CIS527: Active measurements of routes, losses, and delays

Nathan C. Pierce
University of Michigan - Dearborn
4901 Evergreen Rd, Dearborn, MI 48128
piercen@umich.edu

December 4, 2023

Abstract

This paper examines the results of several different measurements taken in computer networks: route stability, packet loss, and round trip times (RTT) of various targets on the internet. Measurements were taken around similar times of day over a period of one week. Targets were selected on their remoteness and for geopolitical factors. Using the *ping* and *traceroute*, data was collected, analyzed, and visualized to make conclusions on network performance and variability across samples using principles discussed over the course.

1 Introduction

Measurements were taken to examine how the internet routes and forwards packets. Inter domain (between Autonomous Systems) is made possible on the internet using Border Gateway Protocol (BGP). The internet is composed of many Autonomous Systems (AS), which provide redundant routing options. Each AS is composed of one speaker router, which advertises local network, other reachable AS, and provides enumerated path information. An AS may also contain other border gateways which receive internetworked traffic, but only the speaker advertises information for routing. Packets enter and exit an AS through the border gateways. There are many paths that a packet can take from any given source and destination. BGP consists of a hierarchy of AS, at the top of which are backbone providers, which supply the principal routes, coming together at internet exchange points (peering points) where regional Internet Service Providers (ISP) can handoff messages.

There are three tiers of network service providers offering connectivity between any two networks connected via the internet. Tier 1 (backbone providers) generally can reach every other major network on the internet. Tier 1 providers can freely exchange packets between peers (other Tier 1 providers). These are generally large telecom companies such at ATT, NTT, British Telecom, Deustsche Telekom, etc. Tier 1 providers build and maintain massive telecommunication infrastructure, allowing connections across oceans and vast distances. Tier 2 providers are the most common, and are regional Internet Service Providers (ISP). Tier 2 providers peer with other networks but also pay to route to Tier 1 providers. Tier 3 providers only purchase transit from Tier 2 and Tier 3 providers – they do not provide any kind of cross-domain transit capabilities, but may contain significant intradomain routing. Autonomous Systems generally offer both inter and intradomain routing, however, in the context of sending a packet across the internet to a destination, what matters is how an Autonomous System knows how to best route a packet through their intradomain network in order to send it on to the next Autonomous System (accomplished by the border router). Tier 3 providers are local to cities, towns, and metro areas. Each tier is composed of AS, which are assigned their own 16-bit AS number. This hierarchical approach to aggregating networks helps to improve scalability of the internet. Traffic passing through an AS is classified as either local (both terminal points are within the intranetwork) or transit (passes through at least one AS). AS are classified into three

Routers at the border of AS determine their forwarding based on the routing tables they construct. The rules for constructing these tables can be complex, and include many different variables. The router must know about the topology of the network in order to build its routing table, as the speaker (advertiser) for the AS will advertise complete enumerated paths between hosts. There is also the issue of security and trust, as trusting other domains to handle your data can have significant consequences both in terms of performance, service, and security. Generally, AS are credible, but there have been examples in the past where implicit trust has resulted in negative consequences. Exactly how these speaker routers build their tables is outside the scope of this experiment, but the core principal is that border gateways optimize reachability over the shortest path – a path that is close to optimally short is an achievement.

The exact path that a packet will take through the internet varies even for a given source and destination. Links and routers are physical components that can fail at any time, severing the paths they enable. In this experiment, several measurements were taken over the course of a week to sample from the distribution of these kinds of failures, which would result in different route paths, delays in packet delivery, and complete packet loss. The measurements were taken using *ping* and *traceroute*.

ping operates by using the Internet Control Message Protocol (ICMP). It sends an ICMP echo request to a specified network endpoint, and receive an echo reply, allowing ping to determine round trip times and dropped packets. ICMP is mainly used for error reporting and has several message types that are useful in network troubleshooting. The *ping* utility is in the application layer, while ICMP is a network protocol. *traceroute* is a tool that prints the path that a packet takes from source to target. It also uses ICMP. *traceroute* sends out a broadcast message over the local network, initially with time to live (TTL) equal to 0. When a router receivers a message with ttl, it checks if ttl is zero (in which case, the message is dropped and an echo reply is returned to the source), or it decrements the ttl and sends it on the next hop. The utility will send out three ICMP requests (by default), incrementing the ttl with each reply, until it receives a reply from the destination. The information made available to us is the exact path a packet will take from source to destination, as well as the round trip time for each hop along the path, as well as packet loss.

While only five targets were selected, a range of different kinds of targets were selected to represent:

- **Remoteness**, that is, relative to other cities geospatially (e.g. separated by large bodies of water or by vast expanses of land and wilderness). The assumption is that core routing infrastructure such as internet exchanges exist in dense urban areas (major cities), so measurement of values that depend on this kind of geospatial separation can be examined. The expectation is that more remote destinations would be prone to changes in networks resulting in different kinds of measurements, and might lead to more dynamic results. No regionally local targets were selected (none within the United States).

- **Geopolitics**, as mentioned, trust is a factor in building routing information for internetwork routers. With the source node being fixed within the United States for all measurements, the geopolitics and relations between the United States and other countries may be a factor in the outcome in the measurements. The assumption is that allied nations will be favored in routing hops over other nations.

- **Time**, different endpoints will be in different timzones, which will affect how full the pipe is between links along any given path. For example, a measurement taken in the United States Eastern Timezone (EST) at 11 AM will be during waking hours when many people are utilizing networks across the internet, affecting the data-bandwidth capabilities of different regional pipes. A measurement at such a time in the United States for a destination in Beijin, China crosses several timezones, and 11AM EST corresponds to midnight in China Standard Time, a time when most people are sleeping and offline. While *when* a measurement is taken will clearly be important in the cross-analysis of different destinations, the specific choice of destination also matters. Some effort was made to select global destinations, in several different timezones.

The selected targets and their endpoints are listed below.

- Tsinghua University
  www.tsinghua.edu.cn (166.111.4.100)
- The City of Yellowknife, Canada
  www.yellowknife.ca (20.104.44.154)
- Wellington International Airport
  www.wellingtonairport.co.nz (52.84.18.97)
- The Falkland Islands Tourist Board
  www.falklandislands.com (51.145.110.210)
- University Centre in Svalbard
  www.unis.no (34.247.153.46)

2 Methodology

Routing, delays and loss were the primary measurements of interest. *ping* and *traceroute* were used to make these measurements from a Linux machine. There was no VPN or tunnel connection used during measurements. The next hop from the source machine's default gateway was the public Comcast gateway, and from there the local ISP forwarded packets to the destination based on their routing schemas.

A simple bash script *measure.sh* was written which would pipe the commands into a text file for each measurement type and each desination (a total of ten text files). The datetime was recorded for each measurement so that we could account for network dynamics relating to date and time. Initially, it was intended that the *measure.sh* would run twice a day at the same time each day, with the script being execute as a cron job on the source machine, but this was not realized due to technical trouble encountered and time constraints, and also that this requires the host machine to be online at those specific times each day, which could not be guaranteed. The result is that some days have more measurements than others, and there are inconsistencies in the time (although some effort was taken to ensure that measurements occurred around the same time of day).

The bash script ran the ping command for each target 120 times, across five targets, for a total of eighteen measurements. No additional arguments were used in the *traceroute* command. For a single execution of the script, a single command would execute sequentially, one after the other.

A python script *convert.py* was written to parse these text files and extract relevant data, as well as for the data analysis.

While *traceroute* does allow us to measure packet loss and round trip times, *ping* is a better option to measure these metrics since we have a large sample size using *ping*, and *ping* specifically measures end-to-end round trip time (RTT) while *traceroute* measures on a per-hop basis.

3 Results and Discussion

The delays and packet loss are displayed in Figure 1. Perhaps the first thing that stands out is that, for each destination, there are inconsistencies in the delay. Recall that eighteen measurements were collected. With this in mind, while the delay data may look largely uniform at first, since the sampling size is small, the exceptions in uniformity are telling. More specifically, they indicate the dynamics of the real world internet, at least in regards to response time. The exceptions in the plot include the third measurement for the Falklands. Interestingly, the delays for the requests to Svalbard (UNIS) for this particular measurement also differ from the rest of the delays, in that they are significantly lower. Examining the route path for third measurement of the Falklands destination compared to the previous, the path looks the same, and the number of hops is also the same, however the ping delays are clearly shorter on average. The delays on this particular day were uniformly low, compared to the prior day where the distribution of delays were much more spread out. It's possible that the requests/repsonses exchanged during this period had a higher priority within the inter domain network compared to other days, although it is unclear why that would be the case.

Interestingly, the location for the final hop in the path to the Falklands site endpoint is actually in London. The Falkland Islands are an overseas territory of the United Kingdom, so it makes sense that the country's official websites are hosted in London where infrastructure is built out. The assumption is that the transatlantic cables connecting the United States to Europe can handle much more traffic than the connections from the United States to South America, so it would make sense that traffic originating from the United States would re-direct to the government's hosting node in London. It's also probable that most of the requests received by the Falkland Islands originate from the Northern Hemisphere. The connection to London is likely much closer for such requests, on average.

Destination	Average	Max	Min	Std. Dev.
Yellowknife	450	38	2.1917	12.159
Svalbard	175.348	1036	98	167.448
Falklands	44.0791	1030	110	158.029
Tsinghua	288.690	1019	240	45.453
Wellington	22.4148	157	16	7.383

TABLE I. All units are in ms. Statistics from the *ping* executions. These delays are end-to-end delays.

A similar anomaly seems to have occurred for the *2023-11-25 10:52:51-05:00* measurement for the Tsinghua University. That is, the average delay in echos is much lower on average, owing to the uniformly low delays. Again, the minimum delay is well within the bounds of the other measurement's minimum delays. Just like with the Falklands, the routing for this uniformly low routing follows the same path as other days where the average delay is higher. Again, it's possible that the priority of these requests were higher, or perhaps the relevant link's delay-bandwidth product was lower than the other days, meaning our requests/responses would spend less time in network queues. Assuming that each ping request and response are of fixed size, a decrease in the delay would correspond to an increase in bandwidth. It's unlikely that the increased bandwidth corresponds to an improvement in the relevant network hardware infrastructure handling our packets, because succeeding measurements to the anomaly are comparable to the preceding measurements. More likely, the pipes were likely just less full during this measurement. 11 AM EST corresponds to midnight China Standard Time. Some of the other measurements in the morning appear to have a lower average delay. 11 AM to 5 PM EST are midnight to 6 AM in China Standard Time. Most likely, our packets likely moved faster between hops once they started being routed in these early morning timezones, when most people are asleep.

One other obvious feature that the Tsinghua measurements share with Falklands is the general uniformity in their delays compared to the other destinations. Yellowknife, Wellington, and especially the Svalbard measurements all have very low uniformity. By visual inspection, Yellowknife seems to have the most consistently low delay in ping echos, with a clear exception during the *2023-11-20 21:10:06-05:00* measurement. This can be concluded by the "flatness" of the rolling average line. The consistency in routing can be seen by looking at Figure 2, which shows the average latency between the three echo requests sent out for each iteration of the traceroute execution. Measurements are uniquely defined by the vertical, green dashed lines. Consistent routing is represented by identical or very similar shapes in the graph. The two destinations with uniformly delays, The Falkland Islands and

Tsinghua University, both have very consistent average delays across hops compirising each iteration of the traceroute command. From these plots, it can be suggested that the routing between the source node in southeast Michigan and Wellington National Airport is quite sporadic across the measurement period. The shape of the average delays are similar across about the first 80 hops, at which point the shape changes, with again another significant spike in the final hops toward hop number 300 and later.

Destination	Average	Max	Min	Std. Dev.
Yellowknife	36.4923	165.4393	2.1917	24.0005
Svalbard	46.5839	137.3570	2.3880	40.1472
Falklands	84.9763	171.2795	2.3580	54.6120
Tsinghua	117.0888	359.1000	2.4030	116.7405
Wellington	23.4061	54.9840	2.2937	9.0164

TABLE II. All units are in ms. Statistics from the *traceroute* executions. The traceroute command sends three echo requests with each iteration – all echo requests are included in the samples. Samples include values measured from the client node to the local area network router, which is expected to be very fast and may slightly skew the statistics here if end-to-end latency is the principal component of interest. *ping* does not have this problem.

Examinations of the traceroute data does in fact reflect the routing change for Wellington, as one additional hop was added near the beginning of the route path, which would affect latencies of all following links. Despite the more sporadic routing, the average delay across all requests are lowest for Wellington. This can be seen by examination of the delay plots in 1 and the statistics in Table II. The mean is the lowest. This is not what was expected, considering how much closer to the source many of the other destinations are (partiuclarly Yellowknife, which ranks second on average in terms of the lowest delay times). However, inspection of the traceroute results for Wellington show that the final hop, for every single measurement, is to a Amazon Web Service (AWS) Cloudfront router. This service provides users with a Content Delivery Network (CDN) that can be used for exactly the purpose here – delivering content to users where there are vast geographical distances between requester and destination. The final hop router is **.ord53.r.cloudfront.net*, which is a router in Chicago, IL. This explains the very low latency as Chicago is very close to the source, relatively.

Svalbard is one of the most remote regions in the world. It is obvious by looking at the plot for the delays in Fig 1 that round trip time (RTT) varies largely for this particular measurement. The rolling average line has many sharp fluctuations as the delay time for many echo requests were very high. Still, there is a clear majority where *most* pings are within about the 100 to 215 ms range. Looking at 2, it seems like the routing path was consistent over all measurements (again, assuming that the shape of the graph corresponds well-enough to the route path e.g. that a consistent shape in the graph indicates a consistent routing path). This measurement variability is also indicated by the standard deviation as seen in Table I. Svalbard has the largest standard deviation, followed closely by Falklands. This pattern applies for the mean latency as well. Examining the consistent routing path for the Svalbard route in the traceroute data, the path from source follows the usual path to the internet exchange in Pontiac, then jumps to a Comcast backbone before directly going to **.ip.twelve99.net*, which are owned by Swedish telecom company, a tier 1 network provider with Autonomous System number AS1299. There are several jumps amongst twelve99 routers before reaching the destination. As seen previously with Tsinghua, this kind of variability in service might be explained by time of day measurements were taken, however there does not seem to be a clear pattern in this case, as there is no consistency in measurements taken that correspond to early morning Sweden time (low data-bandwidth hours) vs. afternoon (high data-bandwidth hours).

Figure 1. The plots for the delays measured using the ping command for all targets. The dashed green line separates measurements on a daily basis, matching the dates on the x-axis. The average (red line) across the plots is the rolling average using a window of ten samples. Destinations are noted in the title of each plot. Notice that the y-scales for the delay are not the same across plots – the minimum value is zero but the increments vary depending on the magnitude of delays seen for each destination measurement. Notably, the increment scale is largely determined by outliers in each plot (as can be seen clearly in the tsinghua plot.

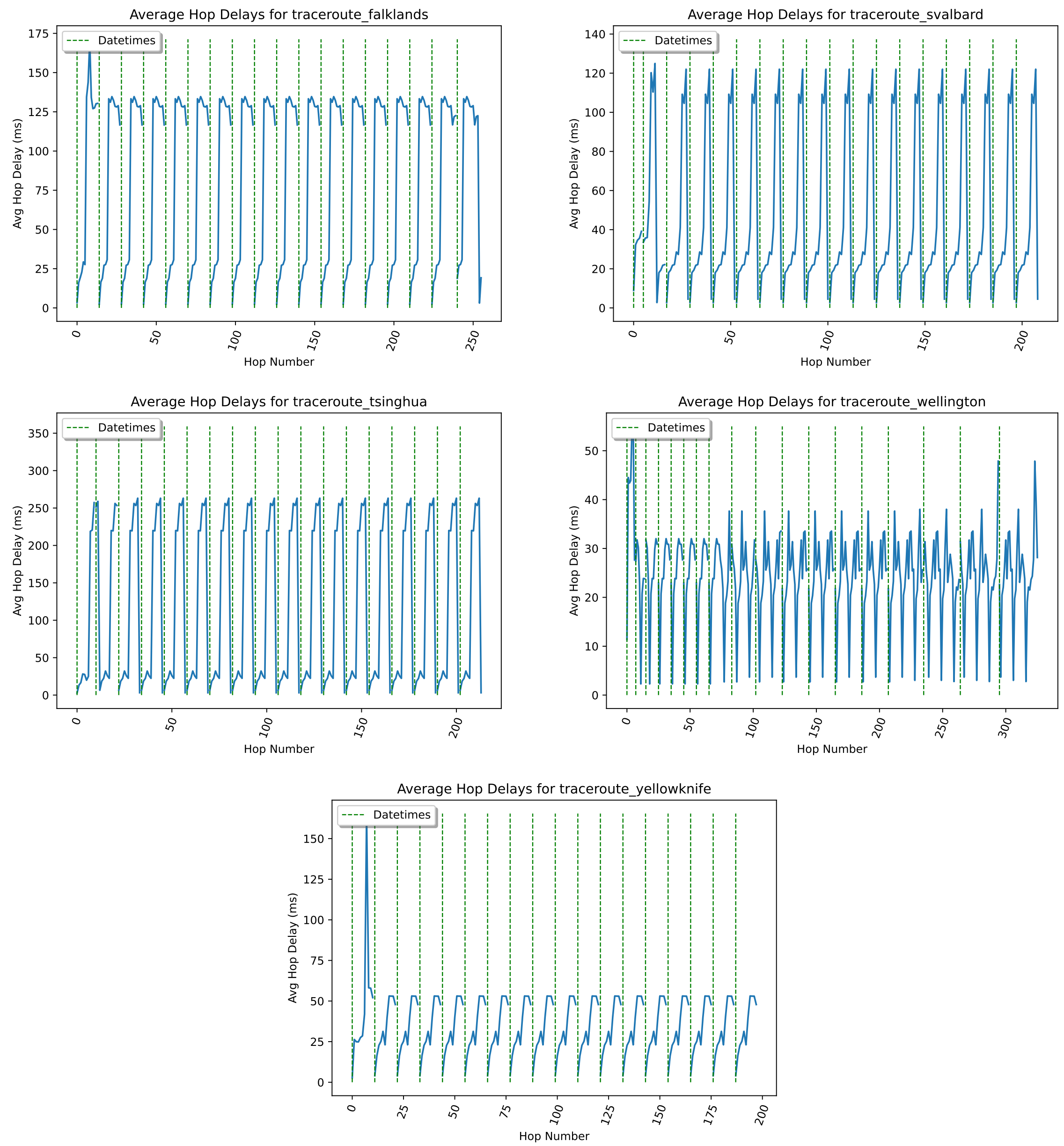

Figure 2. These plots show the average delay of each echo request/response of the traceroute tool for each hop on a given day and destination. The dashed green line separates the datetimes. The average is computed from the three packets that traceroute sends during each iteration of execution (that is, is the average delay amongst three packets sent from source to the hop). One thing this graph shows well is the relative number of hops to unique destinations, since a unique coordinate is well represented here (although the accuracy of the lat/lon as generated from the IP address of the hop may be significantly off. The hop number corresponds to the total hops summed across all measurements for a specific destination. The x-axis limits are not the same across plots, nor are the y-limits.

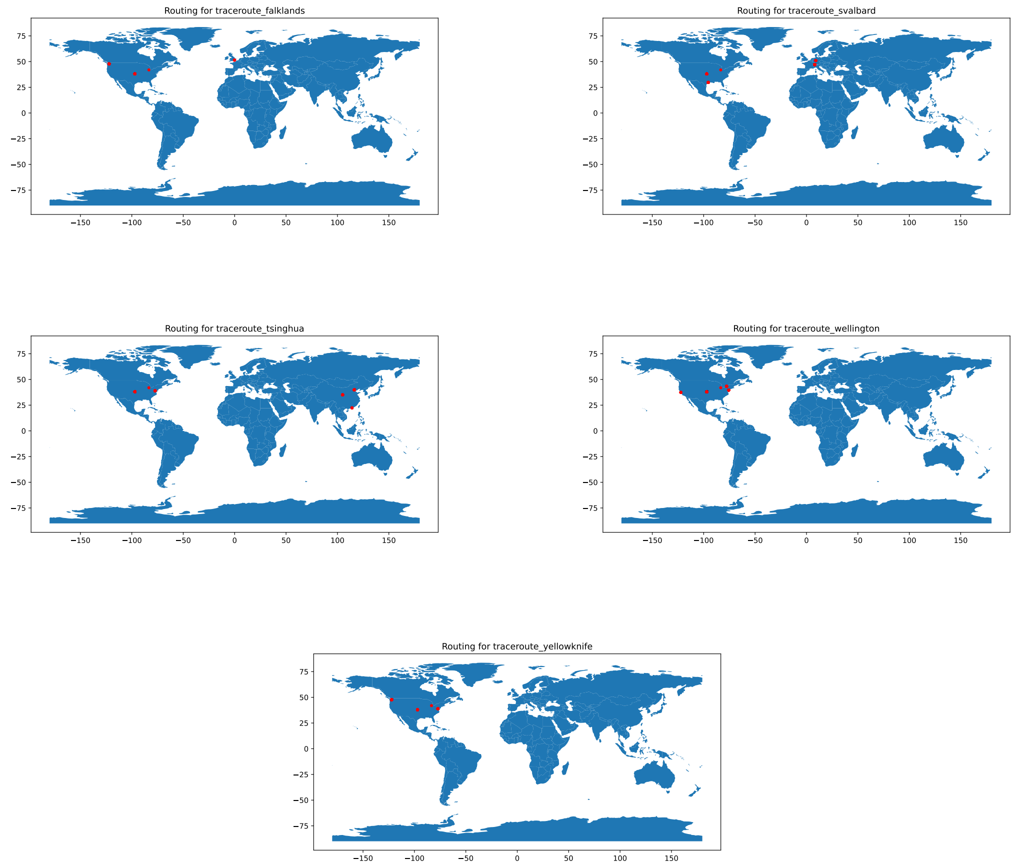

Figure 3. Geospatial plots of next hops in the traceroute for each destination. The target name is in each title's plot. Geolcation based on IP address often has spurious results at best, as IP addresses are constantly changing and lookup tables may not have the latest results. Additionally, ISPs serve their customers over large regions. The graphs here are meant to give a rough idea – as in cases where IP geolcation pairings do change, they usually do not changes which limits accuracy to the region level and not the endpoint level. ignificantly (i.e., the country would not change, but the state or city might). Since the source/target pairings are on an international scale, these figures are included.

4 Conclusion

In this experiment, the variability of network performance was examined using measurements taken using the *ping* and *traceroute* tools. Analysis and visualizations were presented. It was seen how changes in networks that are interconnected over the internet can change on a daily basis, affecting the performance of routing in terms of latency/round trip times and packet losses. Assertions about the data relating to geographical distances between nodes, geopolitical factors, infrastructure capabilities, variability of bandwidth usage for a region over a twentry four hour period, and priority were examined.