AIST: Insights into Queuing and Loss on Highly Multiplexed Links

Maxim Podlesny*, Sergey Gorinsky†, Balaji Rengarajan†
* D.R. Cheriton School of Computer Science, University of Waterloo
200 University Avenue West, Waterloo, ON, N2L 3G1, Canada
mpodlesn@cs.uwaterloo.ca
† Institute IMDEA Networks
Avenida del Mar Mediterraneo, 22, Leganes, Madrid, 28918, Spain
{sergey.gorinsky,balaji.rengarajan}@imdea.org

***Abstract*—In explicit or delay-driven congestion control, a common objective is to sustain high throughput without long queues and large losses at the bottleneck link of the network path. Congestion control protocols strive to achieve this goal by transmitting smoothly in the steady state. The discovery of the appropriate steady-state transmission rates is a challenging task in itself and typically introduces additional queuing and losses. Seeking insights into the steady-state profiles of queuing and loss achievable by real protocols, this paper presents an AIST (Asynchronous arrivals with Ideally Smooth Transmission) model that abstracts away transient queuing and losses related to discovering the path capacity and redistributing it fairly among the packet flows on the bottleneck link. In AIST, the flows arrive asynchronously but transmit their packets at the same constant rate in the steady state. For the link with an overprovisioned buffer, our queuing-theoretic analysis and simulations for different smooth distributions of packet interarrival times agree that queuing under AIST with the target utilization of 1 is on the order of the square root of N, where N is the number of flows. With small buffers, our simulations of AIST show an ability to provide bounded loss rates regardless of the number of flows.**

I. Introduction

Smooth fair lossless transmission at high rates has been an aspiration for many explicit [1]–[6] and delay-driven [7]–[10] congestion control protocols. Smoothness of transmission is particularly important for interactive multimedia and other delay-sensitive applications that would suffer from excessive queuing delays at network links. However, since discovery of a fair sending rate is fraught with packet bursts and other causes of link queuing, smooth transmission conflicts with prompt response to changes in network conditions [11], [12].

Smooth transmission is also relevant to link buffer sizing [13]–[20], which recently attracted significant attention. In the context of TCP (Transmission Control Protocol) congestion control [21], various proposals disagree on how to size the buffer with respect to the number of flows sharing the link. Arguing that aggregate oscillations of TCP traffic subside as the number of flows increases, one view maintains that a small buffer suffices for a highly multiplexed link [14], even if the buffer accommodates only up to 20 packets [16]. However, when the network path restricts each flow to less than several packets per RTT (round-trip time), TCP suffers from high loss rates and frequent retransmission timeouts [22]. Hence, alternative guidelines size the buffer proportionally to the number of flows [13], i.e., argue for an even larger buffer than the traditionally recommended capacity-delay product [23]. The divergent opinions prompted proposals of multiple-buffer forwarding schemes that set the buffer sizes differently for different applications [24], [25]. Furthermore, some studies propose new congestion control protocols for networks with small buffers [26].

978-1-4673-1298-1/12$31.00

The lack of agreement on many issues in congestion control has sound reasons. Practical congestion control protocols tend to be multimodal and complex, e.g., they incorporate smart strategies for packet acknowledgments [27]–[29]. Besides, performance of congestion-control protocols depends on underlying Internet link technologies [30]. Due to these complexities, precise comprehensive analyses of real congestion-control protocols are difficult. For the same reason, experimental evaluations face serious scalability challenges [31], [32]. In particular, while even a single packet-level simulation of transient and steady states for a highly multiplexed link might consume long time, studies of congestion control protocols rarely report reliable results for more than a few hundred flows.

This paper shifts the research focus from protocol engineering to network science and presents an AIST (Asynchronous arrivals with Ideally Smooth Transmission) model to gain insights into steady-state queuing and loss profiles that real congestion control protocols can strive to achieve. AIST abstracts away transient queuing and losses related to discovering the path capacity and redistributing this capacity fairly among the packet flows on the bottleneck link. In AIST, the flows arrive asynchronously and transmit at the same constant rate in the steady state. Such ideally smooth transmissions do not eliminate queuing altogether because packet arrivals from different flows in the steady state are asynchronous [33]. For example, even if the constant-rate flows underutilize the link on average, a queue arises when packets from multiple flows arrive to the link at the same time. The asynchrony of packet arrivals constitutes the main difference between AIST and the perfect TDM (Time Division Multiplexing) which avails the link to each packet immediately upon the packet arrival. Such asynchrony is a ubiquitous property of real congestion control protocols.

Simplicity is a prominent aspect of AIST and renders tractable analysis and scalable experiments. In particular, our simulation methodology exactly captures the steady-state queuing for N concurrent flows by examining only $2N$ packets. The low overhead enables us to assess expected steady-state performance reliably by conducting extensive simulations with up to 5,000 concurrent flows and repeating each experiment 1,000 times.

Our modeling and analysis of AIST are not only novel but also useful for practical congestion control. While simplifying some aspects (e.g., packet size variability) of secondary importance for the studied queuing problem, the AIST model focuses on the time intervals where the number of flows is constant (flows arrive/leave on a longer time scale than gaps between consecutive packets of a flow), i.e., those time intervals where real congestion-control protocols are most effective in performing the queue management. Furthermore, the AIST assumptions of FIFO (First-In First-Out) packet queuing and constant-capacity link represent actual Internet topologies accurately. Our study shows that queuing under AIST at the fully utilized link with the overprovisioned buffer is $O(\sqrt{N})$. This result suggests that bounded queuing delay without packet discard at a fully utilized link is an unrealistic goal for a practical congestion control protocol if the number of flows is arbitrarily high, and the protocol does not incorporate a synchronization mechanism that deliberately avoids bursty arrivals of packets from different flows to the link. Nevertheless, our work also implies that the protocol can overcome this limitation by decreasing the link utilization as the number of flows grows. Among its other results, our paper indicates an exciting option of bounded loss rates at fully utilized links with small buffers.

The rest of the paper is organized as follows. Section II defines our AIST model. Section III analyzes queuing under AIST in overprovisioned buffers. Section IV supplements the analysis with extensive simulations. Section V reports loss rates under AIST with small buffers. Section VI discusses the role of the packet size and link capacity in the AIST model. Finally, section VII concludes the paper with a summary of its contributions.

II. AIST MODEL

AIST models the steady state of a link where the number of packet flows does not change (the steady state ends when one of the flows terminates, or an extra flow begins). N denotes this constant *number of flows*. All flows use the same fixed *packet size* S. The link has a FIFO buffer and constant *capacity* C. *Transmission delay* D of a packet refers to the amount of time spent on transmitting the packet into the link. With equal packet sizes, all packets have the same transmission delay

$$D = \frac{S}{C}. \tag{1}$$

Hence, AIST measures time in units of packet transmission delay, or in packets for short. In particular, if d_i refers to the actual time that a packet of flow i waits in the buffer for its turn to be transmitted into the link, our model captures the queuing with *queuing delay* q_i expressed in packets as

$$q_i = \frac{d_i}{D}. \tag{2}$$

Time 0 denotes the moment when all N flows complete a transient phase and start transmitting their packets at the same constant rate. *Arrival time* t_i refers to the time when the first packet of flow i arrives to the link during the steady-transmission interval ($i = 0, \ldots, N-1$). Without loss of generality, we assume that time $t_0 = 0$ represents the arrival time of such first packet from flow 0. The other $N-1$ flows are also numbered in the order of their first packet arrivals during the steady-transmission interval. We express the respective packet arrival times as

$$t_{i+1} = t_i + \delta_i \tag{3}$$

where $i = 0, \ldots, N-2$, and δ_i represents the *interarrival time* between the first packets of flows i and $i+1$ during the steady-transmission interval ($\delta_i \geq 0$). The distribution of the packet interarrival times has mean $\mu = \frac{D}{U}$ and variance $\sigma^2 = \beta\mu^2$ where β is a fixed factor representing the smoothness of the interarrival process. We refer to parameter U of the model as the *target utilization* of the link, $0 < U \leq 1$.

Throughout the steady state, the transmission pattern within each flow remains stable with consecutive packets arriving to the link with *transmission period* T defined as

$$T = \max\left\{t_{N-1} + \delta_{N-1}, \frac{N \cdot D}{U}\right\}, \tag{4}$$

$$T = \max\left\{\sum_{i=0}^{N-1} \delta_i, \frac{N \cdot D}{U}\right\}. \tag{5}$$

The maximum function in equation 5 ensures the congestion-control property of limited queuing delay even when $U = 1$. Hence, the AIST model represents both deterministic (bounded queuing delay even at a fully utilized link) and random (unsynchronized packet transmission by different flows) features of real congestion control protocols.

In AIST, deterministic target utilization U of the link serves as an upper limit on random *actual utilization* R which is equal to

$$R = \frac{N \cdot D}{T}. \tag{6}$$

In scenarios with $T = \frac{N \cdot D}{U}$ (i.e., when $\sum_{i=0}^{N-1} \delta_i \leq \frac{N \cdot D}{U}$), the actual utilization matches the target utilization: $R = U$. When $T > \frac{N \cdot D}{U}$ (i.e., in scenarios with $\sum_{i=0}^{N-1} \delta_i > \frac{N \cdot D}{U}$), the actual utilization is lower than the target utilization: $0 < R < U \leq 1$.

We examine both overprovisioned and small buffers. While the overprovisioned buffer is large enough to store and forward all arriving packets without loss, packet losses are possible with small buffers. The main metric with the overprovisioned buffer is queuing delay in packets (equation 2). We quantify performance in small-buffer settings with a *loss rate* defined as a fraction of packets discarded in the steady state due to buffer overflow.

III. Analysis for overprovisioned buffers

To analyze queuing under AIST in overprovisioned buffers, sections III-A and III-B apply different queuing-theoretic techniques to the cases of $U = 1$ and $U < 1$ respectively.

A. Target utilization of 1

When the target utilization equals 1 ($U = 1$), queuing scenarios of AIST can be classified into two categories:

- *Fully utilized link* ($R = 1$ and $T = N \cdot D$) where any period of length T during the steady-transmission interval sees arrivals of exactly N packets that consume exactly time T to be transmitted into the link. With the fully utilized link, the queuing converges to a periodic pattern by time T. Figure 1a illustrates such periodic queuing pattern.
- *Underutilized link* ($R < 1$ and $T > N \cdot D$) where any period of length T during the steady-transmission interval sees arrivals of exactly N packets that take strictly less time than T to be transmitted into the link. Whatever queuing delay is at time 0, the link underutilization eventually drains any queuing delay inherited from the transient phase. The queuing converges to a periodic pattern determined exclusively by packets that arrive during the steady-transmission interval. In figure 1b, time interval $[2T; 3T)$ represents such emerged periodic pattern of queuing at the underutilized link.

With either fully utilized or underutilized link the period of the stable queuing pattern equals T. Let τ be such a multiple of T that the queuing converges to its periodic pattern by time τ. Then, the following expression serves as a lower bound on queuing delay q_i of the packet that arrives from flow i during period $[\tau; \tau + T)$:

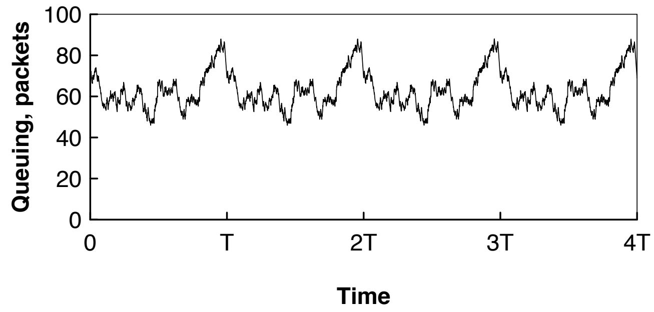
$$q_i \geq h + i - \frac{t_i}{D} \tag{7}$$

where h denotes the queuing delay at time τ, time t_i is $\sum_{j=0}^{i-1} \delta_j$, and $\frac{t_i}{D}$ represents the number of packets that the link is capable of transmitting during time interval $[\tau; \tau + t_i)$. With the underutilized link, the right-hand side of inequality 7 can take negative values. With the fully utilized link, inequality 7 is an equation, i.e., $h + i - \frac{t_i}{D}$ captures queuing delay q_i exactly.

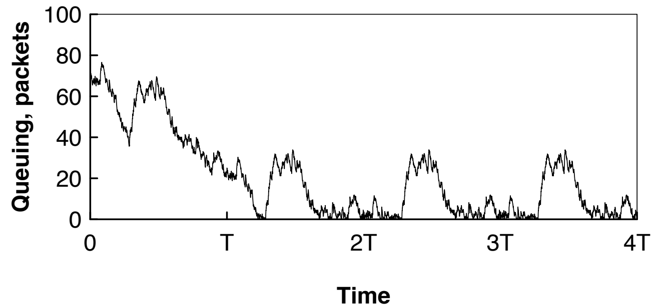
Because we are primarily interested in low bounds on queuing delay, we assume $h = 0$ and express lower bound l_i on queuing delay q_i as

$$l_i = i - \frac{t_i}{D}. \tag{8}$$

Whereas all packet interarrival times δ_j are from the same distribution, and the analysis is for highly multiplexed links with large values of N, the Central Limit Theorem establishes



(a) Fully utilized link with $T = N \cdot D$.



(b) Underutilized link with $T > N \cdot D$.

Fig. 1. Examples of emerging periodic queuing patterns in the overprovisioned buffer under AIST with N = 1,000, S = 1,000 bytes, C = 100 Mbps, and U = 1 (i.e., $N \cdot D$ = 80 ms).

that t_i follows the normal distribution with mean α and variance ω^2:

$$\alpha = i\mu \quad \text{and} \quad \omega^2 = i\sigma^2 \tag{9}$$

where μ and σ^2 are respectively the mean and variance of the interarrival time distribution.

Since t_i is normally distributed, we use equations 8 and 9 to derive probability $\theta(i)$ of $l_i > Q$ as

$$\theta(i) = P[l_i > Q] = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{(Q - i)D + i\mu}{\sigma\sqrt{2i}}\right)\right) \tag{10}$$

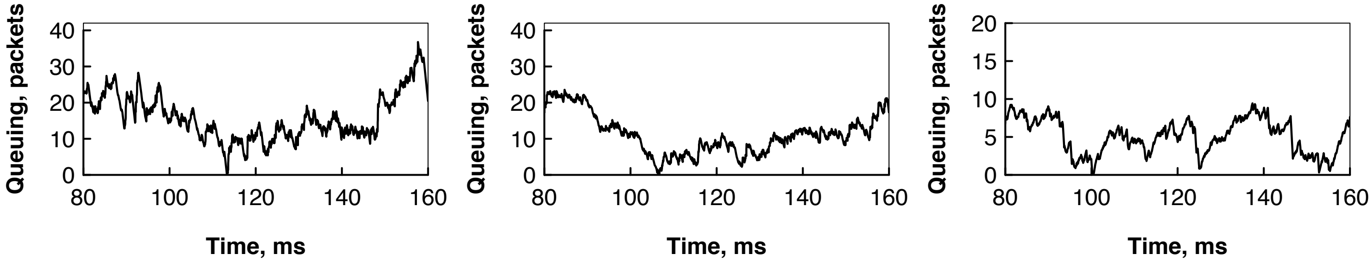
where erf is the error function. Because $\mu = D$ when $U = 1$, and σ is proportional to μ, we simplify equation 10 as

$$\theta(i) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{Q}{\sqrt{2\beta i}}\right)\right) \tag{11}$$

where β is the interarrival process smoothness.

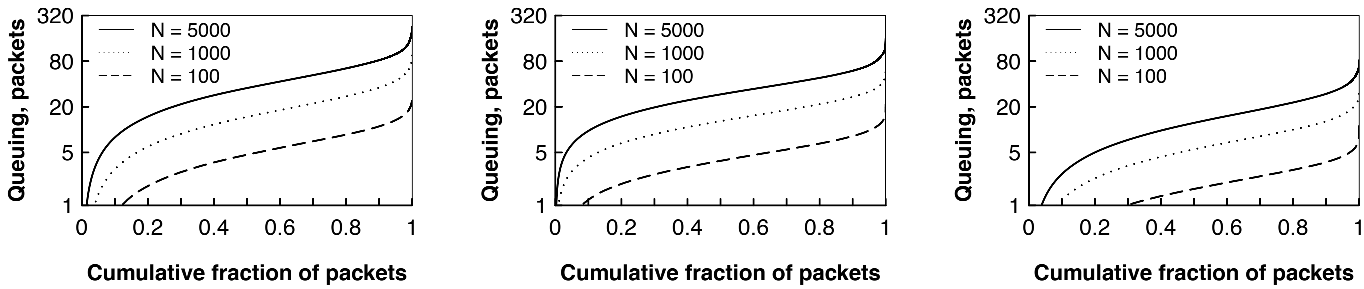
Using p to represent the probability that the queuing delay of a packet exceeds Q, we express this probability as

$$p = \frac{1}{N}\sum_{i=0}^{N-1} P[q_i > Q]. \tag{12}$$



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 2. Sample single periods of stable steady-state queuing in the overprovisioned buffer of a fully utilized link under AIST with N = 1,000, S = 1,000 bytes, C = 100 Mbps, U = 1, and $T = N \cdot D$ = 80 ms.



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 3. Cumulative distributions of steady-state queuing delay in the overprovisioned buffer under AIST with the target utilization of 1 for different numbers of flows: S = 1,000 bytes, C = 100 Mbps, and U = 1.

Because $\theta(i)$ is a nonnegative increasing function, we bound p from below as

$$p \geq \frac{1}{N}\sum_{i=0}^{N-1} P[l_i > Q] \geq \frac{1}{N}\sum_{i=\frac{2N}{3}}^{N-1} \theta\left(i\right) \geq \frac{\theta\left(\frac{2N}{3}\right)}{3}, \qquad (13)$$

$$p \geq \frac{1 - \mathrm{erf}\left(\frac{Q}{\sqrt{\frac{4\beta N}{3}}}\right)}{6}. \qquad (14)$$

To establish a lower bound on queuing delay for the top p packets, we define λ_p to be such that

$$\mathrm{erf}\left(\lambda_p\sqrt{\frac{3}{4}}\right) = 1 - 6p. \qquad (15)$$

λ_p depends only on p, i.e., the fraction of packets. For example, $\lambda_{1\%} \approx 1.6$ and $\lambda_{5\%} \approx 0.8$ for the top 1% and 5% packets respectively. Applying inequality 14, we express lower bound Q_{min} on queuing delay under AIST for the top p packets as

$$Q_{\text{min}} = \lambda_p\sqrt{\beta N} \qquad (16)$$

By deriving equation 16, we have proved the following result:

Theorem 1: Queuing delay for a fixed fraction of packets in the overprovisioned buffer under AIST with the target utilization of 1 is bounded from below by $O(\sqrt{N})$ where N is the number of flows sharing the link.

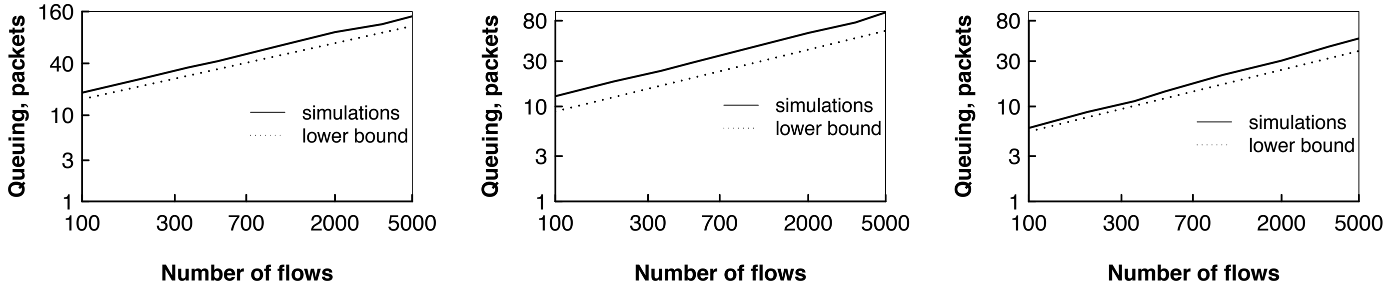
Theorem 1 has important implications for practical congestion control. Similarly to our AIST model, real congestion control protocols do not incorporate any synchronization mechanisms that deliberately avoid bursty arrivals of packets from multiple flows to the bottleneck link on a smaller time scale than the RTTs of the flows. Unless a real congestion control protocol with the target utilization of 1 incorporates such a synchronization mechanism, Theorem 1 suggests that the congestion control protocol cannot provide bounded queuing delay without packet discard if the number of flows can be arbitrarily large.

B. Target utilization below 1

When the target utilization of AIST is smaller than 1 ($U < 1$), the actual utilization is always below 1 ($R < 1$), and we use classical queuing theory to characterize steady-state queuing in the overprovisioned buffer of the underutilized link. In particular, for the Exponential distribution of the packet interarrival times, we apply the M/D/1 steady-state analysis [34] to express probability p that the queuing delay experienced by a packet exceeds Q packets:

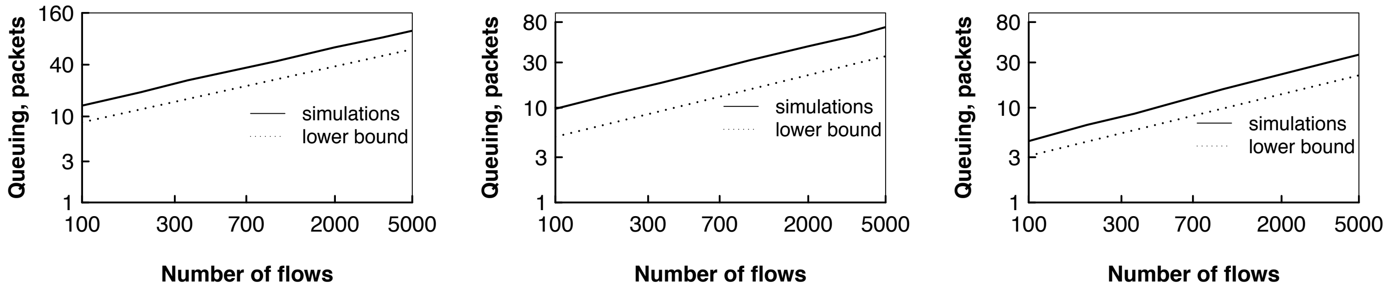
$$p = 1 - (1 - R)\sum_{j=0}^{Q} \frac{((Q-j)R)^j\, e^{-(Q-j)R}}{j!}. \qquad (17)$$

In the particular context of AIST, actual utilization R is at most the target utilization, i.e., $R \leq U$. Our analysis relies on the observation that the sum of independent and identically distributed (i.i.d.) exponential random variables adheres to the



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 4. Analytic and simulation results for the steady-state queuing delay experienced by the top 1% packets in the overprovisioned buffer under AIST with the target utilization of 1: $S =$ 1,000 bytes, $C =$ 100 Mbps, and $U = 1$.



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 5. Analytic and simulation results for the steady-state queuing delay experienced by the top 5% packets in the overprovisioned buffer under AIST with the target utilization of 1: $S =$ 1,000 bytes, $C =$ 100 Mbps, and $U = 1$.

Erlang distribution [35]. The actual utilization equals U in scenarios with $T = \frac{N \cdot D}{U}$ which occur with probability

$$P[R = U] = \frac{\gamma(N, N)}{(N-1)!} \tag{18}$$

where $\gamma(\cdot, \cdot)$ is the lower incomplete gamma function. The actual utilization is lower than the target utilization in scenarios with $T > \frac{N \cdot D}{U}$. In such scenarios, T attains specific value x with probability

$$P[T = x] = \frac{x^{N-1} e^{-\frac{x}{\mu}}}{\mu^N (N-1)!}. \tag{19}$$

Thus, we derive average value $\overline{R}$ of the actual utilization as

$$\overline{R} = \frac{\left(\gamma(N, N) + \frac{N}{N-1}\left(\Gamma(N, N) - N^{N-1} e^{-N}\right)\right) U}{(N-1)!} \tag{20}$$

where $\Gamma(\cdot, \cdot)$ is the upper incomplete gamma function.

IV. Simulations for overprovisioned buffers

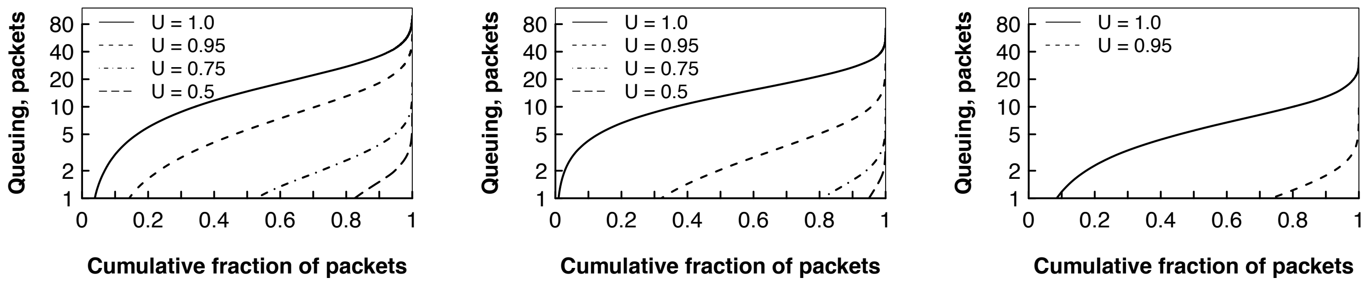
To validate the above analytic conclusions for overprovisioned buffers, this section reports packet-level simulations of AIST. While the focus of our paper is on minimum queuing achievable with AIST in the steady state, we simulate settings where the steady-transmission interval does not inherit any queuing from the transient phase, i.e., queuing delay at time 0 is 0. In these minimum-queuing settings the stable periodic queuing pattern of AIST always emerges by time T with either fully utilized or underutilized link. Thus, each simulation run is for time interval $[0; 2T)$ where $[0; T)$ is a warm-up stage with N packets arriving at times t_i, $i = 0, \ldots, N-1$. By recording the queuing delays of the N packets arriving at times $T + t_i$ during period $[T; 2T)$, we measure the cumulative distribution of the steady-state queuing delay under AIST.

In our simulations, we consider the following 3 smooth instances of the packet interarrival distribution:

- *Exponential distribution* from a Poisson process with average arrival rate $\frac{1}{\mu}$ and smoothness $\beta = 1$;
- *Uniform distribution* between 0 and 2μ with smoothness $\beta = \frac{1}{3}$;
- *Pareto distribution* with index $k = 4$ and smoothness $\beta = \frac{1}{k(k-2)}$.

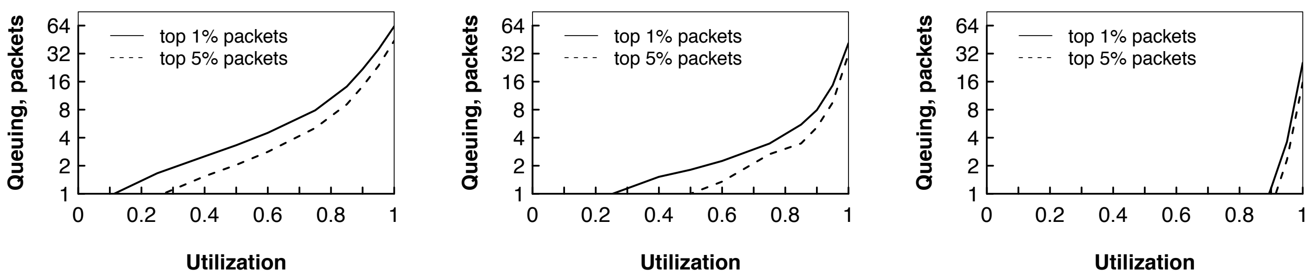
The simplicity of AIST enables extensive simulations. Whereas each simulation run examines only $2N$ packets (i.e., only 2 packets per flow), the simulations capture the expected queuing behavior with high certainty by performing 1,000 experiments for each examined set of parameter settings. Unless explicitly stated otherwise, the parameters take the following default values: $N =$ 1,000, $S =$ 1,000 bytes, $C =$ 100 Mbps, and $U = 1$ (i.e., $N \cdot D = 80$ ms). Figure 2 illustrates patterns of the stable steady-state queuing for these default settings in scenarios where $T = N \cdot D$ (i.e., fully utilized link).

We start the simulations with experiments where the target utilization is equal to 1. To evaluate the dependence of the



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 6. Cumulative distributions of the steady-state queuing delay under AIST in the overprovisioned buffer for different target utilizations: N = 1,000, S = 1,000 bytes, and C = 100 Mbps.



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 7. Impact of the target utilization on the steady-state queuing in the overprovisioned buffer under AIST.

steady-state queuing on the number of flows, we vary N from 100 to 5,000. Figure 3 shows that larger values of N consistently produce longer queues. In particular, while the bottom 10% packets experience no queuing at all with 100 flows and Exponential packet interarrival times, queuing delay for this percentile is 8 packets with 5,000 flows.

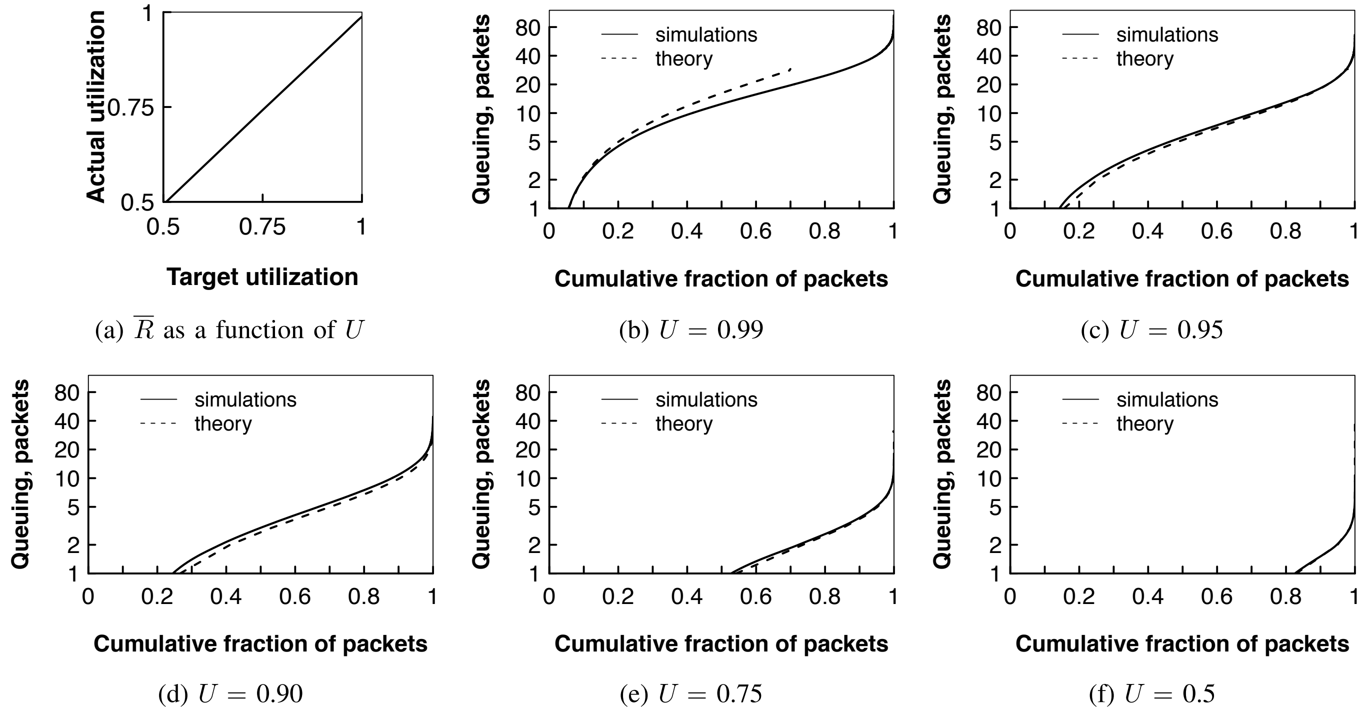
Figure 4 reports analytic and simulation results for the queuing delay experienced by the top 1% packets. The experimental results for all 3 packet interarrival time distributions are consistent with the plotted $O(\sqrt{N})$ lower bounds that depict equation 16. Figure 5 shows a qualitatively similar agreement between the analysis and simulations for the top 5% packets.

Now, we explore whether reducing the target utilization mitigates the above concerns about queuing scalability with respect to the number of flows. In the next set of experiments, U varies from 0.5 to 1. Figure 6 demonstrates that decreasing the target utilization subdues the steady-state queuing substantially. For instance, the bottom 80% packets experience no queuing at all with the target utilization of 0.5. Figure 7 quantifies the significant reductions of the queuing delay as the target utilization decreases. With U = 1, the queuing delay for the top 1% packets is 63, 42, 26 packets under the Exponential, Uniform, Pareto packet interarrival times respectively. Decreasing U to 0.75 and further to 0.5 reduces the queuing delay to 8, 4, 0 packets and 4, 2, 0 packets respectively. The results also reveal that the decrease of the target utilization helps most dramatically under the Pareto distribution. Our findings justify the common practice of operating network links with average utilization of at most 0.5 [36].

In figure 8, we focus on experiments with Exponential interarrival times. For each simulation run, we record actual utilization R. Based on the recorded R values, figure 8a plots average actual utilization $\overline{R}$ as a function of target utilization U and demonstrates that the average actual utilization under AIST is only slightly smaller than the target utilization. Also using these individual R values as input to equation 17, we compute distributions of steady-state queuing delay for the target utilizations of 0.99, 0.95, 0.90, 0.75, 0.5. While figures 8b through 8f plot the respective equation-based distributions as dashed lines, the solid lines in the graphs depict the queuing delay distributions measured directly during the simulations (due to computational challenges presented by equation 17, we report the theoretical results for queuing delays up to around 30 packets). Whereas the simulation and analytic lines are closely aligned in general, their noteworthy deviations occur for top packet percentiles with $U = 0.99$.

V. Loss rates with small buffers

Our derived and validated result of the $O(\sqrt{N})$ steady-state queuing in the overprovisioned buffer under AIST with the target utilization of 1 suggests that avoidance of packet losses with a constant buffer is an unrealistic objective for a real congestion control protocol if the number of flows is arbitrarily large (and the protocol does not incorporate a synchronization mechanism that deliberately avoids bursty arrivals of packets from different flows to the bottleneck link). In this section, we



(a) $\overline{R}$ as a function of U

(b) $U = 0.99$

(c) $U = 0.95$

(d) $U = 0.90$

(e) $U = 0.75$

(f) $U = 0.5$

Fig. 8. Analytic and simulation results for the steady-state queuing delay in the overprovisioned buffer under AIST with Exponential interarrival times for different target utilizations: N = 1,000, S = 1,000 bytes, and C = 100 Mbps.

simulate AIST with small buffers that accommodate either 10 or 20 packets.

First, we experiment in settings with $U = 1$ and vary the number of flows from 100 to 5,000. Figure 9 plots steady-state loss rates with Exponential, Uniform, and Pareto interarrival times. For low values of N, the loss rates increase quickly as N grows. They stabilize after N becomes sufficiently large. The bounded nature of the loss rates opens an exciting prospect that a practical congestion control protocol with the target utilization of 1 can also provide bounded loss rates at links with small buffers.

Then, we simulate setups with the target utilizations below 1. While the loss rate for Exponential interarrival times is the highest among the 3 distributions, this loss rate stabilizes at around 4% and 2% with U = 1 for the 10-packet and 20-packet buffer respectively. If such losses are deemed as too high, reducing the target utilization helps once again: figure 10 shows prompt dramatic reductions in the loss rates for all 3 distributions of interarrival times as U decreases from 1 to 0.5.

VI. Role of the packet size and link capacity

So far, our study focused on the N and U parameters. In this section, we discuss the role of the other two parameters of AIST: packet size S and link capacity C. According to equation 1, the S and C parameters affect transmission delay D. Because AIST measures time in units of transmission delay, and the interarrival time distributions and transmission period T are proportional to D, the S and C parameters do not affect AIST queuing behaviors (queuing delays measured in packets) at all. Hence, the AIST model does not raise any concerns about the ability of real congestion control to scale well with the packet size and link capacity.

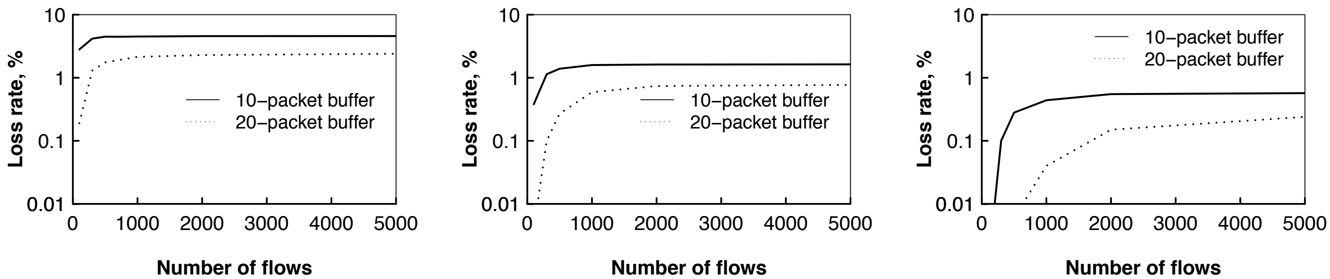
VII. Conclusion

In this paper, we presented and evaluated the AIST model where all flows transmit their packets at the same constant rate in the steady state. The smooth transmission causes queuing at a link due to the asynchrony of packet arrivals from multiple flows, which is a ubiquitous property of real congestion control protocols. Our innovative analysis and simulations examined queuing and loss under AIST with different link buffers.

While simplicity is a prominent aspect of AIST, our simulation methodology exactly captured the steady-state queuing for N concurrent flows by examining only $2N$ packets. The low overhead enabled us to assess the steady-state performance with high certainty through extensive experiments with up to 5,000 concurrent flows and 1,000 runs per experimental setting.

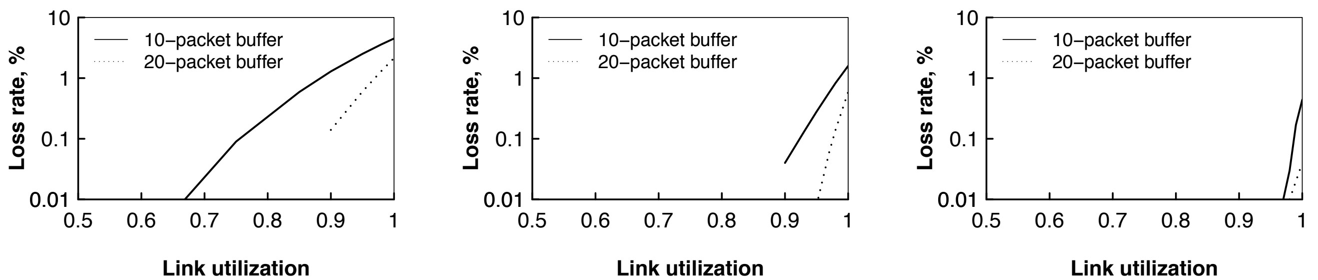
In the simulations, we considered 3 smooth distributions of packet interarrival times. In addition to the Exponential distribution which might be the most realistic assumption for the packet interarrival times, we also examined AIST with the smoother Uniform and Pareto distributions.

Our main result is for AIST with the target utilization of 1. The respective analysis and simulations agree that the steady-state queuing delay experienced by a fixed fraction of packets in the overprovisioned buffer is $O(\sqrt{N})$. Our analysis and experiments for $U < 1$ showed that reducing the target



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 9. Steady-state loss rates under AIST with small buffers and target utilization of 1: S = 1,000 bytes, C = 100 Mbps, and U = 1.



(a) Exponential (b) Uniform (c) Pareto, $k = 4$

Fig. 10. Impact of the target utilization on the steady-state loss rates under AIST with small buffers: N = 1,000, S = 1,000 bytes, and C = 100 Mbps.

utilization alleviates the concerns about queuing scalability with respect to the number of flows. Also, the simulations of AIST with small buffers and target utilization of 1 revealed bounded steady-state loss rates.

The practical utility of AIST is in its insights relevant to real congestion control protocols. Theorem 1 implies that avoidance of packet losses with a constant buffer can be an unrealistic objective for a real congestion control protocol if the number of flows is arbitrarily large (and the protocol does not incorporate a synchronization mechanism that deliberately avoids bursty arrivals of packets from different flows to the bottleneck link). Our study of AIST also indicates that the protocol can overcome this limitation by decreasing the target utilization as the number of flows grows. An exciting alternative is to equip fully utilized network links with small buffers and thereby provide bounded loss rates. Finally, the AIST model does not reveal any limitations on scalability of real congestion control in regard to the packet size or link capacity.

Acknowledgments

This work was supported in part by NSERC (Natural Science and Engineering Research Council) Discovery Accelerator RGPAS-380438 45185 in Canada, Department-of-Education grant S2009/TIC-1468 from the Regional Government of Madrid, Ramon-y-Cajal grant RYC-2009-04660 from the Spanish Ministry of Science and Innovation, FP7-PEOPLE grant 229599 and FP7-ICT grant 288021 from the European Commission.

References

[1] D. Katabi, M. Handley, and C. Rohrs, "Congestion Control for High Bandwidth-Delay Product Networks," in *Proceedings ACM SIGCOMM 2002*, August 2002.

[2] N. Dukkipati, M. Kobayashi, R. Zhang-Shen, and N. McKeown, "Processor Sharing Flows in the Internet," in *Proceedings IEEE IWQoS 2005*, June 2005.

[3] Y. Xia, L. Subramanian, I. Stoica, and S. Kalyanaraman, "One More Bit Is Enough," in *Proceedings ACM SIGCOMM 2005*, August 2005.

[4] M. Podlesny and S. Gorinsky, "Multimodal Congestion Control for Low Stable-State Queuing," in *Proceedings IEEE INFOCOM Minisymposium 2007*, May 2007.

[5] S. Jain, Y. Zhang, and D. Loguinov, "Towards Experimental Evaluation of Explicit Congestion Control," in *Proceedings IEEE IWQoS 2008*, June 2008.

[6] P. Loureiro and E. Monteiro, "Stability and Performance Analysis of the Open Box Transport Protocol," in *Proceedings IEEE ISCC 2009*, July 2009.

[7] R. Jain, "A Delay-Based Approach for Congestion Avoidance in Interconnected Heterogeneous Computer Networks," *ACM Computer Communications Review*, vol. 19, no. 5, pp. 56–71, October 1989.

[8] L. Brakmo, S. O'Malley, and L. Peterson, "TCP Vegas: New Techniques for Congestion Detection and Avoidance," in *Proceedings ACM SIGCOMM 1994*, August 1994.

[9] L. Budzisz, R. Stanojevic, A. Schlote, R. Shorten, and F. Baker, "On the Fair Coexistence of Loss- and Delay-based TCP," in *Proceedings IEEE IWQoS 2009*, July 2009.

[10] L. Budzisz, R. Stanojevic, A. Schlote, F. Baker, and R. Shorten, "On the Fair Coexistence of Loss- and Delay-Based TCP," *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1811–1824, December 2011.

[11] S. Floyd, M. Handley, J. Padhye, and J. Widmer, "Equation-Based Congestion Control for Unicast Applications," in *Proceedings ACM SIGCOMM 2000*, August 2000.

[12] S. Gorinsky, M. Georg, M. Podlesny, and C. Jechlitschek, "A Theory of Load Adjustments and its Implications for Congestion Control," *Journal of Internet Engineering*, vol. 1, no. 2, pp. 82–93, October 2007.

[13] R. Morris, "Scalable TCP Congestion Control," in *Proceedings IEEE INFOCOM 2000*, March 2000.

[14] G. Appenzeller, I. Keslassy, and N. McKeown, “Sizing Router Buffers,” in *Proceedings ACM SIGCOMM 2004*, September 2004.
[15] S. Gorinsky, A. Kantawala, and J. Turner, “Link Buffer Sizing: A New Look at the Old Problem,” in *Proceedings IEEE ISCC 2005*, June 2005.
[16] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, and T. Roughgarden, “Routers with Very Small Buffers,” in *Proceedings IEEE INFOCOM 2006*, April 2006.
[17] S. Gorinsky, A. Kantawala, and J. Turner, “Simulation Perspectives on Link Buffer Sizing,” *Simulation*, vol. 83, no. 3, pp. 245–257, March 2007.
[18] Y. Zhang and D. Loguinov, “ABS: Adaptive Buffer Sizing for Heterogeneous Networks,” in *Proceedings IEEE IWQoS 2008*, June 2008.
[19] R. Stanojevic and R. Shorten, “Trading Link Utilization for Queueing Delays: An Adaptive Approach,” *Computer Communications*, vol. 33, no. 9, pp. 1108–1121, June 2010.
[20] Y. Zhang and D. Loguinov, “ABS: Adaptive Buffer Sizing for Heterogeneous Networks,” *Computer Networks*, vol. 54, no. 14, pp. 2562–2574, October 2010.
[21] V. Jacobson, “Congestion Avoidance and Control,” in *Proceedings ACM SIGCOMM 1988*, August 1988.
[22] L. Qiu, Y. Zhang, and S. Keshav, “Understanding the Performance of Many TCP Flows,” *Computer Networks*, vol. 37, no. 3-4, pp. 277–306, November 2001.
[23] C. Villamizar and C. Song, “High Performance TCP in the ANSNET,” *ACM SIGCOMM Computer Communication Review*, vol. 24, no. 5, pp. 45–60, November 1994.
[24] M. Podlesny and S. Gorinsky, “RD Network Services: Differentiation through Performance Incentives,” in *Proceedings ACM SIGCOMM 2008*, August 2008.
[25] ——, “Leveraging the Rate-Delay Trade-off for Service Differentiation in Multi-Provider Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 5, pp. 997–1008, May 2011.
[26] Y. Gu, D. Towsley, C. Hollot, and H. Zhang, “Congestion Control for Small Buffer High Speed Networks,” in *Proceedings IEEE INFOCOM 2007*, May 2007.
[27] R. de Oliveira and T. Braun, “A Dynamic Adaptive Acknowledgment Strategy for TCP over Multihop Wireless Networks,” in *Proceedings IEEE INFOCOM 2005*, March 2005.
[28] D. Blandford, S. Goldman, S. Gorinsky, Y. Zhou, and D. Dooly, “Smartacking: Improving the TCP Performance from the Receiving End,” *Journal of Internet Engineering*, vol. 1, no. 1, pp. 6–21, January 2007.
[29] R. de Oliveira and T. Braun, “A Smart TCP Acknowledgment Approach for Multihop Wireless Networks,” *IEEE Transactions on Mobile Computing*, vol. 6, no. 2, pp. 192–205, February 2007.
[30] J. Sánchez, J. Domingo-Pascual, and J. Vieira, “AcTMs (Active ATM Switches) with TAP (Trusted and Active PDU Transfers) in a Multiagent Architecture to Better the Chaotic Nature of TCP Congestion Control,” in *Proceedings ICN 2005*, April 2005.
[31] S. Kristiansen and T. Plagemann, “Accuracy and Scalability of ns-2’s Distributed Emulation Extension,” *Simulation*, vol. 87, no. 1-2, pp. 45–65, January 2011.
[32] S. Kristiansen, T. Plagemann, and V. Goebel, “Towards Scalable and Realistic Node Models for Network Simulators,” in *Proceedings ACM SIGCOMM 2011 (poster paper)*, August 2011.
[33] M. Podlesny and S. Gorinsky, “Price of Asynchrony: Queuing under Ideally Smooth Congestion Control,” in *Proceedings IEEE ICNP 2007 (poster paper)*, October 2007.
[34] V. Iversen and L. Staalhagen, “Waiting Time Distribution in M/D/I Queueing Systems,” *Electronic Letters*, vol. 35, no. 25, pp. 2184–2185, December 1999.
[35] D. Montgomery, *"Applied Statistics and Probability for Engineers"*. John Wiley & Sons, 2010.
[36] A. Odlyzko, “Data Networks are Lightly Utilized, and will Stay that Way,” *Review of Network Economics*, vol. 2, no. 3, September 2003.