Network Radar: Tomography from Round Trip Time Measurements

Yolanda Tsang, Mehmet Yildiz, Paul Barford, Robert Nowak*

ABSTRACT

Knowledge of link specific traffic characteristics is important in the operation and design of wide area networks. Network tomography is a powerful method for measuring characteristics such as delay and loss on network-internal links using end–to–end active probes. Prior work has established the basic mechanisms for the use of tomographic inference techniques in the networking context. However, the measurement methods described in prior network tomography studies require cooperation between sending and receiving end-hosts, which limits the scope of the paths over which the measurements can be made. In this paper, we describe a new network tomographic technique based on round trip time (RTT) measurements which eliminates the need for special-purpose cooperation from receivers. Our technique uses RTT measurements from TCP SYN and SYN-ACK segments to estimate the delay variance of the shared network segment in the standard one sender - two receivers configuration. We call this approach *Network Radar* since it is analogous to standard radar. We present an analytic evaluation of Network Radar that specifies the variance bounds within which the technique is effective. We also evaluate Network Radar in a series of tests conducted in a controlled laboratory environment using live end hosts and IP routers. These tests demonstrate the boundaries of effectiveness of the RTT-based approach.

*Y.Tsang is with the ECE Dept. of Rice University and University of Wisconsin – Madison. E-mail: ytsang@rice.edu. M.Yildiz and R.Nowak are with ECE Dept. of University of Wisconsin – Madison. E-mail: yildiz@cae.wisc.edu and nowak@engr.wisc.edu. P.Barford is with the CS Dept. of University of Wisconsin – Madison. E-mail: pb@cs.wisc.edu. This work was partially supported by the National Science Foundation, grants 0335234, 0325653, CCR-0310889, CCR-0325571, and ANI-0099148, DOE SciDAC, the Office of Naval Research, grant N00014-00-1-0966 and by support from Cisco Systems. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Office of Naval Research, DOE, National Science Foundation or Cisco Systems.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IMC'04, October 25–27, 2004, Taormina, Sicily, Italy.
Copyright 2004 ACM 1-58113-821-0/04/0010 ...$5.00.

Keywords

Network Tomography, Delay Measurement, Loss Measurement

1. INTRODUCTION

Network tomography is a powerful method for measuring and analyzing link specific characteristics using end–to–end active probes. This capability is important since link specific information such as delay and loss is otherwise only available to network administrators who have direct access to those links. Prior work has established the basic mechanisms for the use of tomographic inference techniques in the networking context [1, 2, 3, 4, 5, 6, 7, 8]. However, the methods described in prior network tomography studies all require cooperation between sender and receiver end-hosts. This limits both the scope of the paths over which the measurements can be made and wide-spread used of the technique. In this study, we develop and evaluate a new network tomographic technique based on *round trip time (RTT)* measurements which eliminates the need for special-purpose cooperation from receivers. This RTT-based approach can potentially expand the range of paths over which tomographic measurements can be made and enable tomographic tools to be more widely used than prior techniques.

The link delay measurement method that we develop and analyze in this study is based on the idea of sending two closely time-spaced (back–to–back) active probes from a single sender to two separate receivers. If one were to trace the paths of these probes from sender to receiver, they would form a tree with the root at the sender, a common trunk and the leaves at the receivers. The basic tomographic idea is that the two probe packets should experience nearly the same delay on each of the shared links of their paths. If the delays on the shared links are identical, then any differences in total delay measured are caused by the conditions experienced by the probe packets on the unshared links. This simple observation forms the basis for the estimation of the delay characteristics on each link via tomography. By repeating this sort of probing many times to many different pairs of receivers, it is possible to reconstruct the (logical) link delay distributions on all branches connecting the sender to the receivers.

Our method uses RTT measurements of back–to–back packets sent to different pairs of receivers. The important advantage of this approach is that it enables tomographic delay measurement to be conducted widely in the Internet, since special-purpose measurement and cooperation is not required at the receivers. The basic idea is depicted in Figure 1. We send back-to-back packets from the sender 0 to receiver nodes 1 and 2. We then collect response packets from the receivers and measure round trip times. Assuming that the delays experienced on all links beyond the branching point are uncorrelated, it is theoretically possible to determine the delay characteristics on the shared segment from the source to the branching point and on the unshared segments from the branching point, to the receivers, and back to the sender. We call this RTT-based approach *Network Radar* since it is analogous to the idea of standard radar which sends signals into a medium, collects the "echo" and compares signal to echo strength ratio to estimate the distance to the objects. In this paper, we assess the validity and capabilities of RTT-based network tomography through delay variance estimation.

There are several key challenges associated with RTT measurements (in addition to the issues faced by previous tomographic methods which include route stability, identical delays on shared segment, spatial and temporal independence otherwise) which must be considered in order for Network Radar to be practically used. First, extra delays may be incurred due to random return/response generation times at the receivers (ideally the response generation time is zero). Response delays may add a significant "noise" component to the measured RTTs, limiting the accuracy of tomographic methods. Second, a segment of the return paths will be shared by the response packets from the receivers. This could introduce additional correlations into the RTT measurements that are not due to the shared outward segment of interest (ideally the return paths are uncorrelated). We present an initial investigation of the validity of all ideal assumptions as part of this work, and endeavor to determine the robustness of our tomographic methods in realistic, non-ideal conditions.

In building a prototype tool to realize *Network Radar* capability, we had to consider how to gather RTT measurements effectively. The most common tools for measuring RTTs between end-hosts employ the Internet Control Message Protocol (ICMP) using either time exceeded or echo request/reply messages. While tools that use ICMP are useful in network troubleshooting, they have well-known limitations for precise delay measurements including the fact that Internet Service Providers often block or rate-limit ICMP traffic, and that ICMP traffic is often given lower priority in routers. Our solution is the same as has been adopted in other tools which is to use TCP SYN and SYN-ACK connection setup handshaking mechanism to measure RTT. Even this type of RTT measurement can include additional setup time delays that do not occur with more standard RTT measurement techniques, we show that the setup time is quite small and nearly a constant offset which therefore does not significantly affect delay variance measurement.

1.1 Contributions

The contributions of this paper are as follows: (1) development of a tool, *Network Radar*, for estimating shared link characteristics based on RTT measurements from the TCP connection setup mechanism. The tool differs from prior tomography tools in that it does not require special cooperation from receivers or clock synchronization between end hosts, (2) development of a analytical characterization of the delay variance estimator employed by Network Radar, (3) preliminary evaluation the practicality and effectiveness of Network Radar in a controlled network laboratory environment.

1.2 Paper structure

The remainder of the paper is structured as follows. In Section 2 we described related work. In Section 3 we describe the methodology, measurement framework and analytical implications. In Section 4 we describe the details of the experimental environment, the test conducted and the test results. In section 5, we conclude and discuss our future research direction.

2. RELATED WORK

Network tomography based on the use of one-way measurements between cooperating end-hosts has received considerable attention in the networking community [1, 3, 6, 7, 9, 10, 11, 12]. These techniques require synchronization at the end hosts and/or special-purpose measurement capabilities at internal routers. Most of these methods are not widely applicable because of the lack of an available widespread infrastructure for monitoring.

Some recent studies describe measurement tools that attempt to infer path characteristics from RTT measurements are based on the use of Internet Control Message Protocol (ICMP) time-to-live (TTL) [10, 11] and ICMP timestamp options [3]. Our measurement methodology is distinguished from those in our use of the TCP three way handshake mechanism which is essential to the majority of traffic in the Internet, making it widely applicable and easy to deploy. Other tools that use TCP SYN, SYN-ACK for RTT measurements include `Sting` for loss measurement [13] and `Synack` for RTT estimation [14].

The tomographic study conducted by Duffield and Lo Presti in [7] is perhaps the most closely related to ours and serves as a guide for our approach. That paper estimates link delay variance from tomographic measurements in a multicast setting. Specifically, the authors evaluated link delay variance from one-way end-to-end measurements in both an analytical framework and in `ns-2` simulations. The objective of our study is to consider link delay variance as a mechanism for evaluating the robustness of our Network Radar tool. The Duffield and Lo Presti method also assumes the availability of multicast routing, synchronized clocks and the ability to measure at both sender and receiver. Network Radar's RTT-based design mitigates all of these requirements.

3. METHODOLOGY AND MEASUREMENT FRAMEWORK

In this paper, we concentrate on networks comprised of a single source transmitting measurement probes to two receivers. We assume that the topology is fixed throughout the measurement period; i.e., the routing table does not change. For the networks we consider, standard network routing protocols produce a tree-structured topology, with the *source* at the root and the *receivers* at the leaves. The one–sender–two–receiver network is depicted in Fig. 1. The branching node between the source and receivers represents an internal *router*. Connections between the source, router, and receivers are called *links*. Each link between may be a direct connection, or there may be "hidden" routers (where no branching occurs) along the link that are not explicitly shown in Fig. 1.

The basic measurement and inference idea is quite straightforward. Suppose two closely time-spaced (back-to-back) packets are sent from the source to two different receivers. The paths to these receivers traverse a common set of links, but at some point the two paths diverge (as the tree branches). The two packets should experience approximately the same delay on each shared link in their path. The round trip delay consists of

$$y = t_{\text{transmission}} + t_{\text{propagation}} + t_{\text{processing}} + t_{\text{queueing}}.$$

The delay variances are mainly caused by t_{queueing}, and the other terms in the delay can be modeled as a nearly constant quantities.

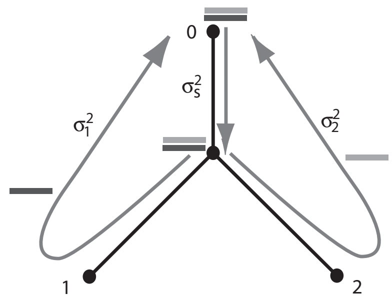

Figure 1: A one sender (0) two-receiver (1, 2) network with delay variances denoted.

For this study we focus on the problem of estimating *delay variance* of the shared network segment. Extending our technique to the problems of delay distribution and loss estimation is beyond the scope of this work although there is nothing inherent in our approach that prevents the estimation of these characteristics. The delay variance estimation problem is easily understood in the case depicted in Figure 1. We index the RTT packet pair measurements by $k = 1, ..., N$. Denote the round trip time measurements to be $\boldsymbol{y} \equiv \{y_1(k), y_2(k)\}_{k=1}^N$ where $y_1(k)$ and $y_2(k)$ denote the kth RTT measurements to/from receiver 1 and 2, respectively. Denote the delay on each link as $d_i, i \in \{s, 1, 2\}$, then $y_1(k) = d_s(k) + d_1(k)$ and $y_2(k) = d_s(k) + d_2(k)$. Note that because the TCP SYN packets are sent back-to-back, the delay on the shared link $d_s(k)$ is assumed to be identical in $y_1(k)$ and $y_2(k)$. Also note that d_i, $i = 1, 2$, refers to the total time that the TCP SYN-ACK packets spend traveling from the branching node to the corresponding receiver, and then back to the sender. Let σ_s^2 denote the delay variance of the shared link and σ_i^2, $i = 1, 2$, denote the delay variances on the unshared paths. Because the delays on the shared and unshared links are assumed to be independent, a straightforward calculation shows that $\sigma_s^2 = \text{var}(d_s) = \text{cov}(y_1, y_2)$ (see Proposition 1 in Section 3.2). Also, from these packet pair RTT measurements, we can resolve the variances on each segment by solving the following equation:

$$\begin{bmatrix} \text{var}(y_1) \\ \text{var}(y_2) \\ \text{var}(y_1 - y_2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \sigma_s^2 \\ \sigma_1^2 \\ \sigma_2^2 \end{bmatrix},$$

but we will focus on the estimation of the variance on the shared path in the remainder of the paper.

3.1 Measurement Methodology

Our method is based on sending back-to-back pairs of TCP SYN packets to different receivers and measuring the delay between the sending time and the time at which the TCP SYN-ACKs are received at the sender. This requires a simple time difference measurement at the sender. Most importantly, this scheme does not require synchronization with the receivers nor special purpose support from any internal network elements. The time-stamping mechanism used at the sender is the `tcpdump [15]` utility, which can be commonly found on most systems. The precision of the timestamp of `tcpdump` is 1μsec and the two packets in a packet pair are sent as close as possible. In our measurements, the average spacing between an outbound packet pair is 10μsec.

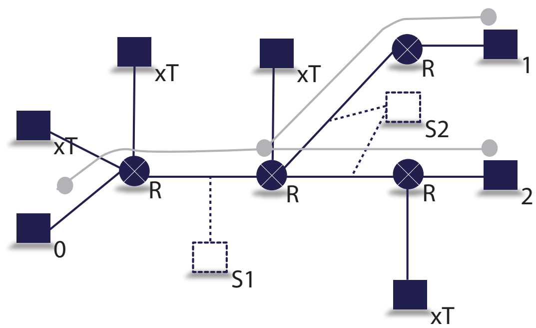

Figure 2: The network under study in WAIL, with 4 Routers and 9 pcs. The sending host is 0 and the receivers are 1 and 2 (logical topology in grey). The boxes xT denote cross-traffic generators and the balls R denote CISCO 3600 series routers. $S1$ and $S2$ denote measurement systems in place to validate the performance of our Network Radar tool.

3.2 Analytical framework

The key statistical quantity in Network Radar is the delay correlation estimator defined as follows. Here, the two endpoints involved in the packet pair probing are denoted simply as 1 and 2.

Definition 1. *Denote the N RTT packet pair measurements $\boldsymbol{y} \equiv \{y_1(k), y_2(k)\}_{k=1}^N$. The RTT covariance is defined as*

$$\widehat{\rho} \equiv \frac{1}{N-1} \sum_{k=1}^{N} (y_1(k) - \bar{y}_1)(y_2(k) - \bar{y}_2) \qquad (1)$$

where $\bar{y}_i$ is the sample mean of $\{y_i(k)\}_{k=1}^N$ for $i = 1, 2$.

Proposition 1. *$\widehat{\rho}$ is an unbiased estimator of the variance on the shared path.*

Proof 1. *Let μ_i, $i = 1, 2$, denote the (unknown) mean RTT in each case. We show first that the true correlation $\rho \equiv E[(y_1(k) - \mu_1)(y_2(k) - \mu_2)]$ is equal to the delay variance on the shared path. Then we show that $\widehat{\rho}$ is an unbiased estimator of ρ. Let $\tilde{y}_i(k) = y_i(k) - \mu_i$, $i = 1, 2$. Then $\rho = E[\tilde{y}_1(k)\tilde{y}_2(k)]$. Each RTT measurement can be written as a sum of the delay on the shared path and the delay incurred on the unshared path:*

$$\tilde{y}_i(k) = \tilde{y}_{i,shared}(k) + \tilde{y}_{i,unshared}(k), \; i = 1, 2.$$

Assuming the delays on the unshared paths are independent, we have

$$\begin{aligned}\rho &= E[(\tilde{y}_{1,shared}(k) + \tilde{y}_{1,unshared}(k))(\tilde{y}_{2,shared}(k) + \\ &\quad \tilde{y}_{2,unshared}(k))], \\ &= E[\tilde{y}_{1,shared}(k)\tilde{y}_{2,shared}(k)],\end{aligned}$$

where we exploit the fact that $\tilde{y}_i(k)$, $i = 1, 2$, are zero mean. Now, assuming the two packets were back-to-back on the shared path, the delays on the shared path are identical and quantity $E[\tilde{y}_{1,shared}(k)\tilde{y}_{2,shared}(k)]$ is precisely the delay variance on the shared path. To show that $E[\widehat{\rho}] = \rho$, verifying that $\widehat{\rho}$ is an unbiased estimator of the delay variance on the shared path, let us consider the expectation of one term in the summation in (1):

$$\begin{aligned}& E[(y_1(k) - \bar{y}_1)(y_2(k) - \bar{y}_2)] \\ = \; & E[(y_1(k)y_2(k)] - \frac{1}{N}\sum_{\ell=1}^{N} E[(y_1(k)y_2(\ell)] \\ & -\frac{1}{N}\sum_{\ell=1}^{N} E[y_1(\ell)y_2(k)] + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E[(y_1(i)y_2(j)].\end{aligned}$$

Noting that

$$E[y_1(k)y_2(\ell)] = \begin{cases} \mu_1\mu_2 & k \neq \ell \\ \rho + \mu_1\mu_2 & k = \ell \end{cases}$$

and substituting into the expression above, shows that the expectation of each term is $(1 - 1/N)\rho$. Therefore, since the terms are identically distributed we have

$$\begin{aligned}E[\widehat{\rho}] &= \frac{1}{N-1}\sum_{k=1}^{N} E[(y_1(k) - \bar{y}_1)(y_2(k) - \bar{y}_2)] \\ &= \frac{1}{N-1} N(1 - 1/N)\rho = \rho. \qquad \Box\end{aligned}$$

We now turn our attention to the accuracy of the estimator $\widehat{\rho}$. The accuracy depends, of course, on the variability of the estimator. The larger the standard deviation of $\widehat{\rho}$, the less confidence we have in the estimated value of the delay variance on the shared path. We examine two issues. First, we provide an expression for the true standard deviation of $\widehat{\rho}$. This expression reveals the various sources of error in the estimation process. Second, we provide a data-based estimator of the standard deviation which can be used to obtain a confidence measure in practice. Calculations for these expressions are somewhat involved, and due to space limitations we simply state the results here.

The variance of $\widehat{\rho}$ is given by

$$E[(\widehat{\rho} - \rho)^2] \;=\; E[\widehat{\rho}^2] - \rho^2,$$

so let us focus on calculating $E[\widehat{\rho}^2]$.

$$\begin{aligned}E[\widehat{\rho}^2] &= \frac{1}{(N-1)^2}\sum_{i=1}^{N}\sum_{j=1}^{N} E[(y_1(i) - \bar{y}_1)(y_2(i) - \\ &\quad \bar{y}_2)(y_1(j) - \bar{y}_1)(y_2(j) - \bar{y}_2)] \\ &= \frac{1}{(N-1)^2}\sum_{i,j} E[\tilde{y}_1(i)\tilde{y}_2(i)\tilde{y}_1(j)\tilde{y}_2(j)] + O(N^{-1})\end{aligned}$$

where $\tilde{y}_m = y_m - \mu_m$, $m = 1, 2$, as in the proof of Proposition 1, and the $O(N^{-1})$ term comes from the fact $\widehat{\rho}$ employs the empirical means rather than the true means. When $i \neq j$ in the sum above, then the expectation of the corresponding term is simply ρ^2. Otherwise, when $i = j$, the expectation is a fourth order cross moment. These moments depend not only on the delay variability on the shared path, but also on the unshared paths. Thus, the "noise" (e.g., delay variability) on unshared paths also impacts the performance of the estimator. From here we can write

$$\begin{aligned}E[\widehat{\rho}^2] &= \frac{1}{(N-1)^2}[N(N-1)\rho^2 + \\ &\quad NE[\tilde{y}_1(i)\tilde{y}_2(i)\tilde{y}_1(j)\tilde{y}_2(j)] + O(N^{-1}) \\ &= \frac{N}{(N-1)}\rho^2 + O(N^{-1})\end{aligned}$$

where we have absorbed the fourth order moment term with the other $O(N^{-1})$ error terms from above. It follows that the variance of $\widehat{\rho}$ is

$$E[(\widehat{\rho} - \rho)^2] \;=\; \frac{1}{N-1}\rho^2 + O(N^{-1}),$$

Thus, we see that the variance of $\widehat{\rho}$ decays like N^{-1}, and it follows that the standard deviation drops off like $N^{-1/2}$. This insures that by using enough probes our estimator $\widehat{\rho}$ should be quite accurate. But, how many probes is "enough"? Unfortunately, as pointed out above, the variance of $\widehat{\rho}$ depends on many unknown quantities, including delay variabilities on the unshared paths, and so it is not possible to analytically answer this question in practice.

However, we can employ a data-based measure of confidence. This is accomplished using the following estimator for the standard deviation of $\widehat{\rho}$:

$$\widehat{\sigma} \equiv \left(\frac{1}{N(N-1)}\sum_{i=1}^{N}\left[\widehat{\rho} - (y_1(i) - \bar{y}_1)(y_2(i) - \bar{y}_2)\right]^2\right)^{1/2}.$$

It can be shown that

$$E[\widehat{\sigma}^2] = E[(\widehat{\rho} - \rho)^2] + O(N^{-2}),$$

which indicates that the estimator of the standard deviation converges to the true standard deviation as the number of probes increases (recall that $E[(\widehat{\rho} - \rho)^2] = O(N^{-1})$). Armed with the standard deviation $\widehat{\sigma}$, one can assess the accuracy of the delay variance estimate $\widehat{\rho}$. For example, if $\widehat{\rho}$ is an order of magnitude larger than $\widehat{\sigma}$, then one can be quite confident in the estimated delay variance. We have not yet implemented this automatic confidence estimator in our experimental work reported here, but plan to incorporate it into the final version of our Network Radar tool.

4. EXPERIMENTAL RESULTS

4.1 Experiment Setup

The experimental validation is carried out in the Wisconsin Advanced Internet Laboratory (WAIL) [16]. The setup consists of 4 Cisco commercial routers (3600 series) and 9 PCs (Redhat Linux). The bandwidth on all connections is $100Mb/s$. The setup is illustrated in Fig. 2. Boxes $0, 1$ and 2 denote the nodes of interests as in Fig. 1. Background (non-probe) traffic is generated using *Harpoon* [17], a flow level traffic generator at boxes denoted by xT. Propagation delays on links are emulated using a special configuration of the Click modular router [18]. During each experiment, background traffic is generated using input distributions derived from `NetFlow` logs captured at the border router at

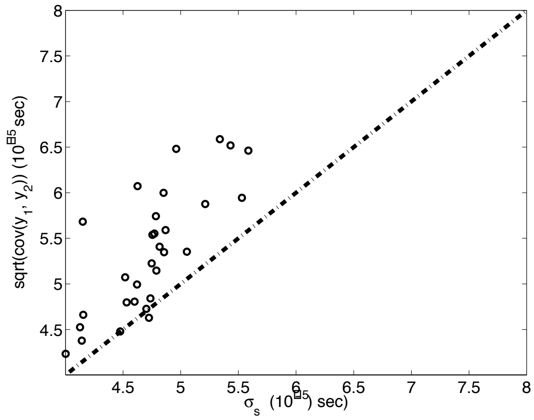

Figure 3: Plot of the square-root of the covariance estimate, $\sqrt{\widehat{\rho}}$, against that of the directly measured delay standard deviation on the shared link.

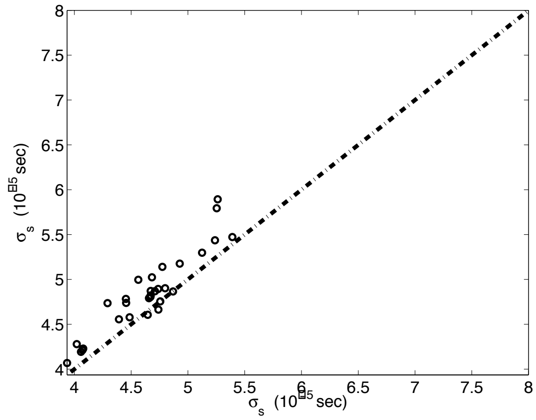

Figure 5: Plot of standard deviation of delay on the shared link measured from sender to the branching node destined to receiver 1 against that to receiver 2.

University of Wisconsin - Madison, while emulated propagation delays on each link are fixed and remain constant.

Each measurement period consist of 1000 packet pairs sent from node 0 (the sender) to receiver nodes 1 and 2. The send rate is fixed at a rate of $10 probes/sec$ ($100ms$ intervals). At the end of each measurement period, we collect `tcpdump` results at the sender (node 0) and at two monitor devices ($S1$ and $S2$) along the path. The monitors, which of course are not pratical outside of the lab, allow us to verify the performance of Network Radar. The monitoring systems take traces of packets traversing the links. The first monitor, $S1$, records the back-to-back packet spacing entering the branching router. The second monitor, $S2$, records outgoing packets from the branching router 2 with extra cross traffic.

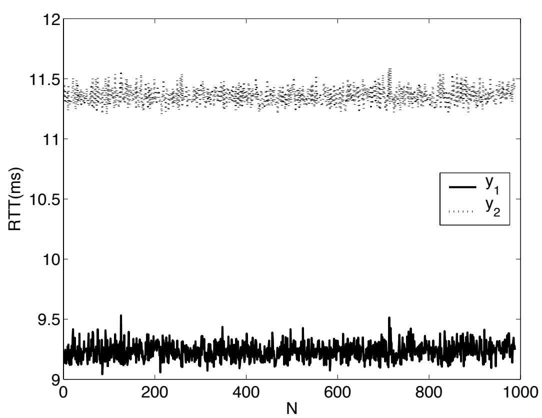

Figure 4: An example of RTT, $\boldsymbol{y}_i, i = \{1, 2\}$ measured at the sender in the testbed.

Figure 3 depicts the square-root of the estimated delay covariance, $\sqrt{\widehat{\rho}}$, against that of the directly measured delay standard deviation on the shared link over 30 measurements. The value $\sqrt{\widehat{\rho}}$ is computed from the RTT measurements of the time difference between the timestamp of TCP SYN and TCP SYN-ACK segments at the sender. An example of RTT measurements in our environment is shown in Fig. 4. Our analysis does not consider packet pairs in which one or both packets are retransmitted or dropped along the forward or return paths. We also ignore packets whose measured round trip time is larger than twice the median RTT on each path. Round trip time greater than twice the median are most likely due to artifacts in the experimental environment such as errors in the time-stamping mechanism. The "true" value for the one way delay on the shared path is the measured time difference of TCP SYN packets at the sender and at the second monitor $S2$. Ideally, these two quantities should be identical and fall onto the 45° line. In practice, however, our estimator slightly over-estimates the true value (over the 45° line). This may result because of deviations from the ideal assumptions of our theory, such as packet pairs that are not perfectly back-to-back and errors in the time stamping mechanism. Nonetheless, the estimator certainly appears to be predictive of the true delay variation, and future work will be aimed at quantifying and improving its accuracy.

The validity of the back–to–back assumption is examined in Fig. 5. If the packets are perfectly back–to–back, then the delay variance, σ_s^2 measured from packets to receivers 1 and 2 should be the same. The offset from the 45° line indicates that packets are not perfectly back–to–back. This can arise from the spacing induced by cross traffic as well as from the discrepencies in the time stamping mechanism. The time stamping mechanism in `tcpdump` as well as those in the devices are known to be imprecise. We further discuss the time stamping issues in the next section.

Finally, we note that the range of the delay variances in our experiments agrees with theoretical predictions. The router queue can be modeled as a M/M/1/K queue. We vary the network load by varying the traffic generator. The delay variance is bounded between 10^{-9} and 10^{-8}. This agrees with our configuration that queue size of Cisco router 3600 series is 40 packets. The maximum delay is in the order of 10^{-6} when the queue is full and the delay variance for a packet if the queue length is one is in the order of 10^{-8}.

5. DISCUSSION AND CONCLUSIONS

Using the WAIL infrastructure, we showed that *Network Radar* is a promising tool for network monitoring. The tool does not require cooperation at the end hosts. We address the challenges associated with RTT measurement. The added variance due to potential extra delays incurred in the response generation times is on the order of 10^{-12}, which is negligible compared to the maximum theoretical variance of the delay variances ($\approx 10^{-6}$) as well as the the variances observed in our experiments ($\approx 10^{-9}$). Our approach assumes that the segment of the return paths will not be shared by the response packets from the receivers. In our scenario, even if the response packets share the portion of the return path, the difference in the RTTs of the packets will typically space them well apart on the return path, and thus they will not incur additional correlation on the return to the sender.

We are aware of the fact that we did not investigate all the possible errors that could affect the effectiveness of the tool. The sender and the receivers have modest CPU load in our experiments. In practice, excessive loads could cause additional delays. The timestamp accuracy in `tcpdump` on a Redhat linux operating system has $1\mu s$ accuracy. It should be enough as variance is in the order of $10^{-9}\mathrm{s}^2$. However, the random effects in timestamping may be a problem. Moreover, `tcpdump` is system dependent and we have only studied under one operating system. The number of probes as well as the probing rate are important elements in delay variances estimation. If we increase the number of probes, we can increase the accuracy of the estimator. However, the probing rate should not be so excessive that it interferes with the normal traffic. The probing period should be larger than the round trip time, so that the packets are approximately independent across pairs.

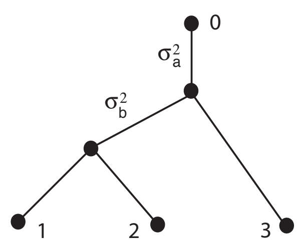

Figure 6: An example of localizing delay variances.

In conclusion *Network Radar* is a tool which will enable network tomography to become much more widely used. In particular, it could be used in case studies of Internet topology to annotate graphs with link specific information. It could also be used as a diagnostic tool by network administrators to isolate and evaluate individual areas of their own network and beyond. In a larger network, a simple extension using our approach to localize link delay variance is illustrated in Fig. 6. We can localize the delay variance σ_b^2 by computing $\mathrm{cov}(\boldsymbol{y}_1, \boldsymbol{y}_2) - \mathrm{cov}(\boldsymbol{y}_1, \boldsymbol{y}_3)$. In this paper, we illustrated the idea of using round trip time measurements to estimate performance (specifically delay variance) on the shared link. However, it can be extended for other type of tomographic studies. In cases where the topology is not known, the delay variances estimated can also be used to infer the topology.

6. REFERENCES

[1] Y. Tsang, M. Coates, and R. Nowak, "Network delay tomography," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2125–2136, August 2003.

[2] N. Duffield, "Simple network performance tomography," in *ACM SIGCOMM IMC*, Miami Beach, FL, Oct. 2003.

[3] K. Anagnostakis, M. Greenwald, and R. Ryger, "cing: Measuring network-internal delays using only existing infrastructure," in *Proceedings of IEEE Infocom*, San Francisco, CA, April 2003.

[4] V. N. Padmanabhan, L. Qiu, and H. Wang, "Server-based inference of Internet link lossiness," in *Proceedings of IEEE Infocom*, San Francisco, CA, April 2003.

[5] M. Coates, A. Hero, R. Nowak, and B. Yu, "Internet tomography," *IEEE Signal Processing Magazine*, May 2002.

[6] A. Adams, T. Bu, R. Cáceres, N. Duffield, T. Friedman, J. Horowitz, F. Lo Presti, S.B. Moon, V. Paxson, and D. Towsley, "The use of end-to-end multicast measurements for characterizing Internet network behavior," *IEEE Communications Magazine*, May 2000.

[7] N. Duffield and F. Lo Presti, "Multicast inference of packet delay variance at interior network links," in *Proceedings of IEEE INFOCOM 2000*, Tel Aviv, Israel, Mar. 2000.

[8] A. Bestavros K. Harfoush and J. Byers, "Robust identification of shared losses using end-to-end unicast probes," in *Proc. IEEE Int. Conf. Network Protocols*, Osaka, Japan, Nov. 2000.

[9] F. Lo Presti, N.G. Duffield, J. Horowitz, and D. Towsley, "Multicast-based inference of network-internal delay distributions," Tech. Rep., University of Massachusetts, 1999.

[10] K. Lai and M. Baker, "Measuring link bandwidths using a deterministic model of packet delay," in *Proc. ACM SIGCOMM 2000*, Stockholm, Sweden, Aug. 2000.

[11] V. Jacobson, "pathchar," 1997, ftp://ftp.ee.lbl.gov/pathchar/msri-talk.ps.gz.

[12] M.Mahajan, N. Spring, D. Wetherall, and T. Anderson, "User-level internet path diagnosis," in *Proceedings of 19th ACM Symposium on Operating Systems Principles*, Lake George, NY, Oct. 2003.

[13] S. Savage, "Sting: a tcp-based network measurement tool," in *Proceedings of USENIX Symposium on Internet Technologies and Systems*, Boulder, CO, Oct 1999.

[14] "Synack," 2004, http://www-iepm.slac.stanford.edu/tools/synack/.

[15] Lawrence Berkeley Laboratories, "tcpdump," http://www.tcpdump.org.

[16] P. Barford and L. Landweber, "Bench-style network research in an Internet instance laboratory," *Computer Communications Review*, vol. 33(3), July 2003.

[17] H. Kim J. Sommers and P. Barford, "Harpoon: A flow-level traffic generator for router and network tests," in *Proceedings of ACM SIGMETRICS '04*, New York, NY, June 2004.

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