Project 3

Jonnathen Ravelo, Ayesha Saleem

4/2/2018

Abstract

In this project, we identified 10 targets and using the network diagnostic tools traceroute and ping, explored the routing stability of the internet and monitored the loss and delay performance characteristics of some Internet paths. We collected data over a period of 7 days and analyzed the data to identify the stability of our routes, the nature of routing changes and, the loss rates of our targets.

The most sites that had the most consistent paths over the trial were www.ebay.com, www.github.com, www.reddit.com, www.vmware.com. Unstable targets appeared to randomly change their routing patterns with no discernable pattern. Hence, we found that our hypothesis wasn't backed up by our experiment. Amazon, a target we picked that has a tremendous amount of traffic, is incredibly inconsistent as well as very unstable. Thus, there isn't a correlation between traffic and stability and consistency. There didn't seem to be any nature to the changes, but rather they were random. In addition, the loss rate for our 10 targets was minimal with an average loss rate of 2%.

1. Introduction

This experiment aims to expose us to the interesting quirks in network routing and packet round trip times using two widely used network diagnostic tools, traceroute and ping.

I. *tracert (traceroute)* is a computer network diagnostic tool for displaying the route and measuring transit delays of packets across an Internet Protocol network. In the scope of this project, we are using traceroute to trace routes from our personal machines to a remote address (either a name or IP address). Our motivation is to monitor the routing stability of the internet. To achieve this goal, we chose 10 targets located all around the world. These targets include Github, Alibaba and Reddit and using traceroute, we both monitored and recorded the routing path for a period of 7 days and analysed the stability of these routes. Before beginning the process, we hypothesized that the routes will all be incredibly stable in the sense that the number of hops won't change. We came to this hypothesis as the addresses we're accessing are widely used and popular.

II. **Ping** is a computer network administration software utility used to test the reachability of a host on an Internet Protocol (IP) network. It measures the round-trip time for messages sent from the originating host to a destination computer that are echoed back to the source (IEEE Internet Control Message Protocol"). Our second motivation in this project was to use ping to monitor the loss and delay performance characteristics of some Internet paths. Samples were collected across various locations including school, home and work. Once the data was collected, it was analyzed and each target was categorized based on their loss rates.

We hypothesized that the more popular a target, the more stable the route, and this experiment is being conducted to gauge the worthiness of our claim.

2. Methodology

The project was split into 2 parts as there are two primary motivations.

Part I. Routing

1. We identified 10 targets we're monitoring: www.alibaba.com, www.amazon.com, www.ebay.com, www.facebook.com, www.google.com, www.github.com, www.reddit.com, www.spotify.com, www.store.steampowered.com, www.vmware.com.
2. Using traceroute (from a local machine), We monitored and recorded the routing path to each target for an entire week (7 days).
3. We performed 1-2 measurements every day, for a total of at least 10 measurements per target for the duration of the week.
4. The results of the tracert were then stored in file that correspond to the date and time the measurement was conducted
5. metrics were gathered from results of the data.

Part II - Stability & Reliability

1. We re-used the same set of 10 targets as we chose for Part I (Routing)
2. Utilizing the "ping" utility, we then conducted 1-2 measurements every day for an entire week (7 days).
3. each measurement requests 120 ping messages to each target.
4. The results of the ping were then stored in a file that correspond to the date and time the measurement was taken.
5. Dates, times, average, minimum, and maximum RTT times, and TTL data were then gathered from the results of the data, and transposed to to an excel spreadsheet

6. Once all measurements were completed, the spreadsheet was analyzed and calculated metrics were taken.

3. Results

Observations on Routing

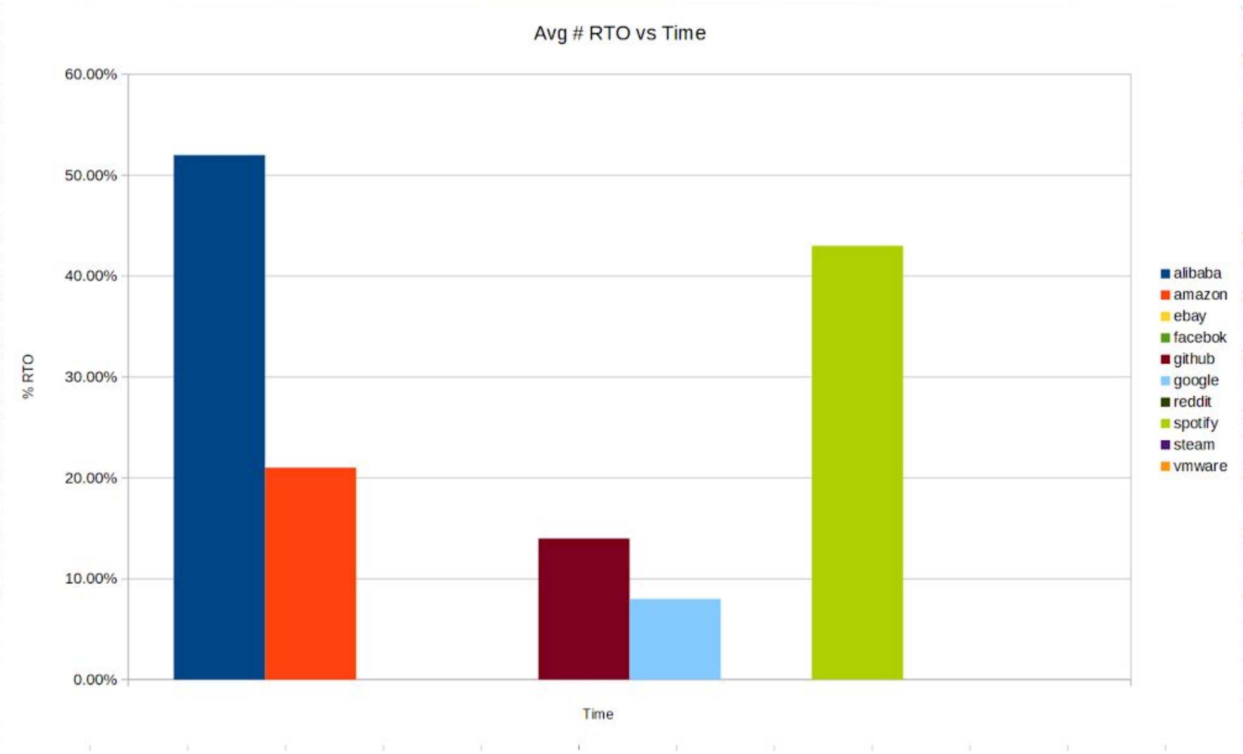
The sites that had the most consistent paths over the trial were www.ebay.com, www.github.com, www.reddit.com, www.vmware.com. Both www.facebook.com, and www.google.com varied minimally in their route paths and hops, and were, for the most part, somewhat-consistent in both areas. Reddit, a social media platform, had nearly the same route pattern every time. Interestingly, www.facebook.com, the largest social media platform, marginally differentiated in it's the route patterns. All of the aforementioned had minimal number of request time outs, and those that did, appeared consistently in the route. This lends to notion that the intermediary node does not allow for ICMP request (not confirmed).

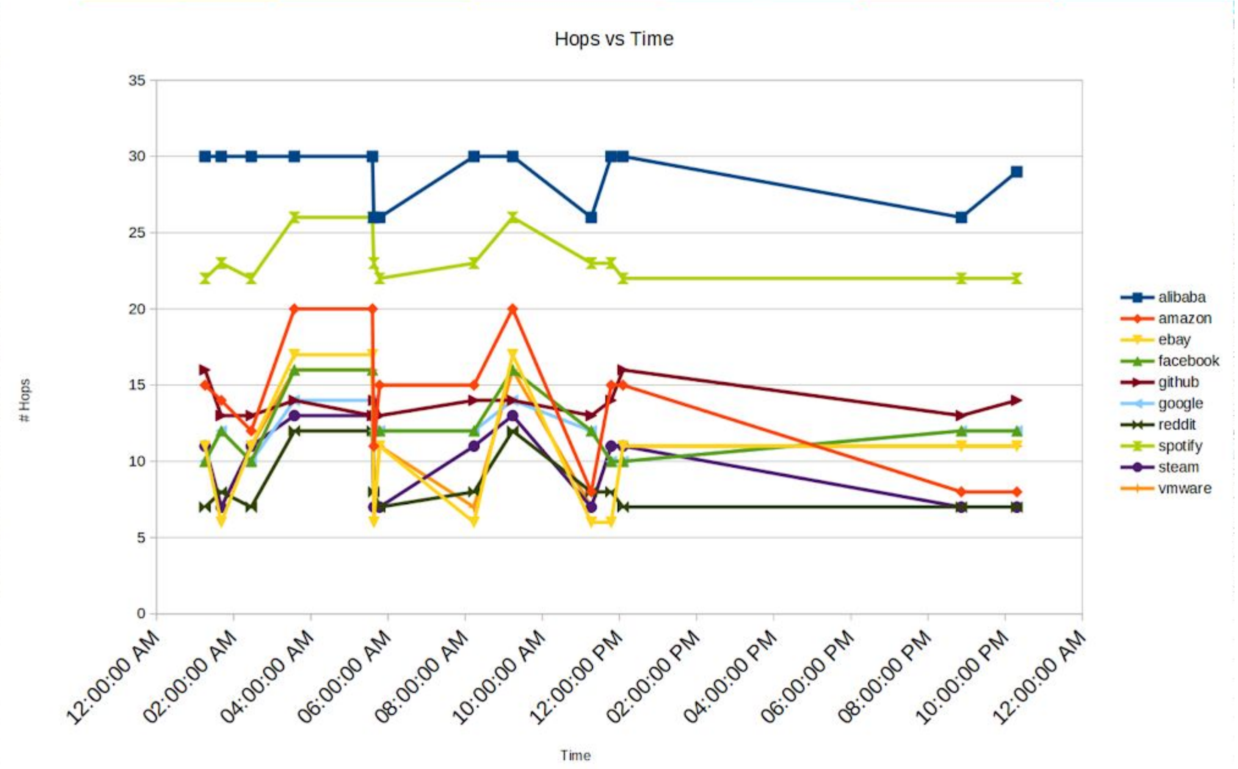
Among those whose routes varied the most included www.alibaba.com, www.spotify.com, store.steampowered.com, and www.amazon.com. Alibaba and Amazon being the top E-commerce sites and Spotify, Steam, and Amazon providing streaming media services. All of these sites varied in the routes they took at any given point in time throughout the day, the number of hops, and rolling destination addresses (different DNS entries). They all suffered from frequent request time outs. Of these, Alibaba was the worst offender, whose patterns would often end in request time-outs and almost nearly approaching 30 hops.

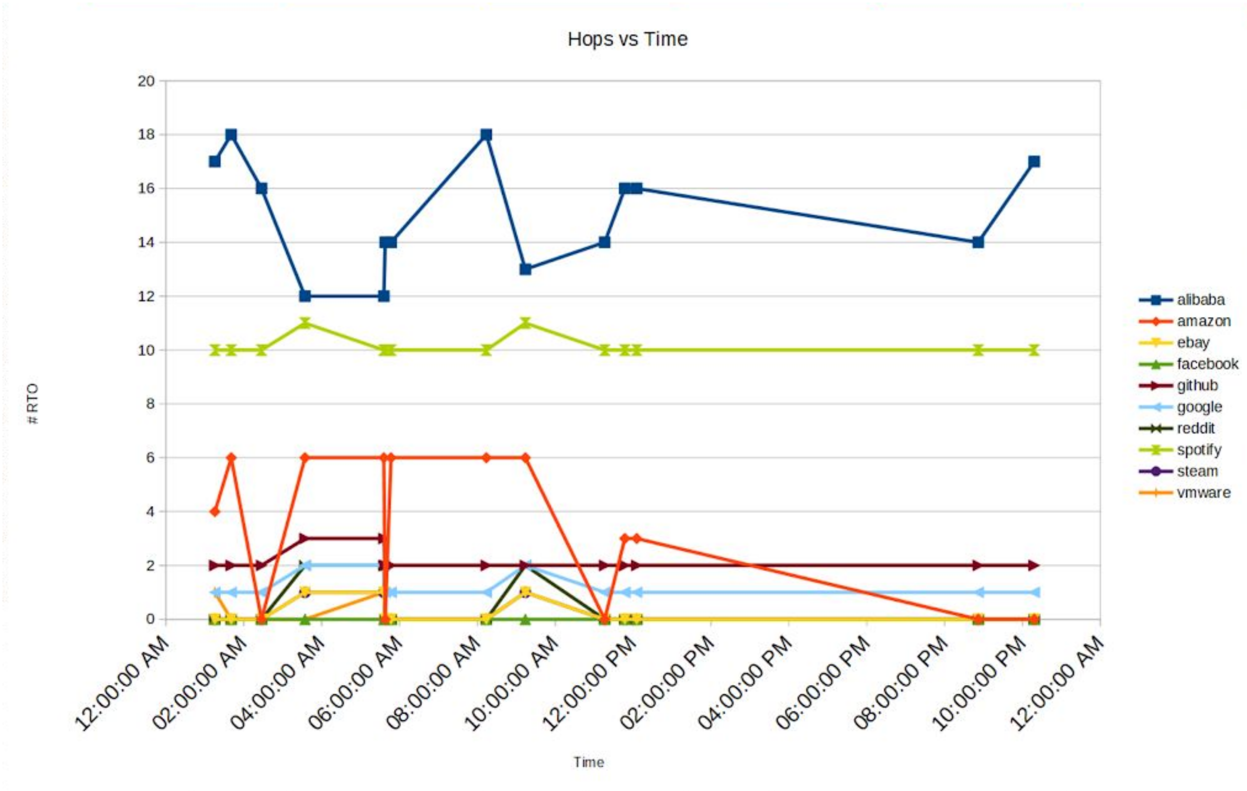
These targets appeared to randomly change their routing patterns with no discernable pattern.

Site	Avg. # Hops	# RTOs	% RTO
Alibaba	29	15	52%
Amazon	14	3	21%
Ebay	11	0	0%
Facebook	12	0	0%
Github	12	1	8%
Google	14	2	14%

Reddit	8	0	0%
Spotify	23	10	43%
Steam	10	0	0%
VMWare	11	0	0%


Avg # RTO vs Time
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
% RTO
Time
alibaba
amazon
ebay
facebok
github
google
reddit
spotify
steam
vmware


Hops vs Time
35
30
25
20
15
10
5
0
# Hops
12:00:00 AM
02:00:00 AM
04:00:00 AM
06:00:00 AM
08:00:00 AM
10:00:00 AM
12:00:00 PM
02:00:00 PM
04:00:00 PM
06:00:00 PM
08:00:00 PM
10:00:00 PM
12:00:00 AM
Time
alibaba
amazon
ebay
facebook
github
google
reddit
spotify
steam
vmware



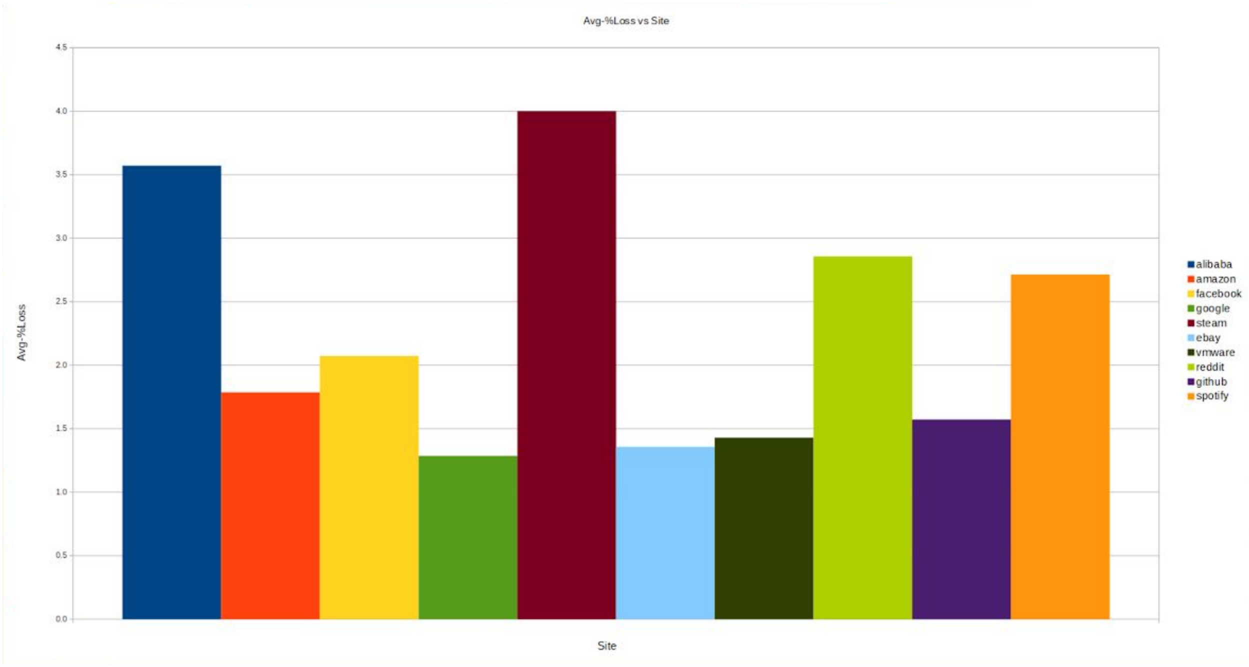
Observations on Stability & Reliability

Using the previous measurements, we classified the targets as `loss free', `minor losses' (0 < loss rate < 5%), `significant losses' (5% < loss rate <10%), and `major losses' (loss rate > 10%).

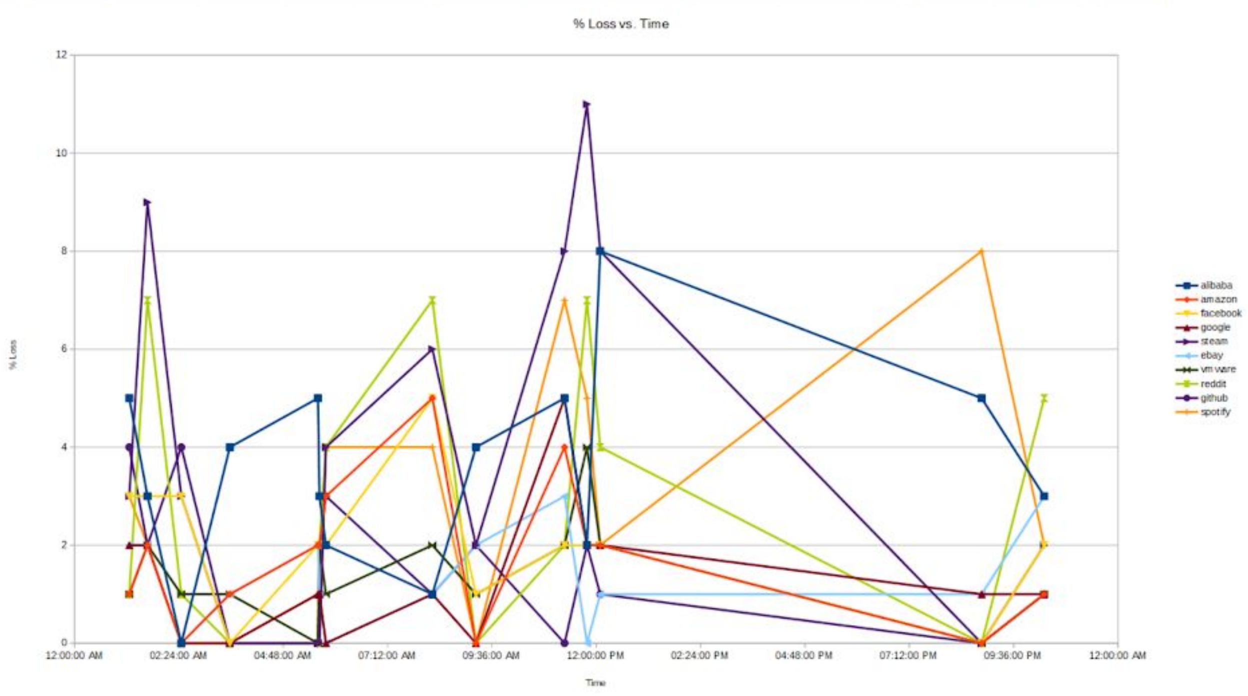
All of our targets are considered “Minor Losses” (loss rate less than 5%) with Steam being the highest, averaging a loss rate of 4% and Google being the lowest with average loss rate of 1.3%.

Target	Avg. % Loss	Loss Type
Alibaba	3.6%	Minor Loss
Amazon	1.8%	Minor Loss

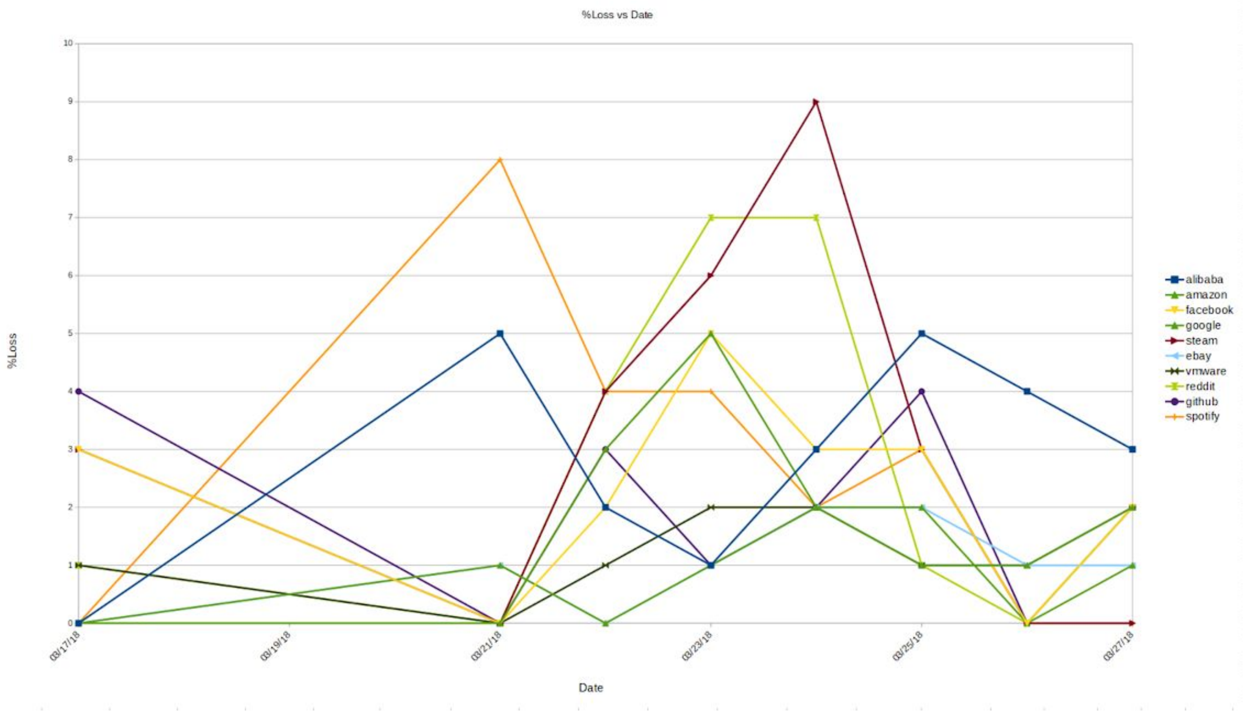
Facebook	2.1%	Minor Loss
Google	1.3%	Minor Loss
Steam	4.0%	Minor Loss
Ebay	1.4%	Minor Loss
VMWare	1.4%	Minor Loss
Reddit	2.9%	Minor Loss
Github	1.6%	Minor Loss
Spotify	2.7%	Minor Loss



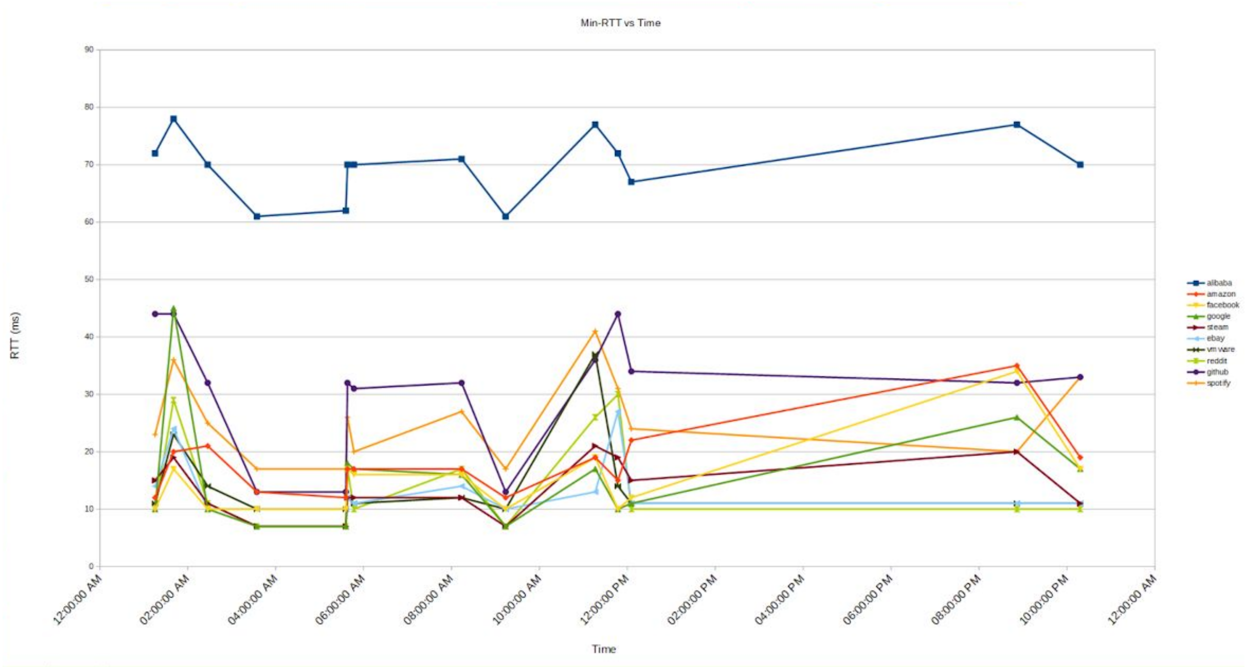
Though every target could largely be considered to be “Minor Loss”, every target did experience “loss free” day (loss rate of 0%) during our 7 day measuring period. Notibibly, Steam.com, Alibaba.com, Reddit.com, and Amazon.com experienced days that fell into ‘Significant Losses’. Steam.com was the only site that experienced “Major Losses” at a peak of 11 requests timed out.

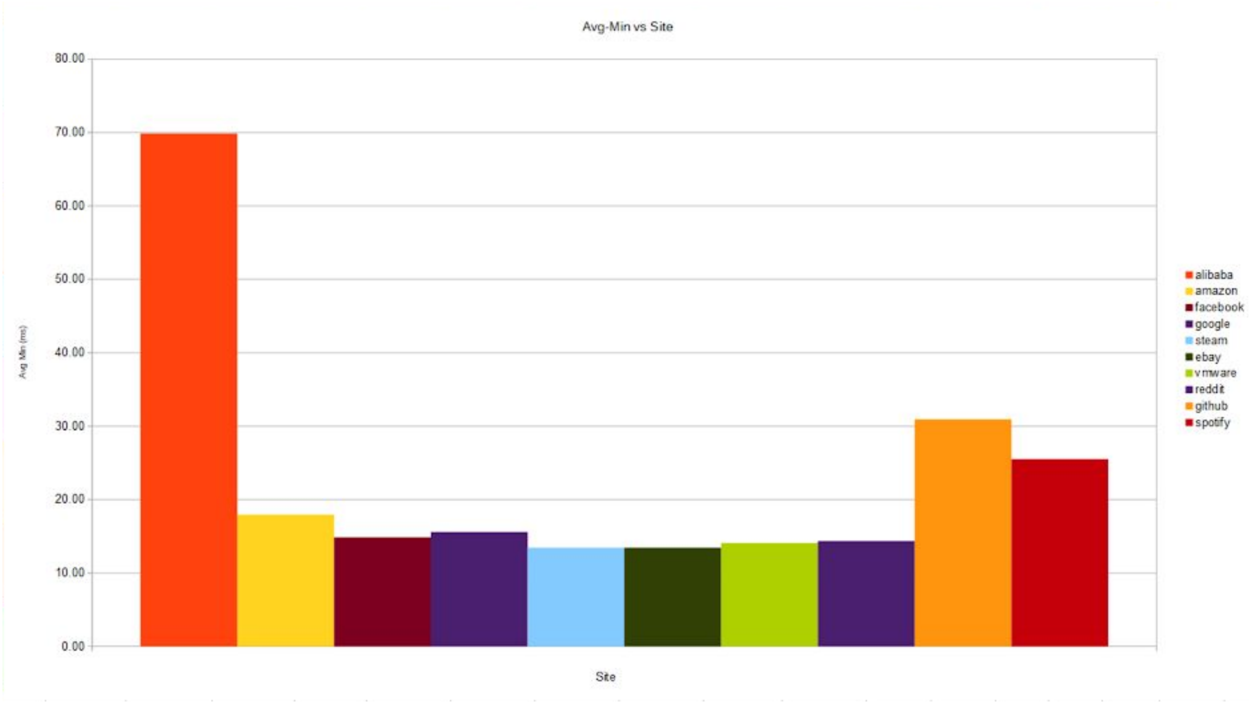


Over the duration of the experiment, there was an increase in the percentage of packet loss during the end of the week and into the weekend. We also noticed decreases in the during the beginning of the week.

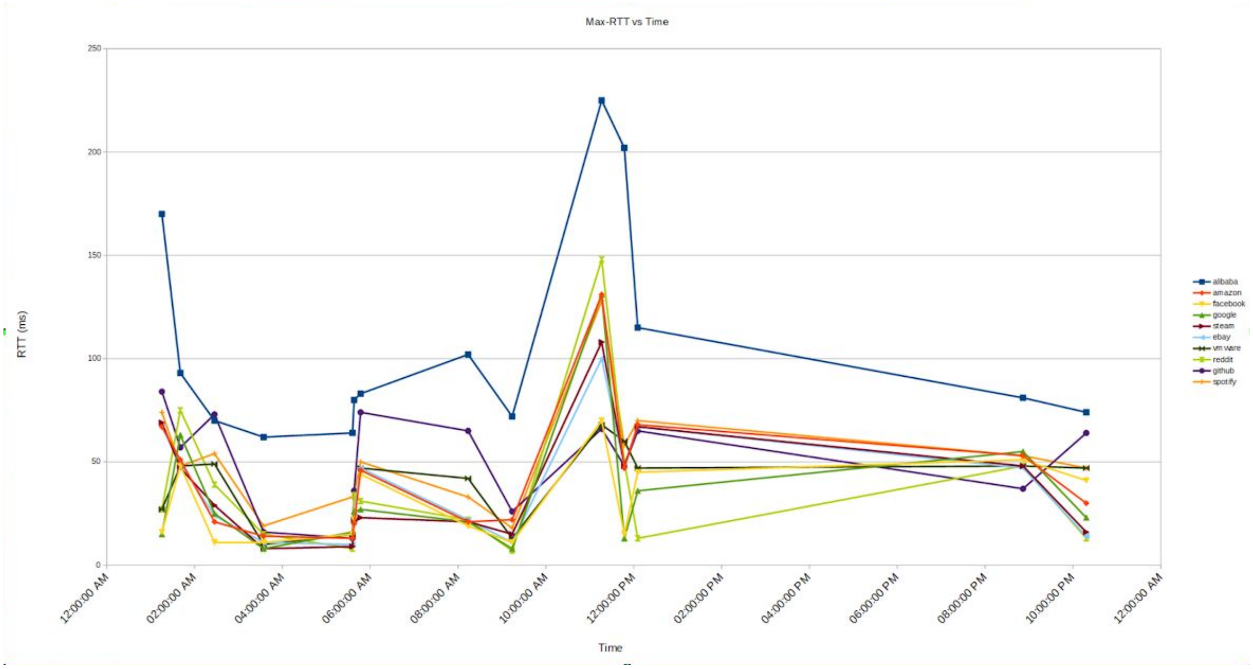


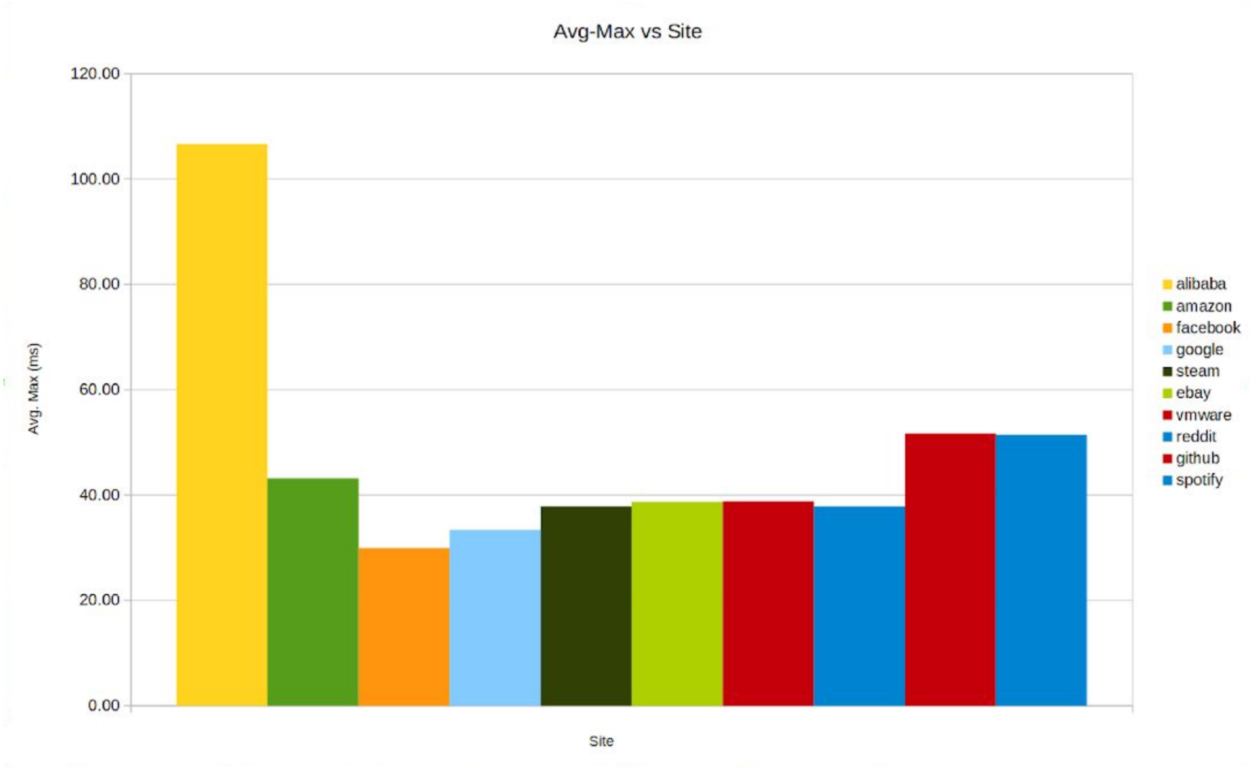
Below we can see the how the minimum RTT time is affected by time of day.



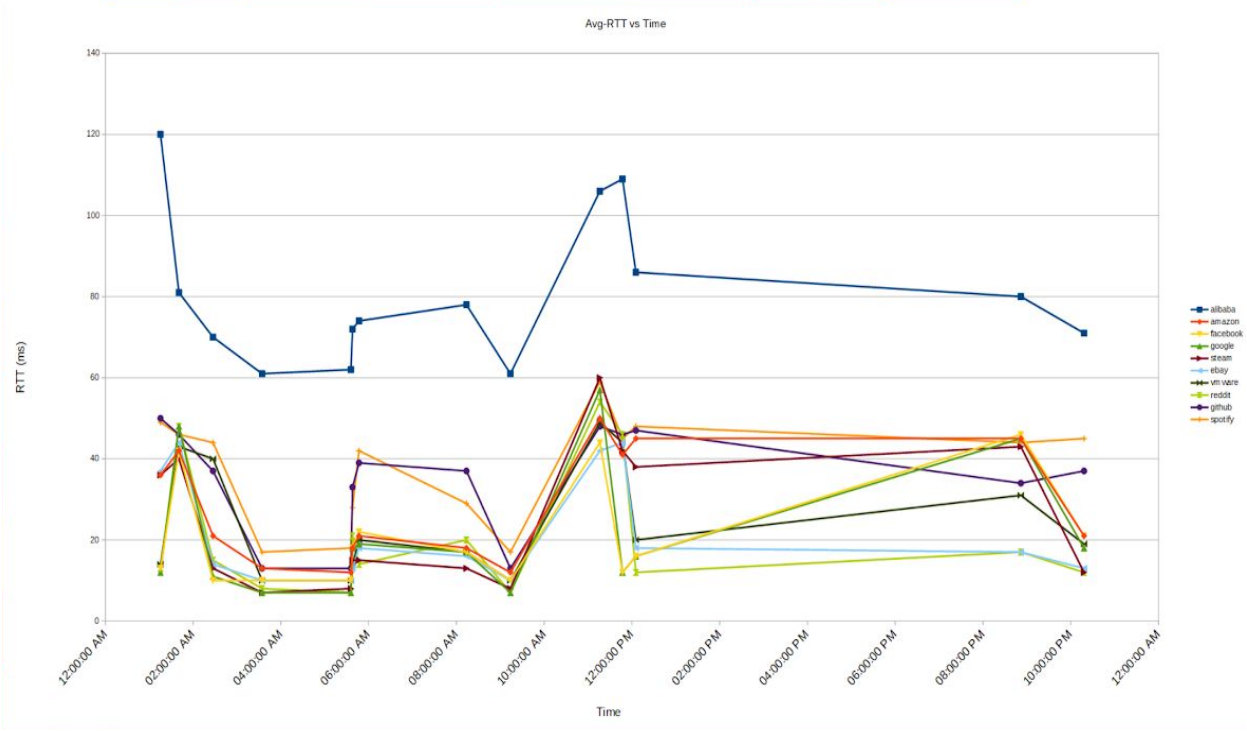


Below we can see trends of the maximum RTT times throughout a 24 hour period. Here we can observe that requests tended to take longer during the late morning to early afternoon (8am - 1pm).

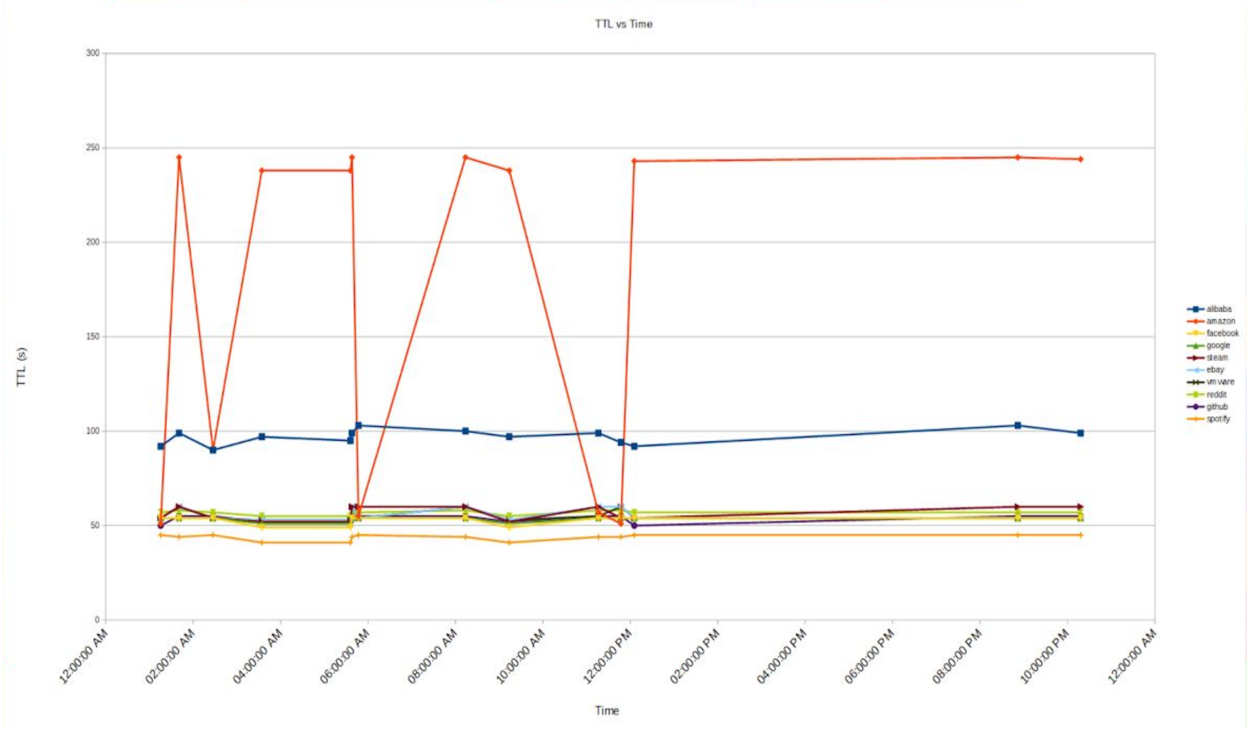


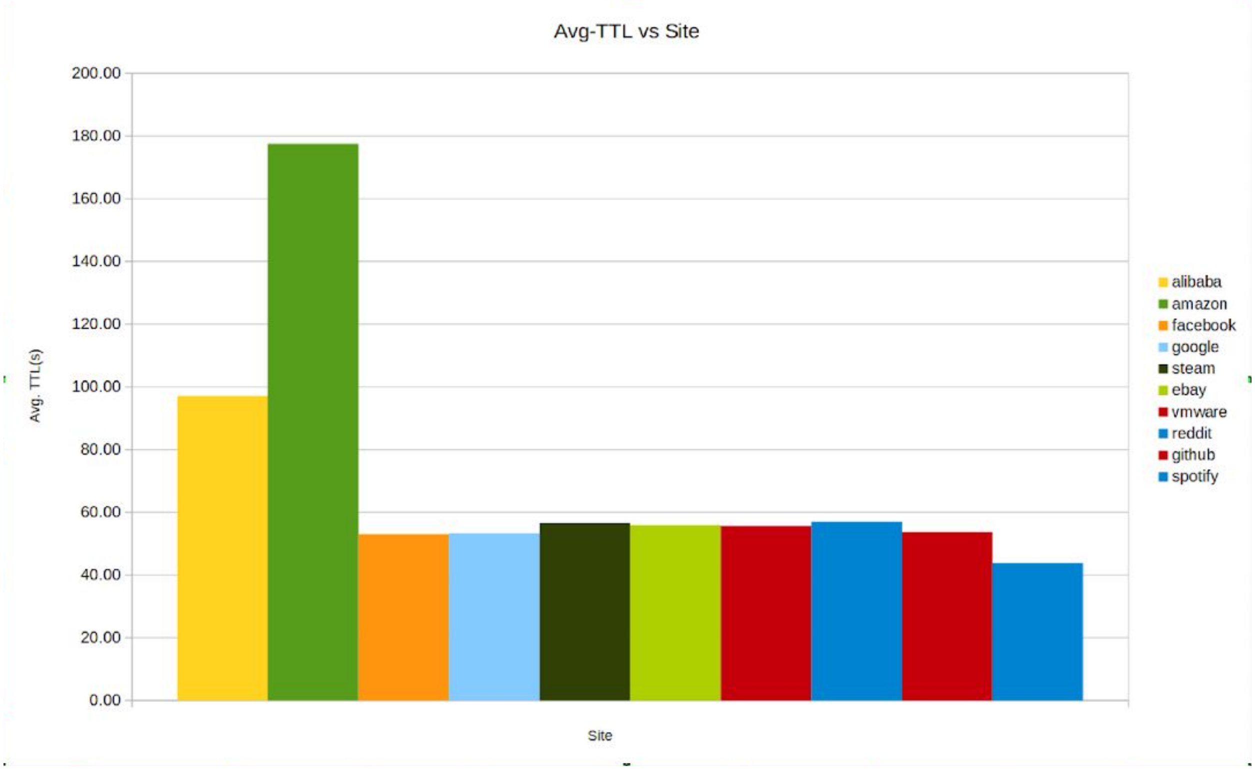


The average RTT times plotted against a 24 period reinforces the trends stated above



Interestingly, major e-commerce sites, Alibaba.com and Amazon.com held higher TTL times throughout the experiment in contrast to the other sites who can be broadly categorized as either Streaming media providers (Steam, Spotify), Social Media platforms (Reddit, Facebook) and Other (VMWare, Google).





4. Conclusion

Our hypothesis wasn't backed up by our experiment. Amazon, a target we picked that has a tremendous amount of traffic, is incredibly inconsistent as well as very unstable. Hence, there isn't a correlation between traffic and stability and consistency. There didn't seem to be any nature to the changes, but rather they appeared random within the confines of the experiment. We also discovered that the loss rate was a lot less than we assumed it would be for our targets. The average loss rate was 2% for our 10 targets; depending on the site this is fairly inconsequential.

5. Lessons Learned

One of our key takeaways, was that networking routing and packet round trip times don't follow any set of rules. There doesn't seem to be any common denominator to determine the path to a target or the RTT for a target. The randomness was definitely an eye opener for us, as we approached the project convinced there would be a pattern. The experiment could have also benefited from more sampling at predefined intervals throughout the day. This would have given more accurate results and our data would have been more well-formed.