NORTHWESTERN
UNIVERSITY

Computer Science Department

Technical Report
NWU-CS-04-48
November 8th, 2004

Can we trust ICMP-based measurements?

Stefan Birrer, Fabián E. Bustamante and Yan Chen

Abstract

ICMP-based measurements (e.g. ping) are often criticized as un-representative of the applications' experienced performance, as applications are based on TCP/UDP protocols and there is a well-accepted conjecture that routers are often configured to treat ICMP differently from TCP and UDP.
However, to the best of our knowledge, this assumption has not been validated. With this in mind, we conducted extensive Internet end-to-end path measurements of these three protocols, spanning over 90 sites (from both commercial and academic networks), over 6, 000 paths and more than 28 million probes in PlanetLab during two weeks.
Our results show that ICMP performance is a good estimator for TCP/UDP performance for the majority of the paths. However for nearly 0.5% of the paths, we found persistent RTT differences between UDP and ICMP greater than 50%, while for TCP the difference exceeds 10% for 0.27% of the paths. Thus, although ICMP-based measurements can be trusted as predictors of TCP/UDP performance, distributed systems and network researchers should be aware of some scenarios where these measurements will be heavily misleading; this paper also provides some hints that can help in identifying those situations.

Keywords: ICMP, TCP, UDP, routers, protocols, trace, end-to-end measurements, PlanetLab.

Can we trust ICMP-based measurements?

Stefan Birrer Fabián E. Bustamante
Yan Chen
Department of Computer Science
Northwestern University, Evanston IL 60201, USA,
{sbirrer,fabianb,ychen}@cs.northwestern.edu

November 8th, 2004

Abstract

ICMP-based measurements (e.g. ping) are often criticized as un-representative of the applications' experienced performance, as applications are based on TCP/UDP protocols and there is a well-accepted conjecture that routers are often configured to treat ICMP differently from TCP and UDP.

However, to the best of our knowledge, this assumption has not been validated. With this in mind, we conducted extensive Internet end-to-end path measurements of these three protocols, spanning over 90 sites (from both commercial and academic networks), over $6,000$ *paths and more than* 28 *million probes in PlanetLab during two weeks.*

Our results show that ICMP performance is a good estimator for TCP/UDP performance for the majority of the paths. However for nearly 0.5% *of the paths, we found persistent RTT differences between UDP and ICMP greater than* 50%*, while for TCP the difference exceeds* 10% *for* 0.27% *of the paths. Thus, although ICMP-based measurements can be trusted as predictors of TCP/UDP performance, distributed systems and network researchers should be aware of some scenarios where these measurements will be heavily misleading; this paper also provides some hints that can help in identifying those situations.*

1 Introduction

Measuring the behavior of network path characteristics is critical for the diagnosis, optimization and development of distributed services. Useful tools of this sort find

application in a variety of contexts, from server selection [3] to the weighting of alternative paths in overlay networks [4]. Unfortunately, performance measurement was not a design goal when the Internet was originally architectured [6] and thus there is limited support available to the system designer.

Over the last few years a renewed interest on measurement techniques [16, 18, 17, 8] have pushed functionality beyond the useful, but rather limited, set offered by tools such as *ping* and *traceroute* [9]. Today's growing toolset includes ICMP-, TCP- and UDP-based instruments such as pathchar [10], sting[17], iperf [19], pathload [16] as well as ping and traceroute [9],

There is, however, a potential dissonance between the application's experience and the view portrayed by the measurement tool. In particular, ICMP-based measurements have been often criticized as un-representative of application performance, as applications often employ TCP or UDP as their transport protocol and there is a well-accepted conjecture that routers are often configured to treat packets from these different protocols differently [17, 7]. While a quick look at the documentation of some of the most popular routers [1] reveals that routers do indeed support protocol based Quality of Service (QoS) policies [5, 15, 11, 14], our research explores how often network administrators make use of this functionality.

In this paper we investigate the dependence of the network characteristics on the higher-level protocol (ICMP, UDP, TCP). This involves identifying anomalies in the measurements with regard to fairness. We note that for most of the path, ICMP performance is a good estimator of UDP and TCP round trip time. However, the average loss rate for ICMP is higher than for UDP and TCP. The hypothesis that UDP traffic has a persistent round trip time penalty of more than 50% holds for 0.45% of all measured paths. We also found 1.76% of the paths with persistent loss anomalies of more than one packet loss per 100 packets.

2 Related Work

Several comparative studies [16, 18, 13] have evaluated existing measurement tools [9, 10, 17, 19], but no work has addressed the effect of the layer-4 protocols on the measured network characteristic.

RON [1] monitors end-to-end path connecting dedicated routers at the entry points of private networks, and it uses these measurement for reactive routing on an overlay. Their work present a relevant detailed evaluation on loss probability. Goyal et al. [7] argues that ICMP-based probes may not be a good estimator for TCP latency and loss rate, since both protocols do not sample network queues in the same way. Our work is complementary, in that we focus on network-path

[1] From brands such as Cisco, Nortel Networks, Juniper Networks and Netgear.

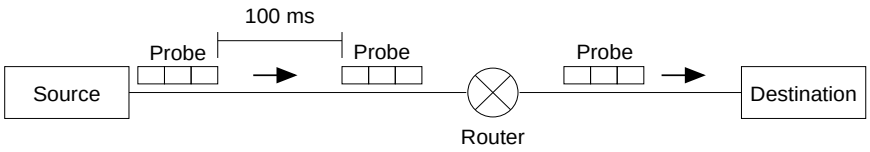

Figure 1: Measurement Methodology: The source sends a probe to the destination and waits 100 ms before sending the next probe. A probe consists of three packets, one of each protocol (ICMP, UDP, TCP), in a random order without spacing.

behavior as experienced by different protocols. Zhang and Duffield [20] look at the over time constancy of Internet path properties and report a loss rate of 0.6-0.9% for TCP (consistent with our findings). Our work, on the other hand, focuses on exploring the constancy of anomalies across protocols.

We borrow the concepts of Global Research and Education Network from Banerjee et al. [2], where the authors look at the interdomain connectivity of PlanetLab nodes. We plan to validate our site classification with theirs (once this becomes available) as part of our future work.

3 Evaluation

3.1 Measurement Methodology

We deployed a ping client/server to PlanetLab, a wide-area test environment. Our measurement client uses a IP-socket and assembles its own ICMP, UDP and TCP packets without using any of the TCP features such as retransmissions. Basically we are comparing the network behavior for IP packets with a different protocol type and a different payload (which conforms to the appropriate standard, i.e. TCP).

Figure 1 illustrates our method of path probing: the client sends 100 probes to the server with 100 ms spacing between them. A probe consists of three packets, one for each of the protocols studied (ICMP, UDP, TCP). These packets are interleaved in random order with no spacing in between. Packets may get lost anywhere in the path from the source to the destination, we account these losses at the destination. Lost rate is then computed as the ratio of packets received to the total number of packets sent.

Every packet forwarded by our client includes a timestamp used to compute an estimate of round trip time (RTT). The server replies immediately to every received packet, including in its reply the client's timestamp. An estimate of RTT is then computed by the client as the difference between current time and the packet's original timestamp.

To remove any possible bias due to packet size on loss probability and queueing delay, we ensure all probes are exactly 100-Byte long (plus IP header). Since the protocols headers are of different size, we pad the packet to 100 Bytes IP-payload.

3.2 Outliers

Outliers are a general problem in real-world measurements. The unpredictable nature of the test environment introduces measurements which lay beyond reasonable boundaries. In a first step, we eliminate RTT outliers for each of our RTT path measurements. We define *outliers* as round trip times which differ more than two standard deviations from the mean of all round trip times for a given protocol. As a few outliers have a strong influence on the mean, we transfer the RTT first into log-space, before proceeding with outlier elimination.

$$
\begin{aligned}
RTT &= \{t_0, t_1, \ldots, t_n\} && (1)\\
RTT' &= \{\log t_0, \log t_1, \ldots, \log t_n\} && (2)\\
t'_x &= \log t_x && (3)\\
mean' &= \overline{RTT'} && (4)\\
std' &= std(RTT') && (5)\\
outlier &= \{t'_x | t'_x > mean' + 2 \cdot std' \\
&\quad \vee t'_x < mean' - 2 \cdot std'\} && (6)
\end{aligned}
$$

If a probe (a triple of packets) is flagged as an outlier for any of the three protocols, the whole probe is discarded in an effort to eliminated possible biases. In addition, all probes which miss one or more of the three packets, are also marked as outliers.

After outlier elimination, the mean RTT and its standard deviation are calculated based on the remaining probes. In the remaining of this paper we use *path measurement* for a given *path*, to indicate the RTT mean of a given connection between two end-nodes across the Internet.

We differentiate between *good* and *bad* path measurements: path measurements with over 50 outliers are marked as bad and not accounted for further analysis. In general, less than 10% of the probes are outliers and thus only a few of our path needs to be eliminated.

3.3 Hypothesis

We employ traditional hypothesis testing techniques from statistics [12]. We validate the well-accepted conjecture that packets from different protocols, due in part

to router configurations, experience different QoS. We use hypothesis testing to estimate the number of path for which we can conclude, with a reasonable confidence, that they hypothesis is true. We employ 95% confidence for all tests.

3.4 Firewalls and other Hurdles

In an early deployment we found that many probes miss round trip times for TCP and UDP connections. Despite PlanetLab deployment policy (stating that firewalls should not filter traffic to and from PlanetLab nodes), we believe our problem were indeed caused by firewalls and port-filtering configurations.

For TCP, we believe the problem was caused by firewalls that only let established TCP traffic pass through. Since we were sending a plain TCP packet without connection establishment (no SYN was sent), these firewalls would just drop our TCP probes. We got around this issue, by setting the SYN bit of all TCP requests and the SYN-ACK for TCP replays. We also found that many PlanetLab sites filter certain ports; in one early configuration we employed port 4000 for all our sockets; this results on UDP packets being dropped at a number of sites.

3.5 Design and Implementation

The ping client/server is implemented in C using the PlanetLab raw socket interface. The code is based on a version of *ping*, dated back to 1983, which we extended to support the three protocols. To coordinate the experiment, we deployed a measurement infrastructure to all the sites, represented by one node each. This infrastructure service is implemented in Java and has the responsibility of coordinating the path probes. No two paths with overlapping sender and receiver sets are probed at the same time, i.e. every node can only be either a source or a destination for one experiment at any given time.

The service will wait for a duration of time, drawn from an exponential distribution with mean of 10 sec. before it tries to lock the destination. If the destination is not currently involved in another measurement, the lock will succeed and each of the peers will start the ping client or server depending on their role in the measurement. The probing lasts for about 10 sec. The client (source) will query the server (destination) for the loss rate and store this information locally with the round trip times. In the case where the lock fails, the client will back-off and restart the procedure by first waiting for some time. When all paths are probed, the server will collect all the traces from all sites and export them into a file for off-line analysis.

We validate our finding using a modified *traceroute* client. This tool works in a way similar to that of our ping client, but it manipulates TTLs so that the routers generate a *time exceeded* message. Following this approach, our traceroute-based

Table 1: Traces Summary

Time	05/02/04 - 05/14/04	
Data Sets		23
Packets		28'741'800
Unique Paths	Total	6'197
Path Measurements	Total	95'806
	Good	69'534
	Bad	22'713
Sites	Total	92
	North America	64
	International	28
	Europe	22
	GREN	71
	Commercial	4

tool allows us to estimate the relative one way delay of TCP and UDP compared to ICMP.

4 Experimental Results

In this section we present our findings based on more than 28 million packets. After outlier elimination we are left with over 20 million packets that we employ for our analysis. Table 1 summarizes the traces we used for the analysis. Note that the European sites are also part of the international sites. The Global Research and Education Network (GREN) combines the academic sites in North America and Europe.

4.1 Connectivity

We eliminated bad path probes by applying the above described outlier elimination technique. However, the majority of bad path probes was caused by a complete outage of one of the four protocols, i.e. only one of the protocols has a loss rate of 100%. Table 2 summarizes bad paths measurements, it is possible for a bad paths to miss all probes for one or more of the protocols, we call this an *outage*. Some of the bad path are caused by infrastructure problems, as PlanetLab nodes may crash or reboot during the probing interval of about 2 hours. TCP outages dominate bad paths, indicating the deployment of firewalls in the PlanetLab test-bed. The negative impact of TCP cannot simply be explained by our measurement technique. Even thought we send duplicate SYN and SYN-ACK packets, the first of these packets is generally expected to pass. We also see that the number of UDP outages is larger than for ICMP, indicating that UDP connectivity is slightly worse than that for ICMP.

Table 2: Bad Paths: The table summarizes the bad paths by outages per protocol. Complete outage means that no packet of the given type is received at the destination.

	Bad Paths	Percentage
ICMP	1'765	0.07
UDP	3'694	0.14
TCP	22'713	0.86
Total	26'272	1.00

Table 3: Percentage of Paths with Persistent RTT Penalties

Protocol	Penalty			
	>5%	>10%	>30%	>50%
UDP	2.07	1.71	0.82	0.45
TCP	0.81	0.27	0.06	0.03
UDP w/o *msu.su*	0.50	0.18	0.02	0.00
TCP w/o *msu.su*	0.83	0.28	0.07	0.03

4.2 Round Trip Time

In our analyzes of round trip time (RTT), we found one site *msu.su* which exhibited persistent anomalies in terms of UDP RTT. Table 3 summarizes the persistent anomalies with 95% confidence. As it can be seen from this table, *msu.su* is responsible for most of the UDP anomalies with over 5% penalty.

Table 4 summarizes the geographical character of the persistent anomalies. International and European sites suffer considerable more UDP anomalies. This is due to the influence of *msu.su*, which causes most of the UDP anomalies.

Beside geographical, the traces also show domain dependency. We classify nodes as part of the GREN or as commercial sites. We left out some of the international sites whose network was not clearly identified. Table 5 summarizes the domain dependency. The GREN suffers from only a few anomalies. Even though we have just a few commercial sites, they account for a significant percentage of

Table 4: Geographical Characterization of RTT Anomalies: The characterization of the anomalies by geographical regions: World (*), North America (NA), International (INTL) and Europe (EU). The values specify the number of anomalies (UDP/TCP) with more than 10% penalty.

Source	Destination			
	*	NA	INTL	EU
*	1.71/0.27	0.56/0.31	4.03/0.19	5.02/0.18
NA	1.52/0.29	0.32/0.43	3.87/0.00	4.87/0.00
INTL	2.11/0.25	1.04/0.07	4.38/0.63	5.34/0.59
EU	2.35/0.25	1.28/0.09	4.62/0.58	5.64/0.74

Table 5: Domain Characterization of RTT Anomalies: The characterization of the anomalies by domain: World (*), GREN (GREN) and Commercial (COM). The values specify the number of anomalies (UDP/TCP) with more than 10% penalty.

Source	Destination		
	*	GREN	COM
*	1.71/0.27	0.50/0.19	1.18/0.79
GREN	1.32/0.21	0.05/0.11	1.08/0.00
COM	2.32/1.16	1.05/0.52	8.33/16.7

Table 6: Temporal Characterization of RTT Anomalies: The characterization of the anomalies by time (CDT/GMT-5h). The values specify the number of anomalies with 10% penalty.

Protocol	Time		
	night	day	evening
UDP	0.00	0.23	0.66
TCP	0.18	0.32	0.97

the identified anomalies, suggesting that the commercial Internet may suffer from a significant amount of anomalies.

Beside domain and geographical dependencies, the time of day influences the characteristic of the network. We use three time intervals: night (0am-8am), day (8am-17pm) and evening (17pm-12am). To give more weight to the time of day pattern, we reduced the set of sites to North America. This analysis is based on 2259 paths for the night-trace, 2170 paths for the day-trace and 1956 paths for the evening-trace. Table 6 shows the anomalies present during the different time intervals. The number of persistent anomalies is substantially reduced during night hours. During this time there is considerable less traffic in the Internet, buffers along the routes are generally less filled, and thus probes are more likely to take the same time independent of their payload. During the day, routers experiencing congestion may prioritize different IP packet payloads differently, thus delaying some of the packet types more than others.

Table 7: Percentage of Path with Persistent Loss Probability Differences

Protocol	Absolute Penalty			
	$<$-1.0%	$<$-0.5%	$>$0.5%	$>$1.0%
UDP	1.76	1.90	0.68	0.29
TCP	1.74	1.92	0.10	0.02

Table 8: Geographical Characterization of Loss Anomalies: The characterization of the anomalies by geographical regions: World (*), North America (NA), International (INTL) and Europe (EU). The values specify the number of anomalies (UDP/TCP) with <-1.0% difference.

Source	Destination			
	*	NA	INTL	EU
*	1.76/1.74	2.13/2.10	1.07/1.02	1.04/0.98
NA	0.81/0.69	1.18/1.04	0.07/0.00	0.09/0.00
INTL	3.82/3.96	4.08/4.30	3.29/3.29	3.16/3.16
EU	4.70/4.76	5.02/5.11	4.04/4.04	3.92/3.92

Table 9: Domain Characterization of Loss Anomalies: The characterization of the anomalies by domain: World (*), GREN (GREN) and Commercial (COM). The values specify the number of anomalies (UDP/TCP) with <-1.0% difference.

Source	Destination		
	*	GREN	COM
*	1.76/1.74	1.27/1.32	1.58/1.58
GREN	2.14/2.02	1.63/1.57	2.15/2.15
COM	0.77/0.77	0.00/0.52	0.00/0.00

4.3 Loss Probability

The loss probability is an important factor for TCP performance. Table 7 shows that a few path experience a persistent lower or higher loss probability for UDP or TCP. These paths have more likely a lower loss probability for UDP and TCP when compared to ICMP. The average loss rate of all measurements for ICMP is 0.93%, while UDP and TCP have a nearly equal loss rate of 0.65%.

Table 8 presents the geographic distribution of the persistent anomalies in loss rate. The results indicate that loss anomalies are actually concentrated within Europe.

The domain characterization indicates that commercial network experiences less anomalies than the academic network. Table 9 summarizes the findings. The loss probabilities are nearly symmetric in terms of the protocol, but asymmetric in terms of the source and destination. Note that we estimate the one way loss probability, so that the measurements indeed picture the asymmetric behavior of the network.

Table 10 shows that the persistent loss anomalies are concentrated during the night and day hours. During the evening, only a few paths experience an decrease in the loss rate for UDP or TCP. However, during night and day hours, substantial more anomalies appear. This might be caused by different resource usage and/or different kind of workloads during the evening hours.

Table 10: Temporal Characterization of Loss Anomalies: The characterization of the anomalies by time (CDT/GMT-5h). The values specify the number of anomalies with an absolute decrease in loss rate of 1.0%, that is one loss in a hundred packets less.

Protocol	Time		
	night	day	evening
UDP	0.53	0.41	0.20
TCP	0.58	0.65	0.40

4.4 Validation

We used a simple TTL-based traceroute client[2] to validate some of the identified anomalies. We successfully validated that either the router *cs.radio-msu.net* or the router *NPI-4700-F0-2.radio-msu.net* must employ different QoS for UDP packets by tracing the path from *northwestern.edu* and *fh-aargau.ch*[3] to *planetlab2.cs.msu.su*. These two routers represent the first two hops of the site with the most anomalies (*msu.su*). Since we only probe the incoming path to the site, we cannot conclude whether it is the outgoing interface of *NPI-4700-F0-2.radio-msu.net* or the incoming interface *cs.radio-msu.net*, which causes the extra delays of UDP packets.

After we ruled out *msu.su* from the traces, we still have about 10 path anomalies left. It is part of our future work to validate these anomalies.

5 Conclusions

ICMP-based measurements are used to estimate TCP and UDP performance. However, there is the possibility of dissonance between the application's performance and the view portrayed by the measurement tool. This paper addressed the question of whether ICMP-based measurements can be trusted. Our measurement-based analysis indicate that for the majority of the paths, ICMP is a good performance indicator for UDP and TCP RTT. However, over 1.7% of the paths experience UDP RTT penalties larger than 10%, while 0.27% of the paths suffer similar penalties for TCP. Further, ICMP has a much higher loss rate than UDP and TCP. The results indicated that there are significant geographical and temporal differences.

Our results seems to argue in favor of ICMP-based measurements as predictors of TCP and UDP performance, of course taking into consideration protocol differences. However, this estimation has some inherited limitations, as the network can

[2]Due to technical limitations, we can only trace from nodes outside of PlanetLab.

[3]FH Aargau is a university in Switzerland (Central Europe).

and sometimes does treat the three protocols differently.

Acknowledgments

We would like to thank Hans-Peter Oser, who kindly loaned us his equipment for some of our validation experiments. We are also grateful to Tamara Teslovich and Kate Solinger for their assistance in evaluating the router documentations and their help in implementing our validation tool. Further, we would like to thank Yi Qiao and Ananth Sundararaj for their helpful comments on early drafts of this paper.

References

[1] ANDERSEN, D., BALAKRISHNAN, H., KAASHOEK, F., AND MORRIS, R. Resilient overlay networks. In *Proc. of the 18th ACM SOSP* (October 2001).

[2] BANERJEE, S., GRIFFIN, T. G., AND PIAS, M. The interdomain connectivity of planetlab nodes. In *Proc. of PAM* (April 2004).

[3] CARTER, R. L., AND CROVELLA, M. E. Server selection using dynamic path characterization in wide-area networks. In *Proc. of IEEE INFOCOM* (April 1997).

[4] CHU, Y.-H., RAO, S. G., AND ZHANG, H. A case for end system multicast. In *Proc. of ACM SIGMETRICS* (June 2000).

[5] CISCO SYSTEMS. (tcp) and (udp) ports used by cisco unity version 4.0. www.cisco.com, March 2004. White Paper.

[6] CLARK, D. D. The design philosophy of the DARPA internet protocols. In *Proc. of ACM SIGCOMM* (September 1988).

[7] GOYAL, M., GUERIN, R., AND RAJAN, R. Predicting TCP throughput from non-invasive network sampling. In *Proc. of IEEE INFOCOM* (June 2002).

[8] GUMMADI, K. P., SAROIU, S., AND GRIBBLE, S. D. King: Estimating latency between arbitrary internet end hosts. In *ACM SIGCOMM Internet Measurment Workshop* (November 2002).

[9] JACOBSON, V. Traceroute: A tool to trace the path of packets in the Internet. ftp://ftp.ee.lbl.gov/traceroute.tar.gz, 1989.

[10] JACOBSON, V. Pathchar: A tool to infer characteristics of Internet paths. ftp://ftp.ee.lbl.gov/pathchar, 1997. A tool to analyze bandwidth, delay and loss rate of every hop between two end hosts.

[11] JUNIPER NETWORKS. Filter-based forwarding. www.juniper.net, 2001. Technology Note.

[12] LARSEN, R. J., AND MARX, M. L. *An Introduction to Mathematicla Statistics and Its Application*, 3rd ed. Prentice-Hall, Upper Sadler River, NJ, 2001.

[13] MURRAY, M. Performance measurement taxonomy. www.caida.org/tools/taxonomy/performance.xml, 2004. This page concerns tools for measuring Internet performance.

[14] NETGEAR. DiffServ. kbserver.netgear.com, May 2004. DiffServ is a method for adding quality of sercie (QoS) using Layer 3 information.

[15] NORTEL NETWORKS. Introduction to quality of service (QoS). www.nortelnetworks.com, 2003. White Paper.

[16] PRASAD, R. S., DOVROLIS, C., MURRAY, M., AND CLAFFY, K. C. Bandwidth estimation: Metrics, measurement techniques, and tools. *IEEE Network 17*, 6 (2003).

[17] SAVAGE, S. Sting: a tcp-based network measurement tool. In *Proc. of USENIX USITS* (1999).

[18] STRAUSS, J., KATABI, D., AND KAASHOEK, F. A measurement study of available bandwidth estimation tools. In *ACM SIGCOMM Internet Measurment Workshop* (November 2003).

[19] TIRMUALA, A., QIN, F., DUGAN, J., FERGUSON, J., AND GIBBS, K. Iperf: The TCP/UDP bandwidth measurment tool. http://dast.nlanr.net/Projects/Iperf/, 2003. A tool to measure maximum TCP bandwidth.

[20] ZHANG, Y., AND DUFFIELD, N. On the constancy of Internet path properties. In *ACM SIGCOMM Internet Measurement Workshop* (August 2001).