Are Smaller Packets Less Likely to be Lost?

Joel Sommers, Victor Omwando, Fred Sisenda
Department of Computer Science
Colgate University
Hamilton, NY 13346
jsommers@colgate.edu ; {vomwando, fsisenda}@students.colgate.edu

Abstract

Despite growth in internet link capacities and the efforts of traffic engineers, packet loss continues to be a reality in today's internet. While many tools for active measurement *of packet loss have been developed over the years, applying these tools to a particular operational network context in order to most accurately estimate loss characteristics has remained a challenge. In this paper we examine the problem of selecting a probe packet size for active loss measurement. Prior work has used either small (*e.g.*, 40 byte) packets in order to minimize probing bandwidth, or larger (*e.g.*, 1000 byte) packets in order to increase the likelihood that a probe will be lost if congestion is taking place along a path. We examine these rules of thumb in a systematic way through a series of controlled laboratory experiments. We find that a—perhaps* the*—key system parameter is whether network queue limits are specified in terms of packets or bytes. Our results suggest that if queue limits are specified in terms of packets, there may be no clear optimal probe packet size that yields the most accurate results. On the other hand, if queue limits are specified in bytes, a packet size chosen based on the average packet size of aggregate background traffic leads to the most accurate loss average estimate.*

1. Introduction

Because of its well-studied affect on transport protocol and application performance, packet loss has been and continues to be a focus of study in the network measurement community. The operational importance of packet loss is reflected in the fact that network service provider service level agreements (SLAs) typically contain clauses guaranteeing no more than a specified packet loss average over a given time period. As a result, packet loss is an important characteristic to measure *accurately*, in order to monitor network performance for SLA compliance. It is also important to measure *efficiently*, in order to quickly detect looming network performance problems.

Over the years, a number of methodologies have been proposed for measuring packet loss, *e.g.*, [1, 3, 6, 7, 9, 11, 12, 16]. These methods fall in two basic categories: *active* and *passive*. While both categories are important for understanding the nature of network packet loss [5], active measurement methodologies that operate by introducing probe traffic have been appealing due to the fact that they can give an end-to-end perspective of packet loss along a path, similar to what a networked application might experience. Furthermore, this end-to-end view is likely to be of more relevance from a network service provider's perspective when monitoring customer traffic for compliance with SLAs.

Unfortunately, there has been a profound lack of published guidelines on how to set various parameters in order to best use active loss measurement tools. For example, the IETF RFC regarding active measurement of packet loss [1] generally lacks guidelines for setting the rate of probing and for setting the packet size (among other parameters). While there are likely to be tradeoffs for any active measurement methodology with respect to the frequency of probing or the probe packet size and the accuracy of the resulting measurements, it is as yet unclear how to set these parameters for a given environment in order to achieve the most accurate results. Given the fact that "ground truth" measures are usually not available for live internet environments, parameters are most often selected based on either default settings, on a given theoretical model, on assumptions based on how network hardware is implemented, or rules of thumb based on experience.

Particularly with respect to the selection of *probe packet size*, there have been two competing interests at play: a desire to minimize the bandwidth consumed by an active measurement probe stream, which suggests the use of small probe packets, and a desire for most accurate loss rate estimation, which may be an argument for the use of larger probe packets. Another consideration has been to use a probe packet size matched to the average packet size used by particular application traffic, *e.g.*, voice over IP traffic.

As examples, in the work of Zhang *et al.* [20], probe sizes of 64 and 256 bytes were used in order to limit probing overhead. In Mahajan *et al.* [11], however, probe sizes of 1080 bytes were used in order to more reliably elicit packet loss during congestion. In our previous work [16], a packet size of 600 bytes was used for similar reasons. To be sure, these latter two works are based on the idea that small packets are less likely to be lost. If however the use of small packets still leads to accurate measurement of packet loss, they must generally be considered a clear winner due to lower bandwidth consumption.

Our focus in this paper is to develop an empirically-grounded basis for selecting probe packet size for active loss measurement. We proceed through a set of controlled laboratory-based experiments in which we use the Harpoon traffic generator [14] running on commodity hardware and operating systems to create a diversity of traffic and loss conditions over a network of systems running the Click [10] modular router software. For a range of traffic conditions and router queue configurations we examine loss rates measured using different probe packet sizes. For our queue configurations, a key decision is whether the queue limit is specified in terms of packets (*i.e.*, the queue can hold a fixed maximum number of packets) or in bytes (*i.e.*, the queue can hold a fixed maximum number of bytes, irrespective of the number of packets). For our experiments, we find that when the queue limit is set in number of packets, the choice of probe packet size has little impact on measurement accuracy. However, if the queue limit is set in bytes, the choice of probe packet size has a clear impact on measurement accuracy. Overall, we find that the choice of how a queue is configured (packets or bytes) is perhaps *the* critical factor when considering the question posed by the title of this paper. Our results lead us to suggest that the probe packet size chosen should be matched, at least roughly, to the average packet size of the background traffic along the measurement path.

The contribution of our work is to explore of the problem of selecting a packet size for active loss measurement in a controlled laboratory setting. Previous work in [8, 11] has either examined this issue in a casual way or only in a simulated setting. In contrast, our focus is on rigorous examination of the impact of key experiment parameters in a controlled laboratory environment. While we consider our present work a first step rather than a definitive statement, we believe that our findings represent a promising start toward establishing a better understanding of how to parameterize tools for active measurement of packet loss.

2. Related Work

Over the years, there have been a number of studies of packet loss in the internet, each with somewhat different emphases and scope. While several studies have focussed on the examination of the sampling process (*i.e.*, the process by which the probes are emitted), others have used current best practices and have sought to understand wide-area loss characteristics. Many early studies used periodically emitted probes (similar to the `ping` tool) as a basis for measurement. For example, the early study by Bolot [7] used a UDP-based tool to periodically emit probes of 32 bytes each for measuring end-to-end packet delay and loss. The work by Yajnik *et al.* used a similar probe process in their study. It was not until the work of Paxson [12, 13] that emitting probes according to a Poisson process in accordance with the PASTA (Poisson arrivals see time averages) principle [19] became *de rigueur* in active loss measurement. This practice was further standardized by the IPPM working group within the IETF [1]. The well-known study by Zhang *et al.* [20] used probes of 64 or 256 bytes sent according to a Poisson process at two different rates.

More recently, the use of the PASTA principle for active loss measurement has come into question, most notably in the work of Baccelli *et al.* [2, 3], in which the authors have shown that Poisson probing is suboptimal and have proposed using a Gamma process instead. Other alternatives to a Poisson process have also been proposed by Sommers *et al.* [15, 16], in which the authors examined a probing process based on a geometric distribution. Thus, the overwhelming focus in past studies of active loss measurement has been the probing process.

Nevertheless, there may be other key aspects of probing that must be carefully studied, in particular the probe packet size. Most prior studies have justified the selection of probe packet size by a desire to limit probing bandwidth overhead. Other studies have cited a need to have somewhat larger probes in order to more effectively elicit loss during periods of congestion [11, 16]. In a somewhat different context, Floyd *et al.* [8] (and also the related work by Widmer *et al.* [18]) have examined a small-packet variant of the TCP-friendly rate control. In purely a simulation setting, they show that small packets are less likely to experience loss when queue capacity is measured in bytes (as compared with queueing in packets), and also when using active queue management mechanisms where the dropping function is in terms of bytes rather than packets. While there are some similarities between their work and ours, our focus is substantially different.

3. Methodology

In this section, we describe our experimental methodology for examining the issue of probe packet size selection for active loss measurement. As motivation for our methodology, consider a sequence of packet arrivals at a router's queue. Assume that the queue operates in a drop-tail discipline. Let the occupancy of the router queue at time t be denoted by $Q(t)$, and the maximum occupancy of the queue as Q_{max}. Now, let the size of an arriving packet be denoted by x. If the queue occupancy is measured in *packets*, then the new occupancy of the queue is $Q(t)+1$, whereas if the queue occupancy is measured in *bytes*, the new occupancy is $Q(t)+x$. Now, consider the indicator function I, which yields 1 (true) if the arriving packet is dropped, and 0 (false) otherwise. Namely, for queue occupancy in packets: $I(x) = [(Q(t)+1) > Q_{max}]$, and queue occupancy in bytes: $I(x) = [(Q(t)+x) > Q_{max}]$.

Clearly, in the queue-by-packet (QbP) case, the decision to drop a packet is made independently of the arriving packet's size, whereas in the queue-by-bytes (QbB) case, the arriving packet's size matters. Thus, one might hypothesize that if the queues in a network operate in a QbP-like manner, the specific size of probe packet used for active loss measurement may not be critical. Providing evidence for such a hypothesis would be a strong argument in favor of the use of small packets. On the other hand, the choice of probe packet size may indeed matter a great deal if network queues are implemented in a QbB fashion.

Another issue to consider is the manner in which queue space is made available for newly arriving packets. For the QbP case, the departure of a single packet (regardless of its size) is sufficient to admit a new packet. If a QbP-based queue contains small packets, the time for it to drain completely will be much less than if it contains large packets, given a fixed transmission rate for departing packets. On the other hand, for a QbB-based queue, it will take the same time to drain completely regardless of the sizes of packets stored (ignoring any per-packet overhead). Thus, one could hypothesize that not only does the packet arrival process play an important role in active loss measurement, but that *one should also consider packet size characteristics of the background traffic* when selecting a probe packet size.

In this work, we are interested in developing an empirical understanding of active loss measurement probes that use different packet sizes. We focus on experiments run in a controlled laboratory setting using commodity hardware and software-based routers. In future work, we intend to expand the range of parameters used and to run experiments in testbeds with standard commodity routers. The parameters we chose to vary were: probe packet size, whether the queue operates in a QbP or QbB mode, the nature of the background traffic, and the packet size distribution of the background traffic.

For the active loss measurement probe, we used a simple geometrically-distributed probe process based on the one described in Sommers *et al.* [16]. We divide time into fixed intervals of 5 milliseconds. At each interval, we send a probe with independent and uniform probability p. In this work, we set p to 0.3. The probe sizes we chose to use were 40, 256, 576, 1152, and 1448 bytes.

Table 1. TCP maximum segment size settings for background traffic scenarios.

Setting	TCP MSS (fraction of connections with that MSS)
A	1448 (100%)
B	1448 (80%), 576 (20%)
C	1448 (60%), 576 (20%), 256 (20%)
D	576 (100%)
E	256 (100%)

In order to create lossy conditions, we used the Harpoon traffic generator to create two different types of background traffic. In one scenario, we created long-lived TCP sources, and in another we transfered files whose sizes are heavy tailed, creating bursty self-similar conditions. In the self-similar scenario, traffic load was tuned to an average of 70% of the bottleneck link capacity.

For each background traffic scenario, we used five different TCP maximum segment size (MSS) settings in order to create different packet size distributions. For each setting, we set some fraction of connections to have a given MSS. Table 1 shows these different setups.

In our network setup, we used the Click software [10] running (in kernel) on commodity hardware for our routers. The standard `Queue` element in Click software distribution only supports QbP mode. We modified the standard `Queue` element to operate either in QbP or QbB mode. In this paper, we only used drop-tail queues; we intend to examine the impact when employing active queue management (AQM) mechanisms in future work.

4. Experiments

Testbed Setup. Our testbed, depicted in Figure 1, consisted of a set of commodity workstation hosts running Linux 2.6 and FreeBSD 7 and a set of commodity servers running the Click software router [10]. Two of the hosts were used to send and receive the loss probes and four hosts were used to generate background traffic conditions using the Harpoon traffic generator [14] across the routers in the topology. Connections were made from each client host to each server host, resulting in four different end-to-end paths for background traffic. The bottleneck link in the topology was a 100 Mb/s link between routers A and B in the testbed. All other links in the testbed were Gigabit Ethernet links. The output queue at router A on the link between routers A and B was the queue that was modified for various experiments we ran. For QbP experiments, we set the limit to 200 packets. For QbB experiments, the limit was set to 256 KB.

Packet traces were captured on the links between the servers and router A and also on the bottleneck link (trace facility not shown in the figure). Using these packet traces, we could identify all packets that were lost due to congestion at the outbound queue onto the Fast Ethernet link at router A. Thus, this trace facility was used to establish ground truth measures for our experiments.

All hosts had a minimum of a Pentium 4 processor. We measured CPU utilization on all hosts during the experiments to ensure that no host was overloaded. Router A, in particular, was equipped with an Intel Xeon quad-core processor running at 2.4 GHz.

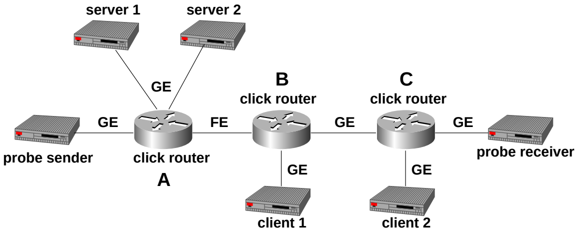


Figure 1. Laboratory testbed.

Results. In our results, we focus on the *loss average* statistic (also referred to as the "loss rate"). This simple statistic is computed as the number of packets loss divided by the total number of packets (*e.g.*, total packets arriving at a queue, or total packets emitted by a probe source). We compute the loss average for both the probes and for the aggregate background traffic. In addition to this general statistic, we also plot the distribution of loss over the range of observed packet sizes. Note that with the loss probes the typical goal is to estimate the aggregate loss average statistic. Thus, an additional measure we calculate for our experiments is the relative error between these measures. If $\hat{L}$ is the loss average estimated by the probes and L is the "true" aggregate loss average (measured using the trace facility described above), the relative error can be calculated as

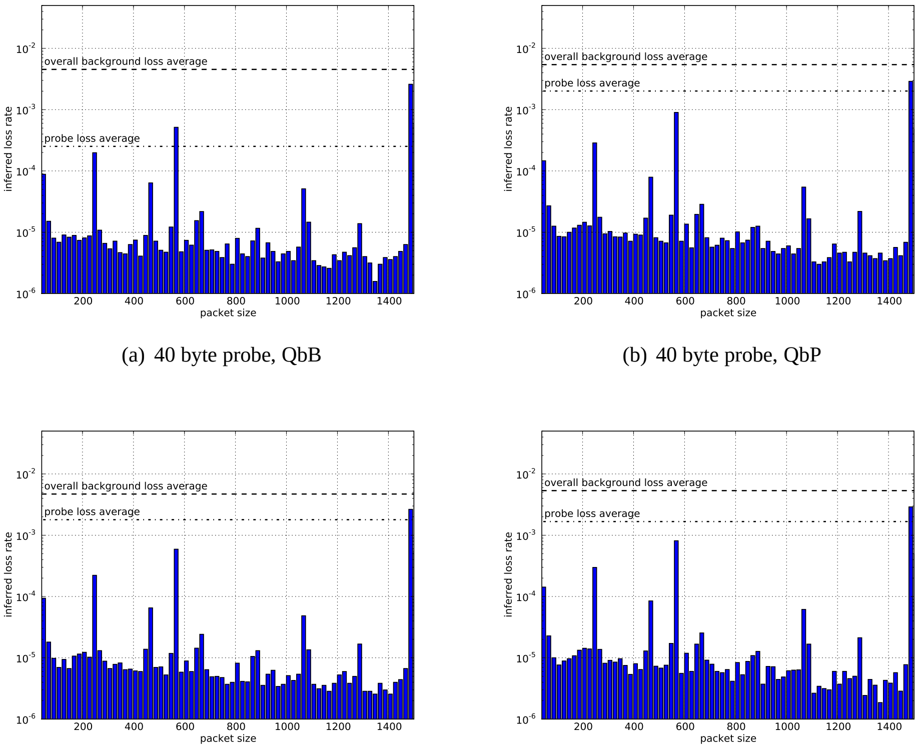
$$\frac{|\hat{L}-L|}{min(\hat{L},L)}$$

Turning to the results, Figure 2 shows background traffic loss distributions for the self-similar background traffic scenario using TCP MSS distribution B. Pairs of plots are shown for the QbB and QbP queue configurations, for probe sizes of 40 and 1152 bytes (results of other experiments are omitted due to space limitations but are consistent with what we describe below). Each plot shows the fraction of losses sustained by background traffic packets of various sizes (each bar covers a range of 20 bytes). Interestingly, we see that, independent of configuration, small background traffic packets tend to be lost at a higher rate than packets of many other sizes, save for a peak around 576 bytes and another set of peaks around 1500 bytes. It is important to note that for these experiments, the most commonly occurring packet sizes are around 40 bytes, 576 bytes, and 1500 bytes. While it may not be surprising that packets with the most frequently occurring sizes are most likely to be lost, it might also be expected that small packets are less likely to be lost, regardless of how frequently they occur. The plots show that at least for the background traffic packets, this is clearly not the case. Note however, that these observations do not necessarily mean that small *probe packets* are more or less likely to be lost. Indeed, a key contributing factor is that multiple small packets are released in close proximity (*e.g.*, back-to-back), resulting in higher losses of small packets (since they may arrive more quickly than queue space can be freed by outgoing packets). Note that this observation holds for both types of background traffic used in our experiments. Thus, it is not enough to look simply at background traffic characteristics in examining the question posed by this paper.

In each plot of Figure 2, we also observe dashed horizontal lines showing the overall loss average measured for the aggregate background traffic and for the probes. (Due to inherent variability in the self-similar background traffic, aggregate the loss averages vary somewhat between experiments.) From these lines we first see that the probe loss averages measured for the QbP scenarios change very little despite changing the probe size. On the other hand, we see that in the QbB configuration with the smallest probe, the loss average measured is far below the aggregate background traffic loss average. We also see that the probe loss average is improved greatly with larger probe size. In fact, the best result (closest to the background traffic loss average) in the QbB configuration is for the 1152 byte probe. We note that these results are consistent across all our experiments.

Finally, in Figure 3 we compare relative error measured between the probe loss average statistic and the aggregate loss average. Thus, tall bars indicate poor estimation by a loss probe. We first examine the 8 groups of bars on the left half of the plot corresponding to the long-lived TCP background traffic scenarios. We see results for TCP MSS distributions A–D, and for QbP and QbB queue configurations. First, observe that for the QbP experiments there does not appear to be any clear advantage for any probe size. On the other hand, for the QbB experiments there are stark differences. We see that the results, in general, are worst for the 40 byte probe. Importantly, we observe that the best results are obtained by choosing the probe size closest to the average background packet size, as shown in Table 2.

For the self-similar results, we again see that for the QbP setups there are no observable patterns between probe size and relative error. For the QbB setups, although the relationship between average background traffic packet size and the probe size with the minimal relative error is not as obvious,



(a) 40 byte probe, QbB

(b) 40 byte probe, QbP

(c) 1152 byte probe, QbB

(d) 1152 byte probe, QbP

Figure 2. Loss distribution across observed packet sizes for self-similar experiments. TCP MSS disribution B was used for these experiments.

Table 2. Comparison of best probe size (probe size that minimizes relative error) with average packet size of background traffic for the long-lived TCP scenario with QbB.

TCP MSS Distribution	Best probe size	Average packet size
A	1152	874
B	1152	730
C	576	678
D	256	389

the pattern established with the long-lived TCP experiments also holds for the self-similar traffic. Namely, that the best results are for probe sizes that are closest to the average background packet size. Finally, we note that in additional experiments we ran (including simulation experiments, and experiments using additional background traffic configurations, not discussed due to space constraints), the results were consistent with what we present here.

5. Discussion and Conclusions

Our results demonstrate that, depending on how a network queue is implemented, the choice of probe packet size may be critical for gathering accurate loss measurements. The natural question that follows is: how do commercial routers work? In particular, are queue limits specified in packets or in bytes? Based on our experience in prior work, the answer to this question is device-specific as well as vendor-specific [4, 17]. Further, along a given path in the internet, it is likely that devices of multiple vendors will be encountered. In future work, we intend to examine the impact on loss probe packet size of different queue implementations along a single path. In addition, we intend to apply our experiments in more complex network topologies, using commercial routers.

Our results also show that when queues operate in byte mode (QbB), a key consideration when selecting probe packet size is the average packet size of the background traffic. Thus, having a knowledge of this traffic characteristic, or the ability to quickly estimate or infer this characteristic, is likely to be important. In future work, we intend to

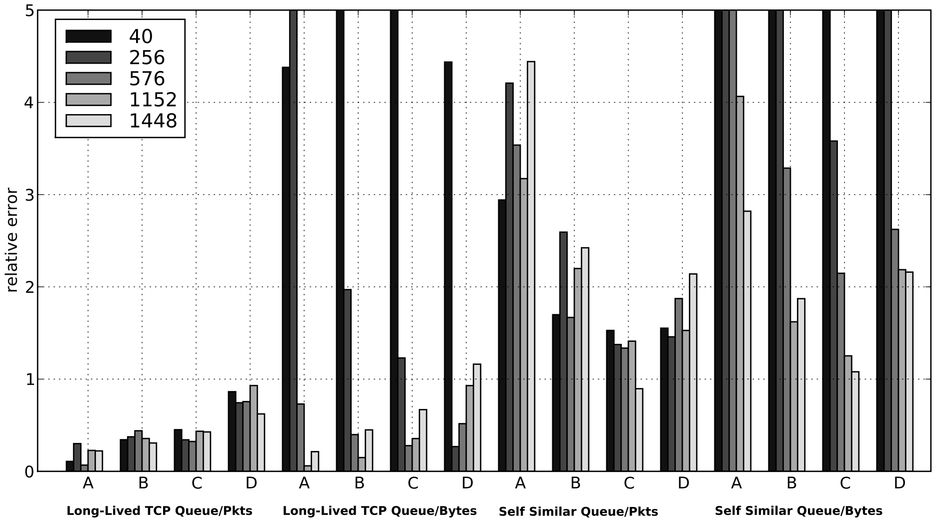


Figure 3. Relative error bars for laboratory experiments. Groups of five bars are shown for each probe packet size. Results for TCP MSS distributions A–D are shown for both the long-lived TCP source background traffic and the self-similar background traffic.

examine how this quantity might be inferred.

Finally, we note that we expect our results to apply to probe processes that are designed to estimate somewhat different characteristics of loss than the loss average, *e.g.*, loss episode frequency and mean loss episode duration [16]. We intend to examine this hypothesis in future work.

References

[1] G. Almes, S. Kalidindi, and M. Zekauskas. A one-way packet loss metric for IPPM. IETF RFC 2680, September 1999.

[2] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot. The role of PASTA in network measurement. In *Proceedings of ACM SIGCOMM*, Pisa, Italy, September 2006.

[3] F. Baccelli, S. Machiraju, D. Veitch, and J. C. Bolot. On optimal probing for delay and loss measurement. In *Proceedings of ACM SIGCOMM Internet Measurement Conference '07*, October 2007.

[4] P. Barford and J. Sommers. A comparison of probe-based and router-based methods for measuring packet loss. Technical report, University of Wisconsin - Madison, 2003.

[5] P. Barford and J. Sommers. Comparing probe- and router-based packet loss measurements. *IEEE Internet Computing*, September/October 2004.

[6] P. Benko and A. Veres. A passive method for estimating end-to-end TCP packet loss. In *Proceedings of IEEE Globecom '02*, Taipei, Taiwan, November 2002.

[7] J. Bolot. End-to-end packet delay and loss behavior in the Internet. In *Proceedings of ACM SIGCOMM '93*, San Francisco, September 1993.

[8] S. Floyd and E. Kohler. TCP Friendly Rate Control (TFRC): The Small-Packet (SP) Variant. IETF RFC 4828, April 2007.

[9] S. Jaiswal, G. Iannaccone, C. Diot, J. Kurose, and D. Towsley. Inferring TCP connection characteristics through passive measurements. In *Proceedings of IEEE INFOCOM*, April 2004.

[10] E. Kohler, R. Morris, B. Chen, J. Jannotti, and M. F. Kaashoek. The Click modular router. *ACM Transactions on Computer Systems*, 18(3), August 2000.

[11] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson. User-level Internet Path Diagnosis. In *ACM SOSP*, October 2003.

[12] V. Paxson. End-to-end Internet packet dynamics. In *Proceedings of ACM SIGCOMM '97*, Cannes, France, September 1997.

[13] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis. Framework for IP performance metrics. IETF RFC 2330, 1998.

[14] J. Sommers and P. Barford. Self-configuring network traffic generation. In *Proceedings of ACM SIGCOMM Internet Measurement Conference '04*, 2004.

[15] J. Sommers, P. Barford, N. Duffield, and A. Ron. Accurate and Efficient SLA Compliance Monitoring. In *Proceedings of ACM SIGCOMM*, August 2007.

[16] J. Sommers, P. Barford, N. Duffield, and A. Ron. A geometric approach to improving active packet loss measurement. *IEEE/ACM Transactions on Networking*, 16(2), April 2008.

[17] J. Sommers, P. Barford, A. Greenberg, and W. Willinger. An SLA Perspective on the Router Buffer Sizing Problem. *ACM SIGMETRICS Performance Evaluation Review*, March 2008.

[18] J. Widmer, C. Boutremans, and J.-Y. Le Boudec. Congestion Control for Flows with Variable Packet Size. *Computer Communications Review*, 34(2), April 2004.

[19] R. Wolff. Poisson arrivals see time averages. *Operations Research*, 30(2), March-April 1982.

[20] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker. On the constancy of Internet path properties. In *Proceedings of ACM SIGCOMM Internet Measurement Workshop '01*, San Francisco, November 2001.