On $\Theta(H \log H)$ Scaling of Network Delays

Almut Burchard
Department of Mathematics
University of Toronto

Jörg Liebeherr
Department of Electrical and
Computer Engineering
University of Toronto

Florin Ciucu
Department of Computer Science
University of Virginia

***Abstract*— A recent result in network calculus theory provided statistical delay bounds for exponentially bounded traffic that grow as $\mathcal{O}(H \log H)$ with the number of nodes on the network path.[1] In this paper we establish the corresponding lower bound which shows that under these assumptions, typical end-to-end delays can indeed grow as $\Theta(H \log H)$. The lower bound is obtained by analyzing the end-to-end delay in a tandem network. A critical assumption is that each packet maintains the same service time at each traversed node. The results of this paper provide conclusive evidence that, in general, delays have a qualitatively different scaling behavior than is suggested by a worst-case analysis or by assuming independence on the service obtained at network nodes.**

I. Introduction

The stochastic network calculus is a tool for deriving statistical delay bounds in networks. In analogy to the deterministic network calculus [1], [2], [3], arrivals are bounded by statistical arrival envelopes, and the service available to different flows at the nodes of the network is bounded from below by statistical (lower) service curves. The framework of the stochastic network calculus permits the consideration of packet networks where traffic and service types are characterized in terms of probability distributions. The probabilistic description of traffic and service also permits to capture and express the benefits of statistical multiplexing gain in packet networks [4].

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 service at each node is described in terms of service curves, the network service can be given as the min-plus algebra convolution of the per-node service curves. This result was established first in the context of a deterministic network calculus, where service at each node satisfies a given lower bound [5], [6]. Finding the corresponding composition result in a stochastic setting turned out to be hard and, for a long time, was limited to special cases and strong assumptions. In [7] it was shown that a straightforward application of the min-plus convolution yields a network service curves that deteriorates with time. Most available network service curves for a statistical setting were derived by either making strong assumptions on the properties of the network, or by modifying the definition of a service curve. Examples of the former can be found in [8], where delays at each node are assumed to satisfy a priori delay bounds, in [9], where it is assumed that a node discards traffic that exceeds a threshold, and in [10], which assumes that service at subsequent nodes is statistically independent. Examples of the latter include [7], which assumes that the statistical service description is made over time intervals, and [11], [12], which assumes sample path guarantees for service.

A recent study [13] presented the construction of a statistical network service curve for a network where traffic arrivals conform to the Exponentially Bounded Burstiness (EBB) model [14]. This coincides with the class of so-called linear bounded envelope processes introduced by Chang [3], which includes as special cases multiplexed regulated arrivals and many Markov-modulated processes but excludes long-range correlated or heavy-tailed traffic models. Traffic arrivals are modeled there as fluid-flow. Under these assumptions, it was shown for a tandem network of H nodes with (EBB) cross traffic at each node (see Figure 1) that delays grow no more than $\mathcal{O}(H \log H)$ with the number of nodes [13]. This scaling behavior is quite different to that obtained with other analytical methods. For example, the deterministic network calculus predicts a linear growth of end-to-end delays in these networks [2]. Delays in product form queueing networks [15] also scale linearly. Finally, a linear growth of delays is also obtained in a stochastic network calculus, when the service at nodes is assumed to be statistically independent [10].

In light of the different scaling properties found by other modeling approaches, the results in [13] raise two questions: *Under which assumptions on the network and the arrivals are $\mathcal{O}(H \log H)$ bounds on the delay valid? 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.

The main result of this paper is an answer to the second question. We show that the $\mathcal{O}(H \log H)$ bound on end-to-end delays from [13] cannot be improved upon without

[1]Throughout this paper we use the big-Oh or Landau notation for the asymptotic comparison of functions. For two sequences A_n and B_n, the notation $A_n = \mathcal{O}(B_n)$ 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 $A_n = \Theta(B_n)$.

additional assumptions. To demonstrate this, we construct an example of a network that satisfies the assumptions for the $\mathcal{O}(H \log H)$ upper bounds on delay and show that typical delays grow with $\Omega(H \log H)$. Concretely, we analyze the delays of packets in a tandem network of H identical nodes with no cross traffic. We prove for the example that the end-to-end delay of packets is bounded from below by $\Omega(H \log H)$. This lower bound on delay remains valid if the flow experiences cross traffic at each node.

We also consider the first question above, i.e., the domain of validity of the $\mathcal{O}(H \log H)$ bound. Different from the fluid flow service assumptions made in [13], we will construct service curves that incorporate packetization effects. We extend the $\mathcal{O}(H \log H)$ scaling bound on end-to-end delays to 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 queuing theory. In contrast with most of the queueing network literature (and also [10]), we do not assume that service times of a packet are independently regenerated at each traversed node.

The significance of the $\Theta(H \log H)$ scaling behavior of end-to-end delays stems from the linear scaling in other network models, in particular, networks with deterministically bounded arrivals and service, and networks with Poisson arrivals and independent exponentially distributed service times. With respect to the $\Theta(H)$ bound when the service satisfies deterministic bounds, 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 network calculus models with independent service, this paper shows that dispensing with the assumption on independent service changes the scaling behavior of network delays.

The paper may also contribute to a discussion when independence assumptions of service are justified. Originally, the independence assumption on service was made to obtain simple and solvable models for packet networks in terms a series of M/M/1 queueing systems [16]. The independence assumption has been empirically justified for networks with high-volume cross traffic, high loads, short network paths, high network connectivity, and other randomizing effects [17]. Conversely, the assumption is known to be optimistic in situations of low loads, long paths, light cross traffic, and heavy-tailed packet-size distributions.

The remainder of this paper is structured as follows. In Section II, we discuss in detail the network model used in this paper. In Section III, we construct an example of a network where delays grow with $\Omega(H \log H)$. In Section IV, we extend the $\mathcal{O}(H \log H)$ upper bound for delays from [13] to packetized arrival models. 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.

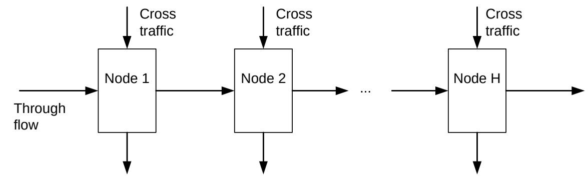


Fig. 1. A tandem network with cross traffic.

II. Network Model

We consider a network with H nodes in series (usually referred to as a tandem network), as shown in Figure 1. Each node represents a workconserving scheduler that determines the order in which arrivals from different flows are processed. Our analysis is valid for any 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. 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, 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.

Arrivals to a flow are modeled by a non-decreasing, left continuous process $A(t)$ with $A(t) = 0$ for $t \leq 0$. Arrivals may be either fluid-flow or packetized. If the arrivals are generated by a process that produces a packet of size $Y_n > 0$ at time T_n ($n = 1, 2, \ldots$), the arrival process is given by

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

which is clearly non-decreasing and left continuous. Departures from a node or a network are described similarly, and will be denoted by $D(t)$. The backlog is defined by $B(t) = A(t) - D(t)$, and the delay is defined by

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

The end-to-end delay experienced by the flow along its path through the network will be denoted by $W_H(t)$, where H signifies the number of nodes on the path. As $t \to \infty$, $W_H(t)$ converges in distribution to the steady-state delay, denoted here by W_H. We will analyze the distribution of W_H through its *quantiles*, defined for $0 < z < 1$ by

$$w_H(z) = \inf\Big\{w \geq 0 : \; P\big(W_H \leq w\big) \leq z\Big\} \,.$$

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

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

The arrivals from the flow to the first node are described by a compound Poisson process [18], where packets arrive at times $T_1, T_2, \ldots$ according to a Poisson process with rate λ, and the size of each packet is independently exponentially distributed with parameter μ. 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 can converge to a steady state. If $\rho \geq 1$, the backlog and delay even at the first node may grow arbitrarily large.

Theorem 1 *Given a network with the assumptions stated at the beginning of the section. Let W_H denote the steady-state end-to-end delay of a packet along the path through the network. There exists for each z with $0 < z < 1$ a constant $\gamma > 0$ that depends only on the load factor ρ and on the value of z such that*

$$Pr\left(W_H \leq \frac{H}{\mu C}\log(\gamma H)\right) \leq z\,.$$

An explicit estimate for γ will be provided in the proof. From the theorem we obtain that the quantiles of the delay satisfy

$$w_H(z) = \Omega(H \log H)\,. \tag{1}$$

In conjunction with Eq. (14) this implies that typical delays grow as $W_H = \Theta(H \log H)$ with the number of nodes.

Note that the end-to-end delay of a packet has two components. The first is the pure processing time, which clearly grows linearly with the number of nodes. The second contribution is the time the packet spends waiting in the queues at nodes $h = 1, \ldots, H$. The theorem indicates that at downstream nodes the waiting time of a typical packet dominates its processing time.

A. Related literature

The tandem network described at the beginning of the section is a variant of an M/M/1/ queueing network where the service times of each packet are identical in each queue. 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. We will give a brief overview of the published work in this area.

Boxma [19] 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 [20] 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 [21], he presents 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 [22], [23], he showed for exponentially 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 [24], [25], [26].

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 in the case of the compound Poisson arrival process described at the beginning of the section, finding the value of the distribution function $P(W_H > w)$ requires, for each w, to solve a transcendental equation and then compute an integral of the solution. In this light, our Theorem 1 adds to the above literature by giving explicit lower bounds that can be numerically evaluated for all values of H.

B. Proof of Theorem 1

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 network by $n = 1, 2, \ldots$. 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 Y_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 Y_n are exponentially distributed with rate 1. For later use, we compute the moment-generating function of $X_n - Y_n$ as

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

Denote by $W_{H,n}$ 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, at time

$$\sup_{j=1,\dots,n}\Big\{\sum_{i=1}^{j} X_i + HY_j + \sum_{i=j+1}^{n} Y_i\Big\}\,.$$

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

$$W_{H,n} = \sup_{j=1,\dots,n}\Big\{HY_j - \sum_{i=j+1}^{n}(X_i - Y_i)\Big\}\,.$$

To derive a lower bound on the distribution of $W_{H,n}$, we split it into two pieces

$$\begin{aligned} W_{H,n} \ \ &\geq \ \ \sup_{j=1,\dots,n}\Big\{HY_j - b(n-j)\Big\} \\ &\quad - \sup_{j=1,\dots,n}\Big\{\sum_{i=j+1}^{n}(X_i - Y_i - b)\Big\} \qquad (2) \end{aligned}$$

that we estimate separately. Here, $b > E(X_i - Y_i)$ is a constant that will be further specified below. Since the network is started with empty queues at time $t=0$, $W_{H,n}$ is stochastically increasing in n and its distribution converges monotonically to the steady-state delay distribution W_H.

For the first supremum in Eq. (2), we use that the processing times Y_j are independent and identically distributed to compute

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

where Y is a random variable with the same distribution as the Y_j. To estimate the product, we take logarithms and use that $\log(1-x) \leq -x$ to obtain

$$\begin{aligned} &\log Pr\Big(\sup_{j=1,\dots,n}\Big\{HY - b(n-j)\Big\} \leq w\Big) \\ &\qquad \leq \ -\sum_{j=0}^{n-1} Pr\Big(Y > \frac{w+bj}{H}\Big) \\ &\qquad \leq \ -\frac{H}{b}\int_{w/H}^{(w+bn)/H} Pr\big(Y \geq y\big)\,dy\,. \qquad (3) \end{aligned}$$

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

$$\begin{aligned} &\log Pr\Big(\sup_{j=1,\dots,n}\Big\{HY_j - b(n-j)\Big\} \leq w\Big) \\ &\qquad \leq \ -\frac{H}{b}e^{-w/H}\big(1-e^{-nb/H}\big)\,. \qquad (4) \end{aligned}$$

Equating the right hand side of Eq. (4) to $\log z$ and solving for w, we obtain

$$\begin{aligned} &Pr\Big(\sup_{j=1,\dots,n}\Big\{HY_j - b(n-j)\Big\} \leq H\log\frac{H(1-e^{-nb/H})}{b|\log z|}\Big) \\ &\qquad \leq z\,. \end{aligned}$$

Estimating the second supremum in Eq. (2) is a classical problem for which many techniques are available. The sum has the same distribution as a random walk consisting of n indepenent steps. If $b > E(X_n - Y_n)$ then the supremum is bounded uniformly in n, since the random walk has a negative drift and almost surely escapes to $-\infty$. Our approach is to consider the Markov chain

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

If b and θ are chosen so that $E\big(e^{\theta(X_i-Y_i-b)}\big) \leq 0$ then $Z_1,\dots,Z_{n-1}$ form a nonnegative supermartingale

$$\begin{aligned} E(Z_{j+1} \mid Z_j) \ &= \ E\Big(e^{\theta(X_{n-j}-Y_{n-j}-b)}Z_j \mid Z_j\Big) \\ &\leq \ Z_j\,. \end{aligned}$$

We invoke Doob's maximal inequality [27] (p. 496) to see that

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

By the definition of $Z_1,\dots,Z_{n-1}$, this implies

$$\begin{aligned} &Pr\Big(\sup_{j=1,\dots,n}\sum_{i=j+1}^{n}(X_i - Y_i - b) > 0\Big) \\ &\qquad \leq \ E\big(e^{\theta(X_n-Y_n-b)}\big) \\ &\qquad = \ \frac{e^{-b\theta}}{(1-\theta/\rho)(1+\theta)}\,. \qquad (5) \end{aligned}$$

Note that the term for $j=n$ in the first line corresponds to an empty sum that does not contribute to the probability.

To complete the proof, fix $0<z<1$ and choose

$$b = \inf_{0<\theta<\rho}\frac{1}{\theta}\Big|\log\big(z(1-z)(1-\theta/\rho)(1+\theta)\big|\,,$$

so that

$$Pr\Big(\sup_{j=1,\dots,n}\sum_{i=j+1}^{n}(X_i - Y_i - b) > 0\Big) \leq z(1-z)\,.$$

We combine Eq. (4) with Eq. (5) to bound the right hand side of Eq. (2) and arrive at

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

The theorem follows by setting $\gamma = (2b|\log z|)^{-1}$ and taking $n \to \infty$. □

C. The role of the packet size distribution

We next investigate how generalizing the packet size distribution impacts the scaling of end-to-end delays. As in the proof of Theorem 1, we work in units where the link capacity is $C = 1$, and the expected packet size $E(Y) = 1$. For simplicity, we will assume in this subsection that arrivals are evenly spaced, i.e., the inter-arrival distance is given by $X_n = 1/\rho$. This is not a serious restriction because under the mild assumption that the second moment of $[X - Y]_+$ is finite, the Strong Law of Large Numbers can be used to bound the second supremum in Eq. (2). 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_H, we use Eq. (2) with b set equal to the packet spacing. Then the second supremum in Eq. (2) is guaranteed to be nonnegative, and we obtain from Eq. (4) that

$$\log Pr\big(W_H \le w\big) \le -\frac{H}{b}e^{-w/H}\,.$$

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

$$w_H(z) \ge H \log\left(\frac{H}{b|\log z|}\right)\,. \tag{6}$$

We will see in the next section that this model, like the Poisson model discussed at the beginning of the section falls within the scope of the $\mathcal{O}(H\log H)$ delay bound discussed in Section IV.

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

$$Pr\Big(Y > y\Big) = \left(\frac{y_0}{y}\right)^{\alpha}, \quad y \ge y_0\,.$$

The parameter α determines the decay of the tail of the distribution, with smaller values of α signifying a heavier tail, while y_0 determines the scale. We assume that $\alpha > 1$ so that the distribution has a finite mean and choose $y_0 = (\alpha-1)/\alpha$ so that

$$E(Y) = \frac{\alpha}{\alpha-1}y_0 = 1\,.$$

For the lower bound, we set $b = 1/\rho$ and insert Eq. (3) into Eq. (2) to obtain

$$\begin{aligned}\log Pr\Big(W_H \le w\Big) \quad &\le \quad -\frac{H}{b}\int_{w/H}^{(w+bn)/H}\left(\frac{y_0}{y}\right)^{\alpha}dy \\ &\underset{n\to\infty}{\to} \quad -\frac{(Hy_0)^{\alpha}}{b(\alpha-1)w^{\alpha-1}}\,.\end{aligned}$$

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

$$w_H(z) \ge \frac{(Hy_0)^{\alpha/(\alpha-1)}}{\big(b(\alpha-1)|\log z|\big)^{1/(\alpha-1)}}\,. \tag{7}$$

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.

As for corresponding upper bounds on the end-to-end delay, we note that the results from [13] can be applied when $\alpha > 2$. We suspect that the resulting delay bounds may be rather pessimistic and are likely to grow with a higher power of H than actual delays. It is an open problem to modify these techniques to obtain any end-to-end delay bounds in the heavy-tailed regime $1 < \alpha \le 2$. A second problem is to find matching upper and lower bounds on the delay for all values of α that capture the exact scaling behavior of the end-to-end delay.

Bernoulli: For our last example, 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.,

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

The mean packet size is given by

$$E(Y) = p\,y_{\max} + (1-p)\,y_{\min}\,.$$

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

$$Hy_{\min} \le W_H \le Hy_{\max}\,.$$

If we choose $b = 1/\rho - y_{\min}$, then the second supremum in Eq. (2) is nonnegative, and we compute

$$\begin{aligned}Pr\Big(W_H \le w\Big) \quad &\le \quad \prod_{j=0}^{n-1} Pr\Big(HY - bj \le w\Big) \\ &\le \quad (1-p)^{\#\{j: w+bj < Hy_{\max}\}} \\ &\underset{n\to\infty}{\to} \quad (1-p)^{(Hy_{\max}-w)/b}\,.\end{aligned}$$

Solving for the quantiles provides the bound

$$w_H(z) \geq H y_{\max} - \frac{b|\log z|}{|\log(1-p)|}\,. \tag{8}$$

This demonstrates that the difference between typical delays and the maximal delay given by the deterministic upper bound remains bounded as the number of nodes becomes 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.

IV. The $\mathcal{O}(H \log H)$ upper bound

In this section we establish the validity of the $\mathcal{O}(H \log H)$ bound for the network from Figure 1. In previous work [13], we derived an $\mathcal{O}(H \log H)$ bound for a network with fluid-flow EBB arrivals of through and cross traffic. However, the lower bounds from the previous section assume a packetized traffic model. Therefore, the derivations of the upper bounds from [13] do not immediately apply, and must be adapted to the network at hand.

We start by recalling some key definitions from the stochastic network calculus. A *statistical envelope* for an arrival process A is a pair of functions, an *envelope function* $\mathcal{G}(t)$ and an *error function* $\varepsilon(\sigma)$ [13], [28], such that for any $0 \leq s < t$ and any $\sigma \geq 0$

$$Pr(A(t) - A(s) > \mathcal{G}(t-s) + \sigma) \leq \varepsilon(\sigma)\,. \tag{9}$$

The service given to a flow by a node or a network is determined by several factors, including the capacity of the node, the characteristics of cross traffic, the scheduling algorithm, and packet size distribution. In the stochastic network calculus, the service is described by a statistical service curve, which gives a lower bound on the departures in terms of the arrivals. A *statistical service curve* [7], [13] is a pair of functions, a *service curve* $\mathcal{S}(t)$ and an error function $\varepsilon(\sigma)$, such that for any $t \geq 0$,

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

Here, the min-plus convolution $f * g$ of two functions is given by

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

and $[x]_+ = \max(x, 0)$ denotes the positive part of the number x. The service given to a flow on its entire path is called a *network service curve*.

We say that an arrival process is EBB [14], if it satisfies a constant-rate statistical envelope with an exponential error function,

$$\mathcal{G}(t) = rt\,,\ \varepsilon(\sigma) = M e^{-\theta\sigma}\,, \tag{11}$$

where r, M, and θ are positive constants. Correspondingly, we say that the service provided to a flow or an aggregate of flows by a node or a network is EBB, if it admits a constant-rate service curve with an exponential error function

$$\mathcal{S}(t) = Rt\,,\ \varepsilon(\sigma) = M e^{-\theta\sigma}\,, \tag{12}$$

where R, M, and θ are positive constants. These definitions are equivalent to the characterization of arrivals and service through moment-generating functions in [3]. The statistical envelope $\mathcal{G}$ will be called an EBB envelope, and the statistical service curve will be called an EBB service curve.

An important example of EBB service is the *leftover* service curve that is available to a flow at a node where the arrivals from all other flows are EBB. More precisely, consider a flow arriving to a workconserving server of capacity C that uses a locally FIFO scheduling algorithm. Assume that the aggregate arrivals from all other flows have an EBB envelope as in Eq. (11) with parameters r, M_g, and θ_g, and that the packet-size distribution of the flow under consideration satisfies an exponential tail estimate

$$Pr(Y \geq \sigma) \leq M_p e^{-\theta_p \sigma}\,.$$

Then it can be shown that for any choice of $R < C - r$, the flow receives an EBB service curve as in Eq. (12), where the exponent $\theta = (\theta_g^{-1} + \theta_p^{-1})^{-1}$ lies between $\theta_g/2$ and $\theta_p/2$ and the constant M can be bounded in terms of the choice of R and the other parameters. For fluid-flow arrivals the formula for the leftover service curve reduces to Theorem 3 of [13]; in the absence of cross traffic one can take $R = C$. We provide a proof sketch for tandem networks as considered in Theorem 1 at the end of the section.

Assume that each node in the tandem network from Figure 1 has capacity at least C, and that the arrivals from each cross flow have EBB envelopes with identical parameters r, M, and θ. Assume furthermore that the through flow has an EBB envelope with a rate $r_0 < C - r$ and is either fluid-flow, or packetized with an exponentially bounded packet size distribution. Since the flow receives a EBB service at each node, the results from Section 4 of [13] imply that end-to-end delay bounds satisfy

$$Pr(W_H > w) \leq H^{\alpha} M' e^{-\theta' w/H}\,. \tag{13}$$

where the constants M' and θ' depend on the parameters of the flows but not on H, and α is a fixed positive power (typically $\alpha \leq 3$). Setting the right hand side equal to z and solving for w yields for the quantiles of the delay distribution the bound

$$w_H(z) = \mathcal{O}(H \log H)\,. \tag{14}$$

We conclude this section by explicitly computing a delay bound for the model considered in Theorem 1. We obtain the EBB characterization of the compound Poisson process from its moment-generating function

$$E\big[e^{\theta A(t)}\big] = e^{\frac{\lambda\theta}{\mu-\theta} t}\,,$$

see [27]. The Chernoff bound implies that Eq. (11) with

$$r = \frac{\lambda}{\mu - \theta} t, \;\; M = 1 \tag{15}$$

defines a statistical envelope that satisfies Eq. (9).

We claim that at the service available to the flow at each node is EBB with

$$\mathcal{S}(t) = Ct, \;\; \varepsilon(\sigma) = \rho e^{-\mu\sigma} ,$$

where $\rho = \lambda/(\mu C)$ is the load factor. To see this, note that if $B(t) > 0$, then $D(t) = A(\underline{t}) + C(t - \underline{t}) - Y^*(t)$, where $Y^*(t)$ is the remaining workload of the packet being served at time t. It follows from the memoryless property of the exponential distribution that for each $\sigma \geq 0$

$$Pr\big(D(t) < A * [\mathcal{S} - \sigma]_+(t)\big) = \rho e^{-\mu\sigma} ,$$

as claimed.

To construct the network service curve, we need to choose two parameters, a time step τ_0 that is used for discretization, and a small rate $\delta > 0$ that is used to relax statistical envelopes and service curves. At each node the service curve is replaced by

$$\mathcal{S}(t) = Ct, \;\; \varepsilon_s(\sigma) = \rho e^{\mu C \tau_0} e^{-\mu\sigma} ,$$

which satisfies in addition to Eq. (10) that

$$Pr\Big(D(t) < A * [\mathcal{S} - \sigma]_+ (t + \tau_0)\Big) \leq \varepsilon_s(\sigma) . \tag{16}$$

From Theorem 1 of [13], we obtain the network service curve

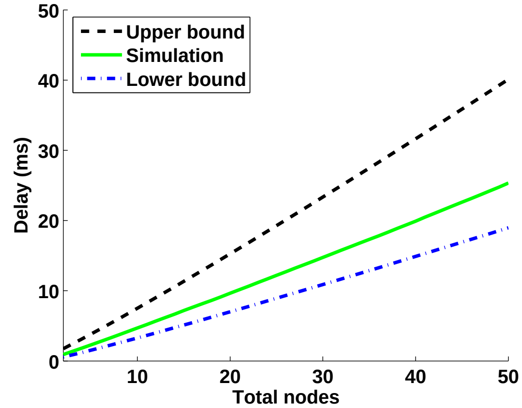
$$\mathcal{S}_{\rm net}(t) = (C - (H-1)\delta)t, \;\; \varepsilon_{\rm net}(\sigma) = H \frac{\rho e^{\mu C \tau_0}}{\mu \delta \tau_0} e^{-\mu\sigma/H} ,$$

which again satisfies Eq. (16). The desired delay bound now follows from Theorem 2 of [13]. After optimizing over τ_0 and setting $\delta = (C - r)/H$, we arrive at Eq. (13) with $\alpha = 2$,

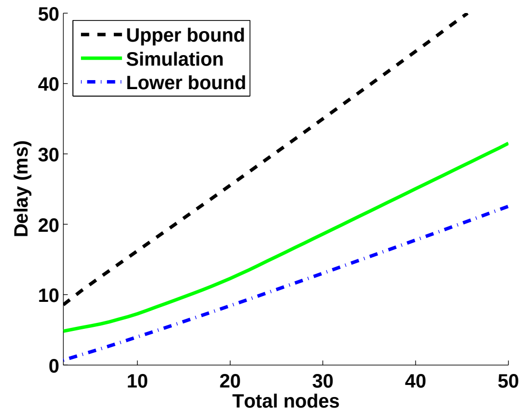
$$\begin{aligned} M' &= e\rho^{H\theta/(\mu + H\theta)} \frac{C}{C - r} , \\ \theta' &= \left(r + \frac{(C - r)}{H} \right) \cdot \frac{\mu\theta}{\mu + H\theta} . \end{aligned}$$

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 III, we consider a tandem network of H nodes, each representing a FIFO server of a fixed capacity C, with no cross traffic. Arrivals are packetized, and the packets $Y_1, Y_2, \ldots$ are independent and identically distributed random variables. The service time of the n-th packet at each node is given by Y_n/C, i.e., it is identical at each node along the path of the flow. Packets arrive at times $T_1, T_2, \ldots$, where the inter-arrival distances $X_n = T_n - T_{n-1}$ are independent and identically distributed in Example 1, and constant in Example 2.



(a) Load factor $\rho = 0.1$



(b) Load factor $\rho = 0.9$

Fig. 2. End-to-end delay $w_H(z)$ 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.)

A. Example 1

We consider the scenario from Theorem 1. 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 [29]. For a given load factor ρ, we determine the arrival rate by $\lambda = \rho\mu C$.

We focus on quantiles $w_H(z)$ where $z = 1 - \varepsilon$ is very close to 1. To simulate $w_H(z)$, 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 Figure 2 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. Note that values of the delays, both bounds and simulations, are quite similar at

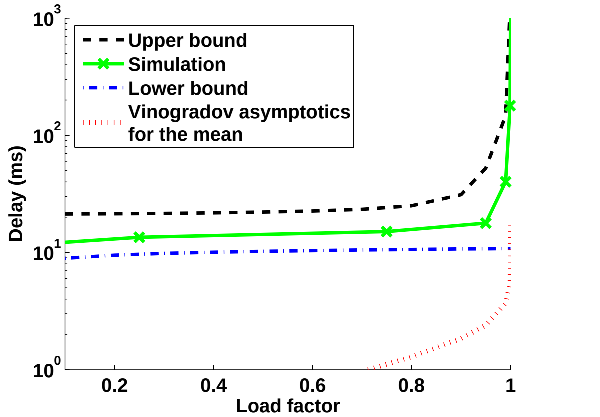


Fig. 3. End-to-end delay $w_H(z)$ as a function of the load factor ρ (path length $H = 25$ nodes, $z = 1 - 10^{-6}$, $C = 100$ Mbps, $\mu^{-1} = 400$ Bytes.)

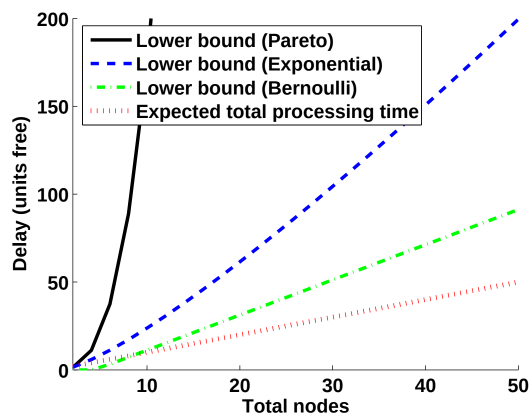


Fig. 4. Lower bounds for the median of the end-to-end delay for different packet size distributions as a function of the number of nodes. ($z = .5$, $C = 1$, $\rho = 0.75$, mean packet size 1, packet spacing $1/\rho$).

low and high load factors. 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 Figure 3 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 $\rho \to 1$, 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 $\rho \to 1$) [22]. 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 $\rho \to 1$ 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 overestimates the rate of blow-up. Thus, the upper and lower bound capture the scaling of delays as $H \to \infty$ but may become loose as $\rho \to 1$.

B. *Example 2*

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 $y_0 = 1/3$, 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. (6), Eq. (7) and (8) 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 $1/\rho = 4/3$. Also included in the plot is the expected value of the pure processing time.

Figure 4 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 sightly 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 growth rate lies well above the average rate $E(Y) = 1$ and is determined by the maximum packet length $y_{\max} = 2$.

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 known $\Theta(H)$ 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 adapting a previously obtained $\mathcal{O}(H \log H)$ upper bound 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.

REFERENCES

[1] R. Cruz, "A calculus for network delay, parts I and II," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 114–141, Jan. 1991.

[2] J. Y. Le Boudec and P. Thiran, *Network Calculus*. Springer Verlag, Lecture Notes in Computer Science, LNCS 2050, 2001.

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

[4] R. Boorstyn, A. Burchard, J. Liebeherr, and C. Oottamakorn, "Statistical service assurances for traffic scheduling algorithms," *IEEE Journal on Selected Areas in Communications. Special Issue on Internet QoS*, vol. 18, no. 12, pp. 2651–2664, December 2000.

[5] R. Agrawal, R. L. Cruz, C. Okino, and R. Rajan, "Performance bounds for flow control protocols," *IEEE/ACM Transactions on Networking*, vol. 7, no. 3, pp. 310–323, June 1999.

[6] J. Y. Le Boudec, "Application of network calculus to guaranteed service networks," *IEEE/ACM Transactions on Information Theory*, vol. 44, no. 3, pp. 1087–1097, May 1998.

[7] A. Burchard, J. Liebeherr, and S. D. Patek, "A min-plus calculus for end-to-end statistical service guarantees," *IEEE Transaction on Information Theory*, vol. 52, no. 9, Sept. 2006, To appear.

[8] C. Li, A. Burchard, and J. Liebeherr, "A network calculus with effective bandwidth," University of Virginia, Computer Science Department, Tech. Rep. CS-2003-20, Nov. 2003.

[9] S. Ayyorgun and R. Cruz, "A service-curve model with loss and a multiplexing problem," in *Proc. 24th IEEE International Conference on Distributed Computing System (ICDCS'04)*, March 2004, pp. 756–765.

[10] M. Fidler, "An end-to-end probabilistic network calculus with moment generating functions," in *Proc. IEEE 14th International Workshop on Quality of Servic (IWQoS 2006)*, 2006, pp. 261–270.

[11] Y. Jiang and P. J. Emstad, "Analysis of stochastic service guarantees in communication networks: A server model," in *Proc. of the International Workshop on Quality of Service (IWQoS 2005)*, June 2005, pp. 233–245.

[12] Y. Jiang, "A basic stochastic network calculus," in *Proc. of ACM Sigcomm 2006*, September 2006, (to appear).

[13] F. Ciucu, A. Burchard, and J. Liebeherr, "Scaling properties of statistical end-to-end bounds in the network calculus," *IEEE Transaction on Information Theory*, vol. 52, no. 6, pp. 2300–2312, June 2006.

[14] O. Yaron and M. Sidi, "Performance and stability of communication networks via robust exponential bounds," *IEEE/ACM Transactions on Networking*, vol. 1, no. 3, pp. 372–385, June 1993.

[15] J. Jackson, "Networks of waiting lines," *Oper. Res.*, vol. 5, pp. 518–521, 1957.

[16] L. Kleinrock, "Message delay in communication nets with storage," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1962.

[17] D. Bertsekas and R. Gallager, *Data Networks*. Prentice Hall, 1992.

[18] F. 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*. Oxford University Press, 1996, pp. 141–168.

[19] O. Boxma, "On a tandem queueing model with identical service times at both counters. Part 1,2," *Adv. Appl. Probab.*, vol. 11, pp. 616–659, 1979.

[20] S. B. Calo, "Delay properties of message channels," in *Proc. IEEE ICC, Boston, Mass.*, 1979, pp. 43.5.1–43.5.4.

[21] O. P. Vinogradov, "A multiphase system with identical service," *Soviet Journal of Computer and Systems Sciences.*, vol. 24, pp. 28–31, 1986.

[22] ——, "A multiphase system with many servers and identical service times," *Stochastic Processes and Their Applications, MIEM, Moscow*, vol. 24, pp. 42–45, 1989 (in Russian).

[23] ——, "On certain asymptotic properties of waiting time in a multiserver queueing system with identical times," *SIAM Theory of Probability and Its Applications (TVP)*, vol. 39, no. 4, pp. 714–718, 1994.

[24] P. W. Glynn and W. Whitt, "Departures from many queues in series," *Annals of Applied Probability*, vol. 1, pp. 546–572, 1991.

[25] P. Le Gall, "The overall sojourn time in tandem queues with identical successive service times and renewal input," *Stochastic Processes and Their Applications*, vol. 52, pp. 165–178, 1994.

[26] F. I. Karpelevitch and A. Y. Kreinin, "Asymptotic analysis of queuing systems with identical service," *Journal of Applied Probability*, vol. 33, no. 1, pp. 267–281, 1996.

[27] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford University Press, 2001.

[28] Q. Yin, Y. Jiang, S. Jiang, and P. Y. Kong, "Analysis on generalized stochastically bounded bursty traffic for communication networks," in *Proc. IEEE Local Computer Networks 2002*, Tampa, Florida, November 2002, pp. 141–149.

[29] S. McCreary and K. Claffy, "Trends in wide area IP traffic patterns," in *Proc. 13th ITC Specialist Seminar on Internet Traffic Measurement and Modeling*, Sept. 2000.