A Comparison of Probe-based and Router-based Methods for Measuring Packet Loss

Paul Barford and Joel Sommers

Abstract— **Probe-based (active) measurements of packet loss have formed the basis for much of our empirical understanding of loss behavior in wide area networks. Router-based (passive) loss measurements via SNMP, while not widely available, offer the potential for a more detailed perspective. In this paper we present a case study that quantitatively assesses and compares the viewpoints provided by both of these methods for measuring packet loss. Our hypothesis was that probe-based and router-based loss measurements should correlate. We investigated this by first comparing SNMP loss measurements on router interfaces to those extracted from packet traces in a series of laboratory experiments. We found these two passive measures of loss to be highly correlated. Next, we evaluated packet loss data gathered over three weeks in a widely deployed infrastructure, using both backbone router interfaces for passive measurements and co-located hosts that sent active probes in a full mesh. We found little correlation between time series of passive, router-based measures of loss and active loss probes for all of the paths in our measurement infrastructure. We also compared the distributional characteristics of these loss measurements including lengths of loss free periods, loss rates during lossy periods, and measures of loss constancy. We found the level of agreement between passive measures and active measures for each of these characteristics to be quite low. Deeper evaluation of our data indicated that current methods for active probing for packet loss suffer from high variance inherent in standard sampling techniques and from effects of end-host interface loss, which we also characterize in this work.**

I. Introduction

Packet loss due to congestion is a fundamental problem in wide area packet switched networks. Great effort has been spent to characterize and model this phenomenon and to design protocols and networks that effectively avoid, control and recover from packet loss. While progress has certainly been made, packet loss and its effects on performance remain a significant issue for network researchers, system designers, and operators.

The evolution of network protocols and network systems has been influenced by basic, empirical studies of packet loss behavior. A number of studies of have had significant effects in this regard. Two of the best examples of protocols that have benefited directly from empirical observations of packet loss behavior are the NewReno [11] and SACK [13] versions of TCP. However, the Internet is a constantly changing environment, and this requires continued evaluation of important phenomena, such as packet loss.

There are two basic methods for measuring packet loss, each with its own set of challenges. The first uses passive monitors that can either be attached to network links or are available from network nodes. A standard means of passive monitoring is the set of Management Information Base (MIB) counters available on network nodes via the Simple Network Management Protocol (SNMP) [5]. These counters track a wide range of activity including packet losses due to congestion. The benefit of passive monitoring systems is that they capture many of the important details of local traffic behavior, such as identifying which subsystem within a router has become overloaded. However, the cost for this detail is often high (*e.g.*, in terms of data storage requirements) and access to links or routers is frequently not possible across administrative domains.

The second means for measuring packet loss is through active end-to-end probing. The simplest active probe used to measure packet loss is the `ping` utility. Like most active probe tools, `ping` sends a series of packets aimed at a target system within the network and measures the response packets returned to the sending system [1]. Lost packets are tracked by the sender through the use of sequence numbers. The benefits of active probes are that they can be run from virtually anywhere in the network, and that they give an end-to-end perspective of network behavior. The difficulty is that the discrete nature of active probing limits the resolution of the measurements (a standard problem in any measurement methodology based on sampling). If more frequent probes are sent into the network then resolution will increase, but if the frequency is too high then the probes themselves can skew the results (analogous to a voltmeter *loading effect*). Despite these difficulties, active probing remains one of the most important methods for gathering packet loss data.

P. Barford and J. Sommers are members of the Computer Science Department at the University of Wisconsin, Madison. E-mail: pb,jsommers@cs.wisc.edu.

[1] Some active probe tools are *one-way* with a separate sender and receiver.

In this paper we address the question, "do probe-based and router-based measurements of packet loss correlate?" To that end, this paper makes the following contributions:

- A quantitative assessment of the accuracy of SNMP measurements of loss for a widely deployed backbone router in a controlled laboratory environment.
- A case study of passive loss measurements taken via SNMP in the Abilene/Internet2 backbone. To our knowledge this is the first broad characterization of packet loss via passive measurements in a live environment.
- A comparison of active loss probes with passive loss measurements via SNMP taken in a live environment.

We examined the accuracy of SNMP loss measurements in order to address common perceptions of the inaccuracy of SNMP loss measurements. Using a series of controlled laboratory experiments we found SNMP to be very accurate in reporting loss. We then took router-based measurements of packet loss via SNMP over three collection periods at all backbone routers in the Abilene/Internet2 infrastructure. We aggregated loss data from all interfaces along each path in the full mesh of paths to obtain end-to-end perspectives on loss behavior. We treated these measurements as the baseline from which we compared a set of active probes for loss taken simultaneously in the same infrastructure.

The active probing tool we used to measure loss is the `zing` utility [12] which sends probe packets at exponentially modulated intervals. This probing method should provide unbiased, time-averaged data for loss conditions along an end-to-end path [2]. We took one-way measurements of packet loss by running `zing` between nodes in the GPS-enabled Surveyor infrastructure [17] that are directly connected to the Abilene backbone routers. This enabled us to probe paths in a full mesh in this backbone without the risk of being unable to account for packet loss at intermediate routers. For our three measurement periods, we set the average probe rate to 10Hz, 20Hz, and 100Hz respectively, and then aggregated the measured loss rates to compare with the SNMP data.

Instead of attempting to develop a single metric for comparison, we evaluated the degree of agreement between the active measurements and the SNMP measurements in a number of ways. First, we compared the correlation coefficients for the time series of loss rates for each end-to-end path. Our results show that there is little correlation between loss rates measured by active probes and loss rates measured by SNMP. Next, we compared distributional characteristics of loss measurements for different loss properties including lengths of loss-free periods, loss rates during lossy periods, and measures of *loss constancy* as described in [22]. In each case we found a low level of agreement between the distributions. This led to our overall conclusion that *probe-based and router-based measures of loss can provide quite different perspectives.*

There are a number of possible explanations for the lack of agreement between the two types of measurements although deeper investigation in this regard is left for future work. One possibility is that the active probing process could affect results. We experimented with the probe process by comparing Poisson modulated `zing` probes with simple `ping` probes sent at the same rates as `zing`. We found negligible difference between the two types of probe processes. We attribute this lack of difference between probe types to the low overall loss rates we observe in our data. Another possible reason for poor correlation is that there may be artifacts in our measurements that bias the results. One such artifact is interface loss on the active probe systems. We see examples of interface loss in our data when there is a loss measured by the active probe but no associated loss measured by SNMP. We attribute these losses to the end-host network interface. We found occurrence of these losses to be rare, and after censoring them from the data we still find very low correlation between measurement methods. The most plausible explanation for lack of correlation is that the sampling rates we employ in our active measurements are too coarse to enable typical loss episodes in this infrastructure to be measured accurately. While three different probe rates were used in our study, we found that correlation did not significantly improve with faster probe rates. This is most likely due to the fact that the overall loss rates as measured by SNMP were so low that it would have required sampling via active probes for very long periods of time before correlations could have been established (see Section III-B.4 for a more detailed discussion of this issue).

Our work has a number of implications. First, new active probe methods for loss may be necessary to get a more accurate picture of loss behavior in low loss environments over short periods of time. Next, infrastructures that use active probes for loss monitoring may need to consider other means for collecting this data. Finally, our study suggests that characterizations and models for packet loss based on active measurements may need to be reevaluated using data from new probing methodologies or from passive measurements.

The rest of this paper is organized as follows. In Section II we discuss work related to this study. Section III presents the details of the data that we collected and evaluated in this work. In Section IV we present passive loss characteristics and compare the active probe loss measurements with the SNMP loss measurements to assess the degree of agreement between the two. We summarize our

study and discuss future work in Section V.

II. Related Work

To our knowledge there has been no prior work that attempts to compare probe-based and router-based measurements of packet loss. A study by Pasztor and Veitch identifies limitations in active measurements, and proposed an infrastructure using the Global Positioning System (GPS) (quite similar to Surveyor [17]) as a means for improving accuracy of active probes [15]. Their work was validated by comparing passive measurements of packet delays *at end-hosts* to delay measured by active probes, but did not address the precision of loss measurements from the perspective of nodes in the network.

There have been many studies of packet loss behavior in the Internet, but to our knowledge, none of these are based on passive measurements. There are a number of passive measurement projects in the wide area including one conducted by Sprint [10] and another by AT&T [8]. Work by Bolot [4] and Paxson [16] used active probe measurements to establish much of the baseline for understanding packet loss characteristics in the wide area. These characteristics include correlation structures on fine time scales and typical loss rates. Yajnik *et al.* evaluated correlation structure on longer timescales and developed Markov models for temporal dependence structures [21]. A study by Zhang *et al.* assesses three different aspects of *constancy* in loss rates in an infrastructure which has many similarities to our own (many of the links traversed by their active probes were in Internet2/Abilene) [22]. That work described an important notion of a loss process called a "change free period", which is a period of time during which a loss rate appears well-modeled as steady. We evaluated change free periods in our data using the methods outlined in that paper.

We used `zing` to measure packet loss in one direction in this study - see RFC 2680 for a description of the methods use by this tool [2]. `zing` relies on coordinated end-hosts which are not always available when taking active measurements. Savage developed the *Sting* [18] tool as a means for solving this problem. Sting uses a clever scheme for manipulating a TCP stream to measure packet loss *in both the forward and reverse direction* from a single end host. Another approach to active measurement of packet loss is to use tomography to infer link-level loss rates [9], [7]. In [9], Duffield *et al.* also discussed the difficulty of acquiring passive link-level measurements and the use of series of packets in active loss probes.

There are a number of widely deployed measurement infrastructures which actively measure wide area network characteristics [17], [3], [14]. These infrastructures use a variety of active probe tools to measure loss, delay, connectivity and routing from an end-to-end perspective. Of these systems, only Surveyor can monitor individual nodes *within* the network.

III. Data

A. Measurement Infrastructure

Our measurement infrastructure is unique in that it consists of widely dispersed end-host measurement stations as is typical in Internet measurement projects, and also includes production routers in the Abilene backbone of Internet2 [1]. We send active probes across the full mesh of end hosts co-located with the routers and collect one way loss measurements at each host. We also periodically query backbone routers via SNMP to collect router interface counters.

Figure 1 depicts the topology of our infrastructure[2]. In all cases but one (New York), our measurement hosts are directly connected to a backbone router. In three cases (Sunnyvale, Cleveland, Washington D.C.), we do not have measurement hosts. In the case of the New York backbone router, there are two hops from the measurement host to the router. In total, we take end-to-end probe measurements from eight hosts (comprising 56 distinct paths) and collect SNMP interface data from eleven routers (roughly 30 interfaces.)

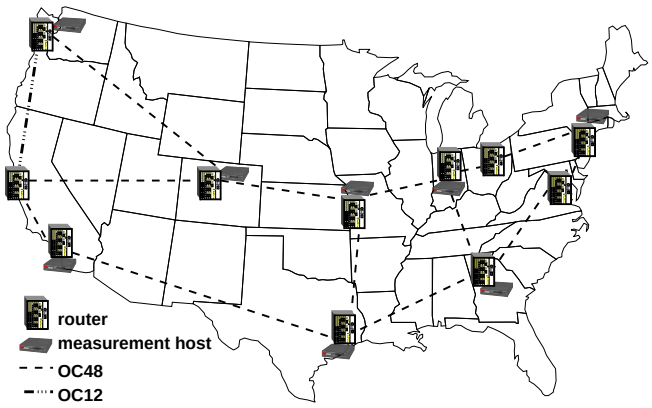


Fig. 1. Map of Abilene backbone nodes (sources of router-based data) and co-located Surveyor nodes used to take active probe measurements.

The end hosts each run BSD/OS version 3.1. Six of the eight hosts have ATM OC-3 (155Mbps) interfaces directly connected to the backbone routers using Fore (Marconi) 200e ATM network interface cards. The remaining two hosts (New York and Houston) are connected to the backbone routers via 100Mbps Ethernet. The reason for distin-

[2]The only Abilene backbone router absent from the picture is in Chicago, which we exclude from our study.

guishing these two types of connections is that the routers under study process incoming packets differently in each case. Data arriving on an ATM interface may take a "fast path" through the router, while data arriving on an Ethernet interface takes the "slow path" in all cases. This difference is discussed further below.

The routers are Cisco 12008 Gigabit Switch Routers (GSRs). The GSRs run a variant of IOS version 12[3]. All backbone links are OC-48 (2.4Gbps), except for the link between Seattle and Sunnyvale, which is OC-12 (622Mbps.) Link utilizations over the period of our study averaged 12%, 8%, and 7% for the 10Hz, 20Hz, and 100Hz measurement periods, respectively. Standard deviations for these periods were 11%, 5%, and 4%. Regarding the representativeness of data collected in this environment, we ascribe to the argument made by Zhang *et al.* in [22] which states that while Abilene's performance is "not likely to be representative of the commercial Internet, we might plausibly argue that our observations could apply fairly well to the better connected commercial Internet of the not-too-distant future, if not today."

B. Data Collection

The data we present and analyze in this paper was collected over the three periods of April 24, 2002 to May 8, 2002 (10Hz probes), July 24, 2002 to July 31, 2002 (20Hz probes), and August 8, 2002 to August 9, 2002 (100Hz probes). Due to the immense amount of data generated from the 100Hz measurements, this data was only collected for two days. In this section we describe the specifics of data collection for the active measurements and for the routers.

B.1 SNMP Router Interface Data

Our router interface data was collected through a process which queried backbone link interface MIBs every 30 seconds. Ingress and egress packet counts, interface drop counts and error counts were collected from counters in the MIB-II `ifTable` and `ifXTable`. In addition, a Cisco enterprise MIB that gives more complete information on interface drop counts was polled. For each measurement, we also noted the last interface change time stamp available in the MIB-II `ifTable` and the operational and administrative statuses to prevent collection of invalid data and to aid in detection of counter wrap-around.

The reason we must poll the Cisco-specific MIB is that the `ifInDiscards` entry in the MIB-II `ifTable` only counts one type of packet discard which can happen on input[4]. Inexplicably, output counts do not have this limitation. While we do not have detailed categorization of the causes of dropped packets, such as can be obtained from the IOS `show interface` command, we have complete identification of packets which are dropped at a given router interface. See Appendix A for more details on MIB loss counters in a GSR.

Polling the MIBs more often than 30 seconds yields diminishing returns. Besides increasing router CPU load, the MIBs are not updated in real time. Individual interfaces propagate local counters to the main processor module approximately every 10 seconds. We decided on 30 seconds as a compromise between increased load on routers and sufficiently detailed data.

Using the SNMP-based loss rates measured at each router, we calculated the loss rates for paths with multiple hops using a union of loss probabilities. Specifically, we calculated loss rate L for a multi-hop path p of length n interfaces for a given 30 second period as $L_p = 1 - \prod_{i=1}^{n}(1 - l_i/t_i)$ where l_i is the sum of packets lost during a 30 second period at interface i and t_i is the sum of packets transmitted and packets lost at the same interface during the same period. The combination of the physical interface layout in Abilene and periodic traceroutes enabled us to easily establish input and output interfaces for the full mesh of paths.

This calculation assumes independence of loss events at each hop in the path. Since there is no simple solution for proving independence, we argue that it is a reasonable assumption in a highly engineered network like Abilene. We appeal to the intuition that it is unlikely that a single flow or even a small group of flows will cause correlated congestion losses at two points on a path. We calculated correlation coefficients for both loss periods and loss rates on all multi-hop paths. We found all coefficients very tightly bunched around zero. While this result does not prove independence, it is consistent with our assumption of independence.

B.2 Evaluation of SNMP Data

In order to effectively compare probe-based and router-based measurements of packet loss, we experimentally evaluated the packet loss counters as implemented in Cisco IOS running on the GSR platform. It is important to question the precision of counters obtained from a rather opaque source, such as from a backbone-class Internet

[3]The specific versions of IOS on the GSRs are a mix 12.0S and 12.0ST. Build revisions are mostly the same for each subversion S and ST at 21 and 19, respectively. The primary difference between the two revisions is that the ST subversion contains support for MPLS.

[4]The `ifInDiscards` counter in MIB-II counts input drops due to lack of buffers, which is distinctly different than lack of input queue space. We have to consult a Cisco interface table in order to obtain input queue drops.

router. Common lore, supported by literature in the networking community (e.g., [19]), holds that MIB counters are of dubious quality depending on the vendor and on the particular unit.

Using the hardware configuration shown in Figure 2(a), we performed three experiments to test the accuracy of the same MIB variables used in our wide-area measurements. We generate traffic from a Spirent AX4000 traffic generator on an OC-12 interface. This OC-12 terminates at a Cisco GSR 12012[5] and the traffic is routed back to the AX4000 over an OC-3. The constraint of going from an OC-12 to an OC-3 forms the bottleneck over which packet loss can be generated. Both links are Packet-Over-SONET, as in Internet2. In each direction, we use optical splitters and connect one image of the light stream to an Endace DAG3.5 capture card. By tuning the packet emission parameters at the AX4000, we can generate varying degrees of packet loss at the router. We compare the loss counters at the router with the two packet traces captured at the DAG cards. In essence, we validate one passive measurement with another. The key is that we have clear visibility into the traces produced from the capture cards. As with our wide-area measurements, we aggregate measurements into 30 second sample intervals.

Our three experiments consist of loss regimes created with the AX4000 to generate approximately 1%, 0.1%, and 0.01% packet loss. We uniformly use 256 byte packets (264 with link-layer framing) and generate packet bursts of varying sizes such that the combination of the average inter-burst time and the average burst length create the desired loss rate. Each experiment is two hours long.

Table I gives correlation coefficients obtained by comparing time series of SNMP loss measurements and measurements obtained from the DAG cards for the three experiments. Note that as loss rate decreases, correlation increases. Lower correlation at higher loss rates is primarily an artifact of edge conditions due to the sampling interval. Figure 2(b) qualitatively shows DAG and SNMP measurement correlation for the 0.1% loss experiment. Over the course of the experiment, the SNMP and DAG measures are well aligned. From these experiments we conclude that these counters are implemented with a high degree of accuracy.

B.3 Active Probe Data

Our active measurement data was collected using a modified version of the `zing` utility installed in end hosts in Abilene. We sent 256 byte probes at exponentially distributed intervals with means of 100ms, 50ms, and 10ms (for 10Hz, 20Hz, and 100Hz probes, respectively.) In analysis, we refer to these traces as `zing` traces. In parallel, we sent 256 byte probes with uniform spacing. This uniform probing methodology is essentially the same as the ubiquitous tool `ping` (though in our case, probe packets flow in only one direction). The probes are sent continuously over each data collection period.

We had to modify `zing` because we were unable to use the packet filter capability[6] of the utility due to practical limitations of the kernels installed on the measurement hosts with ATM interface cards. We also modified `zing` to facilitate data storage in files of reasonable size.

In addition to running probes for packet loss, we took `traceroute` measurements across the full mesh of end hosts every 10 minutes. This data enabled us to determine the sets of router interfaces that were encountered along each end-to-end path in our mesh. The loss data from specific sets of interfaces was then compared with active measurement traces between end hosts. Since `traceroute` only reports the ingress interface to a given router (i.e., as a packet exits a link), we used knowledge of the physical Internet2 backbone layout to consider the other end of the link for SNMP loss calculation[7]. Since our study was conducted in the backbone of Internet2, routes were extremely stable. For example, in the 10Hz data set, there were 122 unique paths observed, yielding 66 route changes. These changes were confined to a three day period - all happened at around 4am UTC at the Denver router. The regularity and specificity of the changes led us to believe that this was a standard maintenance activity on the Denver router.

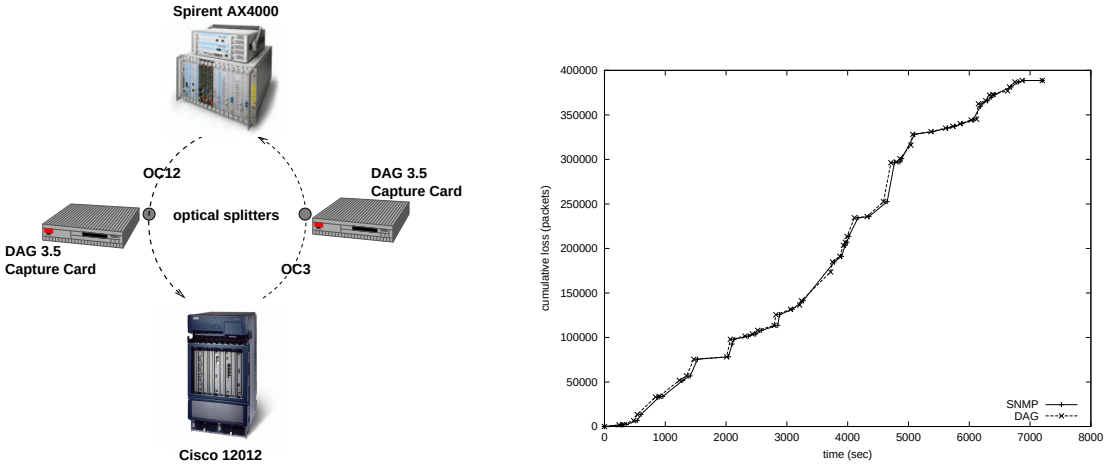
Because we cannot use packet filters at our end hosts, we had no way to determine whether measured loss was due to events internal to the network (uncounted in our router measurements), or at an endpoint. We could not differentiate end host OS buffer overruns or network interface drops (transmission loss) from loss due to network congestion. We can (and did), however, detect interface errors by periodically running `netstat` on end hosts. We return to this issue in our data analysis.

To compare our `zing` and `ping` traces with the SNMP data, we aggregated the probe traces in intervals of 30 seconds to match the SNMP query frequency. The result is

[5]The difference between a Cisco 12008, the GSR model in Internet2, and a 12012 is the number of interface cards that the chassis can accommodate.

[6]Packet filters are in-kernel mechanisms that applications can use for receiving (or less typically, sending) packets at link-layer, thus avoiding any higher-layer protocol processing. Pertinent to our study, they allow an application to find out whether packets have been dropped in the kernel due to buffer limits. Packet filtering capability must be compiled into a kernel, and their use commonly requires executive (“root”) privileges.

[7]The information on the static physical interface layout is openly available from the Abilene NOC [1]



(a) Experimental Configuration.

(b) Cumulative loss over duration of 0.1% loss experiment as measured by DAG cards and SNMP.

Fig. 2. Laboratory evaluation of accuracy of SNMP router-based loss measurements.

TABLE I
CORRELATION BETWEEN SNMP AND DAG LOSS MEASUREMENTS

Experiment	1% Loss	0.1% Loss	0.01% Loss
Coefficient of Correlation	0.73	0.84	0.99

that we have comparable time series at the possible cost of insight into events on smaller time scales for the probe-based measurements. This aggregation causes our analysis to be conservative when reporting *loss events* in the sense that even if they are measured both by SNMP and by an active probe in the same interval, it appears that the active probe has detected the congestion loss event.

B.4 Statistical Issues in Probe-based Measurements of Packet Loss

A fundamental statistical technique for obtaining an unbiased estimate of the average state of a random process is to sample at exponentially distributed intervals. An extension of this approach to queueing systems resulted in the well-known Poisson Arrivals See Time Averages (PASTA) theorem [20]. The theorem, in essence, states that exponentially distributed arrivals at a queue will "see" the average state of the system. This theorem is the basis of a technique for active measurement of packet loss [2], and is the method we employed.

Let X_t be a binary process with states describing whether a packet is lost due to congestion (1) or not (0). We are interested in estimating $p = Pr(X_t = 1)$. Sampling n times at Poisson intervals, we obtain $\bar{X_n}$, the average of the n samples. Thus the expected value $E(\bar{X_n}) = Pr(X_t = 1) = p$. As the number of samples $n \to \infty$, $\bar{X_n} \to p$. Note that this estimate *may have a very large variance*, namely $Var(\bar{X_n}) \approx \frac{p}{n}$. For the standard deviation to be approximately $0.1p$, we need $n \approx \frac{100}{p}$. Thus, for average loss rates on the order of 10^{-4}, we need $n \approx 10^6$ samples.

This simple analysis has a number of very important practical implications. For much lower loss rates than the example above, say 10^{-7} (not uncommon in a highly engineered network like Abilene), even with a relatively fast probe rate of 100Hz we must send probes for nearly 4 months to get an accurate projection of loss rates! Furthermore, if we simply decide to increase our probe rate in order to reduce the time required, we are inevitably forced to make trade-offs because of increasing bandwidth consumption from probe traffic and the potential for skewing the loss measurements. Finally, probe packet size may play a role in probe measurement accuracy since certain routers implement queuing and buffering in different ways. As stated in the introduction, the issue of sampling rate versus loss process is our primary conjecture for the poor correlation results reported in the next section. We do not attempt to offer a set of guidelines for managing the trade-offs between sampling rate, sample volume and accuracy of active measurements for packet loss in this paper.

IV. RESULTS

The first step in our analysis is a qualitative comparison of loss rates for the SNMP and probe measures. We follow this assessment by comparing four distributional characteristics of our data: loss rates, lengths of loss-free periods, loss rates during lossy periods, and loss constancy (based on the notion of change free periods).

In each of our analyses, we first explore the characteristics of packet loss measured at routers over all our paths to provide an understanding of the baseline to which the probe data is compared. To our knowledge, this is the first discussion of wide area loss based on SNMP measurements. Next, we look at the distributional characteristic for all loss measures (SNMP, `zing` and `ping`) along a "canonical path." Finally, we quantify the degree of distributional agreement between `zing` and the router counters and between `ping` and the router counters using the χ^2 goodness-of-fit test with 9 degrees of freedom[8].

We chose the path from Indianapolis to Los Angeles as our canonical path. Our choice was completely arbitrary, but is qualitatively representative of other paths under study. We also note that the end points of our canonical path have direct ATM interfaces to routers. This choice is also arbitrary since we do not see fundamental differences between loss measurements taken between hosts connected by ATM or by Ethernet.

A. Qualitative Comparison

In Figures 3(a), 3(c), and 3(e) we show time series graphs for the router, `zing`, and `ping` data for the canonical path. Note that the y-axis is log scale.

Qualitatively, `zing` and `ping` largely overestimate the lost packets counted by the router interfaces. What is important to note in these graphs is the lower bound of loss rate measured by active probes when we group measurements into sample intervals for the purpose of time series analysis (where lower bound is defined as measuring a single loss event within a specified time interval). This bound is a function of the probe rate and the time interval considered. For example, with our mean probe rate of 10Hz, we sent an average of 300 packets per 30 second sample interval. This set the effective lower bound on loss at a rate of 0.003. For a probe rate of 100Hz, this bound is reduced by an order of magnitude. However, the effective lower bound for SNMP was much lower than that. Assuming an average packet size of 300 bytes, the minimum loss rate over a 30 second period for an OC-48 (2.4×10^9 bps) is roughly 3×10^{-8} which is substantially lower than what can be measured by the active probes[9].

To estimate the effect of transmission loss due to network interface drops we compared the raw data with a "filtered" set. We created this data set from the raw data by removing the losses reported by `zing` or `ping` that were not reported by router interfaces during each 30 second sample interval[10]. Filtered results of the same path are shown in Figures 3(b), 3(d), and 3(f). From the filtered data, we notice that the active probes appear to miss many of the loss events recorded by the router. However, the time scale over which the active probes were taken still effectively overestimated the loss rate during intervals of loss.

There was an apparent periodicity of loss events occurring once per hour present in the 100Hz traces, shown in Figures 3(e) and 3(f). There were no significant background tasks running on the measurement hosts, and the periodicity did not appear in our traces with lower probe frequencies. Furthermore, due to the tremendous amount of data generated we did not initiate download during the 100Hz experiments. We conclude that these figures starkly reveal the effect of transmission loss due to interface drops and do not reflect congestion at a router.

Rows 1 and 2 of Table II help further quantify the effects of transmission loss due interface drops. While the overall loss rate was very low for both raw and filtered data sets, the loss rate of the filtered data was often an order-of-magnitude lower, and occasionally zero. This highlights the shortcomings of active probing for loss even if interface drops could be completely avoided.

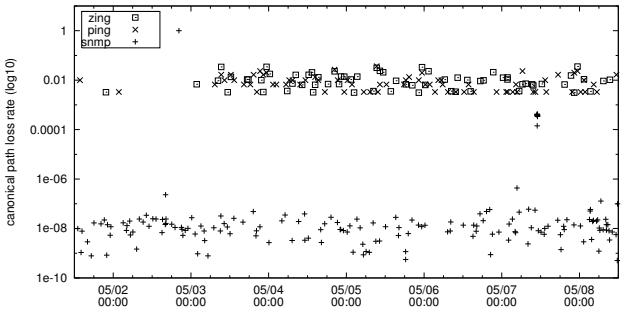
B. Loss Rates

In analysis to follow, we compare the SNMP data with the raw `zing` and `ping` traces from the 20Hz data set. Our reason for continued analysis of the raw data is that many active probing studies have suffered the same restriction of inability to use packet filters to measure interface or operating system buffer drops. Active probe measurements are often taken at remote sites that have volunteered to assist in a particular study, but normally do not grant executive privileges required to use packet filters.

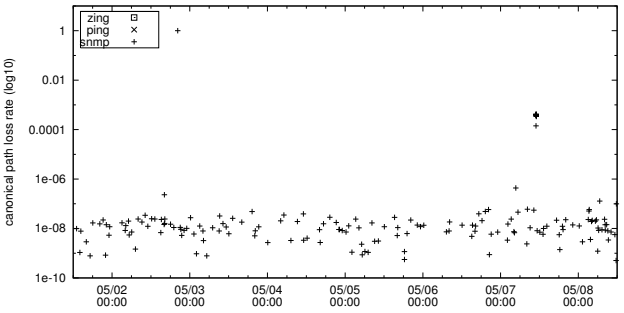
[8]We arbitrarily chose 9 degrees of freedom as a level which conservatively favors finding agreement between two distributions. While other measures of agreement between distributions such as relative entropy could have been employed, our objective was to make a straightforward quantitative comparisons while at the same time demonstrating details of the distributional characteristics.

[9]We see minimum loss rates measured by SNMP on the order of 10^{-9}. These measurements are consistent with the average packet sizes computed from other MIB variables collected during the same sample intervals.

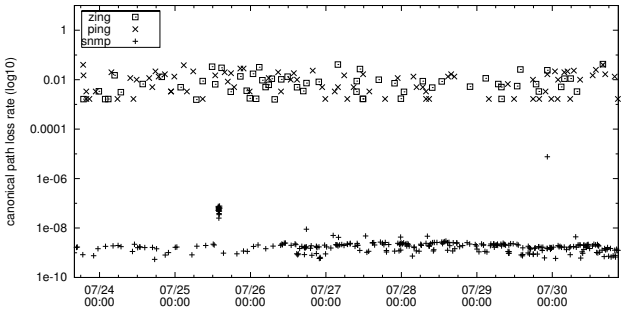
[10]While it could be the case that congestion loss occurred at the router that was measured by the active probe and not by the router counters, we do not consider this to be a significant possibility based on our experiments in §III-B.2.



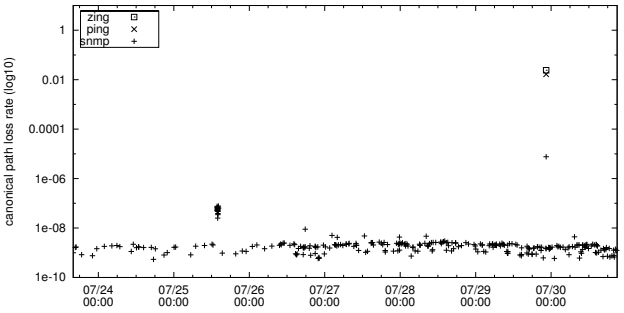
(a) 10Hz Probes - Raw Data



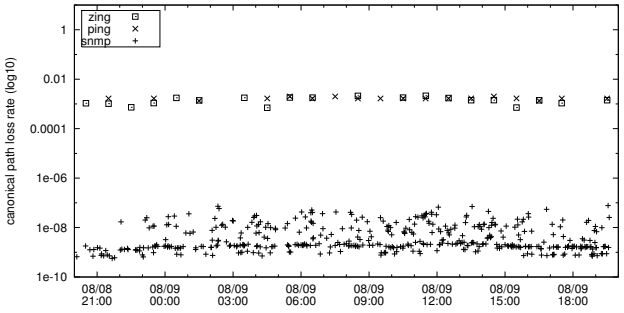
(b) 10Hz Probes - Filtered Data



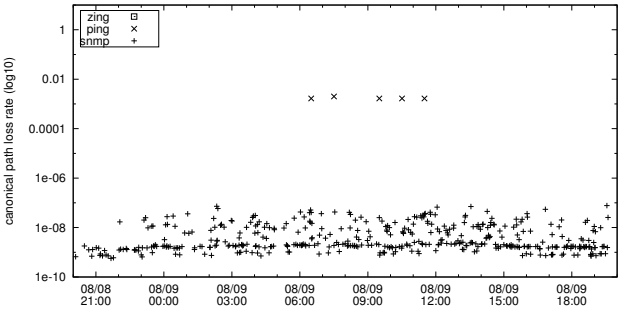
(c) 20Hz Probes - Raw Data



(d) 20Hz Probes - Filtered Data



(e) 100Hz Probes - Raw Data



(f) 100hz Probes - Filtered Data

Fig. 3. Qualitative comparison of loss rates on the Indianapolis to Los Angeles canonical path for each of the raw data sets. Filtered data sets only show active probe events when there was a loss event measured at a router in the corresponding sample interval. The difference between raw and filtered data indicates the effects of transmission loss due to interface drops.

TABLE II
SUMMARY STATISTICS FOR CANONICAL PATH

	Data Set	10Hz		20Hz		100Hz	
		μ	σ	μ	σ	μ	σ
Loss Rate (raw)	SNMP	4.1×10^{-8}	2.4×10^{-6}	4.2×10^{-10}	2.9×10^{-8}	5.2×10^{-10}	2.4×10^{-9}
	ZING	5.9×10^{-5}	3.5×10^{-3}	2.8×10^{-5}	6.7×10^{-4}	9.9×10^{-6}	1.2×10^{-4}
	PING	5.9×10^{-5}	3.8×10^{-3}	3.4×10^{-5}	7.5×10^{-4}	1.1×10^{-5}	1.3×10^{-5}
Loss Rate (filtered)	SNMP	4.1×10^{-8}	2.4×10^{-6}	4.2×10^{-10}	2.9×10^{-8}	5.2×10^{-10}	2.4×10^{-9}
	ZING	0	0	8.2×10^{-7}	1.2×10^{-4}	0	0
	PING	0	0	3.0×10^{-6}	2.1×10^{-4}	3.0×10^{-6}	7.2×10^{-5}
Loss-Free Periods (raw)	SNMP	1.4×10^{2}	1.8×10^{5}	1.9×10^{1}	1.5×10^{3}	6.2×10^{0}	4.4×10^{1}
(seconds)	ZING	5.0×10^{2}	6.4×10^{5}	3.2×10^{2}	6.2×10^{4}	1.4×10^{2}	2.9×10^{3}
	PING	4.6×10^{2}	6.5×10^{5}	2.6×10^{2}	4.8×10^{4}	1.5×10^{2}	5.0×10^{3}
Loss Periods (raw)	SNMP	3.6×10^{-3}	3.4×10^{-3}	1.6×10^{-8}	6.7×10^{-14}	7.6×10^{-9}	1.4×10^{-16}
	ZING	2.8×10^{-2}	5.3×10^{-3}	9.0×10^{-3}	6.5×10^{-5}	1.4×10^{-3}	2.0×10^{-7}
	PING	2.6×10^{-2}	5.9×10^{-3}	9.0×10^{-3}	6.6×10^{-5}	1.7×10^{-3}	3.2×10^{-8}
Change-Free Period	SNMP	2.4×10^{5}	2.7×10^{11}	1.2×10^{6}	0	8.6×10^{4}	0
Duration (raw)	ZING	1.2×10^{6}	0	2.4×10^{3}	3.1×10^{7}	1.2×10^{3}	4.6×10^{6}
(seconds)	PING	1.2×10^{6}	0	4.1×10^{3}	1.3×10^{9}	1.3×10^{3}	7.1×10^{6}
Number of Change-Free	SNMP	5		1		1	
Periods (raw)	ZING	1		511		75	
	PING	1		299		65	

We first considered loss rate distributions for each measurement, including all intervals regardless of whether loss occurred or not. Figure 4(a) shows distributions of loss for all 56 paths for the SNMP data. Figure 4(b) shows the distribution of loss rates for SNMP, `zing` and `ping` (raw data) over the canonical path. We used a complementary cumulative distribution function to highlight the upper tail. Note that the loss rates reported in Table II (rows 1 and 2) appear to be below the minimum discussed above. These loss rates consider *all sample intervals*, some of which are 0, thus pushing the average below the practical minimum for a given sample interval.

Next, we calculated the correlation coefficients for each path between router-based measures and each of the probe traces, and constructed corresponding cumulative distribution functions. Figure 5 shows that for both the raw and filtered traces, correlation is generally poor. The traces responsible for much of the positive correlation suffered little or no loss. Another feature to note is that neither `zing` nor `ping` have distinct correlational advantages.

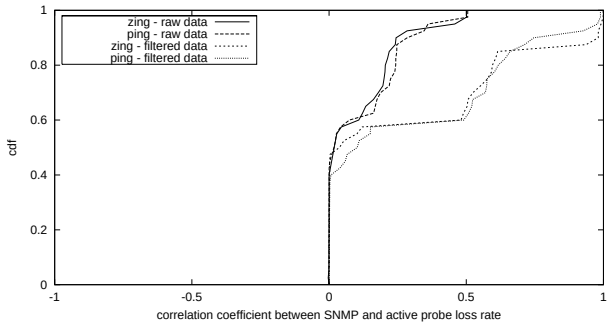


Fig. 5. Cumulative distribution of correlation coefficients between router-based and probe-based measures of loss rates for all paths taken during the 20Hz probe period.

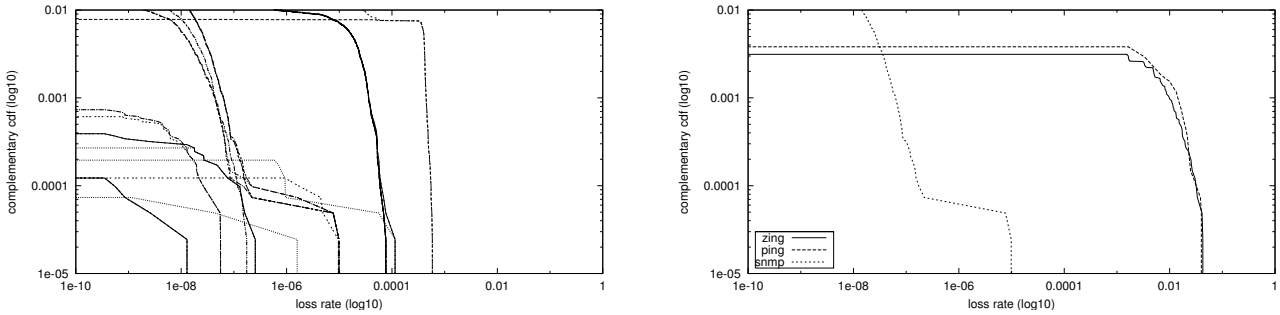
C. Loss-Free Periods

A loss-free period is defined as the maximum number of consecutive 30 second sample intervals during which no loss is measured. Another way of understanding this measure is to think in terms of loss event interarrival times.

Figure 6(a) shows the cumulative distributions of loss-free periods for all paths measured by SNMP. For some paths, it appears that losses occurred frequently, and with apparent regularity. For other paths, however, losses occurred infrequently. This wide variety of loss interarrival times poses a challenge for determining how to best conduct active network measurements for loss while introducing a minimum level of probe traffic into the network.

Figure 6(b) shows the cumulative distributions of loss-free periods for each measurement method along the canonical path. The key feature to notice is that losses were more closely spaced in time as measured by the routers than by `zing` or `ping`.

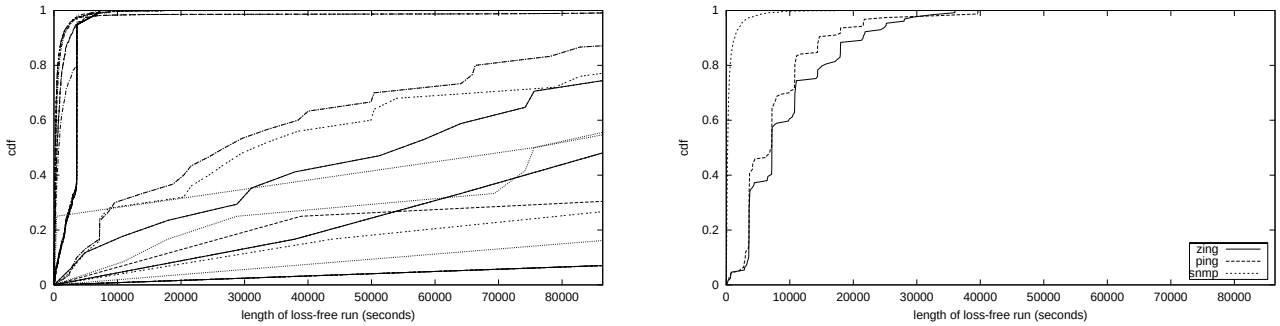
Figure 7 shows the cumulative distributions of the χ^2 goodness-of-fit statistic for `zing` and SNMP, and for `ping` and SNMP. We also plotted vertical lines indicat-



(a) All 56 paths via SNMP

(b) Canonical path, probe vs. SNMP

Fig. 4. Complementary cumulative distributions of loss rates in 30 second sample intervals for all router-based measurements (a) and probe-based measurements on the canonical path (b) taken during the 20Hz probe period.)



(a) All 56 paths via SNMP

(b) Canonical path

Fig. 6. Cumulative distributions of the number of consecutive 30 second sample intervals that are loss free for all paths using SNMP router-based measurements (a) and probe-based measurements on the canonical path (b) taken during the 20Hz probe period.

ing the 95% and 1% acceptance levels[11]. Note that the x-axis is plotted on a log scale. It is immediately clear that even at the 1% acceptance level, `zing` and `ping` were not good fits to the distribution of loss-free periods measured by SNMP.

D. Loss Periods

Next, we assessed the loss rates measured only during the 30 second intervals over which packet loss was detected. Figure 8(a) plots cumulative distributions of loss rates during these loss periods for all paths using the SNMP router-based traces. From the figure, we observe a wide range of loss rates measured by SNMP. This range again poses a challenge for designing how best to actively measure packet loss.

For the canonical path, Figure 8(b) shows that `zing` and `ping` experienced very different loss rates than were measured by routers via SNMP. The lower bound on loss rate measurable by the probes (because of the sampling rate) is obvious from the curves, and for this path `zing` and `ping` measured similar loss rates[12].

[11]The χ^2 goodness-of-fit test is a hypothesis testing procedure. A fit hypothesis is accepted at a given confidence level if the χ^2 metric is less than the χ^2 distribution value with specified degrees of freedom

[12]For the 100Hz loss period average shown in Table II (row 4), note

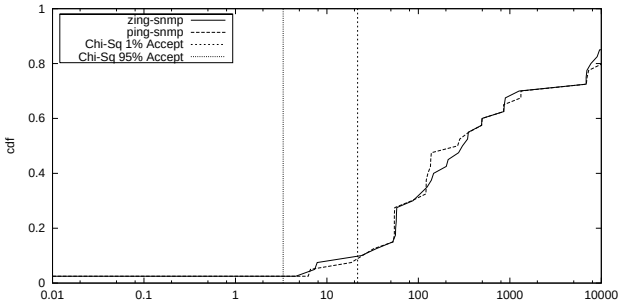


Fig. 7. Cumulative distribution of correlation coefficients between router-based and probe-based measures of loss-free periods for all paths taken during the 20Hz probe period.

We do not plot the results for the χ^2 test on loss period distributions. The reason is that the test falsely indicates that the loss periods measured by `zing` and by `ping` are good fits to the SNMP measurement. The reason for this is simple if we consider the effect of binning when computing the statistic: if we use 10 bins and the maximum loss rate measured is more than 1%, almost all of the measured values for all three data sets will fall in the lowest bins (recall Figure 3(c) and Table II), thus giving a (false) positive indication for goodness of fit with high confidence.

E. Change Free Periods

Finally, we compared how loss constancy is measured by probes versus the loss constancy measured along a path of router interfaces. We used the notion of *change free periods*, as described in [22]. In that work, a time series was defined as "a series of piece-wise steady regions delineated by *change points*. With a parameterized family of models (e.g. Poisson processes with some rate), a time series in a change free period (CFP) is modeled by a particular value of the parameter." As such, the task of identifying CFP's is reduced to identify change points.

We used the bootstrapping method for generating change points. As noted in [22], this method is conservative in the sense that it is more likely to miss actual change points and thus overestimate the duration of steady regions. An area for future work would be to explore other methods for finding change points in the router-based data.

Figure 9(a) shows cumulative distributions of the duration of change free periods for all paths measured using router-based SNMP traces. Analogous to Figure 6(a), it indicates that there was a wide range of durations over which path loss was steady. There were a number of paths for which conditions did not change for days, and there were also a number of paths on which loss conditions changed with much higher frequency.

Figure 9(b), showing cumulative distributions of the duration of change free periods for the canonical path, indicates that `zing` and `ping` both experienced high proportions of short durations of steady loss rates. The view of constancy seen through the router interfaces for this particular path, however, is that the loss rate was steady over the entire collection period.

Figure 10 plots the cumulative distribution of the χ^2 statistic for comparing change free periods seen by `zing` and SNMP and `ping` and SNMP across all paths. Vertical lines are plotted indicating the 95% and 1% acceptance levels. Clearly, neither `zing` nor `ping` are good fits with the router measurements.

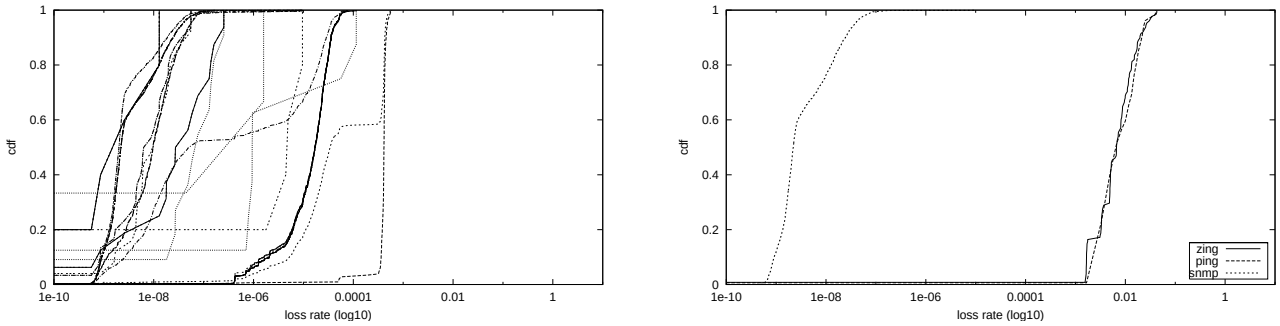
Finally, we plotted the cumulative distribution for the number of change free periods for all the SNMP traces, all the `zing` traces, and all the `ping` traces in Figure 11. Immediately, we notice that there were many fewer change points measured by the router interfaces across all paths. Comparing Figure 11 with Figure 9(a), we infer that there were fewer numbers of change free regions of short duration recorded by the routers, and more, rather long regions of constancy. Our data indicates that `zing` and `ping`, in contrast to the routers, tended to measure more change free periods of short durations.

V. Conclusions and Future Work

In this paper, we present a comparison of active and passive measurements of packet loss. Our study is based on measurements taken in a laboratory and in the wide area over a total of three weeks using the Surveyor infrastructure and the Internet2/Abilene backbone. We gathered passive packet loss data from Abilene backbone routers via SNMP in 30 second intervals. We used `zing` to actively probe for loss at rates of 10Hz, 20Hz, and 100Hz, and then aggregated these values into 30 second intervals to assess how well the two measures correlate with each other. Our comparison considers the degree of correlation between loss rate time series and the degree of agreement between distributions of loss characteristics including lengths of loss free periods, loss rates during loss periods, and the duration of change free regions.

Our laboratory results demonstrate the accuracy of passive SNMP measurements on the equipment used in this study. Our wide area SNMP measurements are the first of their kind to be reported and show that loss rates as measured by active probes are not well correlated with those measured by SNMP. We also show that the distributions

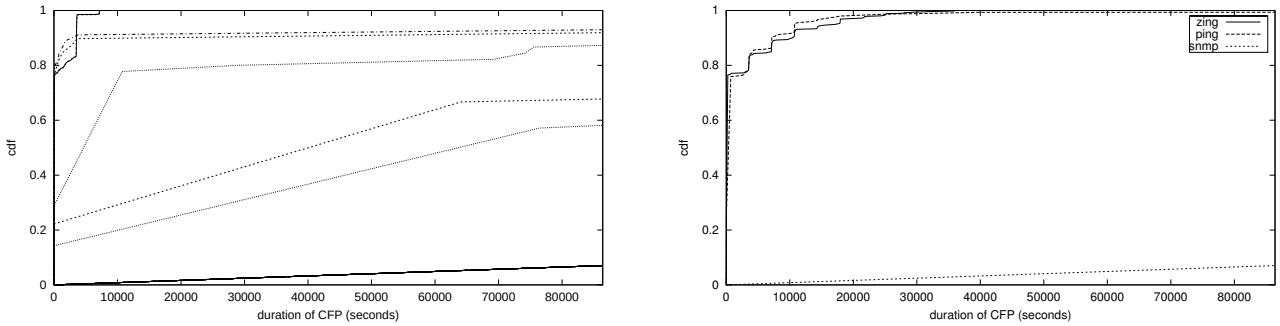
that this very low average loss rate of 7.6×10^{-9} implies an average packet size of approximately 68 bytes.



(a) All 56 paths via SNMP

(b) Canonical path, probe vs. SNMP

Fig. 8. Cumulative distributions of loss rates for sample intervals with loss for all paths using SNMP router-based measurements (a) and probe-based measurements on the canonical path (b) taken during the 20Hz probe period.



(a) All Paths for SNMP

(b) Canonical Path, probe vs. SNMP

Fig. 9. Cumulative distributions of change free periods for all paths using SNMP router-based measurements (a) and probe-based measurements on the canonical path (b) taken during the 20Hz probe period.

of values for loss free periods, loss rates during loss periods, and the duration of change free regions as seen by active probes do not align closely with the distributions of the same values as seen by SNMP.

We also evaluated the differences between loss rates as measured by the Poisson modulated `zing` tool and the more simple `ping` utility which sends out probes at constant intervals. At least for the low loss rates seen in our measurement infrastructure, `ping` provides qualitatively the same level of accuracy as `zing`.

We conjecture that the lack of correlation between passive and active loss measurements is due to the difficulties in sampling very low loss rates. Our next step will be to investigate this conjecture in a controlled laboratory setting. Should this theory prove to be true, then we plan to investigate new methods for probing that are both lightweight and have the ability to detect loss over shorter time frames.

VI. Acknowledgements

We would like to thank Matt Zekauskas, Guy Almes, Chris Robb and the people at Internet2/Abilene for their help in data collection. Without their hard efforts, this study would not have been possible. We would also like to thank Larry Landweber for fruitful discussions, Dave Donoho for his statistical insights, and Jay Rosenbloom of Cisco for help in demystifying IOS. We gratefully acknowledge the generous donation of equipment from

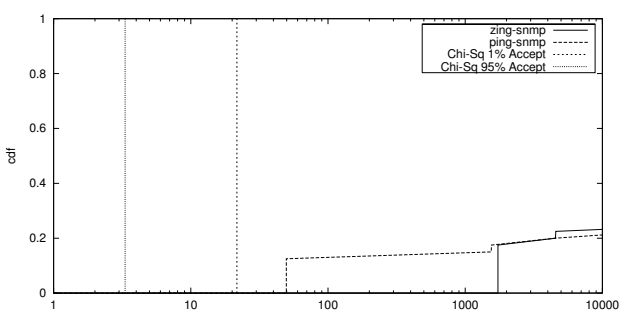


Fig. 10. Cumulative distribution of correlation coefficients between router-based and probe-based measures of change free period duration for all paths taken during the 20Hz probe period.

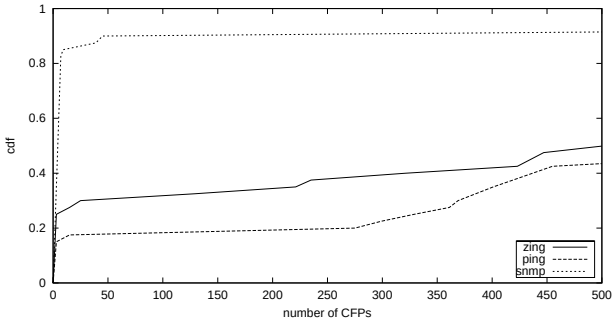


Fig. 11. Cumulative distribution of change free periods durations for all paths taken during the 20Hz probe period.

Cisco Systems, and thank Spirent Communications for use of the AX4000 traffic generator. Finally, we thank Nevil Brownlee, Mark Crovella, Sue Moon, and Yin Zhang for their excellent input and advice on this work.

REFERENCES

[1] Abilene. http://abilene.internet2.edu, 2003.

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

[3] NLANR Active Measurement Program AMP. http://moat.nlanr.net/AMP.

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

[5] J. Case, M. Fedor, M. Schoffstall, and J. Davin. A simple network management protocol (SNMP). IETF RFC 1157, 1990.

[6] Cisco Systems, Inc. Cisco 12000 Series Internet Router Architecture. http://www.cisco.com/public/technotes/arch12000toc_8832.html (requires registration).

[7] M. Coates and R. Nowak. Network loss inference using unicast end-to-end measurement. In *Proceedings of ITC Conference on IP Traffic, Measurement and Modeling*, September 2000.

[8] C. Cranor, Y. Gao, T. Johnson, V. Shkapenyuk, and O. Spatscheck. Gigascope: High performance network monitoring with an SQL interfact. In *Proceedings of SIGMOD '02)*, Madison, WI, June 2002.

[9] N. Duffield, F. Lo Presti, V. Paxson, and D. Towsley. Inferring link loss using striped unicast probes. In *Proceedings of IEEE INFOCOM '01*, Anchorage, Alaska, April 2001.

[10] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockwell, T. Seely, and C. Diot. Packet-level traffic measurements from the sprint ip backbone. In *IEEE Network*, 2003.

[11] J. Hoe. Improving the start-up behavior of a congestion control scheme for TCP. In *Proceedings of ACM SIGCOMM '96*, Palo Alto, CA, August 1996.

[12] J. Mahdavi, V. Paxson, A. Adams, and M. Mathis. Creating a scalable architecture for Internet measurement. In *Proceedings of INET '98*, Geneva, Switzerland, July 1998.

[13] M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow. TCP selective acknowledgement options. IETF RFC 2018, 1996.

[14] W. Matthews and L. Cottrell. The PINGer Project: Active Internet Performance Monitoring for the HENP Community. *IEEE Communications Magazine*, May 2000.

[15] A. Pasztor and D. Veitch. A precision infrastructure for active probing. In *PAM2001, Workshop on Passive and Active Networking*, Amsterdam, Holland, April 2001.

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

[17] The Surveyor Project. http://www.advanced.org/csgippm/, 1998.

[18] S. Savage. Sting: A tool for measuring one way packet loss. In *Proceedings of IEEE INFOCOM '00*, Tel Aviv, Israel, April 2000.

[19] William Stallings. *SNMP, SNMPv2, SNMPv3, and RMON 1 and 2*. Addison Wesley Longman, Inc., third edition, 1999.

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

[21] M. Yajnik, S. Moon, J. Kurose, and D. Towsley. Measurement and modeling of temporal dependence in packet loss. In *Proceedings of IEEE INFOCOM '99*, New York, NY, March 1999.

[22] 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.

APPENDIX

A. PACKET LOSS IN A CISCO GSR

In this appendix, we provide some detail on how packets are actually lost inside the GSRs from an operational perspective, and for understanding the meaning and limitations of our router-based measurements.

Tracing the lifetime of a packet through a GSR [6], the packet may be dropped in the following areas:

Burst buffer: Upon arrival at an interface, the packet is copied into a "burst buffer" of size $2 \times MTU$ where it awaits input buffer allocation. The primary cause for this

type of drop is the inability of the physical interface module to allocate buffer space in a timely manner. This situation can occur with an extremely heavy volume of small packets. Actual buffer space may exist, but it cannot be acquired fast enough.

Input queue drop: In deciding the output interface for a packet, the GSR attempts to make a routing decision in an interface interrupt handler using a cached exact match. This fast path routing decision is the most common path packets take through a Cisco GSR. Packets which cannot be routed using this fast path logic are queued on input awaiting slower processing by the main router processing module. If this input queue exceeds the configured size, packets are dropped. All packets bound for the router itself must take this slow path. Additionally, packets arriving on some interfaces invariably take the slow path. Notably, this slow path is taken for packets arriving on 100Mbps Ethernet interfaces, and other interfaces with relatively slow line rates. The Cisco GSR also employs virtual output queueing to avoid head-of-line blocking. After a fast-path decision is made, a packet is queued in a virtual output queue awaiting transmission across the internal switching fabric. Drops may occur if this virtual output queue is full. These losses cannot be distinguished from other types of input queue drops.

No input buffers: Lack of a properly sized input buffer can cause a packet to be dropped. This drop can occur either on the slow path or fast path of routing decision.

Switching fabric: Internal switching fabric congestion can result in packet loss internal to the router.

Output queue drop: This type of loss is the archetypal situation of congestion in a statistically multiplexed packet switched network. The output line rate is less than the aggregated input source rates. Packets are queued awaiting transmission and are dropped when the output queue is full according to an algorithm such as drop-tail or RED.

Of the above, the only type of drop we cannot measure through SNMP is loss due to congestion in the internal router switching fabric. This type of loss detection requires debugging capability to the router and cannot be gathered from SNMP. Losses of this type are thought to be very rare, although we were not able to find a means for quantifying this phenomenon.