On Optimal Probing for Delay and Loss Measurement

Francois Baccelli*
INRIA-ENS, Ecole Normale Supérieure, France
Francois.Baccelli@ens.fr

Sridhar Machiraju
Sprint, California, U.S.A
Machiraju@sprint.com

Darryl Veitch*†
Dept. of E&E Engineering
University of Melbourne, Australia
dveitch@unimelb.edu.au

Jean Bolot
Sprint, California, U.S.A
Bolot@sprint.com

ABSTRACT

Packet delay and loss are two fundamental measures of performance. Using active probing to measure delay and loss typically involves sending Poisson probes, on the basis of the PASTA property (Poisson Arrivals See Time Averages), which ensures that Poisson probing yields unbiased estimates. Recent work, however, has questioned the utility of PASTA for probing and shown that, for delay measurements, i) a wide variety of processes other than Poisson can be used to probe with zero bias and ii) Poisson probing does not necessarily minimize the variance of delay estimates.

In this paper, we determine *optimal probing processes that minimize the mean-square error of measurement estimates* for *both delay and loss*. Our contributions are twofold. First, we show that a family of probing processes, specifically Gamma renewal probing processes, has optimal properties in terms of bias and variance. The optimality result is general, and only assumes that the target process we seek to optimally measure via probing, such as a loss or delay process, has a convex auto-covariance function. Second, we use empirical datasets to demonstrate the applicability of our results in practice, specifically to show that the convexity condition holds true and that Gamma probing is indeed superior to Poisson probing. Together, these results lead to explicit guidelines on designing the best probe streams for both delay and loss estimation.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Measurement Techniques, Modeling Techniques, Performance Attributes.

General Terms

Design, Measurement, Performance, Theory.

Keywords

Active Probing, Convexity, Auto-covariance, Variance, PASTA.

1. INTRODUCTION

In packet networks, packet loss and delay are two of the most fundamental measures of performance. Their role is even more central when it comes to end-to-end measurements, particularly using *active probing*. Here, loss and delay are not only important performance metrics in their own right, they are *the* information carried by probes about the network 'system' they traverse, information that can be used to infer network parameters such as link capacities and available bandwidth.

Of the two, loss is harder to measure and exploit because it is harder to find in many of today's networks, in particular in the Internet core, and because loss data manifests as a per-packet loss/no-loss indication, a binary variable, which is much coarser than continuous delay, and thereby carries far less information. As a result of these difficulties, with few exceptions, probing for loss has remained straightforward and it has essentially consisted in sending isolated probes and estimating average loss probability by observing their empirical loss rate.

More sophisticated techniques have been proposed for delay measurements involving, for example, probe 'trains' at the sender, and interpreting the spacing between train packets at the receiver. Analogous techniques for loss are hampered by the fact that multiple losses in a train is an event so rare that it would almost never be observed in experiments of reasonable length (excessively intrusive probing rate aside).

In all cases, using active probing to measure delay and loss typically involves Poisson probes or trains, that is exponential probe or train separations, on the basis of the PASTA property (Poisson Arrivals See Time Averages), which ensures that Poisson probing yields unbiased estimates from a single sample path. This property has been used, in particular since Paxson's work [15], to justify sending probes as a Poisson stream. Although the validity of PASTA itself is not in question, Baccelli et al. [2] recently questioned its utility for network probing as a whole. Specifically, results in [2] show that for delay measurements, a wide variety of

*Work funded in part by the "Equipes Associés" INRIA Programme on Internet Probing.

†ARC Special Research Centre on Ultra-Broadband Information Networks, CUBIN is an affiliated program of National ICT Australia (NICTA).

IMC'07, October 24-26, 2007, San Diego, California, USA.

(non-intrusive) processes other than Poisson, which have a *mixing* property, can be used to probe with zero bias.

This paper is motivated by the observation and counterexamples of [2] showing that Poisson probing is not optimal, that is, it does not minimize the variance of delay estimates. This opens up the general question of whether there exist optimal probing strategies. We address this not only for delay as in [2] but also for loss, and not just for simple measures such as average loss and marginal delay distributions, but more complex ones such as loss correlations and delay jitter. We give rigorous results on probing strategies using a network model general enough to include cases relevant to today's network practitioners, including multiple hops where loss mechanisms may vary.

Our first set of results is theoretical. As mentioned earlier, it was shown in [2], in the context of delay, that Poisson sampling can be sub-optimal for variance. We confirm that this holds also for loss, and go much further, by providing a theorem on minimizing variance with wide applicability and giving an explicit family (Gamma renewal) of probing streams, all of which have variance that is better than that of Poisson, and simultaneously zero sample path bias. The optimality result is very general and can be applied to any probing problem assuming that the target process we seek to optimally measure via probing has a convex auto-covariance function. In a probing context, the target process could be a delay process or a loss process, with measures of interest ranging from average delay to average loss, number of packets in a train lost, etc.

In our second set of results, we argue that the convexity property is present in real systems and provide experimental evidence to support our argument. We use empirical datasets collected on routers and on the Internet backbone to demonstrate that the convexity condition holds true and that Gamma probing is indeed superior to Poisson probing.

Together, these two sets of results lead to explicit guidelines on designing the best probe streams for both delay and loss estimation. These guidelines are consistent with the spirit of the probe pattern *separation rule* introduced in [2] as a replacement to the Poisson default.

The remainder of the paper is structured as follows. Section 2 presents background information on sampling for delay and loss. Section 3 proves two theorems on optimal variance for the estimation of delay and loss, and discusses their generality. In Section 4 and Section 5, we provide extensive measurement results primarily based on a unique full-router dataset to support our results. In Section 7 we compare our work to that of others, and conclude in Section 8.

2. OVERVIEW

In this section, we use prior work [2] to provide an overview of the current state-of-art and motivate the problem we study. We first start by discussing the *ground truth* process, which is what we want to measure.

2.1 The Ground Truth Process

Consider the measurement of loss. We begin by defining what it is we actually want to measure, the 'ground truth' for loss. An obvious candidate is the probability p_x that a packet of size x (bytes), injected into the network at some source, would fail to arrive at the receiver. However, despite its importance, p_x is not rich enough to describe the sampling that occurs during measurement. Instead, we take our ground truth to be a binary stochastic process $I_x(t)$, which takes value 1 if a packet of size x, *if it were* injected at time t, would be lost, and 0 otherwise. We assume that this *loss process* $I_x(t)$ is stationary, in which case p_x is well defined as its marginal loss probability $p_x = \mathrm{P}(I_x(t) = 1)$, a constant for any time t.

The definition of $I_x(t)$ is natural and direct: it records whether loss would have occurred or not, without asking *how*. It is therefore a general definition, and as it makes no assumption on loss mechanism or policy, it applies to very general networks.

The following examples help to make this notion of the ground truth as a *process*, more concrete. First, take a simple one hop path model, consisting of a finite FIFO buffer with a continuous time occupancy process $B(t)$ bytes, buffer size K bytes, and droptail dropping policy. In this case, writing $\mathbf{1}$ for the indicator function,

$$I_x(t) = \mathbf{1}\{B(t) + x > K\}, \tag{1}$$

which depends strongly on x. If, on the other hand, the dropping policy was based on the number $N(t)$ of packets, with a maximum of K packets, then

$$I_x(t) = \mathbf{1}\{N(t) + 1 > K\}, \tag{2}$$

which is independent of x. We now extend the first example to a 2-hop FIFO tandem network. To do so, note that the level $B_2(\cdot)$ of the second buffer must be observed when the packet would arrive to it, provided it was not lost at the first. The loss process is therefore

$$\begin{aligned} I_x(t) &= \mathbf{1}\{B(t) + x > K\} \\ &+ \mathbf{1}\{B(t) + x \le K,\, B_2(t + d_x(t)) + x > K_2\}, \end{aligned} \tag{3}$$

where $d_x(t) = (B(t) + x)/C$ is the delay over the first buffer, C is the first hop bandwidth, and K_2 the size of buffer B_2. The above examples show that the actual loss behavior depends strongly on each of packet size, node policies, and network topology. The loss process $I_x(t)$ is general enough to encompass all of these.

It is crucial to note that $I_x(t)$ describes what a probe *would see* if it entered at time t, but the probe *does not actually do so*. The observed loss process is therefore a functional of the system state which is not in any way perturbed by probes. Similarly, we can define the ground truth process for delay, too. Specifically, if $D_x(t)$ is the ground truth delay observed by a packet of size x, then $D_x(t)$ is the delay a packet of size x would have experienced had it been injected into the system at time t. A special case is the *virtual delay* process $D_0(t)$, the delay experienced by a virtual (zero-sized) probe injected into the system at time t.

The process $I_x(t)$ relates to what a single probe would see. We would also like to send in probe patterns, clusters or trains of probes, and ask what losses they would experience. For example, we could ask *whether a probe train with m packets having sizes $\vec{x}$, and inter-probe spacings $\vec{\tau}$, would lose at least one packet if it entered the network at time t.* Alternatively, we could inquire after the distribution of the number of packets lost in a train. Questions such as these cannot be answered using the loss process $I_x(t)$. The time dimension of $I_x(t)$ corresponds simply to the fact that we can ask the question, 'would a packet be lost', at any time. It is not possible to use $I_x(.)$ at times t and t' to determine what losses would be experienced by two packets sent at t and t' simultaneously, because $I_x(t)$ is uninfluenced by the probe at t, and so the second probe, like the first, arrives to an unperturbed system.

To answer questions on losses in probe trains, a more complex notion of ground truth is needed, and one which is specific to train type. For example, we could define the following binary train-loss process $I_{\vec{x},\vec{p}}(t)$, which takes value 1 if at least one probe in the train (defined by the size and spacing vectors $\vec{x}$ and $\vec{p}$), sent into the network at time t, is lost, and 0 otherwise. Note that just as in the case for $I_x(t)$, these probe trains do not actually enter the system. The train-loss process $I_{\vec{x},\vec{p}}(t)$ is a functional of the unperturbed system only. Analogously to before, the variable t is simply a way to ask the question: 'would probes in a train be lost if sent now'.

2.2 Sampling and Intrusiveness

There are two main issues in measurement via probing: *sampling*, and *intrusiveness*. Sampling concerns the fact that probes can only experience, and reveal, the state of the network at particular discrete times. Intrusiveness refers to the fact that real probes perturb the network. This paper focuses on the sampling issue, in particular the question of variance, in a *non-intrusive* context. Before we can do so we must explain what non-intrusiveness really means, and how it is attained.

Consider the case of the delay ground truth process $D_x(t)$ defined above. An ideal observer of the network would know the value taken by $D_x(t)$ over continuous time. In practice, with active probing, continuous observation is not possible. Instead, probe packets are sent at discrete times $\{T_n\}$, and it seems natural to view the observations as the samples $\{D_x(T_n)\}$ of $D_x(t)$. However, this will not be the case because real probes perturb the network, and so their experiences are not described exactly by $D_x(t)$ but by some perturbed form of it. Dealing with these perturbations is a difficult problem that also interacts with sampling. Hence, to study sampling, it is advantageous to somehow avoid perturbation. In [2], this motivated the use of the virtual delay process $D_0(t)$ as a basis of non-intrusive sampling, since zero sized probes do not add to service time, yet will still return meaningful delays. In fact, in many systems, they correspond precisely to the samples $\{D_x(T_n)\}$ of the $x = 0$ ground truth $D_0(t)$.

Although virtual probes were well suited for the purposes of [2], there are several problems with using them as a basis for non-intrusive probing in general. The most obvious one is that for loss, $x = 0$ seems useless, since such probes would never be lost in systems like Equation 1, and therefore would not carry very useful information about the loss experienced by typical packets. Second, virtual probes **are** in fact intrusive whenever network elements act on the number of packets rather than their byte-size (for example in packet based loss systems such as Equation (2)), and this holds equally for loss and delay. The most important reason however is the following. Whereas in [2] $x = 0$ was synonymous with non-intrusiveness, our definition of ground truths for both loss and delay are for arbitrary $x \geq 0$. Hence we need a definition of non-intrusive probing which is independent of probe size. In fact, it must be:
(i) General: applicable to any loss or delay ground truth,
(ii) Faithful: i.e. guaranteed to be non-intrusive for any system,
(iii) Sampling Compatible: amenable to a direct sampling interpretation.
The only way to guarantee each of these properties is to define a purely sampling viewpoint. In other words, relative to a given ground truth process $X(t)$, we *define non-intrusive probing directly as the sampling* $\{X(T_n)\}$, where the $\{T_n\}$ are the times the probes enter the system. Like virtual probes, this notion of non-intrusive probing is theoretical, it cannot be obtained by real probes, but it can be calculated by an oracle with full access to the system state.

For completeness, recall that [2] also dealt with $x > 0$, however this was in an intrusive context for delay only. An extension of that intrusive case to loss is discussed in [3]. Here we do not discuss the intrusive case at all beyond some comments in Section 6.

2.3 NIMASTA

In [2] we coined and proved the NIMASTA property, that *Non-Intrusive Mixing Arrivals See Time Averages*. The context is that of a probe point process which is independent of some ground truth process. The latter is just assumed to be stationary and ergodic, whereas the probe point process is assumed to be stationary and *mixing*. Mixing is a strong form of ergodicity (see [2] for details) which guarantees *joint ergodicity* between the probe and ground truth processes. By 'see time averages', we mean that empirical averages made by the probe observations of a given ground truth tend to the true value, namely to the mean value of the stationary ground truth process:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(X(T_n)) = E[f(X(0))], \tag{4}$$

where the convergence is in the almost sure (a.s.) sense. In this formula, f is an arbitrary positive function and $X(t)$ is the ground truth process of interest. In statistical parlance, this result is one of *strong consistency* of the estimator appearing on the LHS for $E[f(X(0))]$, namely estimates from a single sample path a.s. converge to the true value when the number of samples grows large. Such a property is very desirable in the active probing context where only a *single* sample path is available. In what follows, we will say that when (4) holds the estimator has *no sample-path bias*.

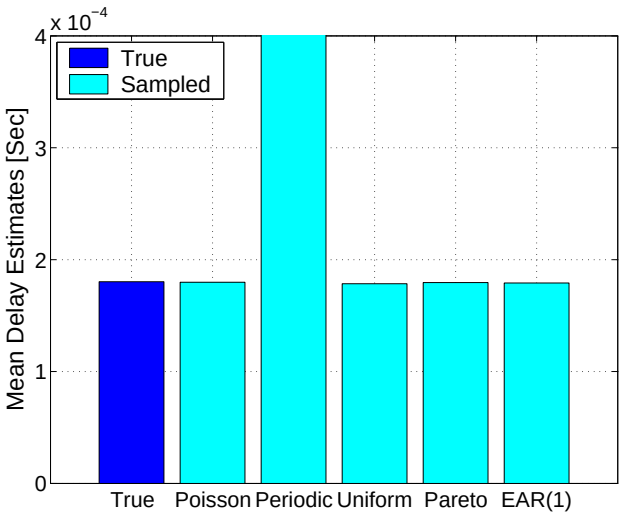

Figure 1: Result from [2] illustrating how mixing processes have zero sample-path bias whereas non mixing streams (such as Periodic) may not.

In [2], Equation (4) was proved in the context of delay for virtual probes, that is $X(t) = D_0(t)$, and three examples of f functions were given corresponding to average delays, the delay distribution, and jitter. In fact, by using the definitions both of ground truth given in section 2.1, and of non-intrusive probing (as defined in section 2.2), the proof of Equation (4) generalizes naturally to arbitrary ground truth processes (see [3]). In this way loss and delay can be integrated into a single framework, and hence the results of [2] can be extended to loss.

The assumption that the probe point process is mixing is necessary for strong consistency to hold. Figure 1 gives an example (taken from [2]) of ergodic probe processes, one of which, the Periodic case, is not mixing. In this case joint-ergodicity is not satisfied, resulting in a phase-lock phenomenon which lead to sample-path bias. Note that sample path bias differs from true bias: it is easy to show that the estimator defined in the LHS of (4) is unbiased for all ergodic probe point processes. In the case of periodic sampling, we have a lack of strong consistency (and hence a non-zero sample path bias as illustrated in the above example) combined with zero true bias.

In this paper, we employ probes satisfying the mixing property so as to guarantee strong consistency, and we examine optimal probing through an analysis of variance within this framework.

3. SAMPLING FOR OPTIMAL VARIANCE

In this section, we explore a class of mixing renewal sampling (or probing) point processes (a point process is a renewal process if it has i.i.d inter-arrival times) in an attempt to determine which probe streams are best to use from the point of view of minimizing estimation variance. Recently, explicit but simple examples were given [2] showing the sub-optimality of Poisson probing (Poisson point processes belong to the class of mixing renewal processes). The main result we prove in this section is that Poisson probing/sampling is provably not optimal (in terms of variance) in this class under very general conditions, provided the auto-covariance function of the relevant ground truth is *convex*. Furthermore, under the same conditions, Gamma probing/sampling processes (which are also mixing renewal) help achieve a variance that is as close to the lowest as possible.

3.1 The Convexity Condition

We denote the ground truth continuous time stochastic process, which we seek to optimally sample, by $X(t)$. We denote its mean by p and its auto-covariance function by:

$$R(\tau) = \mathbb{E}[X(t)X(t+\tau)] - p^2.$$

We work within the same framework of general networks as described in Section 2, but focus on those for which the auto-covariance function $R(\tau)$ exists and is *convex* for $\tau \geq 0$.

The usual sample mean estimator of p using N samples is

$$\hat{p}_1 = \frac{1}{N}\sum_i X(T_i) \tag{5}$$

where $T_0 = 0$ by convention and T_i is the sum of i inter-sample times, which due to stationarity, each have law G with mean μ. Hence T_i has mean $i\mu$, and we denote its probability density by f_i. The variance of $\hat{p}_1$ is given by

$$\begin{aligned}\text{Var}[\hat{p}_1] &= \frac{1}{N^2}\left(NE[X(0)^2] + 2\sum_{i\neq j}\mathbb{E}[X(T_i)X(T_j)]\right) - p^2 \\ &= \frac{1}{N^2}\left(NE[X(0)^2] + 2\sum_{i\neq j}\int R(\tau) f_{|i-j|}(\tau)d\tau\right) - p^2 \end{aligned} \tag{6}$$

which is a function both of the variability of the process $X(t)$ via $R(\tau)$, and that of the sampling stream via the f_k.

As a special case of Equation (5), we pick out the estimator based on periodic samples of period μ, namely

$$\hat{p}_2 = \frac{1}{N}\sum_i X(i\mu), \tag{7}$$

for which the integral $\int R(\tau) f_{|i-j|}(\tau)d\tau$ in Equation (6) degenerates to $R(|i-j|\mu)$.

THEOREM 1. *If $R(\tau)$ is convex, then Var$[\hat{p}_1] \geq$ Var$[\hat{p}_2]$.*

PROOF. Equation (6) holds for all processes. So, to compare the variances it is enough to compare, for all $i \neq j$, the cross terms, namely $\int R(\tau) f_{|i-j|}(\tau)d\tau$ and $R(|i-j|\mu)$. But, if $R(\tau)$ is convex, Jensen's inequality says that

$$\int R(\tau) f_k(\tau)d\tau \geq R\left(\int t f_k(t)dt\right) = R(k\mu) \tag{8}$$

for any k. The result follows. □

We have learnt that under the foregoing assumptions, *no* other sampling process has a variance which is lower than that of periodic sampling. As just one example, by taking G to be exponential in $\hat{p}_1$ and inter-sample times to be independent, we learn that Poisson sampling yields a higher variance than periodic. However, the result is much more powerful than this. It shows that, if $R(\tau)$ is convex, no kind of train or other structure, no matter how sophisticated, can do better than periodic.

Unfortunately periodic sampling does has a disadvantage already discussed: it is not mixing, which makes it vulnerable to sample-path bias due to phase locking effects. Assuming that $R(\tau)$ is convex, we now determine sampling schemes that offer the best of both worlds: mixing to avoid sample-path bias, but with variance close to that offered by periodic sampling.

We will consider sampling using renewal processes that are Gamma distributed. A Gamma law has a two parameter density given by

$$\Gamma_{\alpha,\lambda}(x) = \frac{\lambda}{\Gamma(\alpha)}(\lambda x)^{\alpha-1}e^{-\lambda x}, \tag{9}$$

where $\Gamma(\cdot)$ is the familiar Gamma function, and has mean $\mu = \alpha/\lambda$ and variance $\sigma^2 = \alpha/\lambda^2$. Since $1/\lambda > 0$ is a scale parameter, if $T \sim \Gamma_{\alpha,\lambda}$, then $cT \sim \Gamma_{\alpha,\lambda/c}$. Gamma laws are also stable with respect to the shape parameter α, that is, if $\{T_i \sim \Gamma_{\alpha_i,\lambda}\}$ are independent, then $\sum_i T_i \sim \Gamma_{\sum_i \alpha_i,\lambda}$. The exponential laws correspond to the 1-parameter sub-family $\Gamma_{1,\lambda}$. Another special sub-family are distributions with the **Erlang** law. These have only integral shape values.

We will need one more technical result regarding Gamma laws, the proof of which we leave to the appendix.

LEMMA 3.1. *Let $T \sim \Gamma_{\alpha,\lambda}$, $Z \sim \Gamma_{\beta,\lambda}$ be independent, and set $Y = T + Z$. Then $C = \mathbb{E}[T|Y] = \alpha Y/(\alpha+\beta)$ has density $\Gamma_{\alpha+\beta,(\alpha+\beta)\lambda/\alpha}$, with mean $\mathbb{E}[C] = a/\lambda = \mathbb{E}[T]$.*

We can now prove

THEOREM 2. *The family of renewal sampling processes $G(\beta)$, parametrized by $\beta > 0$, with inter-sample time density $\Gamma_{\beta,\beta\lambda}(x)$, provides, at constant mean sampling rate λ, sampling variance for $\hat{p}_1$ that monotonically decreases with β. The variance is larger (equal or smaller) than Poisson sampling as β is smaller (equal or larger respectively) than 1, and tends to that of periodic sampling in the limit $\beta \to \infty$.*

PROOF. We assume an underlying probability space on which the family of inter-sample variables are defined for each $\beta > 0$. Equation (6) holds for each inter-sample law $G(\beta)$. As the means for each are equal to $\mu = \beta/(\beta\lambda) = 1/\lambda$, proving the variance result reduces to showing that, for each $k > 0$, $\int R(\tau) f_{k,1}(\tau)d\tau \geq \int R(\tau) f_{k,2}(\tau)d\tau$ for any β values β_1, β_2 satisfying $\beta_2 > \beta_1$, where $f_{k,i}$ is the density of the sum $T_{k,i}$ of k inter-sample times, each with law $G(\beta_i)$. We can apply Jensen's inequality to show that

$$\begin{aligned}\int R(\tau) f_{k,1}(\tau)d\tau &= \mathbb{E}[R(T_{k,1})] \\ &= \mathbb{E}[\mathbb{E}[R(T_{k,1})|Y_{k,1}]] \\ &\geq \mathbb{E}[R(\mathbb{E}[T_{k,1}|Y_{k,1})] \\ &= \mathbb{E}[R(T_{k,2})] = \int R(\tau) f_{k,2}(\tau)d\tau\end{aligned}$$

where to show $\mathbb{E}[T_{k,1}|Y_{k,1}] = T_{k,2}$ we identified $(T, Y, \alpha, \beta, \lambda)$ with $(T_{k,1}, Y_{k,1}, k\beta_1, k(\beta_2-\beta_1), \beta_1\lambda)$ and used Lemma 3.1. Since this holds for any β_1, β_2 with $\beta_2 > \beta_1$, we have monotonicity of the variance in β. As β tends to infinity, there is weak convergence of $\Gamma_{\beta,\beta\lambda}(x)(dx)$ to a Dirac measure at $1/\lambda$, as is easily seen using

Laplace transforms. Since the function R is convex, it is continuous, and as it is also bounded (as a second order process), the property

$$\lim_{\beta\to\infty}\int R(x)\Gamma_{\beta,\beta\lambda}(x)(dx) = \int R(x)\delta_{1/\lambda}(dx)$$

follows from the very definition of weak convergence. This shows that the limit of the variances of the Gamma renewal estimators is that of the deterministic probe case, namely the optimal variance. □

This result provides a family of sampling processes with the desired properties. By selecting $\beta > 1$, we can ensure lower (more precisely, no higher) variance than Poisson sampling. By selecting β large, we obtain sampling variance close to the lowest possible, whilst still using a mixing process. Exactly what value should be chosen, however, will depend on other factors. For example, extremely large values might increase vulnerability to 'transient' phase locks in atypical sampling paths or in systems that are not stationary. The important point is that the parameter β can be used to continuously tune for any desired trade-off, and to set the sampling variance arbitrarily close to the optimal case.

This theorem is quite general and applies to any sampling problem provided the selected ground truth process has a convex auto-covariance function. As already explained, in the probing context, $X(t)$ can refer to the simple delay $D_x(t)$ and loss $I_x(t)$ processes and temporal properties thereof, and to train-based metrics too. Theorem 2 can hence be used to decrease the variance of our estimate of p_x, the average loss probability associated with the loss process $I_x(t)$, or any other loss metric such as the train-loss process $I_{\vec{x},\vec{p}}(t)$, or $M_{\vec{x},\vec{p}}(t)$, the number of packets in a train which are lost.

3.2 Known Convex Examples

A natural question is, how likely is it that networks of interest satisfy the convexity property for delay and/or loss? There are simple systems for which exact results are known. For example, Ott [11] showed that convexity holds for the virtual work process (equal to the delay of probes with $x = 0$) of the M/G/1 queue.

We now show that the loss process $I_x(t)$ of the $M/M/1/K$ queue, with the packet-based dropping model of Equation (2), has a convex auto-covariance function. Denote by λ and μ the arrival and the service rates and by $\rho = \lambda/\mu$ the load factor. From [19] (p.13, Theorem 1), the probability that the number of customers in the queue is K at time t, given that it is K at time 0, is

$$P_{K,K}(t) = \frac{1-\rho}{1-\rho^{K+1}}\rho^K + \frac{2}{K+1}\sum_{j=1}^{K}\frac{\exp(-(\lambda+\mu)t + 2t\sqrt{\lambda\mu}\cos(\pi j/(K+1))}{1-2\sqrt{\rho}\cos(\pi j/(K+1))+\rho} \cdot (\sin(Kj\pi/(K+1)) - \sqrt{\rho}\sin(j\pi))^2 \quad (10)$$

in the case when $\rho \neq 1$ and

$$P_{K,K}(t) = \frac{1}{1+K} + \frac{1}{K+1}\sum_{j=1}^{K}\frac{\exp(-(2\lambda)t + 2\lambda t\cos(\pi j/(K+1))}{1-\cos(\pi j/(K+1))} \cdot (\sin(Kj\pi/(K+1)) - \sin(j\pi))^2 \quad (11)$$

in the case $\rho = 1$. In both cases, the auto-covariance function of $I_x(t)$, which is equal to $\pi(K)P_{K,K}(t)$ (with $\pi(K)$ the stationary probability that the queue has K customers) is a convex combination of convex decreasing functions of t and is hence itself convex and decreasing in t.

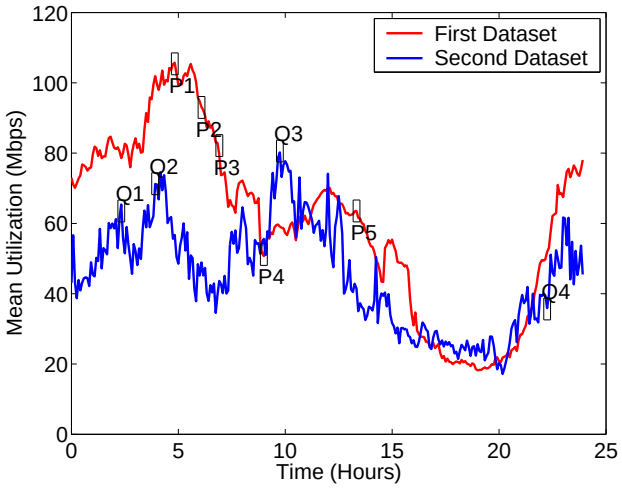

Figure 2: Utilization of the target output interfaces of the first and second datasets. We also mark P1-P5 and Q1-Q4, the representative time intervals used to illustrate the results throughout this paper.

4. FULL-ROUTER RESULTS

Our theorems on the estimation variance of various probing streams are valid if the convexity condition is true. In the previous section, we described two simple systems in which this condition is provably true. It is not possible, however, to prove analytical results for real Internet traffic. Hence, in this section, we use empirical datasets to demonstrate the applicability of our results to real networks, namely, that the convexity condition holds true and that Erlang (Gamma) probing is superior to Poisson probing. We start by describing our dataset.

4.1 The Full-Router Dataset

The dataset we use is the so-called 'full-router' experiment that recorded *all* packets entering and exiting a router in the Sprint IP backbone [12, 7, 13]. This dataset, along with a model of the queueing process inside the router, enables us to compute the continuous time ground truth of the delay/loss metric being measured. We rely on a model-driven approach since today's routers cannot be directly queried to obtain the ground truth in continuous time.

The full-router experiment involved a gateway router in the Sprint backbone network. The input and output traffic from all 6 interfaces of the router were monitored using DAG passive packet capture cards. Two of the interfaces were OC-48 links connecting the router to two other backbone routers. The other 4 interfaces were links to customers - two in Asia (OC-3 links) and two in the United States (one OC-3 and one OC-12 link).

The passive packet capture cards were synchronized with the same GPS signal and generated 64-byte records for each packet. Excluding the layer-2 headers, this provided us with the 20-byte IP headers and the first 24 bytes of the IP payload. In this paper, we use two datasets collected with the full-router experiment. The first of these was collected in August 2003 (and used in [7, 9, 13]) and the second in January 2004. Both datasets captured packets from/to the router for 24 hours.

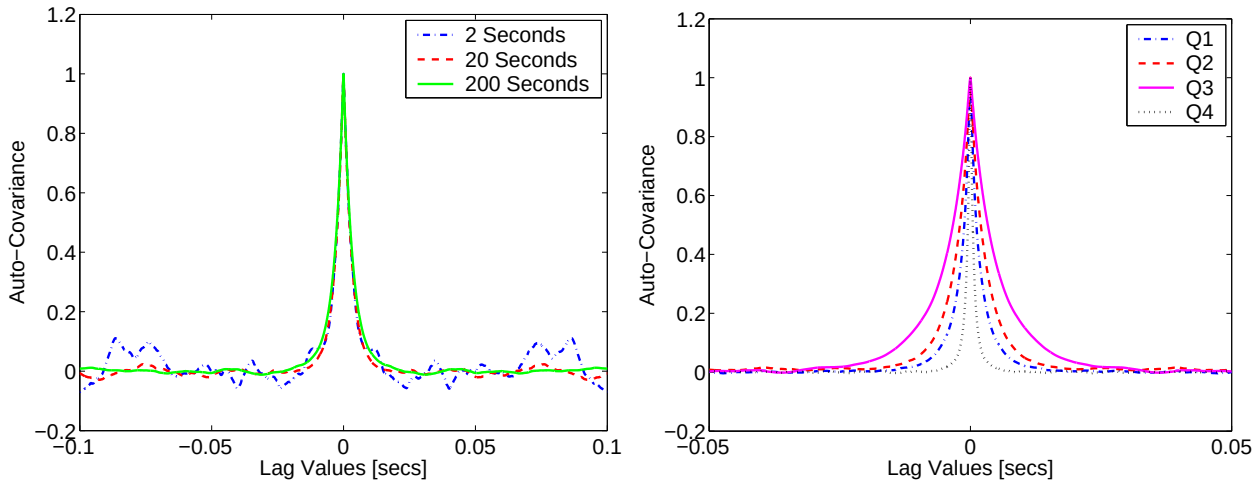

Figure 3: (Left) Plot illustrating the convexity of the auto-covariance of the mean virtual delay for P1. Variance with the estimation of the auto-covariance is significant if we use only a few seconds of P1 to compute the auto-covariance. (Right) Plot illustrating the convexity condition for Q1-Q4. The auto-covariance is computed using 200 seconds of these representative time intervals.

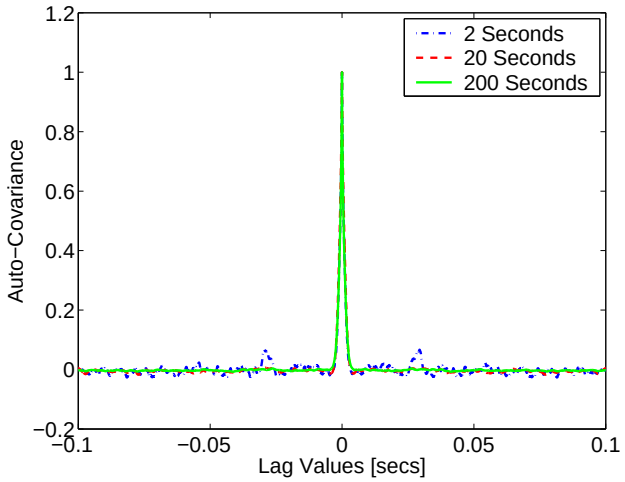

Figure 4: Plot illustrating the convexity condition for P2 using the queueing model involving minimum transit times across the router backplane.

We post-processed the dataset to match packets from the input interfaces to the output interfaces. For details on the matching procedure, we refer the reader to [7]. The matching procedure provided us with the following data on each output interface: all packets that exited that interface, the input timestamp (at the corresponding input interface), the output timestamp and the packet size. Since the timestamps were synchronized, we could also calculate the delay experienced by all packets across the monitored router.

The output links of the chosen router had different levels of utilization. To investigate convexity and the variance properties of various probing processes, we start by choosing a target output interface for each of our datasets. Since a low-utilization interface is not of much interest, we choose two of the most heavily-utilized (OC-3) interfaces of the router. The output interface we choose for the first dataset is the same as the one used in [7]. For each interface, we illustrate our results using a few representative 5-minute intervals. We show the utilization of the chosen interfaces and the representative time intervals in Figure 2.

4.2 Two Router Queueing Models

We start by using our datasets to determine if the auto-covariance of the mean delay across a link is convex. Since our focus is on non-intrusive probing, we focus on the mean virtual delay. To check for the convexity condition, we need to use a model of the queueing process in the router. We start with a simple FIFO model, i.e., the input timestamp of each packet destined to the chosen output interface is considered to be the arrival time of that packet to a FIFO queue. The queue drains at the same rate as the OC-3 link. We plot the (normalized) auto-covariance of the virtual delay function obtained with both our datasets in Figure 3. In Figure 3 (**Left**), we plot the auto-covariance computed using 2, 20 and 200 seconds of P1. There are no visual signs of non-convexity once the variance of the auto-covariance estimates become negligible. We also see similar confirmation of the convexity condition for the second dataset in Figure 3 (**Right**).

Previous work [7] has shown that the simple FIFO model cannot accurately predict per-packet delays. This is primarily because the model ignores the time spent by packets in traversing the backplane. It was shown in [7] that a more detailed model, which uses backplane transit times dependent on packet sizes, can predict per-packet delays more accurately. In Figure 4, we plot the auto-covariance of the virtual delay obtained by applying this detailed model on P2. We find that, even under this model, convexity is not violated. The results presented in the rest of this section are true when either model is used with our datasets. To avoid repetitiveness, we only present the results obtained by applying the simple FIFO model on P2 and the detailed model on P1.

4.3 Comparing Probing Streams

Since our empirical evaluation appears to support the convexity condition, we next investigate if the conclusions of Theorem 2 hold good. To do so, we choose 4 kinds of probing streams - Poisson, deterministic (periodic probing) and two different Erlang probing streams (recall that an Erlang law is a Gamma law with an integral shape value). We choose two different shape parameters - 2 and 4 - for the Erlang probing streams. For each representative time interval, we conduct 50000 experiments with each probing stream. In every experiment, the probing stream starts at a random time from the beginning of the representative time interval. The probing

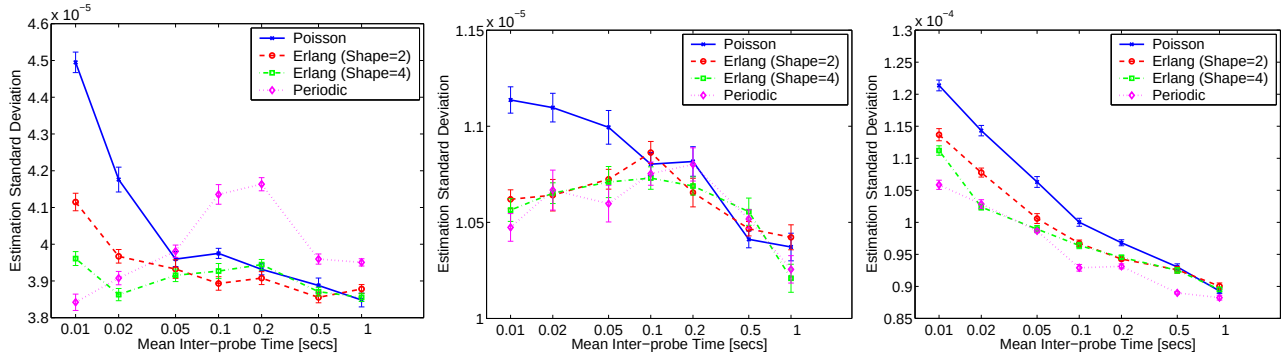

Figure 6: Comparison of the standard deviation of the mean virtual delay estimates achieved with P2 (20 probes) (Left), P4 (20 probes) (Middle) and Q3 (20 probes) (Right).

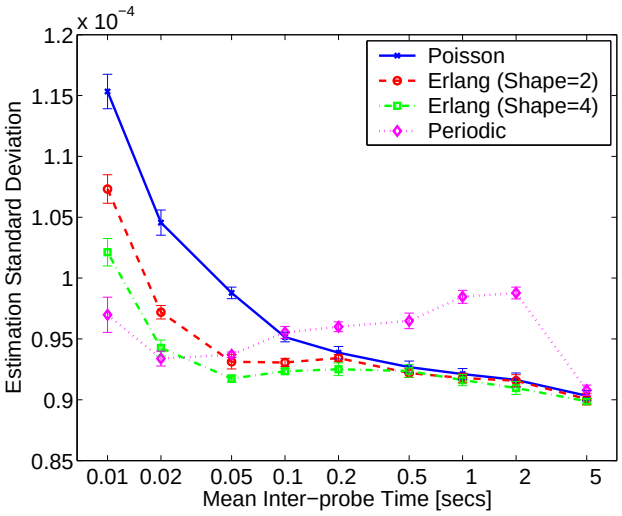

Figure 5: Plot illustrating the higher standard deviation of Poisson probing as compared to Gamma probing with P1 (10 probes). We use a log-scale on the x-axis.

stream consists of virtual probes, whose delays are computed using one of our two queueing models. These virtual delays are then used to estimate the mean virtual delay.

In Figure 5, we plot the standard deviation of the mean virtual delay estimates obtained with P1 (10 probes) for various mean inter-probe times. We make three observations. First, Poisson probing clearly demonstrates higher variance than both types of Erlang probing especially at the low inter-probe times. For inter-probe gaps closer to a second, we find that the difference in variance becomes minimal. Second, at small inter-probe times, the periodic probing stream has the lowest variance. This agrees with Theorem 2. Third, as the inter-probe times increase, the variance of periodic probing actually increases before decreasing at the inter-probe time of 5s.

In Figure 6, we show plots similar to Figure 5 for P2, P4 and Q3 (with 20 probes). The first two observations hold true in all of them - Poisson probing continues to be worse than Erlang probing especially at smaller mean inter-probe times. However, the variance of periodic probing does not always increase. In P4 and Q3, it has the lowest variance, which is consistent with Theorem 2.

4.4 Number of Probes

When we used up to 20 probes, we found that Poisson is distinctly inferior to Gamma probing (Erlang probing being a special case) and periodic probing especially at small mean inter-probe times. We now investigate if this behavior depends on the number of probes used. To do so, we compare the standard deviation of the various strategies by varying the number of probes. We plot the results for P1 and Q3 (for inter-probe time of 10ms) in Figure 7. We find that the number of probes does not affect the relative performance of the probing strategies. Moreover, the results remain consistent with our theorems. Periodic is optimal and Erlang converges to the optimal value as the shape parameter increases.

Slower links are likely to have non-zero auto-covariance at much larger lag values than OC-3 links. Hence, if the convexity condition is true, we expect that Poisson probing will be inferior to Gamma probing on slower links for a larger range of inter-probing times. Hence, in practice, the sub-optimality of Poisson probing is likely to be more significant.

4.5 Periodic Probing

In Figure 5, we saw that the variance of periodic probing anomalously increased before converging to the variance of all other probing strategies at an average inter-probe time of 5s. To investigate if this behavior is consistent, we plot similar results for P2 and Q2 in Figure 8 (**Left** and **Middle**). We find that the variance of periodic probing is not predictable - it shows an anomalous increase for P2 but not for Q2.

There are a few possible reasons behind the anomalous behavior of periodic probing. Though the auto-covariance plots for the virtual delay in either dataset show no visible signs of non-convexity, there might be non-convex effects that are significant only at certain lag values. The anomalous behavior could also be a manifestation of residual phase-locking effects that make the system lack ergodicity when periodic probing is used. Indeed, one of the goals behind Theorem 2 and our use of Gamma probing over periodic probing was precisely to avoid such potential sample-path bias while achieving near-optimal variance (when $R(\tau)$ is convex).

In Figure 8 (**Right**) we compare Erlang probing streams with shape parameters between 2 and 2048. We find that the shape parameters 4 and 16 possess good variance reduction properties. We also see that, with higher shape parameters, the variance of Erlang probing streams does converge to that seen by periodic probing. Thus, empirically, small shape parameters of Erlang between 4 and 16 appear to provide a good trade-off - lack of sample-path bias due to the mixing property and a near-optimal variance.

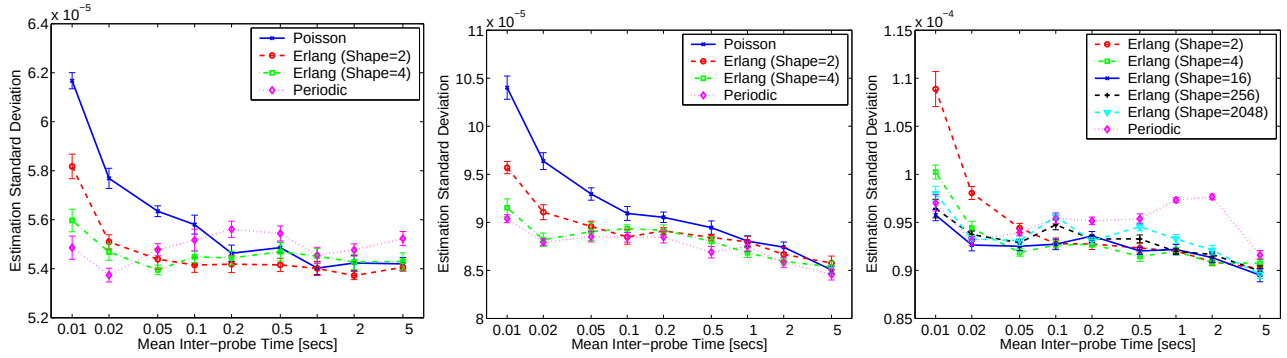

Figure 8: Comparison of the various probing strategies in the case of P2 (Left) and Q2 (Middle) with 10 probes. Comparing different Erlang probing strategies for P1 (Right) with 10 probes.

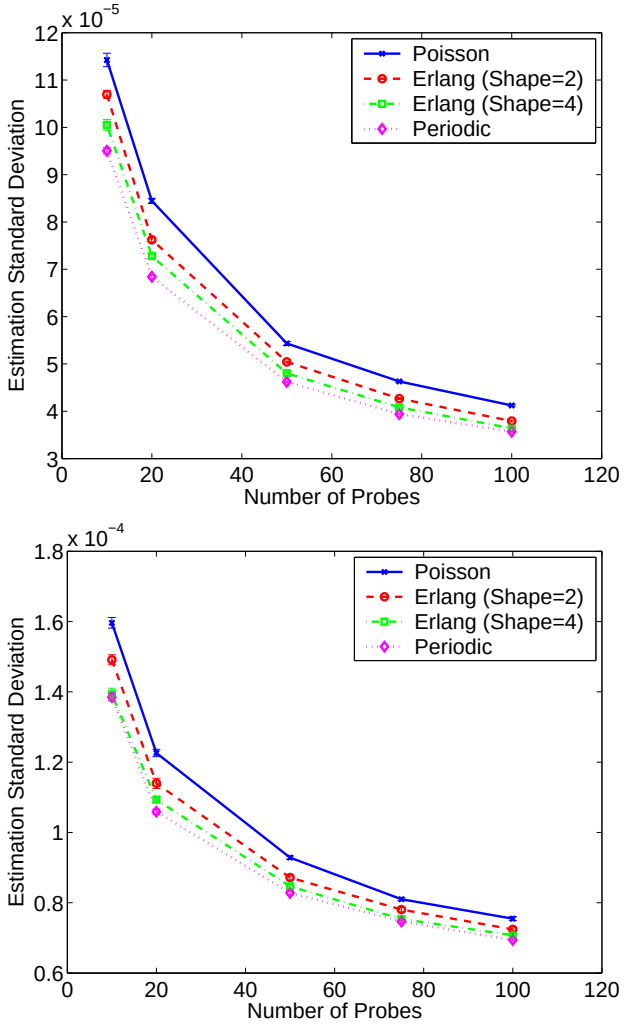

Figure 7: Comparison of the various probing strategies for various number of probes in the case of P1 (inter-probe time of 10ms)(Top) and Q3 (inter-probe time of 10ms) (Bottom).

4.6 Quantile Estimation

As discussed in Section 2, loss can be defined using quantiles of queue size (or virtual delay) at a hop. In particular, the loss process $I_x(t)$ of a probe of size x can be given by:

$$I_x(t) = P(B(t) > K - x). \tag{12}$$

where K is the maximum buffer size. Even when no loss occurs, high delays are a measure of congestion since they trigger TCP timeouts. This has been used in many prior works on loss estimation, for example, [18]. Thus, estimating the indicator functions of delay quantiles is of natural interest to us. Though the router monitored in the full-router experiment had no losses, quantiles defined on its output links can be thought to represent a congestion process. We investigate this now.

As with mean virtual delay, we start by calculating the auto-covariance of the indicator function of virtual delay quantiles in P1 and P2. In Figure 9 (**Left** and **Middle**), we show the auto-covariances of the 0.95-quantile and 0.5-quantile of these datasets respectively. As with mean virtual delay, we find no visual signs of non-convexity. In Figure 9 (**Right**), we plot the standard deviation of various probing strategies for the 0.95-quantile of virtual delay in P1. The results are similar to what we observed before; Poisson is distinctly inferior to Gamma probing especially at higher probing frequencies and periodic probing can be anomalous.

5. ADDITIONAL DATASETS

In the previous section, we computed the virtual delays using the full-router datasets. Our analysis showed that the mean as well as quantiles of the virtual delay process have a convex auto-covariance and that Poisson probing has higher variance than Gamma (or Erlang) probing especially at smaller timescales. We also saw the advantages of Gamma probing over periodic; the latter can have higher variance due to residual phase-locking effects. In this section, we present results performing similar analysis for other datasets. First, we present results analyzing the virtual delay across a 10Gbps OC-192 link we monitored. Then, we use simulation-based experiments to investigate our results for multi-hop paths.

5.1 OC-192 Experiment

In Figures 3 and 4, we showed that the convexity condition appears to be valid for the OC-3 link of a gateway router. We now provide some evidence that the convexity condition is true for higher-speed links in the Internet core. The dataset we use consists of packet traces collected by monitoring one input and one output interface of a backbone router. Both interfaces terminated 10Gbps OC-192 links. These traces were collected for 20 seconds using DAG packet capture cards that had a timestamp accuracy of 100ns. The monitored output interface had an average utilization of 30%.

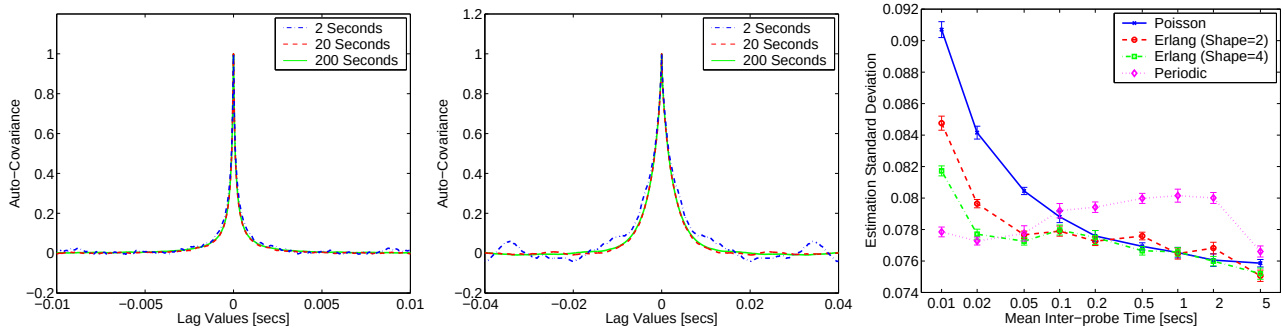

Figure 9: Auto-covariance of the estimated mean value of delay-based quantiles with P2 (quantile value of 0.5) (Left) and P1 (quantile value of 0.95) (Middle). Comparison of the standard deviation of various probing strategies with the latter (Right).

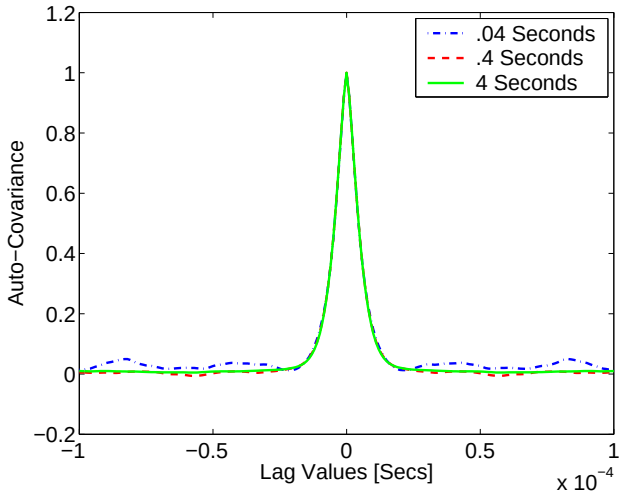

Figure 10: Auto-covariance of the virtual delay process calculated using the OC-192 dataset.

The traces collected at the input interface provide us with the arrival times of a majority (about $60 - 65\%$) of the traffic exiting the output interface. Since our traces do not include the arrival times of the remaining exiting traffic, we cannot apply the queueing models used with the full-router datasets. So we calculate the continuous-time virtual delay process across the output OC-192 interface using the following approach. For each captured packet, we first calculate its delay from the observed input and departure timestamps. Then, we subtract its (size-dependent) transmission time from the observed delay. Assuming a simple FIFO model for the router, the resulting values are essentially samples of the virtual delay process. We then use a linear intrapolation of these samples to calculate the continuous time virtual delay process at arbitrary times. In Figure 10, we plot the auto-covariance of the virtual delay process thus computed. We see that there are no visual signs of non-convexity.

5.2 Simulations

So far, we tested the convexity condition for delays across a single hop. Since our dataset had no loss, we also did not directly analyze the nature of the congestion process except using delay quantiles. In the absence of detailed loss data and datasets spanning multiple hops, we turn to simple *ns-2* [20] simulations to understand these areas better.

We simulate a 3-hop system with link capacities of $6, 20$ and 10Mbps. All hops use FIFO queues with droptail byte-based dropping policies. We use the ground truth calculator described in the appendix to access the loss process $I_x(t)$. We calculate the auto-covariance of the computed loss process $I_x(t)$ for $x = 1540$ bytes. We calculate the auto-covariance with high confidence using multiple simulations each of which is 100 seconds long. We follow this approach for various choices of cross-traffic over the three hops.

In Figure 11 (the **Right** plot is a zoom of the **Left** plot), we plot the auto-covariance (un-normalized to indicate the absolute value) of the loss process in three scenarios. In all three, the middle hop carries traffic generated according to a Pareto process. The last hop carries periodic cross-traffic in the first and Poisson in the rest. In all three scenarios, 3-hop persistent traffic generated according to a Poisson process flows across all three hops. In the first two scenarios, a similar 3-hop persistent TCP flow is present. In each of the three scenarios, the ground truth auto-covariance function is close to convex. This is despite that fact that the TCP flow, which traverses all three hops in two of the three scenarios, creates oscillations, and even despite the periodic traffic on hop three in the first scenario.

We expect feedback, especially from TCP, to generate some negative correlation over a range of lags, which breaks convexity. However, we see this only in a few extreme examples. We illustrate some of these examples in Figure 12. For reference, we keep the same third scenario as in Figure 11. In the first two scenarios, we remove the 3-hop persistent Poisson traffic. We continue to use the 3-hop persistent TCP flow. We find that, some form of non-convexity creeps in due to the significant feedback introduced by this single persistent TCP flow. As the third scenario indicates, this vanishes when the persistent TCP flow shares all of its links with other traffic.

Scenarios such as those in Figure 12, where convexity does not hold, would happen, for example, if covariance functions oscillate. However, a necessary condition for such oscillation is that typical sample paths themselves oscillate in reproducible ways. We expect this to happen only when a very few number of flows with feedback have a significant impact on a path. As our results show, this does not happen in the Internet core. We also note that even if $R(\tau)$ is not convex everywhere, it may be for some domain of interest. For example there may be probing rates that we may be constrained to use, resulting in the non-convex domain of $R(\tau)$ never being sampled in any case. In other words, practical sampling may be above time scales (or below, when stationarity fails at very large scales) where a lack of convexity is a problem. In all such cases, our theorems are still very useful.

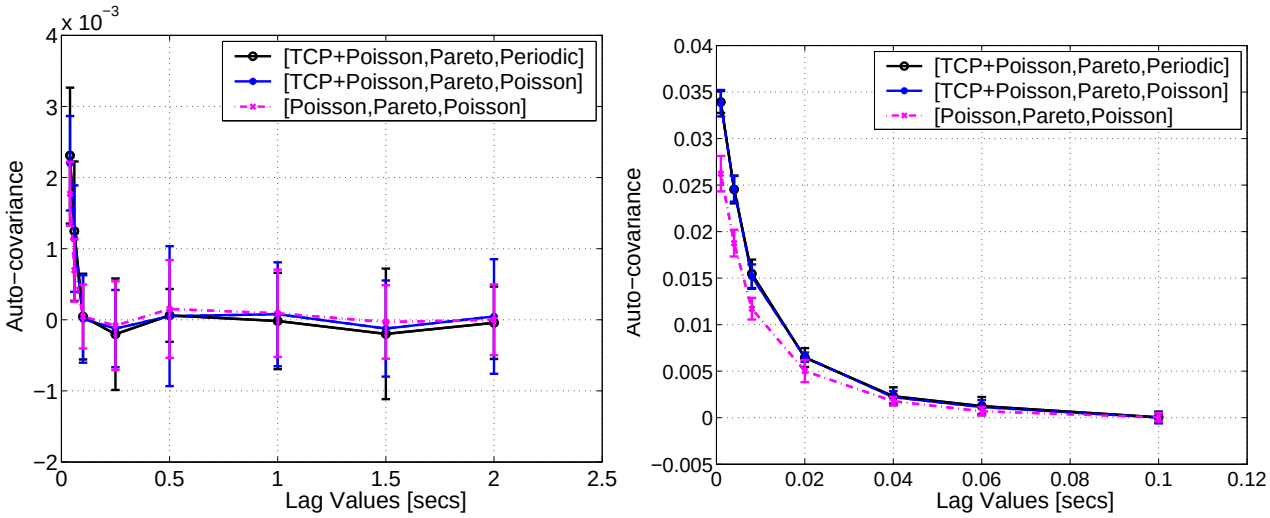

Figure 11: Auto-covariance (un-normalized) function of I_x for a 3-hop system involving a combination of TCP, Poisson and Pareto cross-traffic. The left hand plot is a zoom of the right hand one, showing more points. Three results are shown depending on how many cross-traffic streams have periodic features.

6. PRACTICAL IMPLICATIONS

In [2], we introduced the probe separation rule as a guideline to design probing streams. Although Gamma distributions for all finite β do not have a lower bound above zero, as suggested by this rule, the work in this paper is consistent with its spirit. This is because the probability that inter-probe times are smaller than ϵ goes to 0 when β goes to infinity for all ϵ strictly less then the mean inter-probe time. Apart from being generic (applicable to probe patterns) and mixing (therefore, free of sample-path bias), all Gamma probing streams with $\beta > 1$ are superior to Poisson probing streams in terms of variance when $R(\tau)$ is convex. This endows the Gamma family with an important property that is also suggested by the probe separation rule: tunability to enable different trade-offs. We exploited this property in Section 4.5 to suggest the use of Gamma probing streams with small shape values, which would likely avoid the anomalous behavior exhibited by Gamma probing streams with high shape values and yet be superior to Poisson probing.

Our results in Section 4 and Section 5 show the advantages of Gamma probing for a wide range of metrics including mean delay, delay quantiles and loss indicator functions. We did find extreme cases in which a single dominating TCP flow can introduce non-convexity for loss probing. In future work, we intend to understand better the reasons behind such non-convexity and investigate if it ever occurs in practice.

Throughout this paper we assumed non-intrusive probing and ignored the perturbative impact of probing. As we discussed in [2], non-intrusive results can be applied to the estimation of delay-based metrics in practice as long as probing is rare. In [3], we showed that similar results are valid for the estimation of loss-based metrics, too. The general idea is that, if a system 'forgets it past' fast enough, then probes sent rarely enough emulate well non-intrusive probing. Rare probing is one way in which our results in this paper can be applied in practice. In particular, we can benefit from the low variance of Gamma probing if its perturbative effect is arguably minimal. However, the problem of removing the impact of intrusiveness *without* relying on rare probing is technically challenging and remains open.

7. RELATED WORK

Delay and loss measurement has been the focus of many earlier papers [14, 15, 18, 22]. Early measurement studies [15] focused on end-to-end measurement, primarily using Poisson sampling. Zhang et al. [22] used Poisson sampling to study the time-varying nature of delay, loss and other path characteristics. In all these prior studies, Poisson sampling was justified as an application of the PASTA principle [21]. The IETF IP Performance Metrics (IPPM) Group [8] also recommends Poisson sampling for loss measurement.

Recently, the utility of Poisson sampling for delay was questioned [2, 17]. Empirical studies [18, 10] have also questioned this premise. The importance of variance, in addition to bias, for good measurements was stated in [2]. A few preliminary results showing the non-optimality of Poisson probing was also provided in [2]. Techniques have been proposed for better probing. For instance, Badabing [18] is a tool for loss measurement that proposed a particular probing process assuming that the loss process is Markovian. In contrast, our work is quite general and provides insights into the measurement of any metric.

8. CONCLUSIONS

We have taken a fundamental look at the issue of optimal probing for delay and loss measurement. Given that a large variety of so-called mixing probe processes enjoy the strong consistency property, we investigated the mixing probing process that would minimize estimation variance. In the context of non-intrusive probing, we proved that periodic probing has the least variance if the auto-covariance of the measured metric is convex. But periodic probing is not mixing as it can have phase-locking issues resulting in significant sample path bias. We showed that an alternative family of probing processes, Gamma renewal processes, provide a good middleground - they have no sample-path bias and can achieve variance as close to periodic as possible when the auto-covariance of the measured metric is convex. Using extensive experiments, primarily on a unique set of full-router datasets, we demonstrated the validity of this convexity condition and the sub-optimality of Poisson probing.

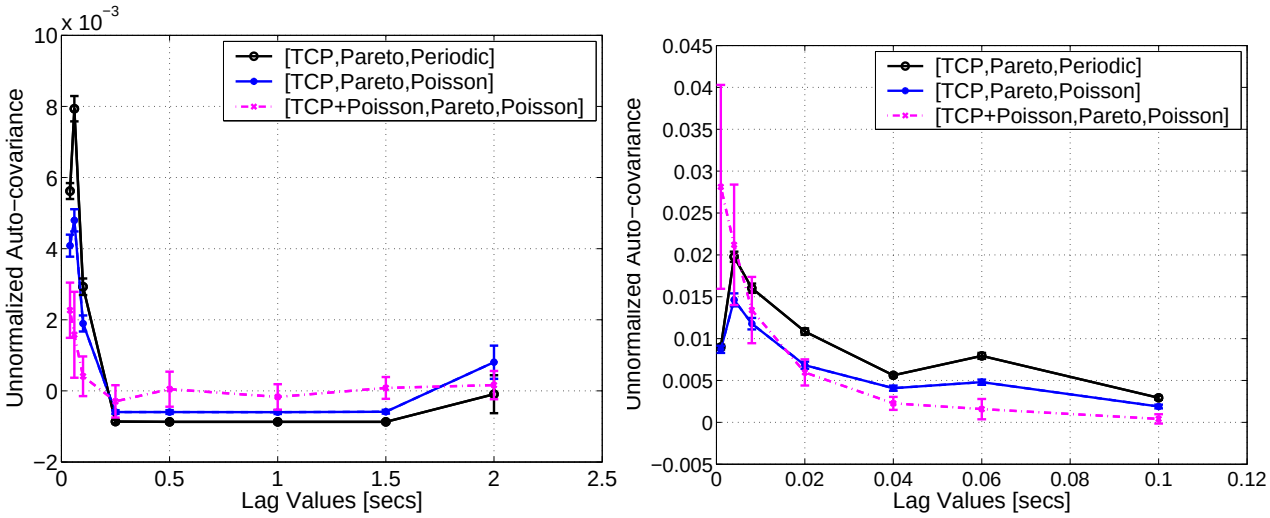

Figure 12: Auto-covariance (un-normalized) function of I_x for 3-hop systems involving a single high-rate TCP flow that can introduce some non-convexity. The left hand plot is a zoom of the right hand one, showing more points.

9. ACKNOWLEDGEMENTS

The full-router data was available following prior work at Sprint ATL. We thank Tao Ye, Gianluca Iannaconne and Nicolas Hohn for their packet matching and analysis code, and Dina Papagiannaki, Ed Kress, Richard Gass, and Jamie Schneider for data collection and access. We also thank our sheperd, Paul Barford, for his comments on earlier drafts of the paper.

10. REFERENCES

[1] F. Baccelli and P. Bremaud. *Elements of Queueing Theory*. Springer Verlag, Applications of Mathematics, second edition, 2003.

[2] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot. The Role of PASTA in Network Measurement. *Computer Communication Review, Proceedings of ACM Sigcomm 2006*, 36(4):231–242, 11–15 Sep 2006.

[3] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot. Loss Measurement via Probing. *Sprint ATL Technical Report No. ATL-020124*, February 2007.

[4] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer-Verlag, 1988.

[5] D. Gaver and P. Lewis. First-order autoregressive gamma sequences and point processes. *Adv. Appl. Prob.*, 12:727–745, 1980.

[6] L. Gradshteyn and L. Ryzhik. *Table of Integrals, Series and Products*. Academic Press, sixth edition, 2000.

[7] N. Hohn, D. Veitch, K. Papagiannaki and C. Diot. Bridging router performance and queuing theory. In *Proc. of ACM SIGMETRICS'04*.

[8] I. I. P. M. (IPPM). `http://www.ietf.org/html.charters/ippm-charter.html`, 2005.

[9] S. Machiraju, D. Veitch, F. Baccelli, and J. Bolot. Adding Definition to Active Probing. In *ACM Computer Communication Review*, April 2007.

[10] M.Tariq, A.Dhamdhere, C.Dovrolis, and M.Ammar. Poisson versus Periodic Path Probing (or, Does PASTA Matter)? In *ACM Internet Measurement Conf.*, pages 119–124, Berkeley, CA, Oct 2005.

[11] T. Ott. The covariance function of the virtual waiting time process in an M/G/1 queue. *Adv. App. Prob.*, 9, 1997.

[12] D. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, F. Tobagi and C. Diot. Analysis of Measured Single-hop Delay from an Operational Backbone Network In *Proceedings of IEEE Infocom*, June 2002.

[13] D. Papagiannaki, D. Veitch and N. Hohn. Origins of Microcongestion in an Access Router. In *Proceedings of Passive and Active Measurement Conference.*, 2004.

[14] V. Paxson. End-to-end routing behavior in the internet. In *Proceedings of ACM SIGCOMM*, 1997.

[15] V. Paxson. End-to-end Internet packet dynamics. *IEEE/ACM Transactions on Networking*, 7(3):277–292, 1999.

[16] K. Petersen. *Ergodic Theory*. Cambridge University Press, Cambridge England, 1983.

[17] M. Roughan. A Comparison of Poisson and Uniform Sampling for Active Measurements. *IEEE Journal on Selected Areas in Communication*, 24(12):2299–2312, Dec 2006.

[18] J. Sommers, P. Barford, N. Duffield, and A. Ron. Improving accuracy in end-to-end packet loss measurement. In *ACM SIGCOMM'05*, pages 157–168, 2005.

[19] L. Takács. *Introduction to the Theory of Queues*. Oxford University Press, New York, 1962.

[20] The Network Simulator - ns-2, 2004. `http://www.isi.edu/nsnam/ns/`.

[21] R. Wolff. Poisson Arrivals see Time Averages. *Operations Research*, 30(2):223–231, 1982.

[22] Y. Zhang, N. Duffield, and V. Paxson. On the Constancy of Internet Path Properties. In *Proceedings of the Internet Measurement Workshop*, 2001.

11. APPENDIX

11.1 Calculating Auto-Covariance of $I_x()$

In Section 5.2, we calculate the auto-covariance of $I_x(t)$, the ground truth loss process. We now explain how we calculate $I_x(t)$. Consider a single hop in a network employing a byte-based drop-

ping policy. By definition, $I_x(t)$ for that hop depends only on the queue size at that hop (and x). Using a complete trace of packet sizes and arrival times to the hop, we can compute the queue size of that hop at any time t and, therefore, compute $I_x(t)$, too. Though $I_x(t)$ is a continuous time process, it can be efficiently stored. To see why, note that there are a finite set of 'changepoints' where the value of $I_x(t)$ changes. These are precisely those instants when the queue size crosses a threshold (above which there are fewer than x bytes available).

We use a similar approach to calculate $I_x(t)$ for multi-hop paths. However, recall that $I_x(t)$ is an indicator function for the event that a probe of size x, had it entered the system at time t, encounters a queue that has less than x bytes available. Hence, in a multi-hop path, $I_x(t)$ depends on the buffer size of intermediate hops at time instants that may be after t. Therefore, to calculate $I_x(t)$, we also compute these time instants. As with single-hop paths, $I_x(t)$ can be efficiently stored; it has a finite set of 'changepoints', which correspond to the time instants when any queue along the path crosses a threshold.

In *ns-2*, queue occupancy does not include the current packet being transmitted. Hence, to obtain the ground truth loss process for our simulations, we calculate 'changepoints' of this alternative definition of queue size.

11.2 Proof of Lemma 3.1

Let $T \sim \Gamma_{\alpha,\lambda}$, $Z \sim \Gamma_{\beta,\lambda}$ be independent, and set $Y = T + Z$. Then $C = \mathbb{E}[T|Y] = \alpha Y/(\alpha+\beta)$ has density $\Gamma_{\alpha+\beta,(\alpha+\beta)\lambda/\alpha}$, with mean $\mathbb{E}[C] = a/\lambda = \mathbb{E}[T]$.

PROOF. From the scaling property of Gamma, $Y \sim \Gamma_{\alpha+\beta,\lambda}$. Since T and Z are independent, the density of $(T|Y=y)$ is

$$\begin{aligned}\Pr(T=x|Y=y) &= \frac{\Pr(T=x,Y=y)}{\Pr(Y=y)} = \frac{\Pr(T=x,Z=y-x)}{\Pr(Y=y)} \\ &= \frac{\Gamma_{\alpha,\lambda}(x)\Gamma_{\beta,\lambda}(y-x)}{\Gamma_{\alpha+\beta,\lambda}(y)} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(y-x)^{\beta-1} y^{1-(\alpha+\beta)}.\end{aligned}$$

Recall the *Beta function* $B(x,y) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$. The required conditional expectation is given by

$$\begin{aligned}\mathbb{E}[T|Y=y] &= \frac{y^{1-(\alpha+\beta)}}{B(\alpha,\beta)} \int_0^y x^{\alpha}(y-x)^{\beta-1}\,dx \qquad (13) \\ &= \frac{y^{1-(\alpha+\beta)}}{B(\alpha,\beta)} y^{\alpha+\beta} B(\alpha+1,\beta) \\ &= \frac{\alpha\, y}{\alpha+\beta} \qquad (14)\end{aligned}$$

using the integral identity 3.191(1) from [6]. Now viewing y as a sample of Y, we have $C = \mathbb{E}[T|Y] = \alpha Y/(\alpha+\beta)$, which is Gamma as stated by the scaling property. □