equation (7) is obtained from equation (6). From the same equation and Lemma 1,

$$\mathbb{E}[\tilde{\sigma}_2^i(s, t)] = \mathbb{E}\left[\sigma_1^i + \frac{\rho_1^i}{C_1}[Q_1^*(s) - Q_1^*(t)]\right] = \sigma_1^i,$$

and equation (8) follows. Also, its second moment is bounded since

$$\begin{aligned} \mathbb{E}[(\tilde{\sigma}_2^i(s, t))^2] &= (\sigma_1^i)^2 + 2\frac{(\rho_1^i)^2}{C_1^2}\mathbb{E}[Q_1^2] - 2\frac{\rho_1^i}{C_1}\mathbb{E}[Q_1^*(t)Q_1^*(s)] \\ &\leq (\sigma_1^i)^2 + 2\frac{(\rho_1^i)^2}{C_1^2}\mathbb{E}[Q_1^2]. \end{aligned}$$

Finally, the convergence follows from Markov Inequality,

$$\mathbb{P}\{|\tilde{\sigma}_2^i(s, t) - \sigma_1^i| \geq x\} = \mathbb{P}\left\{|Q_1^*(s) - Q_1^*(t)| \geq \frac{C_1 x}{\rho_1^i}\right\} \leq \frac{2\rho_1^i \mathbb{E}[Q_1]}{C_1 x}.$$

Since $E[Q_1]$ is upper bounded, $\forall x > 0$, the the left hand side of above equation goes to 0 as ρ_1^i/C_1 goes to 0. This completes the proof.

3.2 Single Output Flow to the $(k + 1)$ th Node

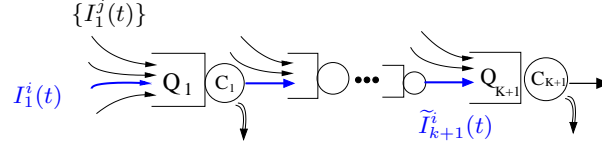


Fig. 2. Scenario S-II

We now come to generalize the observations on the burstiness of a single flow. Consider the case described by Fig. 2 where the tagged i th input flow $I_i(t)$ traverses k nodes and gives rise to $\tilde{I}_{k+1}^i(t)$ for node $k + 1$. We will use similar approaches as in Section 3.1 to characterize burst size $\tilde{\sigma}_{k+1}^i$ of $\tilde{I}_{k+1}^i(t)$.

As before, the Virtual Arrival Time Process $v_k^i(t)$ for the i th stream $\tilde{I}_k^i(t)$ at the k th queue is defined as:

$$v_k^i(t) = \sup_{s \leq t} \{s : \tilde{I}_k^i(s) \leq \tilde{I}_{k+1}^i(t)\} = \sup_{s \leq t} \left\{ s : s + \frac{Q_k(s)}{C_k} \leq t \right\}.$$

and process $v_k^i(t)$ is non-decreasing and right-continuous.

Using the concept of VATP together with the fact that $I_1^i(s, t)$ is regulated, we can obtain Proposition 2. This proposition shows that we can define an envelope process for $\tilde{I}_{k+1}^i(s, t)$ being stochastically regulated with a burstiness process $\tilde{\sigma}_{k+1}^i(s, t)$. Proposition 2 can be proved by mathematical induction. Details of the proof is given in [21] due to limited space.

Proposition 2. *Consider an input stream which is characterized by the parameters $(\sigma_1^i, \rho_1^i, \pi_1^i)$ at the first queue. Then its output from the k^{th} queue is stochastically enveloped as:*

$$\tilde{I}_{k+1}^i(s, t) \leq \min\{C_k(t - s), \rho_1^i(t - s) + \tilde{\sigma}_{k+1}^i(s, t)\},$$

and the burst size process $\tilde{\sigma}_{k+1}^i(s, t)$ for $\tilde{I}_{k+1}^i(s, t)$ is defined by

$$\tilde{\sigma}_{k+1}^i(s, t) = \sigma_1^i + \sum_{m=1}^k \frac{\rho_1^i}{C_m} [Q_m(f_m^k(s)) - Q_m(f_m^k(t))]. \quad (9)$$

where

$$f_m^k(t) = v_m^i \circ v_{m+1}^i \circ \dots \circ v_k^i(t). \quad (10)$$

It is worth noting that $f_m^k(t)$ for the m th queue is a non-decreasing right-continuous process. Moreover, using the knowledge of palm theory [22] as well as mathematical induction techniques, the following lemma shows that the random-time shifting process $Q(f_m^k(t))$ is also a stationary process. Details of the proof is given in [21] due to limited space.

Lemma 2. *If the workload process $\{Q_m(t)\}$ in a work conserving node is a stationary process, for $m = 1, \dots, k$, then $Q_m^{*k}(t) = Q(f_m^k(t))$ is bounded and $\{Q_m^{*k}(t)\}$ is a stationary process.*

We then obtain the properties of the burstiness of $\tilde{I}_{k+1}^i(s, t)$ as the following:

Proposition 3. *(Properties of $\tilde{\sigma}_{k+1}^i(s, t)$)*

The mean value of $\tilde{\sigma}_{k+1}^i(s, t)$ is given by

$$\mathbb{E}[\tilde{\sigma}_{k+1}^i(s, t)] = \sigma_1^i. \quad (11)$$

We also have $\tilde{\sigma}_{k+1}^i(s, t)$ converges to σ_1^i in probability as $\rho_1^i/C_m \rightarrow 0$ for $m = 1, \dots, k$.

Proof. From the stationarity of $Q_j^{*k}(t)$, we have $\mathbb{E}[Q_j^{*k}(t)] = \mathbb{E}[Q_j^{*k}(s)]$ for $j = 1, 2, \dots, k$, from which equation (11) follows. To show the convergence, from equation (9) and Markov Inequality, we have

$$\mathbb{P}\{|\tilde{\sigma}_{k+1}^i(s, t) - \sigma_1^i| > x\} \leq \frac{2\rho_1^i}{x} \sum_{m=1}^k \frac{\mathbb{E}[Q_m^{*k}]}{C_m}.$$

Therefore, $\tilde{\sigma}_{k+1}^i(s, t)$ converges to σ_1^i in probability as $\rho_1^i/C_m \rightarrow 0$ for $m = 1, \dots, k$, since $E[Q_m^{*k}(t)]$ is bounded and k is a finite number. This completes the proof.

At the end of this section, we summarize our main result into Theorem 1 for the burstiness behavior of a single flow in acyclic stable networks.

Theorem 1. *In a stable network with work conserving FIFO scheduling at every node with infinite buffer, if all the input flows at the entrance are regulated with leaky-bucket parameters $(\sigma^j, \rho^j, \pi^j)$, then each flow to node k inside the network is still regulated, with a stochastic burst size $\tilde{\sigma}_k^j(s, t)$. This burst size has the mean value of σ^j and converges exponentially fast to this constant σ^j in probability, as $\rho^j/C_m \rightarrow 0$ for $m = 1, \dots, k - 1$.*

Remark 1. This result implies that in a large network, no matter how the flow is multiplexed with others at each node, as long as the average arrival rate of the flow is much smaller than the capacity of the server, its burst size is almost constant. Indeed, it can be a very good approximation when the network carries a large number of flows, as for instance a core network. Furthermore, in our analysis, we consider a heterogeneous network and do not assume independence between input flows. Thus our main result is a general characterization of single regulated flows inside the network.

4 Performance Discussion

We have theoretically shown the convergence of the burstiness of a single flow in Section 3. However, this result is presented in a limit sense: the convergence holds as ρ_i/C_k goes to 0 for each node on an end-to-end path of stream $I_i(t)$. In real network systems, we need to know the scale of the network for the burstiness converges at a reasonable rate as well as how the distribution of its burstiness deviates from its constant mean as the number of end-to-end hops increases. Furthermore, it is more interesting to measure the distortion of performances, such as mean delay and overflow probability, at an internal node along the end-to-end path due to the burstiness behavior of a single flow.

To answer these concerns, in this section, we will simulate the mean delay and overflow probability in a tandem queue scenario as shown in Fig. 2. In particular, we will illustrate how the values of these performance metrics change when the number of input flows increases at each stage and when the number of end-to-end hops increases. These simulation results reflect the convergence behavior of the burstiness of each flow.

Table 1. Simulation Parameters

N_1	N_k^f	C_k	σ_k	ρ_k	π_k
50	49	1	0.1	$0.7/N_1$	1
200	199	1	0.1	$0.7/N_1$	1
400	399	1	0.1	$0.7/N_1$	1

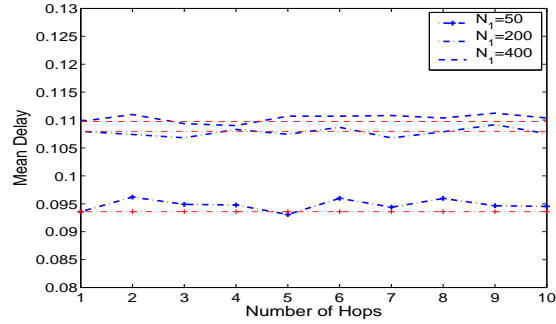
Table 2. Performance Upper bounds with Regulated Inputs

N_1	ED_{INFO}	$P_{Chernoff}$	P_{MD1}
50	0.11433	0.005528	0.013881
200	0.11608	0.012163	0.013881
400	0.11638	0.014004	0.013881

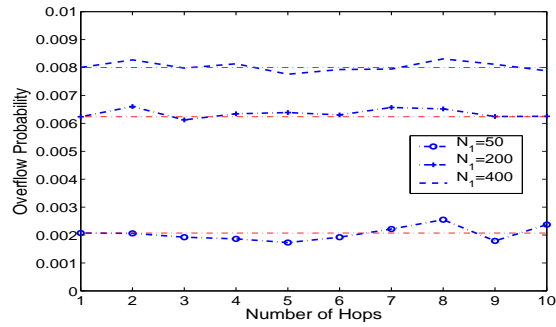
All simulations are carried out on a time-driven fluid simulator. This simulator consists of a number of FIFO queues with infinite buffer size and independent regulated On-Off fresh input streams of chosen parameters. We designed the flows to be homogeneous at each node for simplicity. And we generated independent fresh flows so as to compute the theoretical bounds of the mean delay and overflow probability at each node with available results in the literature ([11, 12]). Note that this independence is required only by computation of the performance metrics. Our results on the burstiness behavior of flows inside the network do not assume independence between flows.

In our numerical simulations, a tandem queue scenario with 10 hops is studied. Parameters about the fresh input flows and the servers are given in Table 1. At each node, we fixed the total load to be 0.7 as the number of sources increases. Therefore, the convergence speed of the burstiness of the i th flow can be measured by the total number of input flows N_k to each node. Furthermore, we set the number of fresh inputs N_k^f , for $k = 2, \dots, 10$ to be equal to $N_1 - 1$. In this way, input scenarios at the latter 9 nodes are equivalent to replacing one fresh input to node 1 with a regulated flow having stochastic burstiness behavior. We thus can easily observe the evolution of the burstiness behavior of a single flow by comparing values of performance metrics at different nodes.

Numerical results on mean delay at each node are given in Figure 3(a); and those on overflow probability are illustrated in Fig. 3(b). These data are reported for a 95% confidence interval. We omit plotting those intervals in the figures because they are too small to be shown compared to the statistical sample mean. If we let m_k denote the value of a performance metric at node k , and the performance distortion be measured by $(m_k - m_1)/m_1 \times 100\%$, then, as seen from these two figures, when there are more than 200 flows entering each node, performance distortions at each node are negligibly small: around 1% for mean delay and 5% for overflow probability. At the same time, increasing the number of hops of an end-to-end path does not distort the performance at each node dramatically. Indeed, each node along the same path provides similar average delay and overflow probability for arrival flows. All these observations imply that the



(a) Mean Delay at Each Node

(b) Overflow Probability at Each Node with Buffer Threshold $B=0.6$ **Fig. 3.** Simulation Results

stochastic burstiness of a single flow converges to its constant mean value at an acceptable speed. In practical applications, we can approximate the stochastic burstiness of a flow inside the network with its mean value, which is the burst size originally obtained at the entrance of the network. Hence, we can obtain the parameters of single flows inside the network and compute the performance of each internal node via single-queue analysis techniques.

At the end of this section, we list in Table 2 the upper bounds of mean delay and overflow probability computed by replacing the stochastic burstiness with its constant mean value. ED_{INFO} is computed from equation (31) in [11]; $P_{Chernoff}$ and P_{MD1} are computed from Chernoff bound and $M/G/1$ queue bound respectively, which are discussed in detail in [12]. It is clear to see that these theoretical upper bounds are still tight to confine the performances of a

node *inside* the network, when the inputs to the node are independent of each other as in our scenario.

5 Conclusion

In this paper, we have characterized the burstiness behavior of single traffic flows as they pass through the network. These traffic processes are initially leaky-bucket regulated at the entrance of the network. By considering the internal input traffic as the output flows of some previous queue, we exploit the multiplexing effects from the many sources at the previous node. In particular we have shown that when single flows are small, the multiplexing does not affect their initial burstiness.

This inheritance of burst-size descriptor achieves scalability in designing shapers at the routers of Differentiated Services networks. Shaping algorithm (Leaky-Bucket Algorithm) can be employed only at the edge of the network for each class of flows while core routers can regard each arrival flow keeping its original burst descriptor and do not have to reshape the traffic.

We have also illustrated that statistical performance upper bounds for deterministically regulated flows can be used to estimate the performance of an internal node in scenarios where input flows are independent of each other. Authors in [19] studied by extensive simulation the inter-source cross-correlation under more general settings. They observed that the inter-source dependence benefits queueing performance and independence assumptions between flows can give conservative estimation. Deeper understandings on the dependence between internal flows are indispensable before the establishment of a general methodology for obtaining statistical end-to-end performance estimates.

References

1. D.Anick, D.Mitra, M.Sondhi. Stochastic theory of a data handling system with multiple sources. *Bell System Technical Journal*, vol.61, no.8, pp1871-1894,1982.
2. W. Leland, M. Taqqu, W. Willinger, D. Wilson, On the Self-Similar Nature of Ethernet Traffic. *IEEE/ACM Trans. on Networking*, Vol. 2, No.1, pp.1-15, Feb. 1994
3. J.W.Roberts. Traffic Theory and the Internet. *IEEE Communications Magazine*, Jan. 2001, pp.94-99.
4. R. Cruz. A calculus for network delay, Part I: Network elements in isolation. *IEEE Trans. Inform. Theory*, vol 37, no.1, Jan. 1991 pp. 114-131
5. R. Cruz. A calculus for network delay, Part II: Network analysis. *IEEE Trans. Inform. Theory*, vol 37, no.1, Jan. 1991 pp. 132-141.
6. J-Y. Le Boudec, P. Thiran. *Network calculus*, Springer Verlag LNCS 2050, June 2001.
7. C.S. Chang *Performance Guarantees in communication networks*, Springer, 2000.
8. V. Firoiu, J.-Y Le Boudec, D. Towsley, Zhi-Li Zhang, Theories and models for Internet quality of service. *Proceedings of the IEEE*, Vol. 90, Issue 9, Sept. 2002, pp. 1565-1591

9. V. Cholvi, J. Echague, J-Y Leboudec. Worst case burstiness increase due to FIFO multiplexing. In *Proceedings of Performance 2002*, Rome, Italy, Sept. 2002.
10. L. Massoulié, A. Busson. Stochastic majorization of aggregates of leaky bucket constrained traffic streams. Preprint, 2001.
11. F. Guillemin, N. Likhanov, R. Mazumdar, C. Rosenberg. Extremal traffic and bounds on the mean delay of multiplexed regulated traffic streams. In *Proc. of INFOCOM 2002*, N.Y., June 2002, pp. 985-993
12. F. Guillemin, N. Likhanov, R. Mazumdar, C. Rosenberg, and Y. Ying. Buffer overflow bounds for multiplexed regulated traffic streams. In *Proc. ITC 18*, Berlin, Elsevier Science, 2003.
13. M. Vojnovic and J.-Y. Le Boudec. Stochastic analysis of some expedited forwarding networks. In *Proc. of IEEE INFOCOM 2002*, New York, NY, June 2002.
14. O. Yaron, M. Sidi. Performance and stability of communication networks via robust exponential bounds. *IEEE/ACM Trans. Networking*, vol.1 , no.3, June. 1993. pp. 372-385
15. D. Starobinski, M. Sidi. Stochastically bounded burstiness for communication networks, *IEEE Tran. Information Theory*, vol. 46, No.1, Jan. 2000. pp. 206-212.
16. D. Wischik. The output of a switch, or, effective bandwidths for networks. *Queueing Systems* 32, 1999, pp. 383-396
17. D. Eun, N.B.Shroff. Simplification of Network Analysis in Large-Bandwidth Systems. *Proc. of IEEE INFOCOM 2003*, San Francisco, CA, March 2003.
18. D. Abendroth, U. Killat. Intelligent Shaping: Well shaped throughout the entire network. *Proc. of INFOCOM'02*, N.Y. June, 2002.
19. W-c. Lau, San-qi Li. Traffic distortion and inter-source cross-correlation in high-speed integrated networks, *Computer Networks and ISDN Systems* 29(1997) pp811-830
20. T. Konstantopoulos, G. Last. On the dynamics and performance of stochastic fluid systems. *Journal of Applied Probability* 37, pp. 652-667, 2000.
21. Y. Ying, R. Mazumdar, C. Rosenberg, F. Guillemin, Analysis of the burstiness of multiplexed regulated traffic flows in networks. *Technical Report, Purdue University*.
22. F. Baccelli and P. Bremaud. *Elements of Queueing Theory*. Springer-Verlag, 2003.