

On Superlinear Scaling of Network Delays

Almut Burchard, Jörg Liebeherr, Florin Ciucu

Abstract— We investigate scaling properties of end-to-end delays in packet networks for a flow that traverses a sequence of H nodes and that experiences cross traffic at each node. When the traffic flow and the cross traffic do not satisfy independence assumptions, we find that delay bounds scale faster than linearly. More precisely, for exponentially bounded packetized traffic we show that delays grow with $\Theta(H \log H)$ in the number of nodes on the network path.¹ This superlinear scaling of delays is qualitatively different from the scaling behavior predicted by a worst-case analysis or by a probabilistic analysis assuming independence of traffic arrivals at network nodes.

I. INTRODUCTION

In this paper, we are concerned with the scaling behavior of end-to-end delays in a packet network for a flow that traverses a sequence of H nodes and that experiences cross traffic at each node. The motivating question is whether end-to-end delays in a network scale well, that is, grow linearly, with the length of the network path, or if delays grow faster than linearly, possibly blowing up when the network path grows large. An important finding of this paper is that the linear scaling of delay predicted by many analytical methods available for end-to-end analysis does not hold, even under relatively general assumptions.

A recent study [11] presented a stochastic fluid-flow analysis of a network where traffic arrivals conform to the Exponentially Bounded Burstiness (EBB) model by Yaron and Sidi [29], which coincides with the class of linear bounded envelope processes introduced by Chang [10]. This class of models includes multiplexed regulated arrivals and many Markov-modulated processes, but excludes long-range correlated or heavy-tailed traffic. In [11], it is shown for a tandem network of H nodes with EBB-compliant cross traffic at each node that the delay of the through flow grows no more than $O(H \log H)$ with the number of nodes. This scaling behavior differs markedly from the scaling of bounds obtained with other analytical methods. For example, product form queueing networks [3,21], where a network path is modeled as a sequence of queueing systems with Poisson arrivals and (for FIFO systems) exponentially distributed service, predict a linear growth of delays. A strong assumption of

these models is that the transmission times of a packet at different nodes are independent, which corresponds to randomly re-sampled the size of the packet at each traversed node. The deterministic network calculus [5, 10], which derives worst-case bounds in a min-plus algebra, has shown that worst-case end-to-end delays also grow linearly in the number of nodes. A linear growth of end-to-end delays also emerges from a stochastic analysis of the network calculus, under the assumption that service at nodes is statistically independent [10, 13], similar as is assumed in product form queueing networks.

In light of the different scaling properties found by these modeling approaches, the results in [11] raise two questions: *Under what assumptions on the network and the arrivals do $O(H \log H)$ bounds hold? And are such bounds ever sharp?* The purpose of this paper is to answer both questions and shed light on the mechanism for the growth of delays in stochastic models for networks.

To address the first question we show that the $O(H \log H)$ bound remains valid even if we replace the idealized fluid flow traffic model from [11] with a traffic model that accounts for the effects of packetization. We discover that the upper bound on end-to-end delays holds in packet networks where the distribution of the size and number of packets arriving in a given time interval has an exponentially bounded tail. This includes in particular the Poisson and related processes frequently studied in queueing theory. In contrast with most of the queueing network literature (and also [13]), we do not assume that service times of a packet are independently re-sampled at each traversed node.

To address the second question, we show that the $O(H \log H)$ bound on end-to-end delays cannot be improved upon. To demonstrate this, we construct a network that satisfies the assumptions for the $O(H \log H)$ upper bounds on delay and show that typical delays grow with $\Theta(H \log H)$. Concretely, we analyze the delays of packets in a tandem network of H identical nodes with no cross traffic, and obtain that the end-to-end delay of packets is bounded from below by $\Omega(H \log H)$. This lower bound remains valid in more general networks where the flow experiences cross traffic at each node.

The derivations for the scaling of the upper bounds are set in the framework of the stochastic network calculus [19], which extends the deterministic network calculus [5, 10] for worst-case analysis in networks to a probabilistic setting. In the stochastic network calculus, traffic arrivals are characterized by statistical arrival envelopes, and the service available to flows at network nodes is expressed in terms of statistical service curves. Different from the deterministic version, the stochastic network cal-

Almut Burchard (almut@math.utoronto.ca) is with the Department of Mathematics at the University of Toronto. Jörg Liebeherr (jorg@comm.utoronto.ca) is with the Department of Electrical and Computer Engineering at the University of Toronto. Florin Ciucu (florin@net.t-labs.tu-berlin.de) is with the Deutsche Telekom Laboratories. Parts of this paper have been presented at the IEEE Infocom 07 conference [7].

¹We use the big-Oh or Landau notation for the asymptotic comparison of functions. For two sequences A_n and B_n , the notation $A_n = O(B_n)$ means that the ratio $\frac{A_n}{B_n}$ is bounded by a constant, while $A_n = \Omega(B_n)$ means that the ratio $\frac{B_n}{A_n}$ is bounded. If both relations hold, we write $A_n = \Theta(B_n)$.

culus can express statistical fluctuations of traffic and capture the statistical multiplexing gain in packet networks. A key technique in the network calculus is to express the service of a flow along a path as a composition of the service received at each node on the path. More precisely, when the service at each node on a path is described in terms of service curves, the service curve for the entire path can be expressed by a min-plus algebra convolution of the per-node service curves. This result was established first in the context of the deterministic network calculus [1]. Finding the corresponding composition result in the stochastic setting turned out to be hard and, for a long time, was limited to strong assumptions on the network or restrictions to the definition of the service curve. Examples of the former can be found in [23], where delays at each node are assumed to satisfy a priori delay bounds, in [2], where it is assumed that a node discards traffic that exceeds a threshold, and in [13], which assumes that service at subsequent nodes is statistically independent. Examples of the latter include [8], which assumes that the statistical service description is made over time intervals, and [17, 18], which assumes sample path guarantees for service. In [11], it was shown that a composition of per-node service curves becomes feasible without such assumptions for a broad class of traffic types, by accounting for a rate penalty at each traversed node. This composition result enabled the derivation of the $O(H \log H)$ bound for delays for fluid-flow traffic. We will show that the same bounds hold when accounting for the packetization of network traffic.

Our derivations of the $\Omega(H \log H)$ lower bound take a queuing theoretical approach. In particular, we consider a variant of an M/M/1 tandem queueing network where the service times of each packet are identical in each queue, i.e., the packet size is not randomly re-sampled at each node. There exists a small (and possibly not widely known) literature that has studied tandem networks with identical service times with no cross traffic. While obviously a niche group of models, they have proven useful for studying scenarios where the independence assumption on the service does not hold. Boxma [6] analyzed a tandem network with two queues with Poisson arrivals and general service times distributions, and derived the steady-state distribution at the second node. Boxma showed that the (positive) correlations between the waiting times at the two nodes are higher than in a network where service times at nodes are independent. Calo [9] showed that in G/G/1 tandem networks with identical service times, the node delay of a packet is non-decreasing in the number of nodes. For a network with Poisson arrivals and a bimodal packet size distribution, he obtained the Laplace transform for steady-state delays. Vinogradov has authored a series of articles on tandem networks with identical service times and no cross traffic. (Some of these papers are only available in Russian language journals.) In [26], he presented an expression for the steady-state distribution of the end-to-end delay in a tandem network with Poisson arrivals and general service time distributions. In subsequent work [27, 28], he showed for expo-

entially distributed service times, that the average per-node delay grows logarithmically at downstream nodes, and hence the average end-to-end delay behaves as $\Theta(H \log H)$. This result provided evidence that the scaling of a tandem networks with identical service times differs from that of a network where service times are independently re-sampled at each nodes. Vinogradov also found the asymptotic scaling behavior of the per-node delay when the load factor approaches one. These results have been extended to general arrivals and to the transient regime [14, 15, 20]. The exact expressions obtained by Vinogradov for the distribution of the end-to-end delay are not explicit and do not lend themselves well to numerical evaluation. Even for a Poisson arrival process, finding the value of the tail distribution function requires, for each w , to solve a transcendental equation and then compute an integral of the solution. Our derivations of a lower bound (in Theorem 3) add to the above literature by giving explicit non-asymptotic lower bounds that can be numerically evaluated for all values of H .

The significance of the $\Theta(H \log H)$ scaling behavior of end-to-end delays stems from the linear scaling predicted by other network models, specifically networks with deterministic arrival envelopes and service guarantees, and networks with independent exponentially distributed interarrival and service times. With respect to the $\Theta(H)$ bound when the service satisfies deterministic guarantees, this paper provides conclusive evidence that, in general, delays scale differently than in the deterministic network calculus. With respect to the $\Theta(H)$ scaling of delays in product form queueing network models and stochastic network calculus models with independent service, this paper shows that dispensing with the assumption on independent service changes the scaling behavior of network delays.

Since the observed scaling behavior extends to the entire distribution of the end-to-end delays, the question arises: What makes delays at nodes downstream a long network path behave fundamentally different than delays at a single node? Our analysis suggests that, on long network paths, end-to-end delays are dominated by short-term congestion; Once a backlog is built up it tends to perpetuate to the downstream nodes of the remaining network path. The longer a path, the more this effect is felt. The phenomenon is reminiscent of shock waves, which are well known to appear in transportation systems [25] but have not been studied in communication networks.

The remainder of this paper is structured as follows. In Section II, we discuss the network model used in this paper. In Section III, we derive the $O(H \log H)$ upper bound for a packetized arrival model. In Section IV, we construct an example of a network where delays grow with $\Omega(H \log H)$, and investigate its scaling properties. In Section V, we give numerical examples that compare the upper and lower bounds obtained in this paper to simulation results. In Section VI, we present brief conclusions.

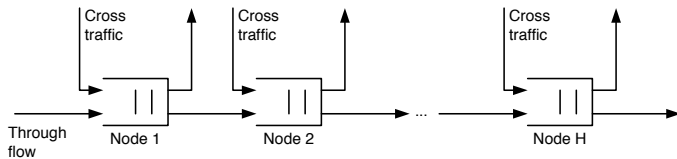


Fig. 1. A tandem network with cross traffic.

II. NETWORK MODEL

We consider a flow, referred to as *through flow*, traversing a network with H nodes in series as shown in Fig. 1. Each node represents a work-conserving output link with a fixed capacity. Our analysis is valid for any link scheduling algorithm that is ‘locally FIFO’ in the sense that it preserves the order of arrivals within a flow or flow class. We assume infinite sized buffers, that is, there are no losses due to buffer overflows. At each node, the through flow is multiplexed with cross traffic. However, we require that statistical bounds on cross traffic are available at each node. Service at different nodes and arrivals from different flows are not required to be statistically independent. We generally assume that a stability condition holds at each node, that is, the average arrival rate to a node does not exceed the service rate. In principle, this model can be applied to rather general network topologies, and flows are not prohibited from looping back on themselves or interacting with each other repeatedly. However, the model does not account for the impact that the through flow has on cross traffic arrivals, e.g., multiplexing of through flow and cross traffic at one node can increase the burstiness of cross traffic at a downstream node.

Traffic arrivals at a node in the time interval $[0, t)$ are modeled by a non-decreasing, left-continuous process $A(t)$ with $A(t) = 0$ for $t \leq 0$. Departures from a node or a network are described similarly, and will be denoted by $D(t)$, with $D(t) \leq A(t)$. The backlog is defined by $B(t) = A(t) - D(t)$, and the delay is defined by

$$W(t) = \inf\{w \geq 0 : A(t) \leq D(t+w)\}.$$

We will use subscripts to denote arrivals, departures, backlog, and delay at a particular node on the path. The end-to-end delay experienced by the flow along its path of H nodes will be denoted by $W_{net}(t)$. As $t \rightarrow \infty$, $W_{net}(t)$ converges in distribution to the steady-state delay, denoted here by W_{net} . We will analyze the distribution of W_{net} through its *quantiles*, defined for $0 < z < 1$ by

$$w_{net}(z) = \inf\{w \geq 0 : P(W_{net} \leq w) \leq z\}. \quad (1)$$

We consider traffic that has *exponentially bounded burstiness* (EBB) in the sense of Yaron and Sidi [29]. In general, we will say that a random variable Y is *exponentially bounded*, if there exist nonnegative constants θ and M such that

$$Pr\{Y > y\} \leq Me^{-\theta y}. \quad (2)$$

An arrival process is EBB, if it satisfies for every $0 \leq s < t$ and all σ

$$Pr\{A(t) - A(s) > r(t-s) + \sigma\} \leq Me^{-\theta\sigma}, \quad (3)$$

where the rate r , and parameters M and θ are nonnegative constants.

EBB arrival processes are closely related to traffic characterizations through moment-generating functions as in [10]. In fact, if the arrival process A satisfies a bound on the moment-generating function of the linear form

$$\frac{1}{\theta} \log E[e^{\theta(A(t)-A(s))}] \leq r(\theta)t + \sigma(\theta),$$

then A is an EBB arrival process with parameters $r = r(\theta)$ and $M = e^{-\theta\sigma(\theta)}$. The class of EBB arrival processes includes many Markov-modulated traffic models, regulated traffic models, as well as packet processes commonly used in queueing theory, however, it does not include self-similar or heavy-tailed processes.

Arrivals may be either fluid-flow or packetized. If the arrivals are generated by a process that produces a packet of size $P_n > 0$ at time T_n ($n = 1, 2, \dots$), the arrival process is given by

$$A(t) = \sum_{n: T_n < t} P_n,$$

which is clearly non-decreasing and left-continuous. In this case, we assume that packet sizes are exponentially bounded.

We will derive a lower bound for delays, which will prove superlinear scaling of delays, for arrival processes $A(t)$ given by a compound Poisson process with intensity λ . Packet sizes are independent and identically distributed with a distribution given by $Pr(P > \sigma) = e^{-\mu\sigma}$. The moment generating function of this arrival process satisfies

$$\frac{1}{\theta} \log E[e^{\theta(A(t)-A(s))}] = \frac{\lambda(t-s)}{\mu - \theta}$$

for any $\theta < \mu$ [22]. It follows from the Chernoff bound that the process is an EBB arrival process, as defined in Eq. (3), where

$$r = \frac{\lambda}{\mu - \theta}, \quad M = 1.$$

We assume that the load factor satisfies

$$\frac{\lambda}{\mu} < C,$$

and we choose $\theta < \mu$ so that $r < C$.

III. THE $O(H \log H)$ UPPER BOUND

In this section we establish an $O(H \log H)$ upper bound on end-to-end delays for the network from Fig. 1 for packetized traffic. For simplicity of notation, we assume that the output link at each node operates at the same constant rate C and that cross traffic at each node satisfies the same EBB bound with

$$Pr\{A_h(t) - A_h(s) > r_c(t-s) + \sigma\} \leq M_c e^{-\theta_c \sigma}, \quad (4)$$

where A_h denotes the cross traffic arrivals to the h -th node. Arrivals of the through flow at the first node also satisfy an EBB bound with

$$\Pr\{A_0(t) - A_0(s) > r_0(t-s) + \sigma\} \leq M_0 e^{-\theta_0 \sigma}. \quad (5)$$

We assume that the packet-size distribution of the through flow satisfies

$$\Pr\{P > \sigma\} \leq M_p e^{-\theta_p \sigma}. \quad (6)$$

We first express the EBB traffic model in the language of the stochastic network calculus. A *statistical envelope* for an arrival process A consists of a nonnegative, non-decreasing function $\mathcal{G}(t)$, which provides a bound on the arrivals in intervals of length t , together with an *error function* $\varepsilon(\sigma)$ [11, 30] that provides a bound on the probability that the envelope is violated on a given interval, such that for all $s < t$ and all σ ,

$$\Pr\{A(t) - A(s) > \mathcal{G}(t-s) + \sigma\} \leq \varepsilon(\sigma). \quad (7)$$

Envelopes characterize the statistical properties of traffic. An envelope for an aggregate of flows can capture statistical multiplexing [4]. The characterization of an EBB traffic flow satisfying Eq. (3) is equivalent to the statistical envelope

$$\mathcal{G}(t) = rt, \quad \varepsilon(\sigma) = M e^{-\theta \sigma}. \quad (8)$$

The network calculus uses the concept of service curves to characterize service available to a flow at a network node. A node offers a deterministic service curve S [12] if it satisfies for all $t \geq 0$ that

$$D(t) \geq A * S(t), \quad (9)$$

where

$$A * f(t) = \inf_{s \in [0, t]} \{A(s) + f(t-s)\}$$

denotes the min-plus convolution operator. Thus, a service curve offers a lower bound on the guaranteed service at a node. A *statistical service curve* [11] consists of a nonnegative nondecreasing function $\mathcal{S}(t)$ and an error function $\varepsilon(\sigma)$, such that for any $t \geq 0$,

$$\Pr\{D(t) < A * [\mathcal{S} - \sigma]_+(t)\} \leq \varepsilon(\sigma), \quad (10)$$

where $[x]_+ = \max(x, 0)$ denotes the positive part of the number x . With a statistical service curve, the probability that the deterministic guarantee from Eq. (9) is violated by more than σ is bounded by $\varepsilon(\sigma)$. To see that statistical service curves are a generalization of the deterministic counterpart, note that Eq. (9) is recovered by using an error function $\varepsilon(\sigma) = 1$ if $\sigma \leq 0$ and $\varepsilon(\sigma) = 0$ otherwise. We say that a node provides an *EBB service*, if

$$\mathcal{S}(t) = R_s t, \quad \varepsilon(\sigma) = M_s e^{-\theta_s \sigma},$$

for some nonnegative constants R_s , M_s and θ_s is a statistical service curve.

The service available to a flow at a node is determined by the capacity of the node, the cross traffic, the link scheduling algorithms, and packetization. We express the available service

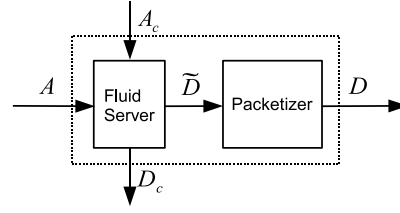


Fig. 2. A network node represented by a fluid server and a packetizer.

in terms of a service curve describing the link capacity that is left unused by the cross traffic. Such a ‘leftover service’ characterization assumes that cross traffic is transmitted with higher priority than the through flow, yielding a conservative bound for any work-conserving locally FIFO scheduling algorithms. In previous work [11], we have, under the assumption of fluid flow traffic, constructed a statistical service curve that characterizes the leftover service at a node. If traffic is packetized, then a packet cannot depart from the node until the entire packet has been processed. This introduces additional dependencies and delays.

A. Single Node Delays

We first derive a leftover service curve for packetized traffic and obtain a single node delay bound. The results will be used in the next subsection for an upper bound on end-to-end delays.

Denote the arrival and departure process of the through flow at a node by $A(t)$ and $D(t)$, respectively. We represent the node as the composition of a virtual fluid server and a packetizer, as shown in Fig. 2. The fluid server experiences through flow arrivals $A(t)$ and cross traffic $A_c(t)$. The output from the fluid server passes through a packetizer that introduces a random delay according to the packet size distribution. The following theorem provides a statistical service curve that accounts for cross traffic as well as packetization.

Theorem 1 Consider a node with a fixed capacity C with a work-conserving, locally FIFO scheduling algorithm. The aggregate cross flow arrivals are characterized by any statistical envelope $\mathcal{G}_c(t)$ with an error function $\varepsilon_c(\sigma)$ that satisfies

$$\int_0^\infty \varepsilon_c(u) du < \infty.$$

The packet size of the through flow be given by a stationary random variable P with

$$\Pr\{P > \sigma\} \leq \varepsilon_p(\sigma).$$

Then, for any choice of $\delta > 0$ and $\tau > 0$,

$$\mathcal{S}(t) = Ct - \mathcal{G}_c(t + \tau) - \delta(t + \tau)$$

is a service curve for the through flow, with error function

$$\varepsilon_s(\sigma) = \inf_{\sigma_c + \sigma_p = \sigma} \left\{ \frac{1}{\delta \tau} \int_{\sigma_c}^\infty \varepsilon_c(u) du + \frac{1}{E[P]} \int_{\sigma_p}^\infty \varepsilon_p(u) du \right\}.$$

The error function of the service curve, $\varepsilon_s(\sigma)$, depends on the error function of the cross traffic, $\varepsilon_c(\sigma)$, as well as the error function for the packet size, $\varepsilon_p(\sigma)$. Setting $\varepsilon_p(\sigma) = 0$ for all $\sigma > 0$ recovers the result for fluid flow traffic from [11]. The theorem requires to choose two parameters $\tau > 0$ and $\delta > 0$. The parameter τ plays the role of a discretization time step, and δ is used to enforce convergence in the estimates for the violation probability. In applications, the free parameters are chosen to optimize the resulting bounds.

Proof: Refer to Fig. 2. If, at time t , the node is serving a packet from the tagged flow, denote by $\tilde{Z}(t)$ the portion of this packet that has already been completed. We interpret

$$\tilde{D}(t) = D(t) + \tilde{Z}(t) \quad (11)$$

as the output from the fluid server. By Theorem 3 of [11], we have the following input-output relationship at the fluid server:

$$Pr\{\tilde{D}(t) < A(t) * [S(t) - \sigma]_+\} \leq \frac{1}{\delta\tau} \int_{\sigma}^{\infty} \varepsilon_c(u) du. \quad (12)$$

The distribution of $\tilde{Z}(t)$ is determined by the distribution of the size of the packet that is being processed in (the packetizer of) the node at time t . Denote the size of this packet by $Z(t)$. The probability density of this distribution is computed by weighting the density of P with the packet size. In other words, if $f_P(y)$ is the probability density of a random packet P , then the density $Z(t)$ is given by $zf_P(z)/E[P]$. By conditioning on the size of $Z(t)$, we compute

$$\begin{aligned} Pr\{\tilde{Z}(t) > \sigma\} &= \int_0^{\infty} Pr\{\tilde{Z}(t) > \sigma | Z(t) = z\} \frac{zf_P(z)}{E[P]} dz \\ &= \frac{1}{E[P]} \int_{\sigma}^{\infty} (z - \sigma) f_P(z) dz \\ &\leq \frac{1}{E[P]} \int_{\sigma}^{\infty} \varepsilon_p(z) dz. \end{aligned} \quad (13)$$

In the second line, we have used that $\tilde{Z}(t)$ conditioned on $Z(t) = z$ is uniformly distributed on $[0, z]$, and in the third line, we have integrated by parts. The claim follows by inserting Eqs. (12) and (13) into Eq. (11):

$$\begin{aligned} &Pr\{D(t) < A * [S - (\sigma_c + \sigma_p)]_+(t)\} \\ &\leq Pr\{\tilde{D}(t) < A * [S - \sigma_c]_+(t)\} + Pr\{\tilde{Z}(t) > \sigma_p\}. \end{aligned}$$

■

The theorem enables us to derive a delay bound where packetized traffic complies to the EBB model. The analysis proceeds in three steps. First, we show that the leftover service curve for EBB cross traffic and exponentially bounded packet sizes yields an EBB service. Second, we show that EBB arrivals with an EBB service yield delays with an exponential decay. Lastly, we combine the results of the first two steps.

To begin, consider that aggregate arrivals from the cross traffic satisfy Eq. (4) and that packets of the through flow satisfy the exponential bound in Eq. (6). We use Theorem 1 with $\mathcal{G}_c(t) = r_c t$, $\varepsilon_c(t) = M_c e^{-\theta_c t}$, $\delta = C - R_s - r_c$ and $\tau = (\theta_c(C - R_s))^{-1}$. This gives the service curve

$$\mathcal{S}(t) = R_s t - \theta_c^{-1},$$

with error function

$$\varepsilon_s(\sigma) = \underbrace{\left(\frac{e\theta_c(C - R_s)M_c}{\theta_s(C - R_s - r_c)}\right)^{\theta_s/\theta_c} \left(\frac{M_p}{\theta_s E[P]}\right)^{\theta_s/\theta_p}}_{=: M_s} e^{-\theta_s(\sigma - \theta_c^{-1})}, \quad (14)$$

where

$$\theta_s = (\theta_c^{-1} + \theta_p^{-1})^{-1}. \quad (15)$$

For the error function, we have used Lemma 1 from the appendix. Since \mathcal{S} with $\varepsilon(\sigma)$ is a statistical service curve if and only if $\mathcal{S} - x$ with $\varepsilon(\sigma + x)$ is a statistical service curve, we can replace σ with $\sigma + \theta_c^{-1}$ to obtain the EBB service

$$\mathcal{S}(t) = R_s t, \quad \varepsilon_s(\sigma) = M_s e^{-\theta_s \sigma}, \quad (16)$$

where $R_s < C - r_c$, and θ_s and M_s as in Eq. (14).

In the second step, we show that the delay distribution of an EBB flow at a node offering an EBB service has an exponential decay. We will apply the delay bound of Theorem 2 in [11], which states for an arrival process A with a statistical envelope as in Eq. (7), and a node offering a statistical service curve as in Eq. (10), a probabilistic delay bound is given by

$$Pr\{W(t) > w(\sigma)\} \leq \varepsilon(\sigma),$$

where

$$w(\sigma) = \inf\{w : \mathcal{S}(t - \tau + w) \geq \mathcal{G}(t) + \delta t + \sigma, \forall t \geq 0\}, \quad (17)$$

and

$$\varepsilon(\sigma) = \inf_{x+y=\sigma} \left\{ \varepsilon_s(x) + \frac{1}{\delta\tau} \int_x^{\infty} \varepsilon_g(u) du \right\}. \quad (18)$$

The bound holds for any fixed choice of the two free parameters $\tau > 0$ and $\delta > 0$. For an EBB traffic envelope (with parameters r_0, θ_0, M_0) at a node offering an EBB service (with parameters R_s, θ_s, M_s) we set

$$\theta_w = (\theta_s^{-1} + \theta_0^{-1})^{-1}, \quad (19)$$

and choose $\delta = R_s - r_0$ and $\tau = (R_s \theta_w)^{-1}$. If $R_s > r_0$, we can compute Eq. (17) as

$$w(\sigma) = \frac{\sigma + \theta_w^{-1}}{R_s},$$

For the violation probability, we evaluate Eq. (18) with Lemma 1 as

$$\begin{aligned} \varepsilon(\sigma) &= \inf_{\sigma_s + \sigma_0 = \sigma} \left\{ M_s e^{\theta_s \sigma_s} + \frac{M_0}{\delta \tau \theta_0} e^{-\theta_0 \sigma_0} \right\} \\ &= e \underbrace{\left(\frac{M_s \theta_s}{\theta_w} \right)^{\theta_w / \theta_s} \left(\frac{R_s M_0}{R_s - r_0} \right)^{\theta_w / \theta_0}}_{=: M_w} e^{-\theta_w \sigma - 1}. \end{aligned} \quad (20)$$

Replacing σ with $R_s w + \theta_w^{-1}$, we obtain the exponential bound

$$Pr\{W(t) > w\} \leq M_w e^{-\theta_w R_s w}. \quad (21)$$

In the third and final step of the single-node delay analysis, we insert the service guarantee from Eq. (16) into the delay bound of Eq. (21) and obtain values for the constants R_s , θ_w , and M_w . We choose the service rate R_s to satisfy

$$r_0 < R_s < C - r_c.$$

Inserting the formula for θ_s from Eq. (15) into Eq. (19) yields

$$\theta_w = (\theta_c^{-1} + \theta_p^{-1} + \theta_0^{-1})^{-1}. \quad (22)$$

It remains to estimate the constant M_w . Using the expression for M_s from Eq. (14) in Eq. (20), and using Lemma 2 to estimate

$$\left(\frac{e\theta_c}{\theta_w} \right)^{\theta_w / \theta_c} \left(\frac{\theta_p}{\theta_w} \right)^{\theta_w / \theta_p} \leq e + 1,$$

we obtain that Eq. (21) holds with

$$M_w \leq e(e+1)\bar{M} \left(\frac{C}{C-r_c-R_s} \right)^{\theta_w / \theta_c} \left(\frac{R_s}{R_s-r_0} \right)^{\theta_w / \theta_0}, \quad (23)$$

where

$$\bar{M} = \max \left\{ M_0, M_c, \frac{M_p}{\theta_p E[P]} \right\}.$$

The free parameter R_s can be eliminated from Eqs. (21) and (23) with Lemma 3 from the appendix.

B. End-to-end Delays

We will use the single node delay results from the previous subsection to derive a delay bound for the packetized through flow in the network in Fig. 1. For this, we construct a statistical service curve, referred to as *network service curve*, describing the end-to-end service for the through flow along its entire path in terms of the service curves at each node on the path. With a network service curve available, we can compute end-to-end performance measures such as end-to-end delay bounds using single-node results.

We will show that, as long each node on the path offers an EBB service, the network also offers an EBB service. For simplicity, we take all per-node service curves to be identical.

Consider more generally a flow that on its path through the network receives at the h -th node a statistical service curve

$S^h(t)$ with error function $\varepsilon_h(t)$. For given values of the parameters $\tau > 0$ and $\delta > 0$, set $\tilde{S}^h(t) = S^h(t - \tau)$ so that for all $t \geq 0$ and all σ

$$Pr\{D(t) < A * [\tilde{S}^h - \sigma]_+(t)\} \leq \varepsilon^h(\sigma)$$

Also set $\tilde{S}_{-\delta}^h(t) = \tilde{S}^h(t) - \delta t$.

With these definitions, Theorem 1 of [11] guarantees that, for every choice of the free parameters $\tau > 0$ and $\delta > 0$, the function

$$S_{net}(t) = \tilde{S}^1 * \tilde{S}_{-\delta}^2 \dots \tilde{S}_{-(H-1)\delta}^H(t + \tau) \quad (24)$$

is a network service curve for the flow that satisfies Eq. (10) with error function

$$\varepsilon_{net}(\sigma) = \inf_{\sigma_1 + \dots + \sigma_H = \sigma} \left\{ \varepsilon^H(\sigma_H) + \sum_{h=1}^{H-1} \frac{1}{\delta \tau} \int_{\sigma_h}^{\infty} \varepsilon^h(u) du \right\}. \quad (25)$$

The network service curve is constructed by a min-plus convolution of the per-node service curves, where at each successive node, the service curve is degraded by an additional rate δ .

If each S^h provides an EBB service with the same parameters R_s, M_s, θ_s , then

$$\tilde{S}_{-(h-1)\delta}^h(t) = (R_s - (h-1)\delta)(t - \tau)$$

is a latency-rate function. Using the min-plus algebra [10], the convolution of these latency rate functions is again a latency-rate function, with rate equal to the minimum of the rates $(R_s - (H-1)\delta)$ and latency equal to the sum of the latencies $(H\tau)$. We choose $R_{net} < R_s$, set

$$\theta_{net} = \frac{\theta_s}{H}, \quad (26)$$

and apply the theorem with $\delta = (R_s - R_{net})/(H-1)$ and $\tau = (HR_s\theta_{net})^{-1}$. Replacing $H-1$ by H whenever it seems convenient, we obtain the network service curve

$$S_{net}(t) = R_{net}t - \theta_{net}^{-1}$$

with error function

$$\varepsilon_{net}(\sigma) \leq \underbrace{H^2 e M_s \frac{R_s}{R_s - R_{net}}}_{=: M_{net}} e^{-\theta_{net} \sigma - 1}. \quad (27)$$

As in the computation of the EBB leftover service curve in Eq. (16), we conclude that

$$S_{net}(t) = R_{net}t, \quad \varepsilon_{net}(\sigma) = M_{net} e^{-\theta_{net} \sigma}$$

provides an EBB network service, with parameters R_{net}, θ_{net} , and M_{net} as given above.

We can now combine the results from this section to establish our first main result, an upper bound on the end-to-end delay of packetized traffic of a through flow in a network, where through and cross traffic satisfy EBB bounds.

Theorem 2 Given a network as shown in Fig. 1 with H nodes where each node has capacity C . Let the arrival process of the through flow be characterized by Eqs. (5) and (6), and let the cross traffic at each node satisfy Eq. (4). Set

$$\gamma = \frac{\theta_c^{-1}}{\theta_c^{-1} + \theta_p^{-1}}, \quad \bar{M} = \max\left\{M_0, M_c, \frac{M_p}{\theta_p E[P]}\right\}.$$

Then, for every choice of R_{net} such that

$$r_0 < R_{net} < C - r_c,$$

the end-to-end delays satisfy

$$Pr\{W_{net}(t) > w\} \leq M_w e^{-R_{net}\theta_w w}, \quad (28)$$

with parameters

$$\begin{aligned} \theta_w &= (H(\theta_c^{-1} + \theta_p^{-1}) + \theta_0^{-1})^{-1}, \quad (29) \\ M_w &= e(e+1)H^2\bar{M}\left(\frac{C}{C-r_c-R_{net}}\right)^{(1+\gamma)(1-\theta_w/\theta_0)} \times \\ &\quad \times \left(\frac{R_{net}}{R_{net}-r_0}\right)^{\theta_w/\theta_0}. \end{aligned}$$

The parameter γ that appears in the theorem describes the relative impact of the cross traffic and the packet-size distribution on the service available to the through flow at a node. In the absence of cross traffic, we have $\gamma = 0$, and if the through traffic is fluid-flow, we have $\gamma = 1$. The free parameter R_{net} can be eliminated with the help of Lemma 3 from the appendix.

Proof: For the decay rate of the violation probability, we compute

$$\begin{aligned} \theta_w &= (\theta_{net}^{-1} + \theta_0^{-1})^{-1} \\ &= (H\theta_s^{-1} + \theta_0^{-1})^{-1} \quad (30) \\ &= (H(\theta_c^{-1} + \theta_p^{-1}) + \theta_0^{-1})^{-1}. \end{aligned}$$

Here, the first line used Eq. (19) with θ_{net} in place of θ_s , the second line used Eq. (26), and the last line used Eq. (15). Similarly, we obtain for the multiplicative constant

$$\begin{aligned} M_w &\leq \left(\frac{H^2 e \theta_{net} R_s}{\theta_w (R_s - R_{net})}\right)^{\theta_w/\theta_{net}} \left(\frac{e \theta_c (C - R_s)}{\theta_s (C - R_s - r_c)}\right)^{H\theta_w/\theta_c} \times \\ &\quad \times \left(\frac{\theta_p}{\theta_s}\right)^{H\theta_w/\theta_p} \left(\frac{e R_{net}}{R_{net} - r_0}\right)^{\theta_w/\theta_0} \bar{M}. \quad (31) \end{aligned}$$

We have used M_w from Eq. (20) with θ_{net} , M_{net} and R_{net} in place of θ_s , M_s and R_s , then expressed M_{net} in terms of M_s with Eq. (27), and finally used Eq. (14) to determine M_s . To keep the formulas manageable, we have not yet inserted the expressions we obtained for θ_{net} and θ_s in Eq. (30).

We need to show that this expression for M_w is no larger than the value given in the statement of the theorem. Since $\theta_w/\theta_{net} < 1$, we can replace the exponent of H by 2. We next collect the terms that involve the exponential decay rates. By

Eq. (30) and Lemma 2,

$$\begin{aligned} \left(\frac{\theta_{net}}{\theta_w}\right)^{\theta_w/\theta_{net}} \left(\frac{e\theta_c}{\theta_s}\right)^{H\theta_w/\theta_c} \left(\frac{\theta_p}{\theta_s}\right)^{H\theta_w/\theta_p} \\ = \left(\frac{e\theta_c}{H\theta_w}\right)^{H\theta_w/\theta_c} \left(\frac{\theta_p}{H\theta_w}\right)^{H\theta_w/\theta_p} \\ \leq e + 1. \end{aligned}$$

Furthermore, we choose

$$R = R_{net} + \frac{1}{1+\gamma}(C - R_{net} - r_c)$$

and apply the arithmetic-geometric mean inequality to estimate

$$\begin{aligned} \left(\frac{R_s}{R_s - R_{net}}\right)^{\theta_w/\theta_{net}} \left(\frac{C - R_s}{C - R_s - r_c}\right)^{\gamma\theta_w/\theta_{net}} \\ \leq \left(\frac{C}{C - R_{net} - r_c}\right)^{(1+\gamma)\theta_w/\theta_{net}}. \end{aligned}$$

The theorem follows by inserting these estimates into Eq. (31). \blacksquare

Let us briefly turn to the question of performing a numerical evaluation. To obtain a bound on the $(1-\varepsilon)$ -quantile of the end-to-end delay, we fix an acceptable violation probability ε with $0 < \varepsilon < 1$. Typically, ε is very small, perhaps 10^{-6} , but we will occasionally also consider larger values, such as $\varepsilon = 0.5$. Setting

$$Pr\{W_{net}(t) > w\} = \varepsilon,$$

and solving for w in Theorem 2 results in the bound

$$w_{net}(\varepsilon) \leq \frac{1}{\theta_w R_{net}} \log \frac{M_w}{\varepsilon}, \quad (32)$$

which can then be minimized over the free parameter R_{net} . Because of the dependence of M_{net} on R_{net} , the minimization problem cannot be solved analytically. We thus propose the following iterative procedure. In the first step, we select the initial value

$$R_{net}^{(0)} = r_0 + \frac{\theta_w}{\theta_0}(C - r_c - r_0).$$

From this, we compute a preliminary bound $w_{net}^{(0)}$ on the quantile $w_{net}(\varepsilon)$, using Eq. (32) with $R_{net}^{(0)}$ in place of R_{net} . To improve this bound, we then compute

$$R_{net}^{(1)} = C - r_c - \frac{(1 - \theta_w/\theta_0)(C - r_c - r_0)}{1 + (1 + \gamma)\theta_w(C - r_c - r_0)w_{net}^{(0)}}, \quad (33)$$

using Eq. (51) from the appendix. We use this value in Eq. (32) to obtain a new bound $w_{net}^{(1)}$. The procedure can be iterated to obtain choices of R_{net} and w_{net} that are arbitrarily close to optimal, but we find that a single iteration generally suffices for our purposes.

C. Scaling of Upper Bounds

We next investigate the scaling behavior of end-to-end delays as a function of the path length and we explore the tail of the delay distribution. In the next section, the upper bounds will be compared to the lower bounds, for the purpose of assessing the quality of our derived bounds.

C.1 Long Paths ($H \rightarrow \infty$)

We analyze how the quantiles of the end-to-end delay scale with the number of nodes H . From Theorem 2, we obtain the explicit bound

$$w_{net}(\varepsilon) \leq \frac{1}{\theta_w R_{net}} \left\{ 2 \log H + \log \frac{e(e+1)\overline{M}}{\varepsilon} + (1+\gamma) \left(1 - \frac{\theta_w}{\theta_0}\right) \log \left(\frac{C}{C - r_c - R_{net}} \right) + \frac{\theta_w}{\theta_0} \log \left(\frac{R_{net}}{R_{net} - r_0} \right) \right\}. \quad (34)$$

Since

$$\theta_w^{-1} = H(\theta_c^{-1} + \theta_p^{-1}) + \theta_0^{-1},$$

it is clear that

$$w_{net}(\varepsilon) = O(H \log H).$$

A closer inspection of Eq. (34) gives us explicit values for the prefactor. For any choice of $R_{net} < C - r_c$ that does not depend on H , we obtain that

$$w_{net}(\varepsilon) \leq 2 \frac{\theta_c^{-1} + \theta_p^{-1}}{R_{net}} H \log H + O(H).$$

It follows that

$$\lim_{H \rightarrow \infty} \frac{w_{net}(\varepsilon)}{H \log H} \leq 2 \frac{\theta_c^{-1} + \theta_p^{-1}}{C - r_c}. \quad (35)$$

C.2 Large Delays ($w \rightarrow \infty$)

We study the tail behavior of the delay bounds for the single-node and the multi-node case. By Lemma 3 from the appendix, we can choose the free parameter R_s in the delay bound of Eq. (21) to obtain for a single node

$$Pr\{W(t) > w\} \leq O(w^{\theta_w/\theta_c}) e^{-\theta_w(C-r_c)w},$$

where θ_w is given by Eq. (22). In particular, the exponential decay rate satisfies

$$\lim_{w \rightarrow \infty} -\frac{1}{w} \log Pr\{W(t) > w\} \geq \frac{C - r_c}{\theta_c^{-1} + \theta_p^{-1} + \theta_0^{-1}}.$$

Likewise, we can choose the free parameter R_{net} in Theorem 2 to obtain for the end-to-end delay

$$Pr\{W(t) > w\} \leq O(w^{(1+\gamma)(1-\theta_w/\theta_0)}) e^{-\theta_w(C-r_c)w},$$

where θ_w and γ are given by Theorem 2. Here, the exponential decay rate satisfies

$$\lim_{w \rightarrow \infty} -\frac{1}{w} \log Pr\{W_{net}(t) > w\} \geq \frac{C - r_c}{H(\theta_c^{-1} + \theta_p^{-1}) + \theta_0^{-1}}. \quad (36)$$

Note that, in both cases, the application of Lemma 3 provides explicit expressions for the tail probability.

IV. THE $\Omega(H \log H)$ LOWER BOUND

In this section we construct an example of a network that satisfies the assumptions of Section II and prove that in this network the end-to-end delay of packets is bounded from below by $\Omega(H \log H)$. The example consists of a tandem network with H nodes, as shown in Fig. 1, where each node is a workconserving FIFO server with an infinite buffer that operates at a constant rate C . The network is traversed by a single flow without any cross traffic at the nodes.

The arrivals from the flow to the first node are described by a compound Poisson process [22], where packets arrive at times T_1, T_2, \dots according to a Poisson process with rate λ , and the size of each packet is independently exponentially distributed with average $1/\mu$. The service time of any given packet is proportional to its size and hence identical at each node. We assume that the load factor $\rho = \lambda/(\mu C)$ satisfies $\rho < 1$. This ensures that the backlog process at each node is stable and the delay distribution converges to a steady state.

A. Main Result

Theorem 3 *Given a network with the assumptions stated above. Let W_{net} denote the steady-state end-to-end delay of a packet along the path through the network. Then, for every $0 < z < 1$,*

$$Pr \left\{ W_{net} \leq \frac{H}{\mu C} \log \left(\frac{H}{2b|\log z|} \right) \right\} \leq z, \quad (37)$$

where

$$b = \inf_{0 < \theta < \rho} \frac{1}{\theta} \left| \log(z(1-z)(1-\theta/\rho)(1+\theta)) \right|. \quad (38)$$

An equivalent statement to the theorem is that the quantiles of the delay satisfy

$$w_{net}(z) \geq \frac{H}{\mu C} \log \left(\frac{H}{2b|\log z|} \right).$$

We can interpret the end-to-end delay of a packet as having two components. The first is the pure processing time, with average $H/(\mu C)$, which clearly grows linearly with the number of nodes. The second contribution is the time the packet spends waiting in the queues at nodes $h = 1, \dots, H$. The theorem implies that at downstream nodes the waiting time of a typical packet dominates its processing time.

The lower bound in Eq. (37) holds for the entire distribution of W_{net} , including the mean delay. The lower bound remains valid in more general networks with cross traffic and with nodes of different capacities C_1, \dots, C_H , provided that the tail of the packet size distribution of the through flow decays no faster than exponentially, the inter-arrival distance has finite mean, and the capacity of a positive fraction of the nodes does not exceed C . The crucial assumption is that the processing time of a given packet at any node is proportional to its size.

Proof: Consider a scenario where the network is started at time $t = 0$ with empty queues, and number the packets in order of their arrival to the first node of the network by $n = 1, 2, \dots$. Let $X_n = T_n - T_{n-1}$ denote the time between the arrival of the $(n-1)$ -st and the n -th packet at the first node, and let P_n be the size of the n -th packet.

For the purpose of the proof, we rescale the unit of traffic by $1/\mu$ so that the average packet size is 1, and the unit of time by $1/(\mu C)$ so that the rate of the server is 1. In these units, the inter-arrival distances X_n are exponentially distributed with rate ρ , and the packet lengths P_n are exponentially distributed with rate 1.

Denote by $W_{net,n}$ the total delay experienced by the n -th packet on its path through the network. The packet arrives to the first node at time

$$T_n = \sum_{i=1}^n X_i,$$

and departs from last node as soon as it and all prior packets have been processed at all nodes, but certainly not before

$$\sup_{j=1, \dots, n} \left\{ \sum_{i=1}^j X_i + HP_j + \sum_{i=j+1}^n P_i \right\}.$$

Subtracting T_n from this expression, we obtain for the end-to-end delay

$$W_{net,n} = \sup_{j=1, \dots, n} \left\{ HP_j - \sum_{i=j+1}^n (X_i - P_i) \right\}.$$

To derive a lower bound on the distribution of $W_{net,n}$, we split it into two pieces that will be estimated separately:

$$W_{net,n} \geq \sup_{j=1, \dots, n} \left\{ HP_j - b(n-j) \right\} - \sup_{j=1, \dots, n} \left\{ \sum_{i=j+1}^n (X_i - P_i - b) \right\}. \quad (39)$$

Since the network is started with empty queues at time $t = 0$, $W_{net,n}$ is stochastically increasing in n and its distribution converges monotonically to the steady-state delay distribution W_{net} .

For the first supremum in Eq. (39), we use that the processing times P_j are independent and identically distributed to compute

$$\begin{aligned} Pr \left(\sup_{j=1, \dots, n} \left\{ HP_j - b(n-j) \right\} \leq w \right) \\ = \prod_{j=0}^{n-1} \left\{ 1 - Pr \left(Y > \frac{w + bj}{H} \right) \right\}, \end{aligned}$$

where Y is a random variable with the same distribution as the P_j . To estimate the product, we take logarithms and use that

$\log(1-x) \leq -x$ to obtain

$$\begin{aligned} \log Pr \left(\sup_{j=1, \dots, n} \left\{ HY - b(n-j) \right\} \leq w \right) \\ \leq - \sum_{j=0}^{n-1} Pr \left(Y > \frac{w + bj}{H} \right) \\ \leq - \frac{H}{b} \int_{w/H}^{(w+bn)/H} Pr(Y \geq y) dy. \quad (40) \end{aligned}$$

In the last line, we have taken advantage of the fact that the terms of the sum decrease with j to estimate the sum by an integral. For the exponential distribution, the integral evaluates to

$$\begin{aligned} \log Pr \left(\sup_{j=1, \dots, n} \left\{ HP_j - b(n-j) \right\} \leq w \right) \\ \leq - \frac{H}{b} e^{-w/H} (1 - e^{-nb/H}). \quad (41) \end{aligned}$$

Equating the right hand side to $2 \log z$ and solving for w , we obtain

$$\begin{aligned} Pr \left(\sup_{j=1, \dots, n} \left\{ HP_j - b(n-j) \right\} \leq H \log \frac{H(1 - e^{-nb/H})}{2b|\log z|} \right) \\ \leq z^2. \quad (42) \end{aligned}$$

Estimating the second supremum in Eq. (39) is a classical problem for which many techniques are available. The sum has the same distribution as a random walk consisting of n independent steps. Since $b > \frac{1}{\rho} - 1 = E[X_n - P_n]$, the random walk has a negative drift, it almost surely escapes to $-\infty$, and the supremum is bounded uniformly in n . To obtain an explicit bound, we consider the Markov chain

$$Z_j = \prod_{i=n-j+1}^n e^{\theta(X_i - P_i - b)}, \quad j = 1, \dots, n-1.$$

We compute the moment-generating function of $X_n - P_n$ as

$$E[e^{\theta(X_n - P_n)}] = \frac{1}{(1 - \theta/\rho)(1 + \theta)}, \quad (\theta < \rho).$$

Let θ be the minimizer in Eq. (38). Then Z_1, \dots, Z_{n-1} form a nonnegative supermartingale

$$E[Z_{j+1} | Z_j] = E[e^{\theta(X_{n-j} - P_{n-j} - b)}] \cdot Z_j \leq Z_j.$$

We invoke Doob's maximal inequality [16] (p. 496) to see that

$$Pr \left(\sup_{j=1, \dots, n-1} Z_j \geq 1 \right) \leq E[Z_1].$$

By the definition of Z_1, \dots, Z_{n-1} , and the choice of b and θ , this implies

$$Pr \left(\sup_{j=1, \dots, n} \sum_{i=j+1}^n (X_i - P_i - b) > 0 \right) \leq z(1-z). \quad (43)$$

Note that the term for $j = n$ on the left hand side corresponds to an empty sum that does not contribute to the probability.

To complete the proof, we combine Eq. (42) with Eq. (43) to bound the right hand side of Eq. (39) and arrive at

$$\Pr\left\{W_{net,n} \leq H \log\left(\frac{H(1 - e^{-nb/H})}{2b|\log z|}\right)\right\} \leq z.$$

The theorem follows by taking $n \rightarrow \infty$. \blacksquare

B. Scaling of Lower Bounds

We now explore the scaling behavior implied by Theorem 3, and compare the obtained lower bound to the upper bounds from Subsection IV-B. To enable a comparison with the upper bounds, we express the compound Poisson arrival process without cross traffic used in this section in terms of the model from Section III. The parameters for the through flow and cross traffic arrivals and the packet-size distribution used in Theorem 2 are given by

$$\begin{aligned} \theta_0 < \mu, & \quad r_0 = \frac{\lambda}{\mu - \theta_0} < C, & \quad M_0 = 1, \\ \theta_c = \infty, & \quad r_c = 0, & \quad M_c = 1, \\ \theta_p = \mu, & \quad E[P] = \mu^{-1}, & \quad M_p = 1. \end{aligned} \quad (44)$$

B.1 Long Paths ($H \rightarrow \infty$)

From Theorem 3 we obtain that the quantiles of the end-to-end delay satisfy

$$w_{net}(z) \geq \frac{H}{\mu C} \log H - O(H). \quad (45)$$

We thus conclude that typical delays grow as

$$W_{net} = \Theta(H \log H).$$

If we compare the lower bound in Eq. (45) to the upper bound from Eq. (35), with the parameters as given in Eq. (44), we see that upper and lower bounds only differ by a factor of two, thus validating the quality of the bounds in Theorems 2 and 3. Note that, since the bounds hold for the entire distribution, the $\Theta(H \log H)$ scaling and the ratio of upper to lower bounds also extends to the mean delay.

B.2 Large Delays ($w \rightarrow \infty$)

For the tail of the delay distribution, Theorem 3 implies that

$$w_{net}(1 - \varepsilon) \geq \frac{H}{\mu C} \log \varepsilon^{-1} - O(1), \quad (\varepsilon \rightarrow 0).$$

It follows that

$$\Pr\{W_{net}(t) > w\} \geq \Omega\left(e^{-\frac{\mu}{H} C w}\right).$$

In particular, the exponential decay rate satisfies

$$\lim_{w \rightarrow \infty} -\frac{1}{w} \log \Pr\{W_{net}(t) > w\} \leq \frac{\mu C}{H}. \quad (46)$$

When we insert the parameters from Eq. (44) into the upper bound from Eq. (36) we obtain that the exponential decay rate must be at least $\frac{C}{H\mu^{-1} + \theta_0^{-1}}$. Since $\theta_0 < \mu$ can be chosen arbitrarily, we can combine the upper and lower bounds to obtain the exact rate of the exponential decay of the end-to-end delays, given by

$$\lim_{w \rightarrow \infty} -\frac{1}{w} \log \Pr\{W(t) > w\} = \frac{\mu C}{H}. \quad (47)$$

C. The Role of the Packet Size Distribution

We next investigate how different packet size distributions impact the scaling of end-to-end delays. For this comparison, we abandon the Poisson arrival process considered earlier in this section in favor of a fixed (deterministic) inter-arrival distance. The constant interarrival distance emphasizes the impact of the packet-size distribution. It also simplifies our computations, and allows a concise comparison of the scaling behavior of different distributions.

As in the proof of Theorem 3, we work in units where the link capacity is $C = 1$, and the mean packet size $E[Y] = 1$. Arrivals are evenly spaced, i.e., the inter-arrival distance is given by $X_n = 1/\rho$. We consider three examples of packet-size distributions: exponential (light-tailed), Pareto (heavy-tailed), and Bernoulli (deterministically bounded).

Exponential: To obtain a lower bound for W_{net} , we use Eq. (39) with b set equal to the packet spacing. Then the second supremum in Eq. (39) is guaranteed to be nonpositive, and can be neglected. For the first supremum, we obtain from Eq. (41) that

$$\log \Pr(W_{net} \leq w) \leq -\frac{H}{b} e^{-w/H}.$$

Setting the right hand side equal to $\log z$ and solving for w leads to

$$w_{net}(z) \geq H \log\left(\frac{H}{b|\log z|}\right). \quad (48)$$

Note that this model falls within the scope of the $O(H \log H)$ delay bound discussed in Section III.

Pareto: Next we consider the situation where the packet size distribution follows a Pareto law

$$\Pr(P > y) = \left(\frac{y_{min}}{y}\right)^\alpha, \quad y \geq y_{min}.$$

The parameter α determines the decay of the tail of the distribution, with smaller values of α signifying a heavier tail, and y_{min} is the minimum packet size. We assume that $\alpha > 1$ so that the distribution has a finite mean and choose $y_{min} = (\alpha - 1)/\alpha$ so that the mean packet size becomes

$$E[P] = \frac{\alpha}{\alpha - 1} y_{min} = 1.$$

For the lower bound, we set $b = 1/\rho$ and insert Eq. (40) into Eq. (39) to obtain

$$\begin{aligned} \log Pr(W_{net} \leq w) &\leq -\frac{H}{b} \int_{w/H}^{(w+bn)/H} \left(\frac{y_{min}}{y}\right)^\alpha dy \\ &\xrightarrow{n \rightarrow \infty} -\frac{(Hy_{min})^\alpha}{b(\alpha-1)w^{\alpha-1}}. \end{aligned}$$

Setting the right hand side equal to $\log z$ and solving for w yields for the quantiles

$$w_{net}(z) \geq \frac{(Hy_{min})^{\alpha/(\alpha-1)}}{(b(\alpha-1)|\log z|)^{1/(\alpha-1)}}. \quad (49)$$

Thus, typical end-to-end delays show at least a power-law growth in the number of nodes. The closer α is to 1, i.e., the heavier the tail of the packet size distribution, the more rapid the growth of the end-to-end delay with the number of nodes. Note that even for large values of α , the $\Omega(H^{\alpha/(\alpha-1)})$ growth observed for the Pareto packet-size distribution always dominates the $\Theta(H \log H)$ growth observed for packet size distributions with exponential tails.

Bernoulli: Lastly, we consider a packet-size distribution that is deterministically bounded. Suppose that there are two packet sizes, $y_{max} > y_{min} > 0$, where large packets occur with some small frequency p , i.e.,

$$Pr(P = y_{max}) = p, \quad Pr(P = y_{min}) = 1 - p.$$

The mean packet size is given by

$$E[P] = p y_{max} + (1 - p) y_{min}.$$

We clearly have the deterministic bound on the end-to-end delay

$$Hy_{min} \leq W_{net} \leq Hy_{max}.$$

If we choose $b = 1/\rho - y_{min}$, then the second supremum in Eq. (39) is nonnegative, and we compute

$$\begin{aligned} Pr(W_{net} \leq w) &\leq \prod_{j=0}^{n-1} Pr(HP - bj \leq w) \\ &\leq (1-p)^{\#\{j:w+bj < Hy_{max}\}} \\ &\xrightarrow{n \rightarrow \infty} (1-p)^{(Hy_{max}-w)/b}. \end{aligned}$$

Solving for the quantiles provides the bound

$$w_{net}(z) \geq Hy_{max} - \frac{b|\log z|}{|\log(1-p)|}. \quad (50)$$

Thus, the difference between typical delays and the deterministic worst-case delay remains bounded as the number of nodes grow large. In other words, typical delays on a long path are essentially determined by the processing time of the largest packets. This finding holds up for bounded packet-size distributions in general.

V. NUMERICAL EXAMPLES

In this section we illustrate the upper and lower bounds on the end-to-end delay by numerical examples. As in Section IV, we consider a tandem network of H nodes, each representing a FIFO link with fixed capacity C , with no cross traffic. Arrivals are packetized, and the packet sizes are drawn from i.i.d random variables P_1, P_2, \dots . The service time of the n -th packet at each node is given by P_n/C . Packets arrive at times T_1, T_2, \dots , where the inter-arrival distances $X_n = T_n - T_{n-1}$ are independent and identically distributed in Example 1, and constant in Examples 2 and 3.

A. Example 1: Lower and Upper Bounds

We consider the scenario from Theorem 3. Packets arrive to the network as a Poisson process. The link capacity is given by $C = 100$ Mbps, and the average size of packets is $\mu^{-1} = 400$ Bytes [24]. For a given load factor ρ , we determine the arrival rate by $\lambda = \rho\mu C$. For computing upper bounds on the end-to-end delay from Theorem 2 we set $\gamma = 0$ and $\bar{M} = 1$.

We focus on quantiles $w_{net}(z)$ where z is very close to 1. To simulate $w_{net}(z)$ for $z = 1 - 10^{-6}$, we start with an empty network and run the simulations until 10^8 packets have completed service at node H , storing the 100 largest observed values of the end-to-end delay at each node. We use the smallest of these values as our estimate for the z -quantile of the end-to-end delay.

In Fig. 3 we show the end-to-end delay bounds as a function of the number of nodes H in the network, when the load factor is low ($\rho = 0.1$) and high ($\rho = 0.9$). The figures illustrate the quantitative relationship between the upper and lower bounds and the simulations. For the chosen range of H , which is already larger than typical routes in a packet network, the graphs appear to grow linearly. This indicates that, for path lengths encountered in practice, a linear growth of delays may be a suitable heuristic, and that analytical models that show a linear growth can be justified.

In Fig. 4 we evaluate the delays for fixed path length ($H = 25$) as the load factor ρ approaches one. In order to capture the blow-up of the delays as $\rho \rightarrow 1$, we use a logarithmic scale on the vertical axis. In addition to the bounds and the simulations, we include Vinogradov's asymptotic formula for the average end-to-end delay ($2H|\log(1-\rho)|$ as $\rho \rightarrow 1$) [27]. The simulations show a significant increase in the end-to-end delay only at values of the load factor well above 90%. Vinogradov's result captures the blow-up as $\rho \rightarrow 1$ rather well, even though it applies to the mean rather than the z -th quantile, but has no useful relationship to the simulations for smaller values of ρ . On the other hand, the upper and lower bounds correctly predict the order of magnitude of the delays seen in the simulations at values of the load factor $\rho < 0.9$, but the lower bound fails to capture the blow-up, while the upper bound over-estimates the rate of blow-up. Thus, the upper and lower bound capture the scaling of delays as $H \rightarrow \infty$ but may become loose as $\rho \rightarrow 1$.

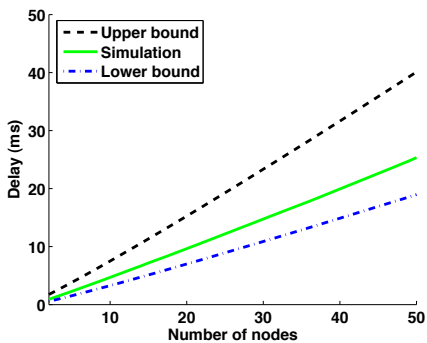
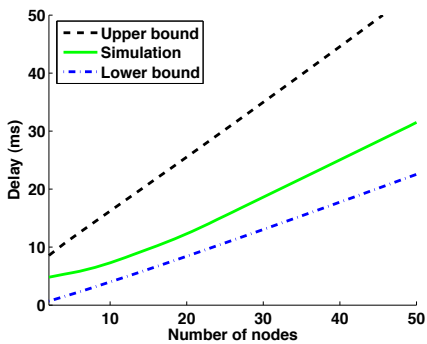
(a) Load factor $\rho = 0.1$ (b) Load factor $\rho = 0.9$

Fig. 3. Example 1: End-to-end delay $w_{net}(z)$ as a function of the number of nodes H for two values of the load factor. (quantile $z = 1 - 10^{-6}$, link capacity $C = 100$ Mbps, mean packet size $\mu^{-1} = 400$ Bytes.)

B. Example 2: Lower Bounds for Packet Size Distributions

In this example we illustrate the impact of the packet size distribution on lower bounds for the median of the end-to-end delays (that is, we set $z = 0.5$). We consider three different packet-size distributions: An exponential distribution ($\mu = 1$), a heavy-tailed Pareto distribution with $\alpha = 1.5$ and $P = 1/3$, and a Bernoulli distribution where a small fraction $p = 0.1$ of packets has size $y_{\max} = 2$ while the remaining packets have size $y_{\min} = 0.8889$. We use the lower bounds in Eq. (48), Eq. (49) and Eq. (50) for varying number of nodes H and fixed load factor $\rho = 0.75$. For the purpose of comparison, we use dimensionless variables, where the link capacity is $C = 1$, the average packet size is $E[Y] = 1$, and the distance between consecutive packets is $\rho^{-1} = 4/3$. Also included in the plot is the expected value of the pure processing time.

Fig. 5 shows that different packet size distributions give rise to fundamentally different scaling behavior. The upper curve shows the power-law growth of the end-to-end delay of the Pareto distribution; here, the power is $\alpha/(\alpha - 1) = 2$. The middle curve shows the slightly superlinear $\Theta(H \log H)$ growth of the delay bounds for the exponential packet-size distribution. For the Bernoulli distribution, we observe linear scaling, caused by the linear growth of the worst-case delay. Note that the

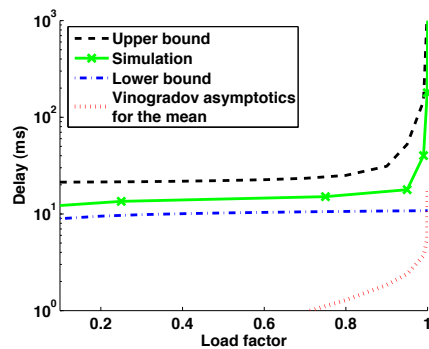


Fig. 4. Example 1: End-to-end delay $w_{net}(z)$ as a function of the load factor ρ (quantile $z = 1 - 10^{-6}$, $H = 25$ nodes, link capacity $C = 100$ Mbps, mean packet size $\mu^{-1} = 400$ Bytes.)

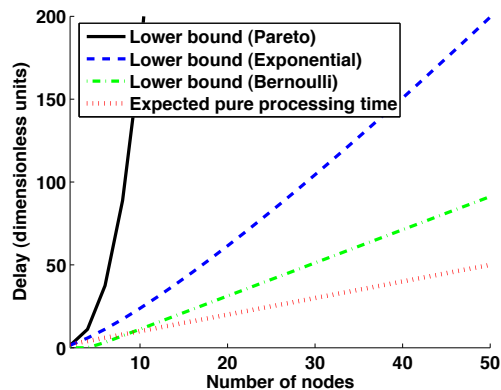


Fig. 5. Example 2: Lower bounds for the median of the end-to-end delay for different packet size distributions as a function of the number of nodes. (quantile $z = 0.5$, link capacity $C = 1$, load factor $\rho = 0.75$, mean packet size 1.)

growth rate lies well above the average rate $E[Y] = 1$ and is determined by the maximum packet length $y_{\max} = 2$.

C. Example 3: Truncated Packet Size

Since packet sizes are limited in practice, we want to evaluate if the superlinear scaling of delays is maintained when a maximum packet size is enforced. We continue with the network configuration and assumptions of Example 2, and set $C = 1$. We consider a constant interarrival spacing of packets of 1 time unit. We only evaluate the exponential packet size distribution with $\mu = 1$. With this choice, the utilization is equal to $\rho = 0.75$. We compare the lower bound for the end-to-end delays with a truncated distribution, where packet sizes cannot exceed k times the average packet size of the original distribution. Specifically, for the truncation, we work with the following packet size distribution:

$$Pr(Y < y) = \begin{cases} 1 - e^{-y}, & \text{if } y < k, \\ 1, & \text{otherwise.} \end{cases}$$

We ignore the impact of the truncation on the mean of the distribution. In Fig. 6 we depict the lower bounds of the median delays ($z = 0.5$) for the original distribution without truncation), and the results for the truncation at $k = 2, 4$, and 6. We

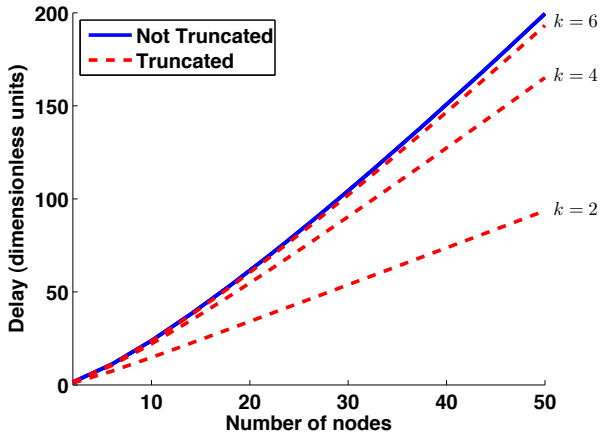


Fig. 6. Example 3: Sensitivity of lower bounds for the median end-to-end delay to truncation of (exponentially bounded) packet size distribution (quantile $z = 0.5$, link capacity $C = 1$, load factor $\rho = 0.75$, mean packet size 1.)

observe that imposing limits on the packet lengths can have a noticeable impact on the scaling of the lower bounds. At the same time, even for a truncation of $k = 4$, the scaling behavior is close to that of the original distribution. As a point of reference, measurements of the average size of IP datagrams report an average around 400 Bytes [24], while the maximum IP datagram size is often set to 1500 Bytes. Thus, networks today meet the conditions under which superlinear delay scaling can manifest themselves.

VI. CONCLUSIONS

We have shown that in a network with exponentially bounded arrivals and service, and where each packet maintains the same service time at each traversed node, end-to-end delays grow as $\Theta(H \log H)$ with the number of nodes. This is quite different from the $\Theta(H)$ scaling obtained when service at nodes is statistically independent. We proved a lower bound for delays in a tandem network without cross traffic where packets arrive according to a Poisson process and have exponentially distributed service times. The $\Theta(H \log H)$ scaling of delays followed by extending a $O(H \log H)$ upper bound for fluid-flow traffic to a packetized arrival description. The $\Theta(H \log H)$ bounds remain valid in networks with cross traffic and with different packet-size distributions, so long as all arrival processes satisfy suitable exponential bounds. An open question is whether there are scenarios with purely fluid-flow arrivals where delays grow as $\Omega(H \log H)$. We believe this to be the case, but suspect it may require to analyze rather subtle correlations between the arrivals from cross flows at different nodes.

ACKNOWLEDGMENTS

The research in this paper is supported in part by the National Science Foundation and the Natural Sciences and Engineering Research Council of Canada.

APPENDIX

I. TECHNICAL LEMMAS

The following lemma from [11] is used in Section III.

Lemma 1 For any positive numbers M_k, θ_k ($k = 1, \dots, K$) and every $\sigma \geq 0$,

$$\inf_{\sigma_1 + \dots + \sigma_K = \sigma} \sum_{k=1}^K M_k e^{-\theta_k \sigma_k} = \prod_{k=1}^K \left(\frac{M_k \theta_k}{\bar{\theta}} \right)^{\bar{\theta}/\theta_k} e^{-\bar{\theta} \sigma},$$

$$\text{where } \bar{\theta} = \left(\sum_{k=1}^K \frac{1}{\theta_k} \right)^{-1}.$$

The next lemma is used to simplify explicit delay bounds in Section III.

Lemma 2 For any positive numbers a_k, x_k ($k = 1, \dots, K$) with $\sum_{k=1}^K a_k \geq e$ and $\sum_{k=1}^K x_k \leq 1$,

$$\prod_{k=1}^K \left(\frac{a_k}{x_k} \right)^{x_k} \leq \sum_{k=1}^K a_k.$$

Proof: Let a_1, \dots, a_n be given. We compute

$$\begin{aligned} \max_{\sum x_k \leq 1} \prod_{k=1}^K \left(\frac{a_k}{x_k} \right)^{x_k} &\leq \max_{0 \leq s \leq 1} \max_{\sum x_k = s} \prod_{k=1}^K \left(\frac{a_k}{x_k} \right)^{x_k} \\ &= \max_{0 \leq s \leq 1} \left(\frac{\sum a_i}{s} \right)^s \\ &= \sum_{k=1}^K a_k. \end{aligned}$$

In the first line, we have divided the maximization over x_1, \dots, x_k into two steps. In the second line, we have used the Lagrange multiplier method to identify the maximizing choice

$$x_k = \frac{a_k s}{\sum a_i},$$

which we then inserted into the objective function. Since $\sum a_i \geq e$ by assumption, the resulting function is increasing in s for $0 \leq s \leq 1$, and so the maximum is assumed for $s = 1$. ■

The third lemma is used to motivate the choice of the free parameters in our study of scaling properties of the upper bounds on delays in Section III.

Lemma 3 Let $R_2 > R_1 \geq 0$ and $\beta_1, \beta_2 \geq 0$ be given constants. Then, for every $x \geq 0$, there exists an R with $R_1 < R < R_2$ such that

$$\begin{aligned} &(R - R_1)^{-\beta_1} (R_2 - R)^{-\beta_2} e^{-Rx} \\ &\leq (R_2 - R_1)^{-(\beta_1 + \beta_2)} \left\{ e^{\frac{(R_2 - R_1)x + \beta_1 + \beta_2}{\beta_2}} \right\}^{\beta_2} e^{-R_2 x}. \end{aligned}$$

Proof: We will show that the inequality holds for

$$R = R_2 - \frac{\beta_2(R_2 - R_1)}{\beta_1 + \beta_2 + (R_2 - R_1)x}. \quad (51)$$

Under the change of variables

$$r = \frac{R - R_1}{R_2 - R_1}, \quad s = \frac{\beta_1}{\beta_1 + \beta_2}, \quad y = \frac{R_2 - R_1}{\beta_1 + \beta_2}x,$$

the left hand side of the claim transforms into

$$\left\{ (R_2 - R_1)^{-1} r^{-s} (1 - r)^{-(1-s)} e^{-ry} \right\}^{\beta_1 + \beta_2} e^{-R_1 x},$$

and the choice of R in Eq. (51) transforms into $r = \frac{s+y}{1+y}$. An elementary manipulation gives $r^{-s} e^{-ry} \leq e^{1-s} e^{-y}$. Plugging in, we obtain

$$r^{-s} (1 - r)^{-(1-s)} e^{-ry} \leq \left(e \frac{y+1}{1-s} \right)^{1-s} e^{-y}.$$

Scaling back to the original variables, we obtain the right hand side of the claim. ■

REFERENCES

- [1] R. Agrawal, R. L. Cruz, C. Okino, and R. Rajan. Performance bounds for flow control protocols. *IEEE/ACM Transactions on Networking*, 7(3):310–323, June 1999.
- [2] S. Ayyogun and R. Cruz. A service-curve model with loss and a multiplexing problem. In *Proceedings of the 24th IEEE International Conference on Distributed Computing System (ICDCS)*, pages 756–765, March 2004.
- [3] F. Baskett, K. M. Chandy, R. R. Muntz, and F. G. Palacios. Open, closed and mixed networks of queues with different classes of customers. *Journal of the ACM*, 22(2):248–260, April 1975.
- [4] R. Boorstyn, A. Burchard, J. Liebeherr, and C. Oottamakorn. Statistical service assurances for traffic scheduling algorithms. *IEEE Journal on Selected Areas in Communications*, 18(12):2651–2664, December 2000.
- [5] J. Y. Le Boudec and P. Thiran. *Network Calculus*. Springer Verlag, Lecture Notes in Computer Science, LNCS 2050, 2001.
- [6] O. Boxma. On a tandem queueing model with identical service times at both counters. part 1,2. *Advances in Applied Probability*, 11(3):616–659, 1979.
- [7] A. Burchard, J. Liebeherr, and F. Ciucu. On $\Theta(H \log H)$ scaling of network delays. In *Proc. of IEEE Infocom*, pages 1866–1874, May 2007.
- [8] A. Burchard, J. Liebeherr, and S. D. Patek. A min-plus calculus for end-to-end statistical service guarantees. *IEEE Transactions on Information Theory*, 52(9):4105 – 4114, September 2006.
- [9] S. B. Calo. Delay properties of message channels. In *Proc. IEEE ICC, Boston, Mass.*, pages 43.5.1–43.5.4, 1979.
- [10] C.-S. Chang. *Performance Guarantees in Communication Networks*. Springer Verlag, 2000.
- [11] F. Ciucu, A. Burchard, and J. Liebeherr. Scaling properties of statistical end-to-end bounds in the network calculus. *IEEE Transactions on Information Theory*, 52(6):2300–2312, June 2006.
- [12] R. L. Cruz. Quality of service guarantees in virtual circuit switched networks. *IEEE Journal on Selected Areas in Communications*, 13(6):1048–1056, August 1995.
- [13] M. Fidler. An end-to-end probabilistic network calculus with moment generating functions. In *IEEE 14th International Workshop on Quality of Service (IWQoS)*, pages 261–270, June 2006.
- [14] P. Le Gall. The overall sojourn time in tandem queues with identical successive service times and renewal input. *Stochastic Processes and Their Applications*, 52(1):165–178, August 1994.
- [15] P. W. Glynn and W. Whitt. Departures from many queues in series. *Annals of Applied Probability*, 1:546–572, 1991.
- [16] G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford University Press, 2001.
- [17] Y. Jiang. A basic stochastic network calculus. In *ACM Sigcomm*, pages 123–134, September 2006.
- [18] Y. Jiang and P. J. Emstad. Analysis of stochastic service guarantees in communication networks: A server model. In *Proceedings of the International Workshop on Quality of Service (IWQoS)*, pages 233–245, June 2005.
- [19] Y. Jiang and Y. Liu. *Stochastic Network Calculus*. Springer, 2008.
- [20] F. I. Karpelevitch and A. Ya. Kreinin. Asymptotic analysis of queueing systems with identical service. *Journal of Applied Probability*, 33(1):267–281, 1996.
- [21] F. P. Kelly. Networks of queues with customers of different types. *Journal of Applied Probability*, 3(12):542–554, September 1975.
- [22] F. P. Kelly. Notes on effective bandwidths. In *Stochastic Networks: Theory and Applications*. (Editors: F.P. Kelly, S. Zachary and I.B. Ziedins) Royal Statistical Society Lecture Notes Series, 4, pages 141–168. Oxford University Press, 1996.
- [23] C. Li, A. Burchard, and J. Liebeherr. A network calculus with effective bandwidth. *IEEE/ACM Transactions on Networking*, 15(6):1442–1453, December 2007.
- [24] S. McCreary and K. Claffy. Trends in wide area IP traffic patterns. In *Proceedings of 13th ITC Specialist Seminar on Internet Traffic Measurement and Modeling*, September 2000.
- [25] P. I. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, February 1956.
- [26] O. P. Vinogradov. A multiphase system with identical service. *Soviet Journal of Computer and Systems Sciences.*, 24(2):28–31, March 1986.
- [27] O. P. Vinogradov. A multiphase system with many servers and identical service times. *Stochastic Processes and Their Applications, MIEM*, 24:42–45, 1989 (in Russian).
- [28] O. P. Vinogradov. On certain asymptotic properties of waiting time in a multiserver queueing system with identical times. *SIAM Theory of Probability and Its Applications (TVP)*, 39(4):714–718, 1994.
- [29] O. Yaron and M. Sidi. Performance and stability of communication networks via robust exponential bounds. *IEEE/ACM Transactions on Networking*, 1(3):372–385, June 1993.
- [30] Q. Yin, Y. Jiang, S. Jiang, and P. Y. Kong. Analysis on generalized stochastically bounded bursty traffic for communication networks. In *Proceedings of IEEE Local Computer Networks (LCN)*, pages 141–149, November 2002.

PLACE
PHOTO
HERE

Almut Burchard received the Ph.D. degree in Mathematics from the Georgia Institute of Technology in 1994. She was on the faculty of the Department of Mathematics at Princeton University (1994-1998) and the University of Virginia (1994-2005). She is currently a Professor of Mathematics at the University of Toronto.

PLACE
PHOTO
HERE

Jörg Liebeherr (S’88, M’92, SM’03, F’08) received the Ph.D. degree in Computer Science from the Georgia Institute of Technology in 1991. He was on the faculty of the Department of Computer Science at the University of Virginia from 1992–2005. Since Fall 2005, he is with the University of Toronto as Professor of Electrical and Computer Engineering and NorTel Chair of Network Architecture and Services.

PLACE
PHOTO
HERE

Florin Ciucu received the Ph.D. degree in Computer Science from the University of Virginia in 2007. In 2007/08, he was a Postdoctoral Fellow at the University of Toronto. Since Fall 2008, he is a Senior Scientist at the Deutsche Telekom Laboratories at the Technical University of Berlin.