Verification of Timeouts: Beware Surprisingly High Delay

Linda Baker
lsbaker@wpi.edu

Sanket Gujar
srgujar@wpi.edu

Andrew Gonsalves
agonsalves@wpi.edu

Abstract

Timeouts are necessary for network probes and outage detection. Timeouts that over- or underestimate the necessary RTT of a system can lead to delays or false negatives in outage detection systems when a host is merely connected to a high-latency, congested network. There have been several studies that have been conducted in order to determine appropriate timeout times for protocols and networking tools, such as the paper "Timeouts: Beware Surprisingly High Delay". We have recreated several experiments originally conducted by Padmanabhan et al. in their paper, and performed various experiments to determine what causes latencies. Based these results, the authors suggested and tested various hypothesis. In this paper, we have recreated these experiments using similarly gathered data as well as data collected from the ISI. We then performed some of our own experiments based on their some of theirs in order to validate their findings work and to analyze some potential aspects of timeouts that the authors may not have observed.

1. Introduction

In the paper Timeouts: Beware Surprisingly High Delay [2], Padmanabhan et al. analyzed the Round Trip Times (RTT) associated with high-latency IP addresses in order to find an appropriate timeout duration for networking tools such as outage detection services. They believed that current timeouts used by many of these detection services were underestimates, which would lead to false negative readings when trying to determine if a high-latency destination was offline. In order to test this theory, the authors collected datasets and ran several experiments to find an appropriate timeout value that would minimize the chance of false negatives without drastically overestimating the necessary timeout to match the RTT of a path to the host. Additionally, they used the results of their experiments to search for common factors between these high-latency IP addresses in order to determine what causes high RTTs.

They started with datasets collected in ISI surveys, and manipulated and filtered these datasets in order to isolate the highest-latency IP addresses and their corresponding RTTs, and analyzed the distributions of these RTTs and how they have changed over time.

Following this ISI analysis, Padmanabhan et al. conducted additional experiments to verify that the high latencies observed in their data were not cause by a flaw in ISIs survey methodology. To accomplish this, they used 2 network scanners called ZMap and scamper to collect their own survey data. Their results showed that both ZMap and scamper supported the conclusions drawn from the ISI datasets, and concluded that the probing scheme does produce discrepancies in the data. The authors also hypothesized that the type of message sent to these IP addresses could contribute to high latencies, if certain network protocols are discriminated against while passing through the internet. Their results of this experiment suggested that different protocols did not substantially contribute to high latency.

Next, Padmanabhan et al. formed several hypotheses that could explain why some IP addresses have such high latencies. They proceeded to test these hypotheses using additional experiments and databases. Using the MaxMind database, they tested whether satellite links, geographic locations or Autonomous Systems correlated with high latencies. They found no correlation with satellites, but found that high-latencies were particularly common in South America and Asia, and that most ASes responsible for high-latency IP addresses were cellular networks.

During some of their experiments, they observed that in many scenarios, the RTTs of subsequent pings to an address decreased from the initial response. Padmanabhan et al. ran one experiment to test the prevalence of this effect, as well as to estimate how much this temporary latency could contribute to RTTs. They found the effect to be extremely common, and found that this "wake-up" time was less than 4 seconds in most cases.

Throughout our project, our goal was to verify the conclusions drawn by Padmanabhan et al. by replicating several of their experiments. For our study, we conducted five different experiments. The first was our ISI analysis, in which we attempted to replicate their dataset modifications so that we could similarly use the ISI datasets. Ad-

ditionally, we ran similar ZMap and scamper experiments to those used by Padmanabhan et al. in order to compare the distributions found in their study. After this, we ran our Traceroute experiment, in which we ran Traceroute on several high-latency IP addresses, in search of common links that may suggest a cause of high latencies. Finally, we conducted our First Ping experiment, in which we replicated the First Ping experiment run by Padmanabhan et al. and tested the prevalence of the First Ping Effect that they described.

2. ISI Survey data

Since 2006, ISI [3] has conducted surveys of the internet several times per year. In these surveys, they select 1% of all allocated /24 address blocks. To each of the selected address blocks, they send ICMP echo request probes to each of the 256 addresses in the block every 11 minutes over a two-week duration. Every echo response received within 3 seconds of the sent request was recorded in the dataset with microsecond precision. Echo requests that were unanswered within 3 seconds were instead recorded without an RTT, and marked as unanswered. If the echo response arrived after more than 3 seconds, the message was saved to a separate PCAP file.

Padmanabhan et al. used these ISI surveys as the basis of their research and experiments in timeouts. They matched the unanswered and timed-out responses to the messages saved in the PCAP records, and were able to estimate the RTT of these high-latency responses with a precision of 1 second. After running several filters on the data to remove duplicated and broadcast responses, they used these datasets to select IP addresses to probe in their experiments.

We planned from the beginning of the project to use our own ZMap experimental results to select IP addresses for our experiments, in order to ensure that our resulting list of IP addresses and their associated RTTs were recent and consistent with our experimental vantage point at WPI. However, we also wanted to replicate some of the CDF analysis of the ISI datasets that Padmanabhan et al. did, in order to test the consistency of our results as well as to analyze how data collected in more recent ISI datasets compared to the conclusions and projections they published in 2015.

Using seven ISI datasets, we attempted to replicate their methods so that we could perform similar analysis of the ISI dataset. We developed a tool that could match the unanswered records to the timed-out records stored in the PCAP file, but we were ultimately unable to calculate the correct RTT for these IP addresses, as we had insufficient information to calculate the exact RTT of each record.

3. CDF Verification

In order to verify the distributions of RTTs that they observed in the ISI datasets, Padmanabhan et al. ran several ICMP internet scans using ZMap and scamper. We replicated their methodology to run ZMap and scamper surveys of our own so that we could compare the distributions of our data. In this section, we describe our ZMap and scamper experiments and analyze the results.

3.1. ZMap

ZMap [4] is a fast, single-packet scanner designed for internet-wide network surveys (cite ZMap site). In their experiment, ZMap was configured to perform internet-wide scans by sending one ICMP echo request to every IPv4-allocated IP address and recording the RTT of the response.

Similarly, we used ZMap to run our own scan, which served as a basis of our following experiments. We ran ICMP probing scans using the WPI networks for 24 hours, collecting a total of 26.77 million samples. We calculated the RTT for each of these samples with millisecond precision. Unfortunately, due to limited computational power we were unable to perform analysis on the complete dataset. Instead, we divided the file and analyzed each fourth separately. We observed that each of these 4 sub-samples had a nearly identical distribution. Figure 1 shows the distribution of RTT samples for one quadrant sample of the data. From this data we can see that a majority of pings were captured within a 2-second timeout window.

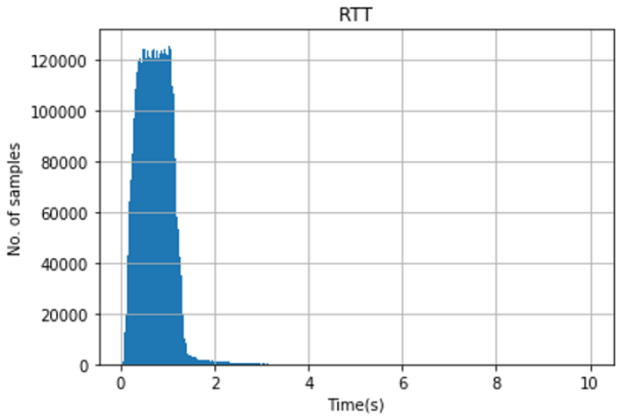


Figure 1: Histogram of ZMap data

Figure 2 shows a CDF of our ZMap data. Our results indicate that 99% of pings can be captured with a 1.75 second timeout value, and 50% pings can be captured in 0.75 secs.

For further analysis, we isolated every IP address that had a latency higher than 10 seconds from our ZMap results, and used an API called ”geoiplookup”, which uses a geo-IP database to match input IP addresses to their approximate coordinates, region, and country.

Using this database, we plotted the location of each of these high-latency IP addresses, as shown in Figure 3. We observed that Europe, Mexico, Brazil and the east-

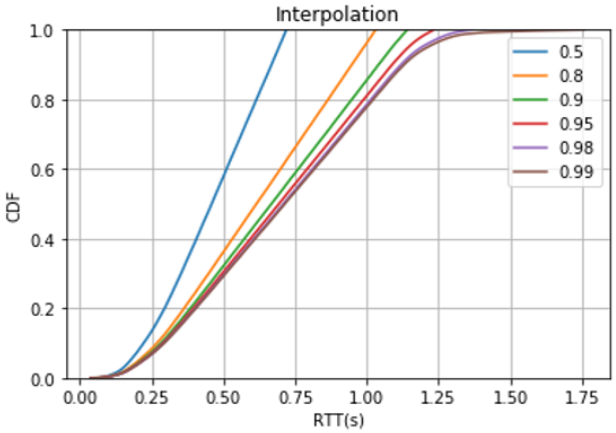


Figure 2: Cumulative Distribution Function Analysis of ZMap data

ern United States showed the highest distribution of high-latency IP addresses.

This seemed to contradict one finding by Padmanabhan et al. They similarly plotted the geographic locations of all IP addresses that had an RTT greater than 1 second from their ZMap scans, and their results suggested that a majority of these high-latency IP addresses were located in South America and Asia. There are several possible reasons why they may have observed more high-latency IP addresses in Asia than our experiments found. We believe this discrepancy was caused by differences in selection criteria. It is possible that many of the Asian IP addresses had latencies between 1 second and 10 seconds, which would include them in their results and not ours. Additionally, our ZMap scans were recorded 3 years after theirs, so it is entirely possible that improvements in infrastructure within the past 3 years have lowered the latencies associated with some of these IP addresses.

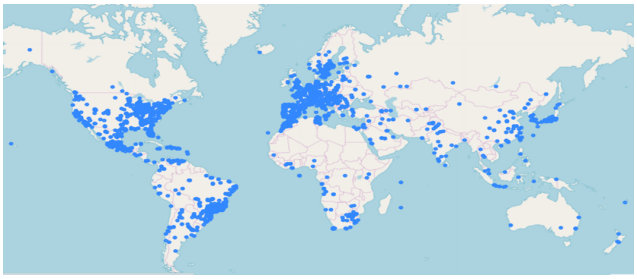


Figure 3: Geograhic locations of IPs having RTT greater than 10 secs

3.2. Scamper

Once our ZMap experiments were completed, our next goal was to verify the distributions observed in ZMap using scamper. Scamper [1] is another network scanner which allows users to send parallelized repeated pings to a set of specified IP addresses. We originally ran this experiment after the ZMap experiment was finished, then we reran the experiment on a larger set of IP addresses after the final presentation, and the results are included in this report.

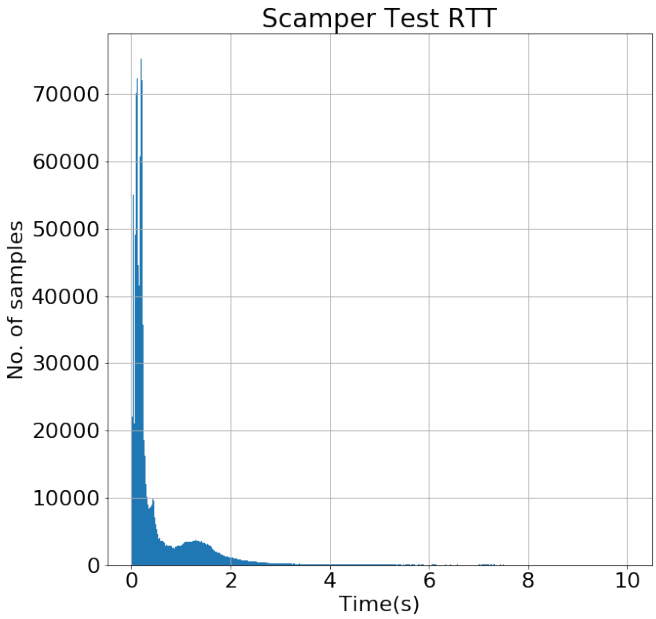
Following the methodology described by Padmanabhan et al., we started the experiment with a list of all IP addresses that had an RTT over 100 seconds in ZMap, and randomly sampled 2000 of them. We pinged each of these IP addresses 1000 times, with a 10 second delay between pings, then analyzed the distribution of their latencies.

Figure 4 shows the histogram of the RTT associated with every ping from our scamper results. As this experiment was focused on IP addresses that showed extremely high latencies, there was a higher proportion of high latencies than were observed in the ZMap experiment. While a majority of the pings were received within 2 seconds, there is still a nontrivial number that arrived after 2 seconds. There were hundreds of responses that had latencies over 100 seconds, although their numbers pale in comparison to the hundreds of thousands of lower-latency responses.

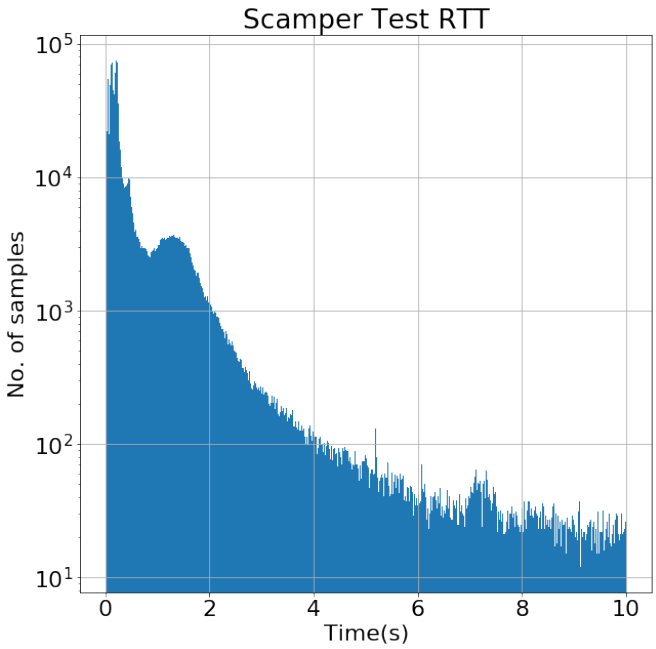
Figure 5 shows a comparison between the scamper CDFs produced by Padmanabhan et al. and us, respectively. Converging near 35 seconds, the 99th percentile of our scamper responses had a higher RTT than the 99th percentile in Figure **??** from our ZMap experiment. However, our scamper CDF has overall lower values than theirs. This discrepancy may have been caused by differences in our selection criteria. They selected IP addresses from ISI records that had at least 5% of RTTs greater than 100 seconds, while we used a less consistent datapoint when selecting IP addresses that had greater than 100 second RTT from our ZMap scan, meaning that our experiment may have been ”polluted” by IP addresses that usually have a low latency, but coincidentally displayed a high latency when our ZMap scan occured. If we were to repeat the ZMap experiment multiple times and isolate the IP addresses that had consistently high latencies, we may have found the RTTs to be generally higher.

4. Traceroute

Diverging from the replication of experiments run by Padmanabhan et al., we designed our own experiment to analyze potential common factors between high-latency IP addresses. We theorized that there may be common links which contribute to high latencies for many IP addresses, more so than the packet's final destination itself. Our original plan was to run Traceroute on a set of high-latency IP addresses and record information about the hop which had the highest latency along the path to the IP address. However, inconsistencies in the structure of the data and DNS naming made it difficult to record the results. Instead, we used a tool called MTR. MTR (My TraceRoute) combines

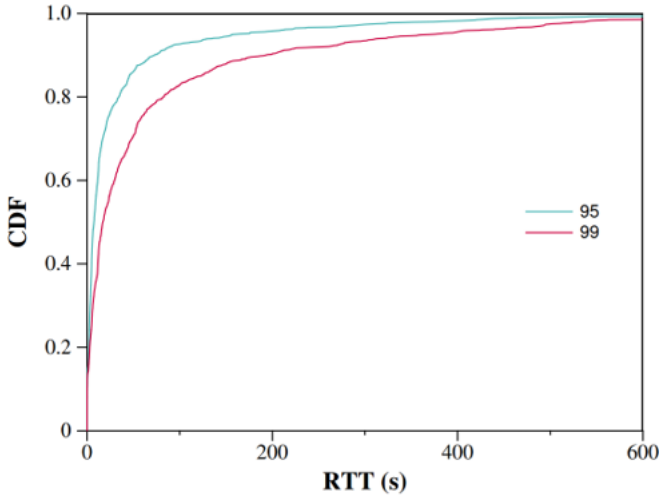


(a) Histogram of scamper data

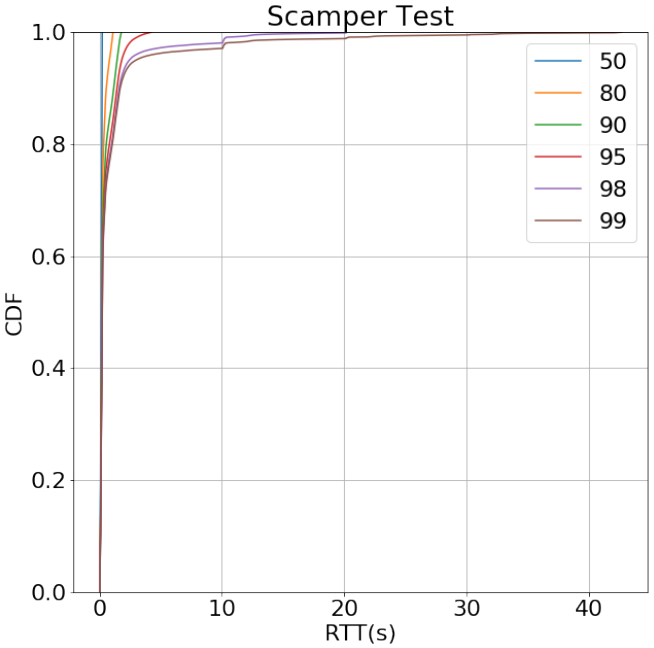


(b) Histogram of scamper data (semi-log scale)

Figure 4: Scamper RTT distribution analysis



(a) CDF of scamper results from Padmanabhan et al.



(b) CDF of scamper our scamper results

Figure 5: Cumulative Distribution Function of scamper data

the functionality of the Linux Traceroute and ping tools, generating a structured and more consistent report by pinging every hop in the link 10 times.

We ran the experiment on the set of IP addresses that had a higher latency than 10 seconds from our ZMap results and found that, as we expected, there were some high-latency hops which were not the final destinations in the route, as shown in Figure 6. The destinations of these IP addresses were recorded in grey, while the locations of the highest-latency hops were recorded in blue.

For further analysis, we plotted a line between each destination and the location of the IP address that had the highest latency in MTR, which can be seen in Figure 7. We observed that many of these paths cross between Europe, the United States of America, and Brazil. We were surprised to see so many links between Brazil and Spain. There are also several seemingly isolated networks, such as the ones in South Africa and China.

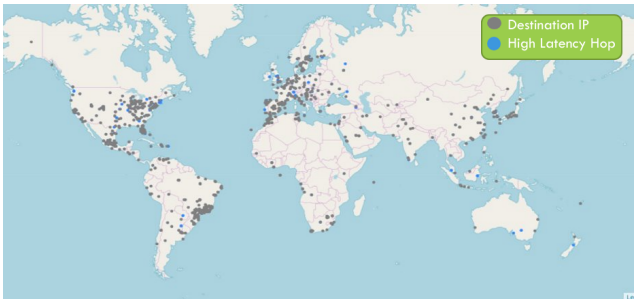


Figure 6: Final destination and high latency hops

Figure 7: Paths from high latency hops to final destination

MTR also reported the AS number of each hop in the route, so we also recorded the Autonomous System associated with the highest-latency hop in each route. We isolated the 10 ASes that were responsible for the most of these high-latency responses, and plotted them in Figure 8. We found that the first and second highest-ranked AS Numbers – which are shown as purple and green in Figure 8 – were owned by Telefnica Brasil and were responsible for a combined 1,184 records out of the total 3,251 Traceroute results. In contrast, the other 8 of these 10 ASes were each responsible for between 75 and 150 records.

Remarably, in their own AS analysis, Padmanabhan et al. similarly isolated the 10 ASes that were responsible for most latencies above 1 second and above 100 seconds, and found that in both cases, Telefnica Brasils ASN was associated with more of these high-latency IP addresses than the rest of the top 10 ASes combined.

A majority of the ASes in Figure 8, such as Chinanet in yellow and SAIX-NET in brown, are mainly localized to their respective nation of origin. There are a few exceptions that cross international boundaries, such as Vodafone Group, in red; Telefonica Wholesale Network, in blue; and Seabone-Net Telecom Italia, in white. These are multinational telecommunication companies based in London, Madrid, and Rome respectively, though they all operate in many other countries. The high latencies observed in these ASes may be a result of the large distances their traffic traveled.

Finally, it is noteworthy that several of these top 10 ASes are cellular networks, as this also supports the findings of Padmanabhan et al.

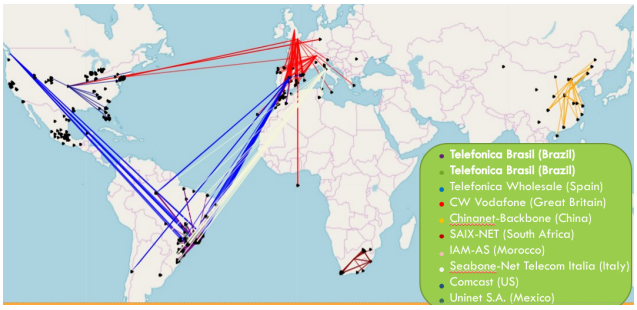


Figure 8: Paths from high latency hops to final destination from top 10 AS

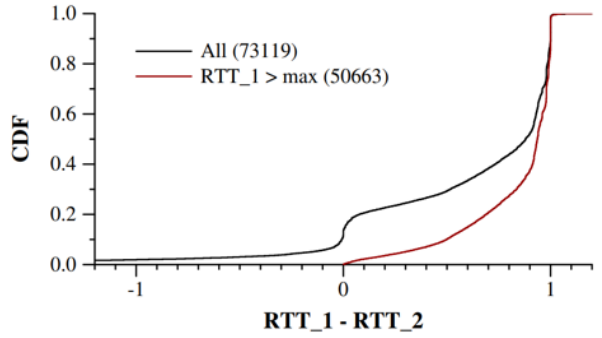
5. First Ping

While analyzing the results of their scamper experiment, Padmanabhan et al. found that the first ICMP echo request sent in a series of ICMP echo requests often had a higher latency than following pings did. They called this the First Ping Effect, and ran an additional experiment to analyze the prevalence of the phenomenon. We followed the experimental procedure provided by Padmanabhan et al. to run our own First Ping experiment to verify their results.

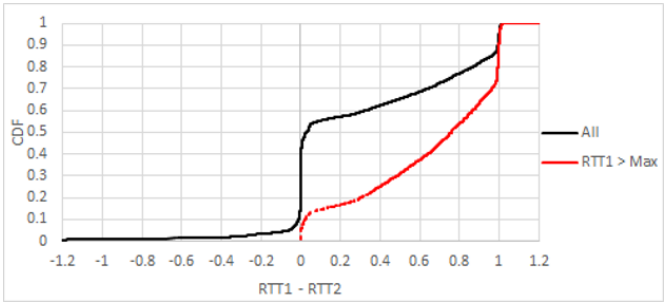
In our experiment, we started with the set of 15945 IP addresses that had an RTT greater than 10 seconds in the ZMap experiment. To filter unresponsive addresses from this list, we pinged each IP address twice with a delay of 1 second and a timeout of 60 seconds. We omitted every IP address that did not respond to either probe, which left 6990 respondents. To each of these addresses, we waited 60 seconds before sending 10 pings, one per second. Once the responses were collected, we filtered out 58 IP addresses that didnt respond to the first probe or at least 4 probes as Padmanabhan et al. did, as these records would skew our calculations.

For 69% of addresses in their experiment, Padmanabhan et al. found that the response to the first of 10 pings had a higher RTT than the maximum of the rest. Furthermore, they found that for 85% of these addresses, the first response had a higher RTT than the average of the rest of the responses. However, our results showed a weaker correlation between the first ping and high latencies. We found that only 49% of our responsive addresses had a first response higher than the maximum of the rest, and 68% percent had a first ping higher than the average of the rest.

Figure 9 (a) shows the difference between the first ping and the second ping in the results found by Padmanabhan et al., and Figure 10 (b) shows a replication using our

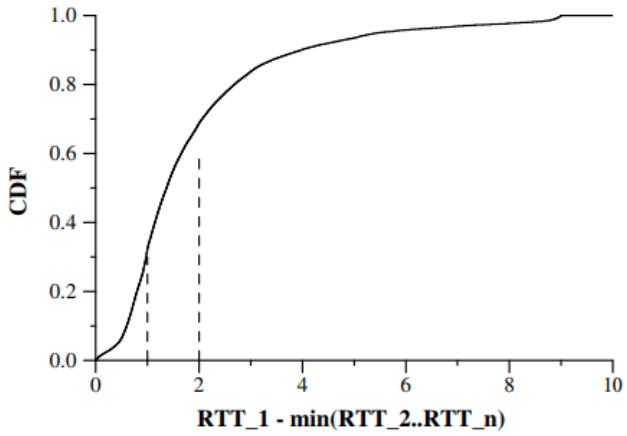


(a) Difference between first probe and second probe latency by Padmanabhan et al.

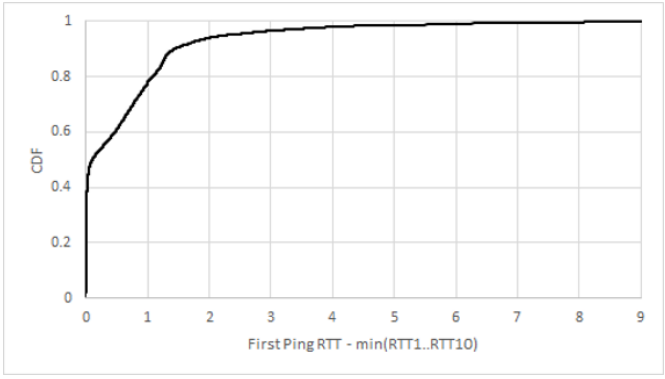


(b) Difference between first probe and second probe latency from our results

Figure 9: Comparison between difference between first probe and second probe latency



(a) Difference between initial and minimum latency by Padmanabhan et al.



(b) Difference between initial and minimum latency from our results

Figure 10: Comparison of Difference between initial and minimum latency

data. Values near zero indicate no significant difference between the first two pings, and values at 1 represent scenarios in which both of the first 2 pings arrived simultaneously. There were a few outliers, which suggested packet reordering when the difference was greater than 1, and the first response was faster than the second in records that were less than 0. The responses near 0 show the greatest difference between Figures (a) and (b). The steep incline at 0 in Figure (a) shows that 40% of IP addresses had their first ping no faster than the second, contrasting Figure (b) which had much fewer records at 0.

Padmanabhan et al. concluded from their results that the First Ping Effect was a result of temporary latency caused by MAC-layer time slot negotiation or device wake-up, and observed that the latency could be estimated by subtracting the RTT of the first ping from the minimum RTT of the remaining pings, which they represented as a CDF in Figure 10 (a). They found that the median of these differences was 1.37 seconds, and that 90% of addresses had a difference below 4 seconds. Figure 10 (b) shows the CDF representing the same analysis run on our data. Again, approximately 40% of all respondent addresses showed no additional wake-up latency. The median difference in our results was a mere 105 ms, and 90% of results were below 1.44 seconds.

Although our results show that the first ping of many pings is prone to having an above average latency, this wake-up latency does not appear as high nor as universally as Padmanabhan et al. suggest. This discrepancy may have been caused by slight differences in our experiments. While we attempted to emulate the experimental methodology as much as we could, our selection strategy differed from theirs. They selected a much larger sample of IP addresses that had an RTT of at least 1 second from one ISI dataset, as opposed to our smaller dataset of addresses with an RTT of at least 10 seconds from our ZMap results. If this is indeed the fundamental difference between the experiments, our combined data could imply that relatively high-latency addresses in the 1-to-10-second range typically have a temporary latency associated with the first ping, while higher-latency addresses have a more persistent and consistent latency that is less likely to decrease over

time.

6. Conclusion

The primary goal of the research conducted by Padmanabhan et al. was to find an appropriate timeout length that would minimize false negative responses from high-latency systems. We attempted to run several transformations on the ISI datasets to follow their example, though we were ultimately unsuccessful; due to a lack of experimentation information about the dataset we were unable to match the timed-out, unmatched responses.

We captured data using ZMap, and after performing our own analysis of the data we found that a timeout value of 1.75 secs can capture 99% of the pings. We were able to extract high latency IP addresses from these ZMap scans, which were used to find the geographic locations of their corresponding hosts. The majority of these high-latency responses came from Europe, Brazil and North America. We found fewer high-latency IP addresses in Asia than the research by Padmanabhan et al. suggested.

Similarly in scamper, we found that a small percentage of echo responses had extreme latencies greater than 100 seconds, but they were much less common than we expected. According to the scamper results that ran on already high latencies, an approximate 35-second timeout should capture 99% of echo responses from these high-latency IP addresses.

The Traceroute experiment confirmed that a hop in the middle of a path can contribute more to high latencies than the destination itself. There were several common trends in our Traceroute AS analysis that support the findings of Padmanabhan et al. Cellular networks and Telefnica Brasil in particular tended to represent many of the high-latency IP addresses that we observed.

Our analysis of the First Ping Experiment suggests that the First Ping in a set often has a higher-than-average RTT, though this difference is not as universal nor as significant as Padmanabhan et al. suggest.

The findings of our Traceroute experiment have the most potential for future development. Potential extensions of this experiment may focus more on the analysis of Autonomous systems in order to find more in-depth common factors in high-latency ASes, or perhaps one could attempt to isolate data centers that represent a disproportionate amout of high-latency traffic.

References

[1] Caida. Scamper. 2018. actively probe destinations in the Internet in parallel, caida.org/tools/measurement/scamper. 3

[2] R. Padmanabhan, P. Owen, A. Schulman, and N. Spring. Timeouts: Beware surprisingly high delay. In *Proceedings of the 2015 ACM Internet Measurement Conference, IMC 2015, Tokyo, Japan, October 28-30, 2015*, pages 303–316, 2015. 1

[3] USC/ISI. Isi internet surveys. 2018. datasets provided by USC/ISI ANT project. , ant.isi.edu/datasets/all.html. 2

[4] Zmap. The zmap project. 2018. Internet Scanner, zmap.io. 2