

Methodological overview to accompany the M-Lab Consortium Technical Report, ISP Interconnection and its Impact on Consumer Internet Performance

Data Overview

The Measurement Lab (M-Lab) dataset includes over five years of global network performance data from user-initiated tests. While M-Lab hosts sixteen separate measurement tools, for the interconnection study we are looking at the data collected by one specific tool, the [Network Diagnostic Tool](#) (NDT). All descriptions and data samples refer to data from NDT, and not to general M-Lab data or data collected by other M-Lab tools.

NDT tests run against geographically-proximate M-Lab measurement points. These measurement points are located in well-peered (representative) transit networks, such that each user-initiated test crosses and measures performance over at least one network (autonomous system) boundary. Measuring across this boundary allows M-Lab to assess the impact of ISP interconnection on consumer Internet performance. In order to analyze the effects of interconnection on consumer performance, it was necessary to segment the data according to several key parameters:

- Time of the test,
- Client geographical location - where the test comes from,
- Measurement point ISP - the ISP in which a measurement point is located,
- Client access ISP - the ISP to which the user running the test is connected.

To segment the database on these parameters, we developed a tool called [Telescope](#). To visualize the collected data, we created a web application called [Internet Observatory](#). This document describes how M-Lab researchers implemented both tools, what analytical choices we made, and how an external researcher can reproduce the results for independent review.

Accessing Raw Data

To access the M-Lab dataset, we used BigQuery. BigQuery is a cloud-based SQL engine that contains all of the NDT data from the M-Lab dataset and allows the user to retrieve data very quickly via SQL queries, without the requirement to download the entire dataset (see [M-Lab BigQuery documentation](#) for more information).

Alternatives

It would also be possible to access the data by processing the raw NDT result files (available from M-Lab via [BigStore](#)). We decided that this strategy was not ideal, as it would require analysts to download and parse terabytes of data in order to reproduce the results of the study, and thus limit validation and extension of this work to those with vast computing resources.

Selecting Data

Specifying Time Range

Selecting data within a given timeframe is trivial with BigQuery. The M-Lab BigQuery dataset is split into tables by month. Each table has a column, `web100_log_entry.log_time`, which specifies the time at which a particular NDT test began (in Unix time format). To select data from a particular date range, we must construct a query that includes appropriate `FROM` and `WHERE` clauses, joining tables from multiple months when necessary. For example, to retrieve results for the first quarter of 2014 we could add the following clauses:

```
FROM
```

```
[measurement-lab:m_lab.2014_01],
```

```
[measurement-lab:m_lab.2014_02],
```

```
[measurement-lab:m_lab.2014_03]
```

```
WHERE
```

```
((web100_log_entry.log_time >= 1388534400) AND -- 2014/01/01 in UNIX time
```

```
(web100_log_entry.log_time < 1396310400)) -- 2014/04/01 in UNIX time
```

Specifying Geographical Location

M-Lab NDT clients use [mlab-ns](#) to find an appropriate M-Lab measurement point.¹ mlab-ns uses Google AppEngine's IP geolocation to match M-Lab clients to the geographically closest M-Lab measurement point. We therefore assume in the study, *ISP Interconnection and its Impact on Consumer Internet Performance* (hereafter, the Interconnection Study) that NDT clients and the measurement points they locate and test against generally are in the same geographical region. Because we are looking at relative changes in a self-contained dataset, making comparisons between performance during one time period and another remains a valid signal even in cases where clients from outside a region test against a measurement point. To find results that represent users in a particular geographical area, we examine the results from M-Lab measurement points in that area. We describe how to select results by measurement point in the section "Specifying Measurement Point ISP."

Alternatives

It would be possible to perform geolocation individually on client IP addresses of test results and then filter out results where the clients are geographically distant from the server. We chose not to perform such filtering due to the fact that there are no highly accurate IP geolocation data sources that are also freely available for external researchers to reproduce our results. At the time of the study, M-Lab had measurement points in 9 different major metropolitan areas across the US, which we considered to be sufficient coverage for any US client to reach a geographically proximate measurement point.

Specifying Measurement Point ISP

Discovery

To select data by measurement point ISP, we first need to discover the available M-Lab measurement points. To do this, we can use mlab-ns, the M-Lab name server, which clients use to find M-Lab measurement points.

Using the [map of M-Lab measurement points](#) provided by mlab-ns, it is possible to create a list of all M-Lab measurement points in the United States. Clicking the tagged measurement point locations, we can construct hostnames for each measurement point. For example, in New York, we can see the following entry (note that each

measurement point includes three servers, and a given metro may host more than one measurement point):

```
M-Lab Site: lga01 | iupui_ndt | ndt | mlab1 | online | |:-----|:----|:-----|:-----| |
iupui_ndt | ndt | mlab2 | online | | iupui_ndt | ndt | mlab3 | online |
```

From this, we can construct the measurement point hostnames for M-Lab's NDT testing service and resolve their IP addresses using DNS:

Hostname	IPv4 Address
ndt.iupui.mlab1.lga01.measurement-lab.org	74.63.50.19
ndt.iupui.mlab2.lga01.measurement-lab.org	74.63.50.32
ndt.iupui.mlab3.lga01.measurement-lab.org	74.63.50.47

A search of [public BGP information](#) reveals that these IPs are announced by AS29791 - Voxel Dot Net, Inc, which is an AS that transit ISP Internap owns ([Internap acquired Voxel in 2012](#)). By repeating this process for all US M-Lab sites, we can determine each measurement point's NDT-specific IP address, and the ISP in which the measurement point is connected.

Data Selection

Armed with the NDT-specific IP address from a given measurement point, we can select NDT test results that match a particular measurement point by adding a WHERE clause to our query such as:

```
WHERE
(web100_log_entry.connection_spec.local_ip = '74.63.50.19' OR
 web100_log_entry.connection_spec.local_ip = '74.63.50.32' OR
 web100_log_entry.connection_spec.local_ip = '74.63.50.47')
```

The output rows will be limited to results collected from the measurement point at M-Lab site lga01. Since lga01's ISP is Internap, results associated with this measurement point

reflect interconnection performance across boundaries between Internap and the test client's access ISP.

Specifying Test Client Access ISP

Discovery

In order to select data based on a test client's access ISP, we must determine which results are associated with particular access ISPs. To do this, we need a data source that maps ISPs to IP address blocks. We chose the [MaxMind GeoLite ASN](#) database, as it is widely used in Internet measurement research and freely available, allowing other researchers to reproduce or extend our analysis.

As ISPs change names or merge with other ISPs, it is common for a portion of the ISP's AS names to differ from the ISP's "official" company name. For example, the access ISP CenturyLink has expanded its operation through the acquisition of other consumer providers such as Qwest. The registration and naming of this network infrastructure has not been consistently updated, and may still reflect prior ownership. Unfortunately, there is no freely available data source that maps ISPs to all of the ASes they control. We invite the creation of such a source. To address this, it was necessary to construct these mappings manually by investigating AS ownership for the set of major US ISPs included in the study.

After creating these mappings, we constructed regular expression searches for each access ISP that would match all AS names associated with that access ISP. For example, to find AS names associated with CenturyLink, we search for AS names that match the following case-insensitive regular expression: "(Qwest)|(Embarq)|(Centurylink)|(Centurytel)". Using these regular expressions, we can find all IP blocks associated with a particular ISP in the MaxMind database.

Data Selection

To identify results in BigQuery that are associated with a particular access ISP, we must add WHERE clauses to our query to select NDT tests whose test client IP addresses fall within one of the ISP's IP blocks. In order to filter addresses in an efficient manner, Telescope generates a BigQuery query that converts the IP from Internet standard format (dotted string) to the integer-based network address format. The query then compares the integer address against the first and last addresses in ISP-owned IP blocks.

For example, the following entry in the MaxMind database describes an IP block associated with CenturyLink:

```
1068332544,1068333055,"AS209 Qwest Communications Company, LLC"
```

This translates to the following BigQuery WHERE clause:

```
PARSE_IP(web100_log_entry.connection_spec.remote_ip) BETWEEN 1068332544 AND  
1068333055
```

By adding such a clause for each IP block associated with a particular ISP and joining all such clauses with logical ORs, we can extract the NDT test results associated with that provider.

Limitations

Telescope is currently limited in that it does not take into account changes in IP subnet ownership over time. In other words, Telescope uses a single MaxMind database snapshot (specifically, the snapshot from 2014/08/04) and applies this data to all tests, regardless of the date the test took place.

We assume in the Interconnection Study that changes in IP subnet ownership over time were not significant enough during the period studied to affect analysis. A more precise solution would be to translate ISP names to IP addresses using the MaxMind database snapshot closest in time to when the individual test was performed.

Deriving Metrics

NDT is a test of network performance that collects information about data streamed over a sustained TCP connection. NDT results contain very fine-grained details about network performance across a given connection, made up largely by TCP state variables gathered within a web100-instrumented server kernel². We derive meaningful metrics from these variables using simple arithmetic, detailed below:

Metric	Derivation from BigQuery fields
---------------	--

Minimum Round Trip Time	= web100_log_entry.snap.MinRTT
Average Round Trip Time	= web100_log_entry.snap.SumRTT divided by web100_log_entry.snap.CountRTT
Download Throughput	= web100_log_entry.snap.HCThruOctetsAcked divided by (web100_log_entry.snap.SndLimTimeRwin + web100_log_entry.snap.SndLimTimeCwnd + web100_log_entry.snap.SndLimTimeSnd)
Upload Throughput	= web100_log_entry.snap.HCThruOctetsReceived divided by web100_log_entry.snap.Duration
Packet Retransmit Rate	= web100_log_entry.snap.SegsRetrans divided by web100_log_entry.snap.DataSegsOut

Filtering Invalid Tests

M-Lab archives all collected data, and makes it all available publicly. This means that the M-Lab dataset includes tests whose results are erroneous. There are several reasons why a test result could be invalid: the test could be interrupted before completion or it never achieved equilibrium on the network, which prevents an accurate measure of performance. Telescope uses the same validation methods to eliminate invalid results as [M-Lab Public Data Explorer](#) (PDE). The [PDE wiki page](#) describes these validation steps in detail.

While the PDE documentation implements its filters entirely in BigQuery SQL, Telescope implements these filters in client-side code. Telescope can be configured to output its generated BigQuery query, but it is important to note that this query yields both valid and invalid results. Telescope performs additional processing on the query results to remove invalid results before it outputs data.

Filtering Repeated Tests

M-Lab does not place limits on how frequently users can run NDT tests. The data therefore includes instances where a single IP is associated with multiple tests in one day. We considered filtering the data to limit the number of times a single IP could be included in the results per day. Upon evaluation of the numbers for recurrent testing from single sources, we concluded that repeated tests were not significant enough to affect our analysis. In a typical day, over 80% of M-Lab's tests came from IPs that perform four or fewer NDT measurements. There are outliers that submit more than 100 tests per day from a single IP, but these generally make up less than 2% of the results.

Visualization

To create visualizations of the collected data, we aggregated results based on what effects we wished to observe. To illustrate diurnal effects, we aggregated results by hour of day; to show day-by-day changes over time, we aggregated results by day.

Choice of Measurement Statistic

Observatory's visualizations represent results by displaying the median value of each aggregated unit. We chose to use the median value because we believe it most accurately reflects the experience of the typical end user, is most understandable to the largest possible audience, and is an acceptable statistical method for basic normalization. We considered using other statistics, such as the mean value or the value at different significant percentile points (e.g. 90th percentile, 95th percentile), but ultimately decided against these, as outliers in the data affect these other statistics more significantly, obscuring the results for typical usage.

Excluding Datasets with Insufficient Samples

Not all datasets have a significant number of samples for each aggregation unit. There are a variety of reasons why this can occur, such as:

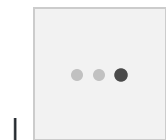
- there is low network activity in the time window,
- the access ISP has a low number of users in the specified region,
- the access ISP or M-Lab measurement point experienced downtime and did not collect test results for a given time period.

Because the visualizations are intended to make M-Lab data clear to a non-technical audience, it was important to exclude results that were misleading due to low sample sizes. For this reason, we excluded graphs from the Internet Observatory when over 20% of data points were derived from fewer than 50 samples. In the graphs that the Internet Observatory displays, it represents data points with a lighter, dashed line if the sample size is below 50. We chose this cutoff as it was a statistically-conservative minimum and allowed us to visualize a large portion of the collected data.

Analysis of Collected Data

Interpreting Patterns in the Data

Under ideal performance conditions, all of the metrics NDT collects should be essentially uniform over time. In a network without congestion, these metrics should be consistent with no variations by time-of-day or from day-to-day (demonstrated in Figure 1). Furthermore, we should expect an upward trend over the very long term reflecting various service and equipment upgrades. If there is insufficient capacity in some shared portion of the network, there will be performance degradation during peak daily load. The data rate will drop, while the round trip time and/or packet retransmission rate will rise. This effect is most easily observed in a diurnal plot (Figure 2), where data from a period of time is aggregated by hour of the day and plotted on a 24-hour scale.



(Figure 1) Example of diurnal performance patterns in ideal network conditions (based on artificial data). Network performance is not affected by the ebbs and flows of other users' traffic, since all shared links having sufficient capacity.



(Figure 2) Example of network performance in conditions where insufficient network capacity leads to differences between periods of high-user traffic and periods of low-user traffic (based on artificial data). What is important in this example is the way in which degraded performance is most acute during the times when the most people are online and using the network.

If the load continues to rise across the network(s), and ISPs managing the network(s) fail to upgrade their capacity, the periods of performance degradation will intensify. As more users attempt to retrieve more content over a fixed amount of network capacity, the duration of degraded performance will become longer. Quality of service will decline and download throughput rates will fall lower across more hours of the day, until eventually the diurnal performance plots have no intervals of normal performance, even during off-peak loads. If the ISP is not upgrading capacity, the best performing hours of the day (off-peak) typically remain fairly stable until the situation is very bad, while the worst performing hours of the day (during peak load) will become steadily worse. Daily aggregates (e.g. mean, median) also drop.

It is important to note that diurnal variations nearly always reflect human behavior³. For some metrics, such as network load (not measured by M-Lab), these variations are natural and expected. For others, such as performance, the variations are symptoms of insufficient capacity, and suggest some level of harm to the users.

Inferring Interconnection-related congestion

Key to studying the effects of interconnection between ISPs is ruling out congestion in one or another of the interconnecting ISPs' networks that by *itself* explains performance variation. All of the data in this study comes from test traffic that traversed at least two different ISPs (one transit, one access) and cross an interconnection between.

If non-congested tests are present when testing from a given access ISP, we can rule out the access ISPs network as the sole cause of problems. The same logic applies to tests against a given transit ISP. Often, the pattern of observed performance degradation is quite distinctive: the start and end time can be unique relative to

congestion elsewhere in the network, or exhibits synchronized changes – the same hills and valleys.

Large, national ISPs nearly always have multiple interconnections. It is possible that some of the interconnections have sufficient capacity and others do not. For M-Lab, each client request is routed to the closest M-Lab measurement point by mlab-ns (M-Lab's name server). To the extent that ISPs are minimizing path lengths (best practice), client server pairs that are physically close to each other are likely to use the closest interconnection. Future study could interrogate the routes taken, and understand more clearly which routes between access ISP/transit ISP pairs appear to cross problematic interconnection points and where these points are. For now, we are including the possibility of circuitous routing as an artifact of consumer performance (circuitous routing would impact M-Lab and non-M-Lab traffic alike), and are looking simply at the pairwise relationship between specific access ISPs and specific transit ISPs, without pinpointing the location of a given problem to one or another specific interconnection point between the two.

It's important to note that while we can pinpoint interconnection relationships between ISPs as a factor in observed congestion, we cannot know who did what/when, or what non-technical factors are involved. The multitude of potential causes – legal, logistical, or economic – are beyond the reach of network measurement. This opacity extends into why ISPs interconnect at particular locations, or whether they interconnect at all, or what third-parties may or may not be involved. As a result, attribution and identification of motives are outside of the scope of this research.

This said, there is one extremely important underlying factor that can be deduced: if a pair of ISPs have multiple parallel interconnects that are in separate geographical locations and share no physical infrastructure, but exhibit similar patterns of degradation, then it is very unlikely that the root cause is technical. The most likely causes are business, financial or political, e.g. one of the ISPs is refusing to upgrade the interconnects, presumably because the other can't (or won't) meet the proposed contract terms.

Footnotes

¹ It is possible for NDT clients to manually select a geographically non-proximate server, though most M-Lab clients find NDT servers by using mlab-ns.

2 “Web100: Extended TCP Instrumentation” -

<http://www.web100.org/docs/mathis03web100.pdf>

3 The exceptions are computer-to-computer applications that run on a fixed schedule, usually during off-peak times to minimize interference with human activity.