About the Author

Chris Ritzo is a senior technologist with the Open Technology Institute at New America. Ritzo provides technical support, web development, information architecture, research, partner engagement, and project management to OTI for Measurement Lab and other initiatives. Ritzo hold Masters degrees in Information Science and Communication from the University of Illinois.

Broadly interested in technology’s impact on and use within communities, Chris considers his work grounded in his past experiences with the maker movement, community organizations, grassroots media, cooperatives, community access television, and indymedia. Chris’ work examines the intersections of technology and public institutions, particularly libraries and schools. Ritzo is a proponent of open science and open source projects as a mentor and participant in the global Mozilla community. Prior to joining New America, Ritzo has worked in teaching, training, and support roles in the higher education and K12 sectors.

Acknowledgments

The author would like to thank current and former colleagues who contributed to the success of this work. Jordan McCarthy and Fernanda Lavalle contributed initial hardware research and testing, software development and field support. Georgia Bullen, Matt Mathis, and Peter Boothe were essential in the development of the research design and analyses, and provided iterative feedback and support throughout the project. Lindsey Tepe was an awesome collaborator throughout the pilot, bringing a depth of knowledge and context which helped inform and improve this work. An immense thanks to Alison Yost for her meticulous review, editorial, layout and general communications leadership. And lastly, a huge thank you to Elizabeth Hoover, John Crites, Dan Foreman, and all the IT staff at ACPS, whose trust and support on this project was absolutely integral to our success.

About New America

New America is committed to renewing American politics, prosperity, and purpose in the Digital Age. We generate big ideas, bridge the gap between technology and policy, and curate broad public conversation. We combine the best of a policy research institute, technology laboratory, public forum, media platform, and a venture capital fund for ideas. We are a distinctive community of thinkers, writers, researchers, technologists, and community activists who believe deeply in the possibility of American renewal.

Find out more at newamerica.org/our-story.

About OTI

The Open Technology Institute (OTI) works at the intersection of technology and policy to ensure that every community has equitable access to digital technology and its benefits. We promote universal access to communications technologies that are both open and secure, using a multidisciplinary approach that brings together advocates, researchers, organizers, and innovators.
Contents

Executive Summary 2
Introduction 5
Pilot Study Design 7
  Partner Selection and Research Design 7
  Qualitative Data Collection 8
  Quantitative Data Collection 8
Quantitative Data Analysis Findings 13
  Bandwidth Caps Used for Quality of Service (Qos) 13
  Comparing Measurements from QoS Limited vs. Non-QoS Limited Connections 16
  Relationship Between Bandwidth and Latency 19
  Latency of the Path to Content and Differences in Upstream Peering 22
Key Findings 27
  School Districts Lack Network Measurement Tools 27
  School Networks Present Unique Technical Challenges for Network Measurement 27
  Network Management Practices Should Be Considered in Any Measurement Program 27
  Upstream ISP Peering May Affect School Network Performance 28
  Performance Measurements Should be Compared with Data on Network Capacity 28
Directions for Future Work and Research 29
  Technical Challenges 29
  Refining the Research Methodology and Data Collection System 30
  Next Steps: Community Collaboration 32
Appendix: About M-Lab’s Broadband Performance Measurement Data 33
  About Measurement Lab 33
  About the Network Diagnostic Tool 33
Notes 34
The Measuring Broadband in Schools pilot study was conducted in 2015–16 by staff at New America’s Open Technology Institute (OTI) and Education Policy program, in partnership with the IT staff at Alexandria City Public Schools (ACPS) in Alexandria, VA. Using the open source tools and infrastructure provided by Measurement Lab (M-Lab), 16,696 Network Diagnostic Tool (NDT) measurements were conducted from ACPS school connections over a two-month period during the spring of 2016.

Quantitative data analysis was completed using R Studio and published openly in a GitHub repository. Initial analysis examined basic upload, download and other metrics across classrooms and non-classrooms, and confirmed the use of quality of service (QoS) rules on the network for classrooms, limiting download speeds to ~100 Mbps and upload speeds to ~70 Mbps. Further examination of the bandwidth and latency of QoS limited and non-QoS limited connections revealed clusters of high and low latency in the test results. Possible causes of latency such as upstream peering arrangements were explored by segmenting data by the upstream transit provider of each M-Lab server.

Summary of Key Outcomes

- **School districts lack network measurement tools.** New America researchers hypothesized that school districts lack much-needed network measurement data, and this was reinforced by IT staff at ACPS. A system to conduct network testing and methods to analyze the data would be very useful.

- **School networks present unique technical challenges for network measurement.** This pilot helped uncover technical hurdles to measuring network health within a highly managed network. From opening firewall ports to testing device drivers, the study has enumerated several lessons learned that will be helpful for future work.

- **Network management practices should be considered in any measurement program.** Initial data analysis found Quality of Service (QoS) network management rules had placed bandwidth limits on most classroom connections. Inventorying the network and understanding what network management practices are in place is paramount in the design of any measurement system.

- **Upstream ISP peering may affect school network performance.** The issues teachers or students experience
when accessing online content, testing platforms, or other online services are typically thought to be caused by poor performance of the school network or available bandwidth. However, our findings confirm that issues accessing content online can be caused by factors beyond the control of the school. The path to content and the hosting/peering arrangements of content providers and Internet Service Providers (ISPs) also affect the delivery of content, notwithstanding the considerations schools make in provisioning their networks.

- **Performance measurements should be compared with data on network capacity.**

  While measurements of internet speeds are important, understanding them in the context of network capacity is still lacking. Client measurements AND data on the utilization capacity of the network would help to demonstrate how client measurements relate to overall capacity.

The outcomes of the pilot, as well as the challenges encountered in conducting it, will contribute to future research and iterative design of the data collection system. Recommendations for improvements to the system largely concern scalability. Ideas to refine the prototype data collection system were identified, both for the device placed in schools and the server that coordinated the test schedule. The direction for future work on the measurement system largely centers on device management, and remote administration, but also includes ideas for additional tests to conduct and data to collect for comparison and analysis. A more full-featured, cloud-based service to replace the server component used in the pilot is recommended. Finally, New America Researchers discuss the formation of a community of interest around a network measurement and analysis toolkit for schools and public institutions, which New America researchers are engaging in with members of the academic community and with the Internet2, U.S. UCAN program.²

---

**Alexandria, Virginia**

*A suburban district located five miles south of Washington, DC*

<table>
<thead>
<tr>
<th>Total students</th>
<th>14,216</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total teachers</td>
<td>1,185</td>
</tr>
<tr>
<td>Number of schools</td>
<td>21</td>
</tr>
<tr>
<td>Students on FRPL</td>
<td>59 percent</td>
</tr>
<tr>
<td>Internet access</td>
<td>1 Gbps lit fiber connection</td>
</tr>
</tbody>
</table>

In schools across the United States, IT departments are routinely tasked with supporting teachers as they move toward more technology-centric instructional environments. It may seem obvious that this can only be done with a foundation of robust broadband infrastructure. In practice, however, schools don’t always know the state of their infrastructure, or how to best improve it. The challenges that school administrators face when budgeting for and deploying technology vary widely, as do their approaches to supporting its use within their schools. Measuring and assessing network health is a critical challenge facing public schools as they plan for both today’s and tomorrow’s broadband needs.

During the 2015–16 school year, New America’s Open Technology Institute (OTI) and Education Policy program partnered with the IT staff at Alexandria City Public Schools (ACPS) in Virginia to conduct a pilot study as a first step toward addressing this challenge. The study, Measuring Broadband in Schools, looked at the complexities of understanding network capacity in education institutions, and sought to better understand the challenges of measuring network capacity at the point of use in schools.

The study leveraged the open source tests, tools, and open data provided by Measurement Lab (M-Lab) to measure broadband speeds in Alexandria’s classrooms. M-Lab is an open data platform for internet measurement that is run by OTI in partnership with Google and Princeton’s Planet Lab. New America researchers prototyped a device-based measurement system, and, over a two-month period during the spring 2016 school semester, developed and installed small computers to collect network performance measurements from classroom connections in multiple ACPS school buildings and classrooms. Using a well-known internet measurement test, the Network Diagnostic Tool (NDT), the research team gathered quantitative data on the performance of the network. The goal was to pilot a low-cost, open-source, and reproducible method for assessing the broadband infrastructure in an educational setting. This work built upon OTI’s previous policy recommendations to the Federal Communications Commission (FCC), advocating for open and accessible measurement and evaluation tools within schools and libraries receiving E-rate funding. The E-rate program, which was authorized with the Telecommunications Act of 1996, subsidizes some of the cost of telecommunications and information services for schools and libraries. In 2014, the
FCC increased funding to $3.9 billion annually to schools and libraries to meet their goal of high-speed connectivity, along with another $2 billion in multi-year funding to support Wi-Fi upgrades. To ensure these connectivity goals are met, school districts and other public institutions need data and methods of analysis to understand how their networks perform and where capacity may need to grow in order to support the needs of students and instructional staff.

Within school districts, IT managers and administrators often have the dual role of architecting and implementing technology solutions as well as assessing, planning, and budgeting for the future. Balancing day-to-day infrastructure and instructional technology support with data gathering and assessment is challenging, even for well-staffed districts. Data about network performance, however, are critical to the annual funding requests districts make to the E-rate program. The challenge of building, maintaining, and upgrading schools’ broadband networks and their connection to the internet is no small task. In addition, consistently and accurately assessing capacity in order to balance instructional needs with available technology budgets is always a challenge.

As the use of technology to supplement instruction in PreK-12 education has become more common, the need for network performance data at the classroom level has intensified. The internet is used for instruction in increasingly diverse ways, for everything from assessment and student performance tracking, to one-to-one initiatives that support flipped classrooms and blended learning instruction. As teachers come to depend upon these new technologies for daily instruction, inadequate service can significantly disrupt teaching and learning. This pilot project is intended to be a prototype toolkit to address the need for classroom level speed data, that can be extended and improved through future collaboration with relevant communities of interest.

OTI has long advocated that the FCC “improve data collection and measurement tools so that E-rate participants, researchers and policymakers can better understand and analyze the program.” The challenge of building, maintaining, and upgrading schools’ broadband networks and their connection to the internet is no small task. In addition, consistently and accurately assessing capacity in order to balance instructional needs with available technology budgets is always a challenge.

As the use of technology to supplement instruction in PreK-12 education has become more common, the need for network performance data at the classroom level has intensified. The internet is used for instruction in increasingly diverse ways, for everything from assessment and student performance tracking, to one-to-one initiatives that support flipped classrooms and blended learning instruction. As teachers come to depend upon these new technologies for daily instruction, inadequate service can significantly disrupt teaching and learning. This pilot project is intended to be a prototype toolkit to address the need for classroom level speed data, that can be extended and improved through future collaboration with relevant communities of interest.
In designing this pilot study, the New America team researched potential partners, designed qualitative data collection instruments, and prototyped and tested a measurement device to collect quantitative data about network performance. This section describes the study design and details the processes used for data collection.

**Partner Selection and Research Design**

In selecting a school partner for the pilot, several districts in Maryland, Virginia and Washington, D.C. were considered. Some of the key attributes for selecting a pilot partner included:

- District level partnership with administrative support
- Middle or high school with scheduled internet use
- Small to mid-sized school, with 1,000-4,000 students
- On-site technology staff who could assist with support
- Physical architectures representative of typical schools
- Proximity to New America offices in Washington, D.C.

Based on these criteria, Alexandria City Public Schools (ACPS) was selected as a viable partner to approach. This choice seemed ideal since the Commonwealth of Virginia has been an early adopter of online testing and ACPS has been a leader in integrating technology and instruction. An important component of this research was a strong partnership with the district. ACPS was particularly interested in participating to explore new ways to assess the capacity of its network. In the 2016 school year, ACPS purchased 2Gbps internet service to its wide area network (WAN), which serves 14,216 students in 21 schools.

New America researchers obtained institutional review board exemption for this research from a third party, Chesapeake IRB, based in Columbia, MD, per the research requirements of ACPS’s Office of Accountability.

New America researchers worked with ACPS staff to identify three schools to host the measurements: one elementary school, one middle school, and one high school. Within each school, three locations were selected to have the network speed and health measured periodically over the course of two months. At each of these locations a small computer
running the NDT measurement test was connected to the network and configured to run scheduled tests automatically. New America researchers coordinated with ACPS IT staff to ensure the devices functioned properly and were accessible remotely for monitoring progress, and made regular visits to the schools to diagnose issues that arose during the course of data collection.

During the testing period, the devices functioned just as any computer in use in the schools. NDT tests were scheduled to occur randomly throughout a 24-hour period, to distribute data collection throughout the day, allowing data to be captured during peak and off-peak times. The NDT test uses synthetic data to collect information about the connection quality, and specifically does not collect information about internet traffic, such as emails, web searches, videos viewed, or any personally identifiable information. Supporting information about NDT and M-Lab can be found in the Appendix.

Over the course of the pilot, many other schools expressed interested in participating in the next iteration of the project, suggesting that there is a clear need and appetite for this work if there are resources in the future to scale to more districts, schools, and locations.

**Qualitative Data Collection**

While this case study details the findings of the quantitative measurement data described in the previous section, qualitative research was also carried out to provide context to the quantitative measurement of the network. ACPS staff sent out a survey to teachers, librarians, and other instructional staff district-wide. The surveys and other informal conversations and interviews with ACPS staff provided deeper context to the real experiences of school network use and complement the broadband performance measurement data. Comparing quantitative and qualitative data sets made it possible to understand network performance trends in the context of instructional and student need. This paper focuses on the quantitative data analysis, and was produced in tandem with the analysis of the qualitative data.11

**Quantitative Data Collection**

The data collection system was designed to use a set of small computers running NDT and a cloud-based server that scheduled tests at randomized times throughout the day. These measurement devices were placed in classrooms at three schools within the district, and collected regular speed and response time data over the data collection period (two months during the spring 2016 semester). The intention was to prototype a data collection system that could measure network performance over time, rather than just providing a one-time snapshot. The diagram on the next page illustrates how M-Lab’s tests are conducted. Note that M-Lab’s servers are connected to the internet via transit providers. This provides M-Lab with a unique vantage point for network measurement.
Diagram of an M-Lab test

STEP 1
A user, connected to the Internet by one or another access ISP, runs a test.

STEP 2
The test sends traffic from their device to the closest M-lab measurement point, and back.

STEP 3
This measures performance from the access ISP into the Internet (not just within the access ISP's network).

STEP 4
The measurements generated by a given test are shown to the user, and put into the public domain.
**Client Hardware and Software**

The device selected to conduct the measurements for the study was the ODROID C1+, running Ubuntu Linux. Each of these small computers were configured with a static IP address and connected to the school network, allowing an ACPS network administrator to provide remote access to each device.

The data collection design required automated, scheduled testing at randomized times, with the goal of gathering a good distribution of tests across each hour of the day. It was important that tests were run a specified number of times throughout a given time increment (hour) and across all hours of the day, and that researchers had a way to easily query the M-Lab dataset for the measurements that the test devices had completed.

Running the tests also required the development of a prototype Server-Client system, mlab-governor-client. The governor server provided a simple means of authenticating clients from a given partner, location, and device; scheduled randomized times for NDT tests to be run for authenticated clients using an exponential distribution; and stored test ID information provided by each client in a MongoDB database. The client script provided a “wrapper” for the NDT test on each client device. This script contained functions to request a randomized schedule of tests from the governor server once a day, and ran NDT on that schedule. For each test attempt, the client script generated a random test ID, ran an NDT test, and saved the successful test ID to a log file on the device. When the schedule for that day was completed, the script reconnected to the governor server to report completed test IDs and requested a new schedule.

**Data Collection Method**

Automated NDT tests from each device were run on a daily randomized schedule provided by the server for a period of two months. After the period of data collection, all devices were collected and the test ID logs were gathered from them. Since the tests run from each device represented a specific classroom and school, the log files from each device were saved separately into a series of folders corresponding to the school location and classroom. For example:

/ tcwilliams/client-3_TC_testIDs.log
/ hammond/client-5_hammond_testIDs.log

A standard set of queries was defined for the metrics: download throughput, upload throughput, round trip time and packet retransmission rate. Two additional queries for less frequently used metrics were also included: network limited ratio and client limited ratio. A Python script was used to query the M-Lab NDT dataset in BigQuery for each of the test IDs in each of the log files, automating the query and output of test results for each metric. The result was a set of six comma separated values (CSV) files for each device:

/ tcwilliams/client-2_TC_download_query.csv
/ tcwilliams/client-2_TC_upload_query.csv
/ tcwilliams/client-2_TC_rtt_query.csv
/ tcwilliams/client-2_TC_packet_retrans_query.csv
/ tcwilliams/client-2_TC_net_lim_ratio_query.csv
/ tcwilliams/client-2_TC_client_lim_ratio_query.csv

In the initial data analysis, the CSV files for each metric were combined into a series of spreadsheets, one for each metric, with additional metadata columns added to keep reference to the original device ID, the school and classroom in which the device was placed, and the type of media used to connect the device to the school network (wired/wireless). Additional columns were added, weekday and hour, calculated from the timestamp for each test. Weekday is a numeric value from 1-7 for the day of the week, beginning with Sunday; and hour is the hour of the day in which a test was run, from 1-24. The analysis of the results was focused on the download throughput, upload throughput, round...
trip time and packet retransmission rate metrics. While the data collected for the network limited ratio and client limited ratio metrics were retained, it was determined that these metrics would likely not be as useful without additional information from the point of data collection.

All statistical analysis and visualization was conducted using R Studio.\(^{19}\) The CSV files for each metric were imported as data frames into an R Studio project, which then was placed into a GitHub repository to provide version control, enable team collaboration, and to provide replicability.\(^{20}\) Charts and graphs generated in subsequent sections and appendices reference files contained in the R Studio project within the GitHub repository.

**Data Cleaning Process**

During the initial data analysis, we uncovered several issues which necessitated querying for the data a second time. First, while examining each metric separately was fine for the initial exploratory analysis, it complicated the ability to compare variables—for example, comparing latency and speed metrics. Additionally, the initial queries calculated each metric during the query itself, and as such did not retrieve the individual fields making up each metric. Lastly, time values in M-Lab data are stored as Unix timestamps in Universal Coordinated Time (UTC). The initial data did not adjust for the fact that the data for this study originated in Eastern Standard Time. Our measurements were also taken before and after the 2:00am change to Daylight Savings Time on Sunday, March 13, 2016. Additionally, after the start of the project, M-Lab added a new field to its NDT data, blacklist_flags, which tagged any test data that could have been affected by peak loads at M-Lab server measurement points.\(^{21}\)

All of these factors necessitated re-querying for all raw data, and the following data cleaning process resulted in the final dataset for analysis.

1. All raw data fields that make up each computed metric were queried, returning NDT tests between January 1 and May 1, 2016, matching for the public IP addresses for ACPS’ wide area network.\(^{22}\) The following fields were calculated from the raw data values using calculations provided by M-Lab documentation, and added as columns to each row:\(^ {23}\)

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Datetime(__)UTC</td>
<td>Date and time conversion in UTC time zone</td>
</tr>
<tr>
<td>Datetime(__)EST</td>
<td>Date and time conversion in EST time zone</td>
</tr>
<tr>
<td>Hour</td>
<td>Hour of the day calculated from datetime(__)EST</td>
</tr>
<tr>
<td>Weekday</td>
<td>Numeric day of the week derived from datetime(__)EST</td>
</tr>
<tr>
<td>DownloadThroughput</td>
<td>The measured download speed in Mbps</td>
</tr>
<tr>
<td>UploadThroughput</td>
<td>The measured upload speed in Mbps</td>
</tr>
<tr>
<td>MinRTT</td>
<td>The minimum Round Trip Time, a measure of connection latency without traffic</td>
</tr>
<tr>
<td>AvgRTT</td>
<td>The average Round Trip Time, a measure of connection latency during the test</td>
</tr>
<tr>
<td>PacketRetransRate</td>
<td>The ratio between re-transmitted packets and all transmitted packets</td>
</tr>
</tbody>
</table>
NetworkLimRatio

During the server-to-client portion of the test, the amount of time in which the network was congested

ClientLimRatio

During the server-to-client portion of the test, the amount of time in which the client device limited the data that could be received

ReceiverWindowScale

The maximum amount of received data that can be buffered at one time by a client test

2. Rows with a malformed or missing client_testid were removed to a separate tab, no-client-testid. These were likely tests run from alternate M-Lab NDT clients from within the school district, such as the M-Lab Measure Chrome Extension, but were not part of the official study. The previously separate upload and download test results were imported into additional tabs for comparison: orig-download-test-results and orig-upload-test-results.

3. A lookup function was written on the field mlab_test_id, which occurs in both data sets, to extract metadata from the originally queried data into the newly queried raw results data. This associated the newly queried result rows with the metadata from their originating client testid, device, room, and grade level.

4. After this matching process, rows without matching metadata (device, school, room, etc.) were moved into a separate tab, no-device-metadata.

5. The remaining data were then sorted by the column datetime_EST, and for rows where the date was greater than 03/13/2016 02:00:00, the timezone offset was changed from -5 to -4 to account for the Daylight Savings Time change. To indicate which rows were adjusted, an additional column was added, dst_flag. Adjusted records with test dates after 03/13/2016 02:00:00 were marked with a 1. Non-adjusted rows were marked with a 0.

6. The remaining rows were exported as a CSV file to be used in the analysis.24 This included 9088 distinct upload tests and 7608 download tests between 2/18/2016 and 4/29/2016.
QUANTITATIVE DATA ANALYSIS
FINDINGS

As a pilot research project meant to identify, test, and refine methods for analyzing network performance test data, the initial data analysis started with a search to find any skewed or unusual distributions in the data. To encourage replicability and transparency in the project, the work done in R Studio was published in a GitHub repository.25, 26

Initial analysis confirmed that ACPS was using bandwidth caps for quality of service (QoS), and provided a means to compare connections using QoS and non-QoS caps. Differences between bandwidth and latency were then examined, and finally other causes of latency such as upstream peering arrangements were explored.

**Bandwidth Caps Used for Quality of Service**

The initial data analysis confirmed the presence of network QoS bandwidth caps on devices connecting to the wired network in classrooms. In Figure 1 on the following page, showing download measurements faceted by classroom and school location, there are caps in the measurements for all elementary school classrooms. Download speeds appear limited to ~100 mbps and upload speeds limited to ~70 mbps. In high school classroom B323, there were two groupings of tests: one mirroring the caps seen in elementary schools, and another group of tests above the 100 mbps threshold. The team suspects that the measurement device in B323 was moved between student and staff network ports during the data collection period, resulting in a mix of QoS limited and non-QoS limited measurements. The measurements collected at the student help desk in the high school did not appear to have any limits in place.
Figure 1: Download Speed Measurements by Hour, Faceted by Grade Level and Room

26 | ELEMENTARY

47 | ELEMENTARY

9 | ELEMENTARY

B323 | HIGH SCHOOL

HELP DESK | HIGH SCHOOL

Download Throughput (Mbps)

Hour of the Day
Figure 2: Download Speed Measurements, High School Classroom B323, by Hour

Figure 3: Upload Speed Measurements by Classroom, by Hour, Limited to Measurements Below 100 Mbps
That QoS limits were found to be in place on most classroom connections confirms that ACPS’ network management practices were successfully being applied. For this study, it means that most of the data collected were measurements of the network controls, not measurements of the total capacity of the connection. However, since data collected in two locations did not have bandwidth controls applied to all measurements, those data reflect the uncapped speeds of the district’s internet connection. Subsequent analysis will compare QoS limited connections and non-QoS limited connections.

- **QoS Limited**: Includes data from all classrooms except high school classroom B323, where a mix of QoS limited and non-QoS limited tests were collected

- **Non-QoS Limited**: Includes data collected from the high school help desk, where no QoS limits were determined to be in place

Data from high school classroom B323 were excluded from the final analysis since they contained a mix of QoS and non-QoS limited measurements.

### Comparing Measurements from QoS Limited vs Non-QoS Limited Connections

Next, the team compared measurements from non-QoS limited connections to those that did have QoS limits in place. The tables below and on the next page show the numeric summaries for comparison. The median values for download and upload speeds show the difference in speeds, as well as the median values for latency metrics.

---

**Numeric Summary: QoS Limited Measurements**

*Values are in Megabits per second (Mbps), milliseconds (ms), or percentage (%) as indicated.*

<table>
<thead>
<tr>
<th></th>
<th>Download Speed (Mbps)</th>
<th>Upload Speed (Mbps)</th>
<th>Min RTT (ms)</th>
<th>Avg RTT (ms)</th>
<th>Packet Retrans. Rate [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.328</td>
<td>1.36</td>
<td>1.000</td>
<td>3.799</td>
<td>0.000</td>
</tr>
<tr>
<td>1st Quantile</td>
<td>78.710</td>
<td>62.92</td>
<td>3.000</td>
<td>6.662</td>
<td>0.002</td>
</tr>
<tr>
<td>Median</td>
<td>88.247</td>
<td>63.20</td>
<td>4.000</td>
<td>7.475</td>
<td>0.004</td>
</tr>
<tr>
<td>Mean</td>
<td>81.252</td>
<td>63.20</td>
<td>4.843</td>
<td>8.073</td>
<td>0.004</td>
</tr>
<tr>
<td>3rd Quantile</td>
<td>91.030</td>
<td>63.38</td>
<td>8.000</td>
<td>10.221</td>
<td>0.005</td>
</tr>
<tr>
<td>Maximum</td>
<td>92.446</td>
<td>68.55</td>
<td>28.000</td>
<td>28.432</td>
<td>0.225</td>
</tr>
</tbody>
</table>
These data show that for tests from non-QoS limited connections, the median speeds of ACPS’s wide area network were 234 mbps down and 283 mbps up. These findings emphasize that when designing a measurement initiative for assessing managed networks, an inventory of network constraints such as QoS limits should be conducted in advance, and the locations from which measurement tests are run should be chosen based on the intended measurement goals. As an example, a school might conduct measurements from a device connected to a switch near its internet connection, on an un-managed port, in order to collect baseline measurements of its network’s total capacity. And it might place multiple measurement devices inside classrooms to measure different parts of the network with potentially different QoS controls.

However, a more detailed examination of the data identified clustering patterns indicating that there may have been additional factors influencing all connections, even those not limited by QoS network controls. Some of this clustering can be seen in Figures 4-7 on the next page, primarily two clusters of tests (low/high) within the hourly distributions of non-QoS limited tests. This clustering is explored in more detail in the next section, *Relationship Between Bandwidth and Latency*.
Figure 4: QoS Limited Download Speed Measurements by Hour

Figure 5: Non-QoS Limited Download Speed Measurements by Hour

Figure 6: QoS Limited Upload Speed Measurements by Hour

Figure 7: Non-QoS Limited Upload Speed Measurements by Hour
Relationship Between Bandwidth and Latency

Whether accessing online testing program websites, or simply using web content during the course of classroom instruction, teachers and students are using valuable time to connect to the internet and access materials that have been deemed necessary for quality education. While popular metrics of internet connection performance focus mostly on download and upload speeds, the relative latency of the connection is important as well, particularly with real-time applications like video streaming. Schools regularly use online testing applications, subscribe to online content, or use rich media in the course of research and instruction, all of which require robust connection speeds and low latency.

The measurements in the previous section comparing test data from QoS limited and non-QoS limited connections suggest that classroom connections (QoS limited), in addition to having throttled speeds, also have higher general latency. MinRTT is the minimum latency measured in transmissions from the server to the client, reported in milliseconds. In comparison, AvgRTT is the average latency of data transfers from the server to the client and is calculated as the sum of round trip times sampled during the test against the total number of samples, also reported in milliseconds. Consequently AvgRTT is a latency measurement when the connection is under load during the test, and MinRTT is a latency measurement when the connection is not under load. In the numeric summaries in the previous section, we referenced the median values for the latency measurements, Average Round Trip Time (AvgRTT) and Minimum Round Trip Time (MinRTT). These values are repeated below for reference.

<table>
<thead>
<tr>
<th></th>
<th>Non-QoS Limited</th>
<th>QoS Limited</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median</strong></td>
<td>4.000 ms</td>
<td>3.000 ms</td>
</tr>
<tr>
<td><strong>Minimum RTT</strong></td>
<td>3.750 ms</td>
<td>7.475 ms</td>
</tr>
</tbody>
</table>

The charts on the following pages show download and upload measurements indexed by MinRTT values, for QoS limited tests. Very few upload measurements were below 45 mbps and most MinRTT measurements were below 20ms, so upload tests below 45 mbps are eliminated as outliers, along with all tests where MinRTT is above 20ms. For comparison, the same graphs are also shown for speed measurements indexed by MinRTT for non-QoS limited tests. Finally, a comparison of AvgRTT and MinRTT for download measurements where either metric is below 20ms is shown in Figure 12.

Students and teachers are using valuable time to connect to the internet and access materials that have been deemed necessary for quality education.
Figure 8: QoS Limited Download Measurements Indexed by MinRTT <=20ms

Figure 9: Non-QoS Limited Download Measurements Indexed by MinRTT <=20ms
Figure 10: QoS Limited Upload Measurements Indexed by MinRTT <=20ms

Figure 11: Non-QoS Limited Upload Measurements Indexed by MinRTT <=20ms
An interesting observation from the charts on the previous pages is the clear clustering of tests when download or upload throughput is paired with MinRTT. When AvgRTT is compared to MinRTT (Figure 12 above), there are similar clusters on both metrics, between 2-5ms and between 7.5-10ms. If a connection has decent speeds but high latency, the user experience of accessing content that requires low latency would be affected. For example, if a user was streaming a video, high latency on the connection would likely cause buffering, even if the speed of the connection was fast.

There are many factors influencing latency on the path to content, making cause difficult to assess. A school district purchases an internet connection, but how well the ISP is connected or “peered” with other ISPs can affect the user experience by introducing additional latency between the user and the content they are trying to reach. The next section further examines the clustering of tests seen above, and assesses potential sources of higher latency measurements by grouping these measurements by the M-Lab server the test was conducted against.

**Latency of the Path to Content and Differences in Upstream Peering**

The internet is an interconnected network of networks—all running common protocols that make the experience of browsing relatively seamless. Home users, businesses, and schools purchase an internet connection from an ISP. These ISPs provide access to the internet, by connecting their networks to transit ISPs. Transit ISPs provide connection to the internet itself, where most content we use as consumers is hosted. For example when a student conducts research online, their requests to view various websites move from their computer.
over their access ISP’s network, into transit ISP’s networks, and back. The illustration above is from M-Lab’s website, and attempts to visually describe the interconnectedness of the internet.

To provide diversity in connectivity, M-Lab deploys multiple redundant “server pods” in most metropolitan regions, with each pod connected to a different upstream transit provider available in that region (see the Appendix for more information on M-Lab). Examining M-Lab test data segmented by a named M-Lab server can show differences in measurements solely on the peering connectivity between the client ISP and the transit ISP connected to the M-Lab server.
By noting the unique server IPs for all tests conducted in the study, and combining them with the upstream transit provider for those servers, the data can be segmented by transit provider in a column called “transit.”

A table containing the unique server IPs, M-Lab server names, locations and the transit provider those servers are connected to is shown above.

<table>
<thead>
<tr>
<th>M-Lab Server IP</th>
<th>M-Lab Server Name</th>
<th>Location</th>
<th>Connected Transit Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>216.156.197.139</td>
<td>IAD01</td>
<td>Reston, VA</td>
<td>XO</td>
</tr>
<tr>
<td>38.90.140.139</td>
<td>IAD02</td>
<td>Ashburn, VA</td>
<td>Cogent</td>
</tr>
<tr>
<td>66.198.10.139</td>
<td>IAD03</td>
<td>Ashburn, VA</td>
<td>Tata</td>
</tr>
<tr>
<td>173.205.4.11</td>
<td>IAD04</td>
<td>Ashburn, VA</td>
<td>GTT</td>
</tr>
<tr>
<td>173.205.4.24</td>
<td>IAD04</td>
<td>Ashburn, VA</td>
<td>GTT</td>
</tr>
<tr>
<td>4.35.238.203</td>
<td>IAD05</td>
<td>Ashburn, VA</td>
<td>Level 3</td>
</tr>
</tbody>
</table>

The internet is an interconnected network of networks—all running common protocols that make the experience of browsing relatively seamless. Home users, businesses, and schools purchase an internet connection from an ISP. These ISPs provide access to the internet, by connecting their networks to transit ISPs. Transit ISPs provide connection to the internet itself, where most content we use as consumers is hosted.
Figure 13\textsuperscript{41}: QoS Limited Download Speeds and MinRTT Faceted by M-Lab Server/Upstream Transit Provider

Figure 14\textsuperscript{41}: Non-QoS Limited Download Speeds and MinRTT Faceted by M-Lab Server/Upstream Transit Provider

Download Throughput
This finding suggests that at the time of testing, tests conducted between the ACPS network and Level 3 had higher latency than those tests conducted through other peered transit providers. It is also an indicator that the issues teachers or students experience when accessing online content, testing platforms, or other services can be caused by other factors outside of the school’s network and available bandwidth. The path to content and the hosting/peering arrangements of content providers and ISPs affect the delivery of content, notwithstanding the considerations schools make in provisioning their networks.
School Districts Lack Network Measurement Tools

The prototype system to measure network capacity at the point of use in the classroom was validated by ACPS, confirming that tools to provide these data meet an unfilled need. ACPS staff members brainstormed other ways that they could use the pilot system, beyond the tests conducted in this pilot. For example, they suggested running four to five (or more) tests concurrently in the same classroom on the same network to see how many concurrent tests could run in a given location before the network showed signs of high load. As future projects are developed, some potentially valuable additional features to the prototype would allow partners to integrate network capacity or utilization data from vendor systems. For example, network switches use a common protocol called Simple Network Management Protocol (SNMP) and an SNMP plugin for our system could be integrated to receive switch utilization data. Similarly, a vendor system plugin could provide data on WiFi utilization. These additional data could be displayed alongside performance measurements to provide a more nuanced understanding of performance, capacity, and utilization.

School Networks Present Unique Technical Challenges For Network Measurement

This pilot study provided an opportunity to prototype new tools for running automated performance tests from multiple vantage points within a network. The proof-of-concept measurement prototype used Internet2's NDT command line client, running on a small computer connected to the school network. An important goal of the pilot was to prototype a measurement system consisting of a central server to provide scheduling and data collection from one or more devices within a school network. Placing small, unobtrusive devices in different places within a school network allowed us to measure at several different points, including the point of student use, simulating a computer in the library, a wireless laptop or tablet, or a desktop computer in the classroom.

Network Management Practices Should Be Considered In Any Measurement Program

The initial analysis found that Quality of Service (QoS) network management rules had placed bandwidth limits on most classroom connections. This is normal networking practice in a large...
institution, but this meant that most of the data for this study consisted of measurements of the network controls, not measurements of the total capacity of the connection. The result does confirm that ACPS’ network management practices were successfully being applied. However, understanding what these controls are when determining the points from which measurement takes place is paramount to measuring what one intends to measure. In this study, measurements from both QoS limited and non-QoS limited connections were collected, which was useful for comparison.

**Upstream ISP Peering May Affect School Network Performance**

The analysis found clusters of tests with higher and lower latency and higher and lower bandwidth measurements. To understand the potential cause of this clustering, the data were segmented by transit provider. By segmenting the data this way, the analysis showed that tests conducted between the ACPS network and Level 3 had higher latency than those tests conducted through other peered transit providers. Transit providers connect networks of ISPs to the internet itself, through “interconnection” or “peering” arrangements. Because M-Lab hosts its servers in the data centers where these transit provider to ISP interconnections are made, the data collected at ACPS can be segmented by different transit providers.

This finding indicates that the issues teachers or students experience when accessing online content, testing platforms, or other online services can be caused by factors beyond the control of the school, like its network and available bandwidth. The path to content and the hosting/peering arrangements of content providers and ISPs also affect the delivery of content, notwithstanding the considerations schools make in provisioning their networks.

**Performance Measurements Should Be Compared With Data On Network Capacity**

It is clear from interactions with ACPS staff and other educators that simply measuring performance of a network connection, whether at the classroom and device level or from the egress point of a school network, only provides part of the picture of whether a network is functioning at needed capacity.

ACPS showed the New America team the vendor systems it uses to provision and manage WiFi access points, as well as the utilization graphs of network hardware. M-Lab uses similar tools to monitor its server infrastructure, for example to understand the capacity of its switches over time. ACPS could see total bandwidth used over time from a particular WiFi access point, and had data on the total network utilization over time at the district and school levels. But connecting data from disparate tools and with speed/performance measurement data from client tests is no small task. The pilot identified the need to add features to the prototype measurement system, including the ability to interface with commonly used network management protocols like SNMP, as well as support plugins for vendor provided systems or at the very least provide data import capabilities. Client measurements AND utilization capacity of the network would help to demonstrate how client measurements relate to overall capacity.
**DIRECTIONS FOR FUTURE WORK AND RESEARCH**

**Technical Challenges**

The pilot study uncovered several issues which limited the scope of the collected data. Some of these challenges resulted from using a prototype data collection device and test that was not typical for devices operating within school networks. However, other issues were potentially related to school infrastructure and possibly WiFi device driver incompatibility. IT staff were able to assist in solving some connectivity issues, but as in many school districts the staff’s time and expertise varies, and additional requests for troubleshooting were limited so as not to interfere with the day-to-day operational needs of ACPS. The small Linux computers running the tests were non-standard for the district IT staff, and were solely managed by New America staff, either remotely or during several on-site visits.

Here are five technical challenges that future research should consider:

- **Ethernet ports were not always live.**
  In one school location, the ethernet ports that the devices were connected to were not active.

- **Requiring static IP addresses for measurement devices introduced added complexity.**
  Assigning static IP addresses for the devices was accommodated by ACPS IT staff, but not having this requirement would make the measurement system more flexible. It would also make this system easier to implement in a variety of settings with staff of varying degrees of expertise.

- **WiFi connections for measurement devices were not persistent.**
  WiFi connected devices presented a significant issue in this pilot. Six measurement devices were placed in each school: three wireless and
three wired. While WiFi connections to the internet worked for as many devices as possible, it is possible that the ports required to run NDT were not open to the WiFi networked devices. After a power surge or outage, the WiFi devices did not automatically reconnect. In some cases the devices would not connect to the WiFi at all, though they were very near to the access point. This indicates that there may have been device driver issues with the device’s network cards, or other network management practices may have limited the ability of the devices to connect in all locations.

- **Remote access was not available for all measurement devices.**
  Network administrators in schools typically implement more restrictive firewalls than most home networks, so asking for specific ports to be opened for remote access was non-standard. This is related to the challenges of the measurement prototype, which required static IPs. Future iterations of this collection system need a strong remote management component to allow for flexible deployments within managed networks.

- **Ports needed for running NDT were not opened for all measurement devices.**
  Services provided on the internet all use specific “ports,” which are the communication channels that allow users to access those services. When a webpage like [https://www.newamerica.org](https://www.newamerica.org) is loaded, the web browser uses port 443. When conducting an NDT test, the client (code on the device) uses ports 3000-3010 as a kind of handshake with the server. The client requests a test and the server responds, providing a random port between 49152-65535 to conduct the actual test. While this is typically not an issue in most home networks, school networks are more actively secured through firewall rules, and the ports NDT needed were not initially open. ACPS staff were able to open the required ports, but this request is outside of the norm for many IT staff managing school networks. It would be advantageous if NDT could run on standard ports that are typically open in school networks, but this might not be possible. In any case, opening ports will be necessary in most school networks and tests should be confirmed to be working end-to-end.

**Refining the Research Methodology and Data Collection System**

The challenges identified in the initial pilot will inform the next stages of research and development. The data collection system, described in detail above, consisted of the device or client, the software running on it, and the server which provided the devices with scheduling information. There are several potential improvements that could be made to the devices running tests, as well as improvements to the server component of the data collection system.

**Potential Improvements to the Client / Device**

Here are three potential improvements that could be made to the client/device:

- **Use a dynamic IP address.**
  Our prototype data collection system had networking requirements that would not be scalable to large deployments. Though the choice of a standard hardware device from which to run tests was advantageous for numerous reasons, the operating system and software running on the measurement device needs to have further consideration and testing. Rather than requiring a static IP, it would be preferable for the devices to receive a dynamic IP address, like all other computers on the school network.

- **Use an overall device management utility for scalability and repeatability.**
  In the pilot, everything on the measurement device was enabled manually: the static IP, the startup client script, and the configuration. This included setting up the devices to begin with, as well as the means of accessing them remotely. In the future, the software on the device which coordinated scheduled testing might itself
provide the remote management capability over standard ports that are already enabled on most networks. The core need here is to know status of each deployed device, which could be accomplished by modifying the client script to send additional information about its status to the server. The server then becomes the place where device configurations are stored, where the client requests that configuration, where the client sends its test IDs, and where the client sends information about its state for monitoring purposes. Remote management is needed for scaling up to more devices, or to undertake larger studies using this method. A device management platform would need to provide everything from initial setup, configuration, and upgrading, to remote management and data visualization.

- **Choose alternate tests or multiple tests to provide more data.**
  The pilot project used the NDT command line client, which provided some flexibility in coding an initial prototype. However, using more standard browser-based tools like the M-Lab Measure Chrome extension, or supporting a variety of test versions within the system would provide more flexibility in device choice. Supporting multiple tests from the same device would also provide more data for comparison and analysis, allowing for a more detailed analysis as to connectivity, network management, or other infrastructure issues.

**Potential Improvements to the Server**

Enhancements to the server side would provide additional features to make a complete system of evaluation more meaningful. For this pilot, the server only provided the most basic information on scheduling to each client upon request.

School administrators might wish to deploy the server on their own, perhaps within their own networks, which would necessitate the need to provide custom configuration options. For example, some districts might want to use static IP addresses, and others might not. Some might wish to place their server outside their network, and others inside the school WAN. There are additional tools available on M-Lab beyond the NDT that might be useful for future projects. For example, the Perfsonar suite of tests has a strong user base in the research and education community, and another M-Lab hosted test, Bismark, provides several tests which could also be used for comparative performance data.

**Here are five potential improvements that could be made to the server:**

- **Use a robust device provisioning and management platform.**
  The manual provisioning and remote management methods used in this pilot study are not scalable. A platform or service that would allow both device setup and ongoing remote management is needed to truly support broad adoption and use. This device provisioning and management platform could be used just to setup and manage devices, but would be more robust if it also was a central service for additional features, analysis, etc. as described below.

- **Provide a partner authentication and authorization.**
  A service is needed to allow multiple school districts to create accounts where they would centrally manage configurations for devices to be deployed in their schools. Configurations would include scheduling and other required information for the measurement devices, as well as customizable metadata allowing for each device, groups of devices, or locations to be tagged for later use. This service would also provide a means to view and manage already deployed devices.

- **Design a data gathering process to get test information.**
  Having a central place for the data submitted by test devices is key. In the pilot, New America researchers queried the M-Lab database for data directly. An updated server platform would
obtain the test data submitted by devices and store it in a database for visualization, analysis, and download.

- **Deploy a visualization dashboard that presents data analysis.**
  In this pilot, all visualization and analysis was done separately. In order to be more broadly useful, the methods of analysis and visualization used in this pilot should be built into the server itself. This would provide schools an immediate method of analysis on the data their devices submit.

- **Add data import functions, support for additional tests, and a plug-in architecture to connect additional data to performance measurements.**
  In addition to network performance tests run from devices in this system, the server would also need to provide means to upload or import data from other IT systems in use. For example, the network capacity data from switches on a school network would allow comparison of measured performance with overall network capacity at the same time. Another example would be data from various enterprise WiFi controllers. The system could use a plugin architecture that would allow schools to use data from a variety of vendor provided systems.

### Next Steps: Community Collaboration

To build a truly viable network measurement and assessment product that schools and other public institutions could use to assess their own connectivity is no small task. New America’s intent was to build a proof-of-concept measurement system and methodology, and build on it with others in a community of practice. In early 2017, Internet2 and New America began an initiative called the Community Anchor Institutions Internet Measurement in Schools Working Group. By coordinating with interested participants from the K-20 community nationwide, Internet2 and New America’s OTI are planning for a next stage prototype measurement system based on the work from this pilot and the ideas generated in monthly working group calls. The group aims to build a minimum viable product for a second and broader pilot measurement program as a toolkit that could be scaled up, one that could grow sustainably through a working group and community of practice around network measurement and assessment.
APPENDIX: ABOUT M-LAB’S BROADBAND PERFORMANCE MEASUREMENT DATA

About Measurement Lab
Measurement Lab (M-Lab) is the largest open source Internet measurement effort in the world. (Open source refers to source code that is made available for free so that anyone can use or modify it to fit his or her own purposes.) M-Lab provides performance tests that help consumers develop an accurate picture of their Internet service by offering a state-of-the-art server platform that supports many measurement needs. These measurements are collected and then released to the public for use by policymakers, researchers, and others who are interested in Internet issues.

About the Network Diagnostic Tool
As a single-threaded TCP throughput measurement, the Network Diagnostic Tool provides an assessment of the overall quality of service of a broadband connection. NDT assesses the ability of a connection to support particular applications, such as high-definition video or voice over IP services, and it detects potential problems with access. NDT measures a rich set of properties of a connection over a fixed period of time based on a highly-instrumented kernel stack on the server, which records latency, retransmissions, timeouts, and throughput, among nearly a hundred other metrics. NDT runs for ten seconds for both the download and upload segments of the test, and therefore the amount of data consumed during the test will be a product of the throughput rate of the client. M-Lab’s early adoption of the BBR congestion control mechanism will also reduce test time and improve accuracy, matching the results of multi-threaded tests.

NDT is a highly-portable measurement tool due to the simplicity of its design, which has minimal client requirements. All of the forms of NDT deployments, whether in the browser, on mobile, or through a hardware device, are functionally similar and produce comparable data. NDT can be run in the browser without the need for additional plugins due to the availability of an HTML5/Javascript client. NDT clients are available in several programming languages (such as Java, C, and C++), and it has been independently integrated into various pieces of consumer software. In order to support the development of other desktop software and mobile applications that leverage NDT, M-Lab also supported its inclusion in MeasurementKit. The library enables any software developer to run M-Lab tests without having to write the test themselves, nor having to know the details of how it works.
Notes


3 Concurrently we conducted a district-wide survey and asked classroom teachers to complete internet usage logs, to provide context around instructional use with the raw measurement data. See: Lindsey Tepe, “From Online Testing to Online Learning in VA,” Education Policy Program, New America, June 6, 2017, https://www.newamerica.org/education-policy/policy-papers/online-learning/


27 “Broadband in Schools Github Repository - Figure 1,” Open Technology Institute, accessed May 9, 2017, https://github.com/opentechinstitute/bb-schools-analysis/blob/master/initial-data-analysis.R#L104
29 “Broadband in Schools Github Repository - Figure 3,” Open Technology Institute, accessed May 9, 2017, https://github.com/opentechinstitute/bb-schools-analysis/blob/master/initial-data-analysis.R#L112
33 “Broadband in Schools Github Repository - Figure 7,” Open Technology Institute, accessed May 9, 2017, https://github.com/opentechinstitute/bb-schools-analysis/blob/master/initial-data-analysis.R#L127
39 The transit provider for each M-Lab server can be obtained by emailing support@measurementlab.net
41 “Broadband in Schools Github Repository - Figure 14,” Open Technology Institute, accessed May 9, 2017, https://github.com/opentechinstitute/bb-schools-analysis/blob/master/initial-data-analysis.R#L177
44 This is in reference to the server used to provide devices with randomized test schedules, and not M-Lab's servers through which tests are automatically conducted.
This report carries a Creative Commons Attribution 4.0 International license, which permits re-use of New America content when proper attribution is provided. This means you are free to share and adapt New America’s work, or include our content in derivative works, under the following conditions:

- **Attribution.** You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

For the full legal code of this Creative Commons license, please visit [creativecommons.org](http://creativecommons.org).

If you have any questions about citing or reusing New America content, please visit [www.newamerica.org](http://www.newamerica.org).

All photos in this report are supplied by, and licensed to, [shutterstock.com](http://shutterstock.com) unless otherwise stated. Photos from federal government sources are used under section 105 of the Copyright Act.