Where Has the Time Gone? Examining Over a Decade of Broadband Latency Measurements

Jason Schnitzer
University of Colorado Boulder
jason.schnitzer@colorado.edu

David Reed
University of Colorado Boulder
david.reed@colorado.edu

Ashutosh Gandhi
University of Colorado Boulder
Ashutosh.Gandhi@colorado.edu

***Abstract*—The "Measuring Broadband America" (MBA) program created by the Federal Communications Commission (FCC) recently paused a large-scale thirteen-year study of access network performance in the United States. Throughout the program, continuous active measurement across a diversity of service providers, network technologies, and testing topologies rendered a rich set of observations for examining the evolution of fixed broadband access speed and latency performance. This paper examines the MBA latency survey methodology followed by a rigorous longitudinal analysis of the multi-year data set. Our evaluation compares the corpus of historical MBA network observations to evolving broadband latency performance expectations and measurement interpretations. Following our longitudinal analysis, we discuss opportunities to clarify consumer perception of latency measurement methodology. We conclude the study by discussing how different latency measurement methodologies and descriptive interpretations of these observations may impact consumer understanding of the FCC's proposed broadband label. This study's data products and software artifacts are made available to the research community at https://github.com/UCBoulder/bclear/tree/main.**

***Index Terms*—Broadband communication, internet measurement, latency, longitudinal analysis**

I. Introduction

Network latency plays an increasing role in the evaluation of the broadband-connected experience [1]–[6]. Traditionally characterized in terms of service speed (throughput) alone, broadband consumer perception now includes some expression of transmission delay as a critical performance metric [4], [7]–[11]. Rising to a level of regulatory concern in the US, a latency metric is now included in the FCC's recent rule-making. The proposed Broadband Consumer Label is designed to help consumers make better-informed choices when evaluating ISP offerings [12]. The label includes a quantity of "Typical Latency", aimed at providing consumers with visibility into ISPs' claims of timely access network response alongside the more conventional measures of download and upload speeds (service data rates).

Beginning in February 2011 and concluding July 2023, the FCC's MBA program collected observations of broadband performance using an active measurement methodology based on the SamKnows platform [13]. The system applied two forms of test (Idle and Latency under load (LUL)) that documented a process to validate collected observations and produce derived data products for interpretation and presentation [14]. A total of twelve annual reports were completed analyzing measurements over a month of a year's total data collection [1].

In this paper, we present a rigorous longitudinal analysis of network delay measurement collected over 138 months of the MBA program. Our study encompasses all data in the latency measurement corpus while providing an overview of the MBA active testing methodology, measurement topology, and information model.

This research is motivated by a recognized need for a common interpretation of latency performance to inform consumer expectations of connected broadband wireline services [10], [16]–[18]. As high-speed data service has become the source of subsidy to establish broadband universal service, policymakers have taken steps to establish a definition of broadband that establishes Broadband Internet Access Service (BIAS) performance requirements to meet public policy goals [1]. These requirements typically include performance thresholds associated with key broadband service attributes impacting consumers' online quality of experience associated with service data rate or speed, latency, and data use allocation (e.g., monthly data caps). For example, discussion by the FCC noting the purpose of speed testing is to "determine if the network is properly provisioned to furnish the required speed", and the goal of latency testing is to indicate "whether there is sufficient capacity in the network to handle the level of traffic, which is of particular importance when the network is experiencing high traffic load" [2]. In this paper, we confine our attention to the performance metric of network latency while discussing emerging US standards and policies that interpret it.

Our study embodies the entirety of the FCC MBA dataset with specific consideration for delay observations. Over 3.5 billion latency measurements were evaluated in pursuit of the following contributions:

[1]For example, the 2021 report includes data from September and October totaling a non-contiguous four weeks [15].

[2]See para. 33, Federal Communications Commission, In the Matter of Connect America Fund, WC Docket No. 10-90, Order of Reconsideration, Released October 31, 2019.

978-3-903176-64-5

- Where previous studies focused on FCC MBA speed test measurements, this work provides an examination of wireline network latency observations. We demonstrate a novel data processing pipeline for reproducible analysis of FCC MBA raw data and make these materials available to the research community.
- We present a longitudinal analysis of the entirety of the body of data, uncovering specific insights into the trends of latency over the FCC's thirteen-year study. Where existing works evaluate windows of data in reduced temporal ranges (e.g. 1 month), this study provides an analysis addressing the entirety of the MBA project.
- Our analysis compares the performance of three wireline broadband technologies (DSL, Cable, and Fiber) to emerging definitions and standards for latency performance to examine sensitivities to different descriptive statistical interpretations. Specifically, we compare current performance standard definitions for broadband latency to historical FCC MBA data to demonstrate the effect of these interpretations on latency performance outcomes.

Section II of this paper establishes the background and motivation for establishing a standard definition of broadband network latency. In section III, we describe the MBA latency data set used in this study followed by an overview of our data validation methodology and longitudinal analysis. Section V discusses results.

II. Background

Previous studies have referenced the FCC's MBA data to examine performance in broadband wireline networks in the context of throughput (network speed) [19] [20]. However, few focus specifically on the metric of latency over an extended timescale and with attention to measurement methodology and consideration for statistical description.

Speaking to the growing importance of latency performance on broadband subscriber Quality of Experience (QoE) [21], [22], overall wireline performance trends are examined with attention to technology and operational factors that may impact overall service quality. Central to the study of network delay measurements is the definition of latency itself [4] [17] [23]. An abundance of terms have emerged in an attempt to characterize the amount of time observed between the sending and receiving of data over an IP network path; "latency", "delay" and "lag" are often used interchangeably [24] [4].

A retrospective overview of the MBA program is provided in [25], where the goals of a reliable longitudinal data collection for consumer broadband performance are examined. In [19], a deeper evaluation of the FCC MBA data set presents a data validation methodology along with a consumer cost and availability analysis based on speed measurements over a one-month window.

The largest longitudinal study of FCC MBA includes an eight-year analysis of FCC MBA data encompassing a view of reliability, throughput, and latency [20]. An analysis of both Idle latency and LUL is presented but limited to a single year (2019). For that year, they observe decreasing overall trends in both median latency and LUL with greater reductions found in LUL performance.

Though the relationship between operator investment in network infrastructure and increased speed is direct, the response of network latency is not. Conventionally, technology upgrades to access networks increase data rates (e.g., 1Gbps to 5Gbps) but without consideration for latency reduction (e.g., 50ms to 10ms). Although a relationship between increased throughput and reduced latency has been observed [22], direct investments are seldom aimed at improved latency performance alone. More recently, broadband operators have invested in technologies aimed specifically at latency reduction as public funding programs have started to define latency performance requirements for broadband infrastructure [3].

III. Data Description and Validation

In our longitudinal examination of MBA latency data, we first summarize highlights of the FCC's measurement methodology and provide a description of the data corpus [4]. We then present our data validation process, followed by our analysis in the sections below.

A. Description

Within the FCC's study, there are two environmental conditions under which latency measures are considered. Table I summarizes the behaviors of two MBA latency tests evaluated in our analysis[5] while applying the following terms and methodologies.

- **Idle Latency.** A latency observation conducted in the absence of any other network traffic present at the Whitebox location in the home. It is useful to distinguish that this network inactivity would refer to the protocol layer in which the traffic was measured. For example, a network could be considered quiescent at layer-3 (and above) while still active at layer-2 in order to express Media Access Control (MAC) management messages and related protocol operations. It is also important to note that Idle does not refer to the state of the broadband access network traffic load during the time of test execution, but only to the subscriber network in which the Whitebox test agent resides.
- **Latency Under Load (LUL).** A latency observation conducted in the presence of some other significant amount of network activity, and is, therefore, considered more representative of real-world performance [4][6]. In

[3]Source: NTIA, Notice of Funding Opportunity for Broadband Equity, Access, and Deployment Program, May 12, 2022. Available at https://broadbandusa.ntia.doc.gov/sites/default/files/2022-05/BEAD%20NOFO.pdf

[4]For a detailed overview of the SamKnows platform and open methodology employed by the FCC, see FCC Website Measuring Broadband America – Open Methodology, available at https://www.fcc.gov/general/measuring-broadband-america-open-methodology.

[5]The FCC MBA G.711 and ICMP tests are not addressed in this study.

[6]In the MBA testing, the load is HTTP traffic generated during co-incident speed tests. The HTTP traffic load used to as the speed test is generated across 8 concurrent TCP connections during the same 10-second interval as the LUL test.

FCC MBA nomenclature, the term “Downstream LUL” is used to describe a latency test conducted when load traffic is sent from the test server down to the Whitebox, conversely “Upstream LUL” represents a measurement conducted when traffic load is transmitted from the Whitebox up to the test server.

B. Validation

This section describes in detail our methodology to validate the FCC MBA latency measurement data. Validation is necessary since the raw data collected in the measurement effort can contain erroneous test results or otherwise unverified information. The primary data used in our study originates from the MBA program idle latency and LUL measurements as described in Table I. Our data validation pipeline consists of three distinct stages.

1) Resolve Measurement Topology and Access Technology

Each test Whitebox hardware agent is assigned a unique identifier (unit_id) used as an index for measurements associated with the device. We establish the correlation between all unit_id participating in the MBA consumer panel and the ISP, as well as the ISP networking technology (DSL, Cable, Fiber). These unit_ids are randomly generated, which preserves the volunteer panelists’ anonymity.

To resolve Whitebox unit_id to ISP, the FCC publishes two files and the yearly reports – *unit_profile.csv* and *excluded_units.csv*. The *unit_profile.csv* file identifies various details of each active test unit or Whitebox, including ISP, technology, download and upload service tier, and geo-location. The excluded_units.csv file compiles a list of Whitebox units omitted from the yearly report analysis. Our analysis included whiteboxes from the *unit_profile.csv* and *excluded_units.csv* files. This decision aligns with our primary focus on latency-based tests, and we didn’t require units to strictly adhere to their speed tiers as described by the FCC.

All measurements are then classified into either “on-net” or “off-net” [7]. Units within the on-net test set of an ISP are categorized as belonging to that specific ISP. Based on both on-net tests and the data from the two files, the mapping of units results in the formation of the unit_id to ISP relationship.

To resolve each test Whitebox unit_id to network technology, we reference *unit_profile.csv* and *excluded_units.csv* data. For those units lacking a network technology assignment, we refer to the Technical Appendix Report which describes the technology type of each ISP participating.[8].

For ISPs supporting multiple wireline broadband technologies, the maximum transmission speed of the ISP technology is correlated to the speed tiers supported. The technology’s speed tier is then compared to the average monthly download speed calculated for each unit_id to determine the technology is likely to be used.

The resolution of each Whitebox to ISP, network technology, speed tier, and timezone offset is consolidated into a single data product and stored for future analysis (*unit_id_map.csv*).

2) Validation Filters

The filters applied to latency measurements are summarized in Table II. These filters are designed to align with the validation criteria utilized by the FCC.

3) Validated Data Amendment

Following the data filtering process, we enhance the raw dataset by amending four new fields, each explained below:

- **test_type:** Denotes whether a test is categorized as an “on-net” or “off-net” test using the procedure described above.
- **validation_type:** Indicates whether the unit_id in the raw data is part of the FCC’s validated units list. If the unit_id is found in the *unit_profile.csv* files, it is classified as “fcc” Otherwise, it is labeled as “bmc” (representing excluded and “off-net” Whiteboxes).
- **dtime_local:** Local time conversion using timezone_offset and timezone_offset_dtc fields from the *unit_profile.csv* and *excluded_units.csv* files. If the time falls within the daylight saving period, it is adjusted to the local daylight saving timezone. This field plays a crucial role in categorizing the “time_category” field, described below.
- **time_category:** This field classifies timestamps into one of three day-parts used by the FCC: “peak hours” (7PM-11PM weekdays), “off-peak hours” (weekdays), or “sat-sun” (all weekend hours).

Upon the application of all filters, these fields are merged with the unit_id_map.csv file. In this way, eight new fields to the original dataset: 1) ISP, 2) ISP technology, 3) download speed tier, 4) upload speed tier, 5) test_type, 6) validation_type, 7) dtime_local, and 8) time_category. The validated data products are now prepared for longitudinal analysis.

IV. Analysis

This section presents a longitudinal analysis of FCC MBA latency data FY2011-2023 produced using the validated data described in Section III. Additional presentation filters are applied to the validated data to conduct the analysis in this section [9]. The following data was removed in advance:

- All “on-net” testing was removed, leaving only measurements conducted using “off-net” testing servers. Further discussion of “off-net” testing infrastructure is provided in section V.

[7] In the MBA measurement topology, “on-net” refers to tests executed using target servers located within the ISP network, where ”off-net” servers are those hosted outside of the ISP management domain.

[8] Published each year, these appendices are available on the FCC website at https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-broadband-america-program-fixed.

[9] Discontinuities are due to gaps in the MBA data from the FCC website. Monthly gaps in data include Jan’11, Sep/Oct’2020 and 2022 and Mar’2023. FY2023, data is only available up to Jul’2023, as the FCC paused the MBA program at this time.

TABLE I: Description of MBA Latency Measurement Tests

MBA Test	Description	IDLE/LUL	Data Model	Schedule
UDP Latency (Quiescent)	Observes and records the Round Trip Time (RTT) of UDP datagrams transmitted between a Whitebox and test server in the absence of other test traffic. Each test datagram consists of a 8-byte sequence number and a 8-byte timestamp. Any datagram arriving outside of a 3 second timeout window is treated as lost. The test operates consistently in the background, sending a set of UDP datagrams randomly distributed over a 1-hour measurement interval. A 99th percentile filter is applied to discard the top 1 percent of measurement values observed during the interval. Of the lower 99 percent values left from the 1-hour measurement interval, the minimum (RTT_{min}), maximum (RTT_{max}), average (RTT_{avg}) and standard deviation (RTT_{std}) are calculated across the measurement set and recorded. units: microseconds	IDLE	curr_udplatency	Hourly, Continuous
UDP Latency (Working)	Observes and records the RTT of UDP datagrams transmitted between a Whitebox and test server in the presence of other test traffic load. The test operates concurrent with 10-second upload or download speed tests. Results for each direction are recorded separately. While speed test load is running, a UDP stream of datagrams spaced at 100ms is sent within the interval. Each test datagram consists of a 8-byte sequence number and 8-byte timestamp. Any datagram arriving outside of a 3-second timeout window is treated as lost. For each 10-second measurement interval, the minimum (RTT_{min}), maximum (RTT_{max}), average (RTT_{avg}) and standard deviation (RTT_{std}) are calculated across the measurement set and recorded. units: microseconds	LUL	curr_ulping, curr_dlping	Once within each daypart window: 12am-6am, 6am-12pm, 12pm-6pm, 6pm-8pm, 8pm-10pm, 10pm-12am 12am-12pm, 12pm-12am

- All measurements from satellite and fixed wireless networks were filtered from the validated data, leaving only data from ISPs using DSL, Cable, and Fiber wireline access technologies.
- Data from ISPs with a few panelists (under 50 Whiteboxes) or other inconsistencies participating in the MBA program were removed.

As a result, test data from the following ISPs is excluded: RCN, T-Mobile, Fluidata, Sky, Rogers, Hargray Communications, WideOpenWest, PenTeleData, and EnTouch. Also, WildBlue/ViaSat and Hughes are removed as they used Satellite technology.

The remaining data used in the analysis includes the following ISPs: Comcast, CenturyLink, TimeWarner, Charter, AT&T, Verizon, Cox, Qwest, Windstream, Cablevision, Mediacom, Frontier, Clearwire, Insight, Brighthouse, Optimum, Hawaiian Telecom, and Cincinnati Bell.

1) Idle Latency

Our presentation of the FCC's MBA data begins by plotting Idle latency across the 13 years of study FY2011–2023. Figure 2a shows the median of measurements for all ISPs and technology types considered in our study. As described in III, the Idle latency calculations evaluate only the bottom 99 percent of observations from the test interval to remove the top 1 percent of RTT measurements.

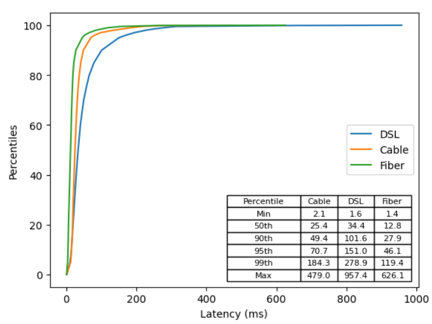
Overall, the median value for maximum Idle latency measurements per month has decreased from up to 38ms through 2014 to below 29ms by 2020. The graph superimposes a linear trend line based on this data that suggests Idle latency reduced at approximately 1.4ms per year throughout the study, representing an improvement of approximately 3-5 percent per year. Though evident from visual inspection, the application of a Mann-Kendall test [26] ($p_{value} = 0.0$, $\tau = 0.89$) confirms a decreasing trend, or continuing ongoing Idle latency performance improvement over time FY2011–2023.

Figure 2b shows variations in these median Idle latency observations based on ISP technology types of DSL, Cable, and Fiber. In this chart, each bar signifies the median of the maximum (RTT_{max}) 99th percentile measurements for a specific month for each technology type. The range of Idle latency values for Fiber over this time are lowest ranging between 12–28ms, compared to the ranges for Cable between 25–40ms and DSL between 35–50ms. While not shown in Figure 2b, the linear trend lines generated by these results indicate trends in Idle latency decrease of approximately 0.9ms per year for DSL, 1.2ms per year for Cable, and 0.7ms per year for Fiber. Overall, this breakdown of Idle latency results by technology makes intuitive sense as they generally correspond to improved or lower latency measurements as service speeds increase by each technology type.

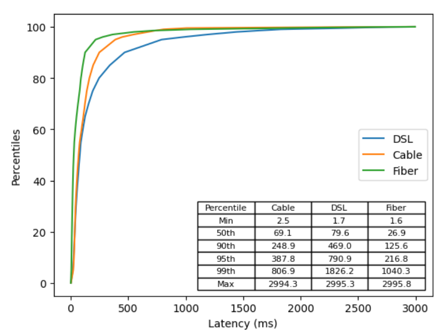
Figure 1a illustrates a Cumulative Distribution Function (CDF) for the 99th percentile of maximum (RTT_{max}) Idle latency measurements between Jul'1-31, FY2023. During this

TABLE II: Inventory of Data Validation Filters Applied to MBA Latency Data

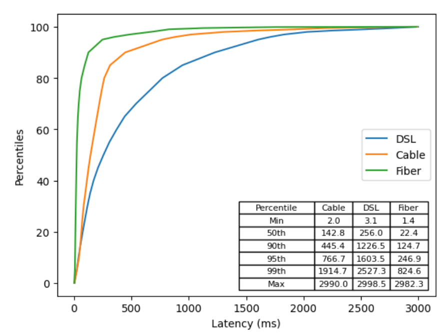
Filter Name	Description	Formula	MBA File Name
$zeros$	Remove null value entries for all columns.	$col\_value == 0$	curr_udplatency.csv curr_ulping.csv curr_dlping.csv
$RTT_{delta} > 300ms$	Remove test instances where the range of units of individual round trip times exceeded 300ms.	$RTT_{max} - RTT_{min} > 300ms$	curr_udplatency.csv
$RTT_{min} < 0.05ms$	Remove test instances where any round-trip time was reported as 0.05ms or lower.	$RTT_{min} < 0.05ms$	curr_udplatency.csv curr_ulping.csv curr_dlping.csv
$packet\_success < 50$	Remove test instances with less than fifty successful packets.	$success < 50$	curr_udplatency.csv curr_ulping.csv curr_dlping.csv
$packet\_loss > 10\%$	Remove test instances where packet loss exceeds 10%.	$failures/(success + failures) > 10\%$	curr_udplatency.csv curr_ulping.csv curr_dlping.csv



(a) Idle

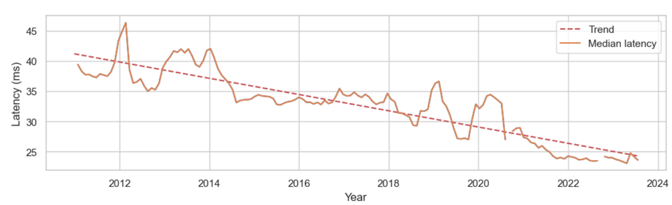


(b) LUL Downstream

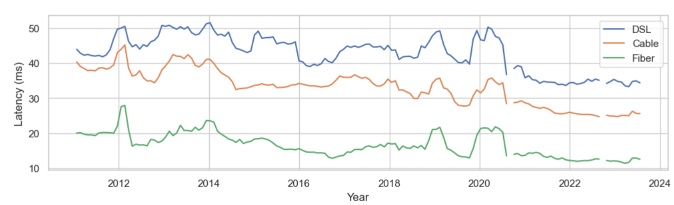


(c) LUL Upstream

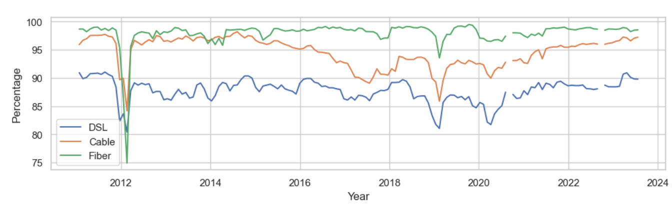
Fig. 1: **Cumulative Distribution Functions (CDFs) RTT_{max} (Jul'2023):** (a) Idle Latency; (b) Downstream LUL; (c) Upstream LUL.



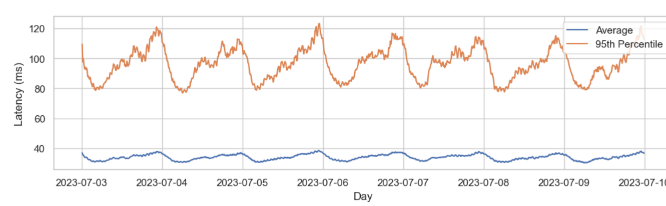
(a) Idle Median - All Technology



(b) Idle Median - By DSL, Cable, Fiber



(c) Idle 99th Percentile ($< 100ms$) - All Technology



(d) Idle 95th percentile vs. average - All Technology

Fig. 2: **Idle latency results RTT_{max} (FY2011–2023):** (a) median (RTT_{max}) all technology; (b) median (RTT_{max}) DSL, Cable, Fiber; (c) 99th percentile ($RTT_{max} < 100ms$); (d) average vs. 95the percentile (RTT_{max}) all technology.

window, 98.5% of all RTT_{max} observations conducted on Fiber networks fell below the 100ms threshold, while the figures for DSL and Cable networks falling below this threshold are 89.8% and 97.2%, respectively. The 100ms threshold is a useful figure of merit since it is often used as a latency performance requirement that ISPs must meet to qualify for broadband infrastructure subsidies[10]. To examine this performance standard further, Figure 2c plots the percentage of Idle latency measurements below 100ms over 30-day increments FY2011–2023. These Idle latency results indicate a 100ms Idle latency threshold at the 95th percentile would be met by ISPs using Fiber and Cable technologies in recent years but not by ISPs using DSL.

2) Latency Under Load (LUL)

Next, we examine working latency tests FY2011–2023. Figure 3a shows the median of all (RTT_{max}) measurements for each upstream and downstream LUL test across all ISPs and technology types. The chart illustrates a substantial initial disparity between upstream and downstream LUL in 2011 that has gradually declined to the present. Upstream LUL shows a continuous decrease, dropping from a peak of 950ms in late 2011 to 200ms in 2018, down to 120ms by mid-2023, which almost matches the values of downstream LUL. The improvement of downstream LUL is less evident, demonstrating a more gradual decline from 200ms in late 2011 to 100ms by mid-2023. Figure 3a includes trend lines to further quantify LUL improvement over the time of the study. A quadratic trend line best fits upstream LUL results, showing a decrease of approximately 133ms per year from 2011 to 2014 (approximately 20% per year improvement), 73ms per year from 2014 to 2020 (approximately 15% per year improvement), and 8.5ms per year from 2020 to 2023 (less than 10% per year improvement). Similarly, based on the linear trend line, downstream LUL has improved at an overall decrease of about 11ms per year (roughly 5–10% per year improvement).

Figure 3b provides the downstream LUL test results by access technology type. This illustration shows median downstream LUL over Fiber has improved from 2012 to the present. Given that latency typically decreases with increased service speed, and Fiber typically offers much higher speeds than most Cable and DSL services, it is notable that Fiber technology provided the highest test values between the second half of 2011 and the first half of 2012. Starting from 2016, however, a general trend emerged wherein Fiber LUL values consistently appear lower or equivalent to Cable and DSL. Specific to Cable, downstream LUL performance has fallen between 70-110ms except for increased LUL values up to 200ms FY2012-2017. While not shown in Figure 3b, the quadratic trend lines generated by these results indicate the improvement in downstream LUL is roughly 10.5ms per year for DSL, 4ms per year for Cable, and 12ms per year for Fiber.

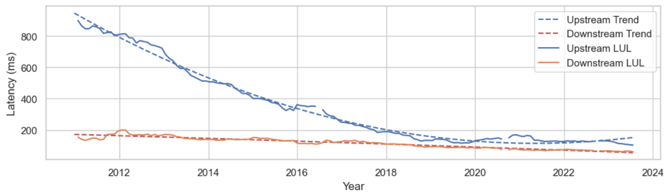
Figure 3c similarly presents the median upstream LUL data with a breakdown for each technology type. Unlike downstream LUL results, however, the upstream LUL results illustrate similar trends observed in the Idle latency charts, where Fiber values consistently provide the lowest upstream LUL values, falling from 140ms in 2011 to 22ms in 2023. In addition, after the elevated values for DSL and Cable experienced early during 2011-2012, DSL and Cable show a consistent decline to the present. However, the rate of improvement of Cable is greater than DSL. While not shown in Figure 3c, the quadratic trend lines generated by these results indicate the trend in downstream LUL improvement for DSL is roughly 72ms per year, for Cable is 54ms per year, and for Fiber is 10ms per year. For DSL and Cable, these rates of improvement in upstream LUL are substantially higher than those calculated for downstream LUL over this period, while Fiber is the same (perhaps reflecting the symmetric nature of fiber service tiers concerning speed).

Figure 3d presents the 95th percentile values for downstream and upstream LUL. As anticipated, the downstream LUL values are lower than upstream, ranging from 770ms in 2012 to 270ms FY2023. In contrast, upstream LUL ranges from 1520ms FY2012 to 600ms FY2023. Notably, both exhibit a consistent decline, which decreases significantly from 2020 onwards. Under these 95th percentile LUL values, however, there will be significant impairment to real-time services that can typically show substantial impairment when latency exceeds 100–250ms.

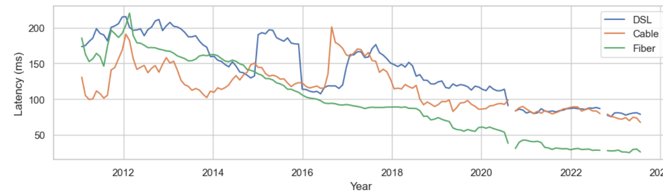
Figure 3e illustrates the 95th percentile for the (RTT_{max}) downstream LUL, segmented by technology type. The pattern remains consistent with previous graphs: Fiber values are lower than Cable, which, in turn, are lower than DSL. Notably, Cable and DSL values are very close after mid-2014 with Fiber having significantly lower values. Given that we are plotting the 95th percentile, the latency values are approaching the 2-second mark for DSL. While not shown in Figure 3e, quadratic trend lines generated by these results indicate the trend in downstream LUL improvement for DSL is roughly 50ms per year from 2012-2018, followed by a steep fall of approximately 125ms per year from 2018-2023, for Cable is 70ms per year, and for Fiber is 26ms per year.

Figure 3f illustrates the 95th percentile for the (RTT_{max}) upstream LUL, segmented by technology type. The pattern remains consistent with previous graphs: Fiber values are lower than Cable, which, in turn, are lower than DSL. It's important to highlight that upstream values are notably higher

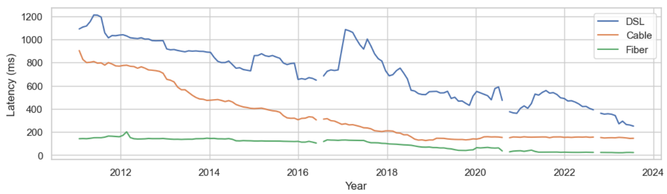
[10]The FCC adopted detailed performance measure requirements in 2018 and 2019. See Connect America Fund, Order, WC Docket No. 10-90, 33 FCC Rcd 6509 (WCB 2018) (First Performance Measures Order). Stating in paragraph 4: "To show compliance with latency obligations, a price cap carrier must certify that 95 percent or more of all peak period measurements (also referred to as observations) of network round trip latency are at or below 100ms when measured during the peak period between the customer premises and the nearest Internet Exchange Point (IXP) in an FCC-designated metropolitan area (FCC-designated IXP)...carriers participating in the Measuring Broadband America (MBA) program may use the results from that testing to support their certification that they meet the latency requirement, so long as they deploy at least 50 Whiteboxes to customers within the CAF Phase II-funded areas within each state and certify that 95 percent or more of the measurements taken quarterly during peak periods for two weeks were at or below 100ms." See also, Connect America Fund, Order on Reconsideration, WC Docket No. 10-90, 34 FCC Rcd 10109 (2019).



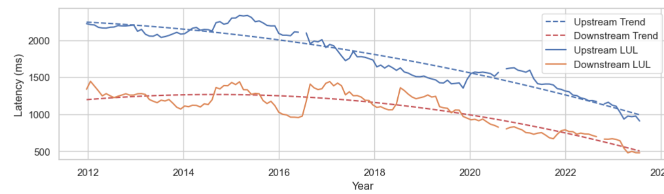
(a) LUL Median Upstream & Downstream - All Technology



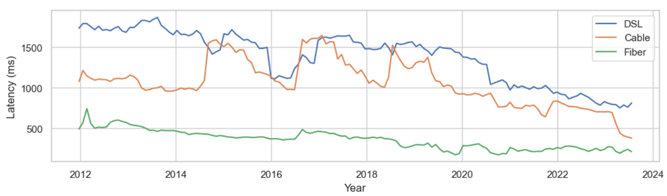
(b) LUL Median Downstream - DSL, Cable, Fiber



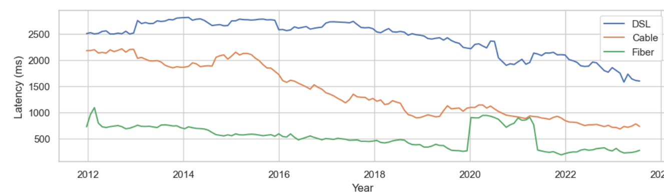
(c) LUL Median Upstream - DSL, Cable, Fiber



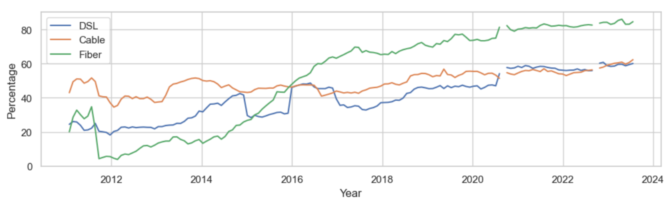
(d) LUL 95th Percentile Upstream and Downstream - All Technology



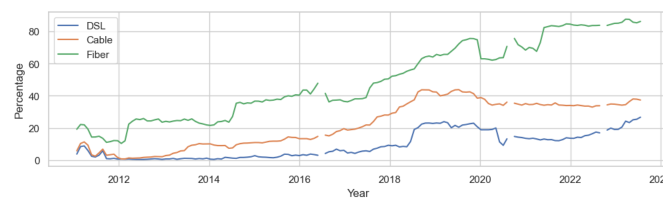
(e) LUL 95th Percentile Downstream by Technology



(f) LUL 95th Percentile Upstream by Technology



(g) LUL Percentage of Downstream $< 100ms$ by Technology



(h) LUL Percentage of Upstream $< 100ms$ by Technology

Fig. 3: **Latency Under Load Results RTT_{max} (FY2011–2023):** (a) median Downstream & Upstream LUL for all technology types; (b) median Downstream LUL by technology type; (c) median Upstream LUL by technology type; (d) 95th percentile Upstream & Downstream for all technology types; (e) 95th percentile downstream by technology type; (f) 95 percentile upstream by technology type; (g) Percentage of Downstream LUL ($RTT_{max} < 100ms$) by Technology; (h) Percentage of Upstream LUL ($RTT_{max} < 100ms$) by Technology.

than downstream values. While not shown in Figure 3f, the quadratic trend lines generated by these results for DSL indicate an increase of about 65ms per year FY2012-2016 and then a decrease of approximately 153ms FY2016-2023 for Cable is 125ms per year, and for Fiber is roughly 63ms FY2012-2020 followed by a spike to 900ms for the period 2020 to mid-2021 and then a subsequent decrease until 2023.

Figures 1b and 1c illustrate when real-time services are not impaired due to downstream and upstream LUL across technology types. To do so, we assume below 100ms establishes an acceptable performance threshold to support most real-time applications in each direction. Focusing on a single month, Figures 1b and 1c illustrate the technology LUL distributions relative to the 100ms performance threshold by direction. In the downstream, Figure 1b shows that the percentile meeting the threshold is 84.6% for Fiber, 62.3% for Cable, and 60.1% for DSL. In the upstream, 1c shows the percentile meeting the 100ms threshold in the upstream direction is 86.2% for Fiber, 37.4% for Cable, and 26.7% for DSL.

On a broader timescale, Figure 3g plots the percentage of Downstream LUL (RTT_{max}) measurements below 100ms over 30-day increments between FY2011–2023, where Figure 3h represents Upstream LUL. One general observation from Figure 3g is that the trend in Downstream LUL performance to meet a 100ms threshold has shown a steady improvement over the years for all technologies with a significant increase for Fiber in the period from 2014 to mid-2017. A similar trend is seen for Upstream LUL (Figure 3h) with the Fiber values meeting the 100ms mark at the 90th percentile by 2023.

While our examination has evaluated an extended timeline using the 95th percentile of the RTT_{max} observation, it is worth noting the sensitivity of our results to different statistical interpretations of the MBA data set. Figure 2d plots the 95th percentile and average of Idle latency across all access technologies for the week of July 3-9, FY2023. The 95th percentile line shows diurnal variation with a consistent pattern of off-peak latency between 2-8am between 80–90ms and peak latency between approximately 6pm-midnight between 100–120ms. In comparison, displaying the average of RTT_{max} over the same week renders an interpretation entirely below the 100ms target.

V. Results and Discussion

We discuss four results from our analysis of the latency measurements obtained over the entire duration of the FCC's MBA program, while highlighting the limitations of interpreting MBA data in our discussion.

First, the data collected for both the Idle and LUL latency metrics provide a general, quantitative measure of improvement in latency on the public Internet FY2011-2023. In writing this statement, we recognize that analysis of the MBA Program data cannot support conclusions regarding Internet performance or broadband subscriber experience. However, to the extent that the MBA Program provides some representation of general Internet performance, imperfect as it may be, it is a useful and encouraging finding that there has been substantial improvement in the latency performance of residential broadband services over this time based upon data obtained from this large sample of major ISPs serving the United States. While we cannot formulate a "law" to characterize this pace of improvement (e.g., Moore's law for semiconductors), FY2011-2023 the analysis identified consistent Idle latency improvements of 3-5 percent per year as well as improvements in LUL or working latency in the upstream and downstream between 10-20 percent and 5-10 percent per year, respectively. We also note that the pace of improvement in latency appears to be slowing in the past four years, which could reflect increasing costs to lower latency from current levels.

Second, these improvements in Internet performance have occurred across all three technology types (DSL, Cable, and Fiber) included in our analysis. For example, these technologies showed significant improvement in downstream LUL from roughly 200ms in FY2011 to below 100ms FY2021. This does not imply that these technologies' performance was the same. Usually, the latency of Fiber was lowest, followed by Cable and then DSL. While our analysis was constrained to "off-net" test servers, we note that many changes in MBA infrastructure are continuous across the multi-year study. Table IV shows the significant changes to the count of "off-net" servers participating in different years of the study.

Third, to the extent that the latency measurements collected in the MBA program are representative of real-world trends in residential broadband services, the results of this analysis do have implications for the requirements associated with the definition of broadband services associated with government funding and grants for broadband infrastructure. In particular, we applied a 100ms threshold to the results to investigate the circumstances under which the different technologies achieved performance that met this requirement.

Table III provides a further sensitivity analysis of the results to the 100ms latency threshold. These figures indicate that both the Fiber and Cable technologies satisfy the 95th percentile 100ms threshold for the Idle metric based upon RTT_{max} and RTT_{avg} (even when considering that the FCC only provides Idle latency data representing the 99th percentile of measurements collected). The DSL technology only meets this threshold for the Idle test between the 85th-90th percentile for RTT_{avg}. Highlighting the sensitivity of performance conclusions drawn from different measurement methodologies, none of the three access technologies would meet the 100ms standard using the LUL test measurements until well below the 95th percentile criteria.

While a detailed discussion of the best latency test measurement methodology and metrics to apply for a definition of broadband is beyond the scope of this analysis, we do assert the Idle metric to be a more beneficial option than LUL. The Idle metric provides a measure of latency over the access network from the Whitebox in the home and the test server in the core network. The test assumes there exists no other traffic originating within the home network (quiescent state). Given that the broadband access network is the segment of interest, the Idle methodology used in the MBA study does propose to isolate the measurement from con-incident traffic and other impairments within the home. Moreover, requiring the 95th percentile of the RTT_{max} metric versus RTT_{avg} could further refine the definition to reflect a useful maximum threshold impacting consumers' quality of experience. As noted above, both Fiber and Cable (in aggregate) would meet a 95th percentile Idle latency requirement of 100ms for RTT_{max} for broadband service, while DSL would not meet this performance standard. Given that DSL is becoming an obsolescent technology, this outcome would be consistent with current market conditions of investment in broadband infrastructure.

The LUL test, in contrast, could lead to significant complexities if included as a requirement for broadband service. This measure applies a significant load to the local access network as part of a speed test, with the intent of creating an upper bound on the latency likely to be encountered during heavy usage between the Whitebox and the test server. Though active testing methodologies have proposed that working latency observations better approximate actual user experience [4], our analysis suggests that MBA LUL results raise concerns about consistency and test topology. Given the topological context of a Whitebox relative to its broadband access network, introducing load into the measurement may bias the interpretation of an ISP's performance by including home network effects beyond the operational demarcation [23]. In doing so, additional queuing delays (also known as "buffer bloat") in home router or gateway implementations might exaggerate working latency measurement values [27]. For the purposes of compliance with latency requirements in the access network, the Idle test could be viewed as a more transparent and reasonable measure of the actual latency in the broadband

TABLE III: **Summary of Latency Measurement Results (Jul'2023).** Organized by row representing the observation presented in MBA test data (Idle, LUL Downstream (DS), LUL Upstream (US)), the minimum (min), maximum (max), and selected percentiles represent a field of latency results calculated across Cable, DSL, and Fiber access technologies. The 95th percentile column is highlighted to support the discussion in section V.

	CABLE (ms)						**DSL** (ms)						**FIBER** (ms)					
Test	**min**	**50th**	**90th**	**95th**	**99th**	**max**	**min**	**50th**	**90th**	**95th**	**99th**	**max**	**min**	**50th**	**90th**	**95th**	**99th**	**max**
IDLE																		
RTT_{min}	0.1	12.3	24.5	29.8	57.8	188.5	0.1	23.9	44.9	53.3	81.7	679.8	0.2	7.2	16.8	23.2	55.7	453.9
RTT_{max}	2.1	25.4	49.4	70.7	184.3	479.0	1.6	34.4	101.6	151.0	278.9	957.4	1.4	12.8	27.9	46.1	119.4	626.1
RTT_{avg}	1.9	16.8	31.1	37.1	66.5	920.7	1.3	26.8	50.8	60.4	98.3	843.3	1.3	9.7	18.7	27.7	62.5	590.7
LUL DS																		
RTT_{min}	0.1	13.2	25.7	30.7	60.4	1100.7	1.0	25.3	46.5	55.1	94.9	1558.4	0.9	8.1	17.3	23.7	51.6	970.7
RTT_{max}	2.5	69.1	248.9	387.8	806.9	2994.3	1.7	79.6	469.0	790.9	1826.2	2995.3	1.6	26.9	125.6	216.8	1040.3	2995.8
RTT_{avg}	1.9	37.2	119.6	174.6	411.9	1791.3	1.3	56.8	323.1	469.1	1113.9	2477.4	1.3	18.1	81.6	106.7	372.3	2022.0
LUL US																		
RTT_{min}	1.0	15.0	27.4	32.6	62.6	864.3	1.0	23.8	45.1	53.6	92.4	1836.6	0.8	8.3	17.3	22.3	50.3	833.8
RTT_{max}	2.0	142.8	445.4	766.7	1914.7	2990.0	3.1	256.0	1226.	1603.5	2527.3	2998.5	1.4	22.4	127.4	246.9	824.6	2982.3
RTT_{avg}	1.6	47.0	236.6	391.1	962.3	2524.8	2.0	179.0	864.8	1143.6	1857.9	2629.9	1.5	16.3	74.2	123.3	570.8	1652.7

TABLE IV: Post Validation Count of Off-net Servers and Panel Whiteboxes Included in Analysis

	2011	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
Number of Off-net Servers	78	71	163	182	136	62	72	93	110	22	14	16	62
Number of Whiteboxes	9092	10171	6465	7598	7540	7443	5948	6037	6909	7086	5990	5119	4157

network under normal use as compared to the LUL test.

Fourth and finally, our analysis of the LUL data indicates that despite the ongoing improvement in Internet performance, there remains room for improvement in broadband service performance to minimize impacts on real-time services. Recent advances in latency reduction mechanisms offer promising performance improvements that were not available for ISP deployment earlier in the MBA study [28]–[30]. Whether the threshold requirement for adequate performance of a real-time service is 100ms or 250ms, the LUL figures in Table III show that network performance across all technology types can often operate between the 50th and 90th percentiles.

VI. Future Work

This paper focused on the longitudinal analysis of latency data and the sensitivity of interpretation based on derived metrics and emerging performance policy thresholds. Opportunities for future work beyond this study offer promising contributions based on the FCC's MBA data corpus, including:

- In limiting our study to organizational ISP level analysis to focus on analysis over the entirety of the MBA data set, the scope did not incorporate geospatial information included in the MBA data set. By evaluating latency measurement sensitivities to both physical and network server topologies, further insights can be derived regarding the evolution of broadband wireline network delay.
- A comparative analysis of latency observations collected by other active performance measurement platforms, including different data sources such as Measurement Labs [31] and RIPE Atlas [32].
- A deeper examination to determine if there is a relationship to be found between speed and latency in the MBA data.

VII. Conclusion

In this paper, we provided a detailed longitudinal analysis of network delay measurement collected over the 138 contiguous months of the US FCC MBA program from 2011 to 2023. Our study encompassed both Idle and Working latency observations in the MBA data corpus while examining active testing methodology, measurement topology, and information model. Though an overall decreasing trend in latency was shown across the three broadband technologies studied, it was highlighted that sensitivity to MBA testing methodology and data interpretation can result in different outcomes when evaluated against various latency performance definitions and statistical interpretations. The results of this analysis were summarized to examine long-term latency trends in the public Internet and to motivate consideration for methodologies used to evaluate emerging broadband latency performance policies. As regulatory frameworks such as the FCC's Broadband Label evolve, a more precise technical definition of latency merits careful examination for consumers to make informed decisions based on ISP performance claims.

References

[1] D. Reed, “Is speed enough? examining the definition of broadband and its implications for public policy,” in *Examining the Definition of Broadband and Its Implications for Public Policy (August 3, 2021). TPRC49: The 49th Research Conference on Communication, Information and Internet Policy*, 2021.

[2] M. Candela, V. Luconi, and A. Vecchio, “Impact of the covid-19 pandemic on the internet latency: A large-scale study,” *Computer Networks*, vol. 182, p. 107495, 2020.

[3] M. Clark. (2023) The quiet plan to make the internet feel faster. [Online]. Available: https://www.theverge.com/23655762/l4s-internet-apple-comcast-latency-speed-bandwidth

[4] “Latency explained,” Broadband Internet Technical Advisory Group, Tech. Rep., 2022. [Online]. Available: http://www.bitag.org/documents/BITAG_latency_explained.pdf

[5] K. Sundaresan, G. White, and S. Glennon, “Latency measurement: What is latency and how do we measure it?” in *NCTA 2020 Fall Technical Program*. NCTA, 2020.

[6] T. Høiland-Jørgensen, B. Ahlgren, P. Hurtig, and A. Brunstrom, “Measuring latency variation in the internet,” in *Proceedings of the 12th International on Conference on emerging Networking EXperiments and Technologies*, 2016, pp. 473–480.

[7] D. D. Clark and S. Wedeman, “Understanding the metrics of internet broadband access: How much is enough?” *Available at SSRN 4178804*, 2022.

[8] K. MacMillan and N. Feamster, “Beyond speed test: Measuring latency under load across different speed tiers,” 2022. [Online]. Available: https://internetequity.cs.uchicago.edu/wp-content/uploads/2022/04/Beyond-speed-test.pdf

[9] Z. Shmeis, M. Abdullah, P. Nikolopoulos, K. Argyraki, D. Choffnes, and P. Gill, “Localizing traffic differentiation,” in *Proceedings of the 2023 ACM on Internet Measurement Conference*, 2023, pp. 591–605.

[10] K. MacMillan, T. Mangla, M. Richardson, and N. Feamster, “Best practices for collecting speed test data,” *Available at SSRN 4189044*, 2022.

[11] F. C. Comission. (2024) 2024 section 706 report - inquiry concerning the deployment of advanced telecommunications capability to all americans in a reasonable and timely fashion. [Online]. Available: https://docs.fcc.gov/public/attachments/DOC-400675A1.pdf

[12] F. C. Commission. (2024) Broadband consumer labels — federal communications commission. [Online]. Available: https://www.fcc.gov/broadbandlabels

[13] S. Knows. (2024) https://samknows.com/. [Online]. Available: https://samknows.com/

[14] “Measuring broadband america. technical appendix to the twelfth mba report.” Federal Communications Comission, Tech. Rep., 2022. [Online]. Available: https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-broadband-america-program-fixed

[15] “Twelfth measuring broadband america. a report on consumer fixed broadband performance in the united states,” Federal Communications Comission, Tech. Rep., 2022. [Online]. Available: https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-fixed-broadband-twelfth-report

[16] D. Gino. (2021) Focusing on latency, not throughput, to provide a better internet experience and network quality. [Online]. Available: https://www.iab.org/activities/workshops/network-quality/

[17] B. Briscoe, G. White, V. Goel, and K. De Schepper. (2021) A single common metric to characterize varying packet delay. [Online]. Available: https://www.iab.org/activities/workshops/network-quality/

[18] K. Claffy and D. Clark, “Challenges in measuring the internet for the public interest,” *Journal of Information Policy*, vol. 12, pp. 195–233, 2022.

[19] A. Dunna, Z. S. Bischof, and R. Fontugne, “Sanitizing a view of consumer broadband in the united states,” in *TMA*, 2020.

[20] J. Fries, T. V. Doan, R. Gupta, and V. Bajpai, “An eight years perspective on the internet broadband infrastructure in the usa,” in *2022 IFIP Networking Conference (IFIP Networking)*. IEEE, 2022, pp. 1–9.

[21] A. Flickinger, C. Klatsky, A. Ledesma, J. Livingood, and S. Ozer, “Improving latency with active queue management (aqm) during covid-19,” *arXiv preprint arXiv:2107.13968*, 2021.

[22] D. Reed and L. Perigo. (2021) Measuring isp performance in broadband america: A study of latency under load. [Online]. Available: https://www.iab.org/activities/workshops/network-quality/

[23] T. Høiland-Jørgensen, B. Ahlgren, P. Hurtig, and A. Brunstrom, “Measuring latency variation in the internet,” in *Proceedings of the 12th International on Conference on Emerging Networking EXperiments and Technologies*, ser. CoNEXT ’16. New York, NY, USA: Association for Computing Machinery, 2016, p. 473–480. [Online]. Available: https://doi.org/10.1145/2999572.2999603

[24] Amazon. (2023) What is network latency? [Online]. Available: https://aws.amazon.com/what-is/latency/

[25] E. W. Burger, P. Krishnaswamy, and H. Schulzrinne, “Measuring broadband america: A retrospective on origins, achievements, and challenges,” *ACM SIGCOMM Computer Communication Review*, vol. 53, no. 2, pp. 11–21, 2023.

[26] M. Hussain and I. Mahmud, “pymannkendall: a python package for non parametric mann kendall family of trend tests.” *Journal of Open Source Software*, vol. 4, no. 39, p. 1556, 7 2019. [Online]. Available: http://dx.doi.org/10.21105/joss.01556

[27] P. Harrison, “Buffer-bloated router? how to prevent it and improve performance,” *Commun. ACM*, vol. 66, no. 6, p. 73–77, may 2023. [Online]. Available: https://doi.org/10.1145/3564262

[28] B. Briscoe, K. D. Schepper, M. Bagnulo, and G. White, “Low Latency, Low Loss, and Scalable Throughput (L4S) Internet Service: Architecture,” RFC 9330, Jan. 2023. [Online]. Available: https://www.rfc-editor.org/info/rfc9330

[29] “Low latency docsis: Technology overview”,” CableLabs, Tech. Rep., 2019. [Online]. Available: http://www.bitag.org/documents/BITAG_latency_explained.pdf

[30] F. Baker and G. Fairhurst, “IETF Recommendations Regarding Active Queue Management,” RFC 7567, Jul. 2015. [Online]. Available: https://www.rfc-editor.org/info/rfc7567

[31] Measurement Lab, “The M-Lab NDT data set,” https://measurementlab.net/tests/ndt, (2009-02-11 – 2015-12-21), bigquery table `measurement-lab.ndt.download`.

[32] R. Atlas. (2024) https://atlas.ripe.net/. [Online]. Available: https://atlas.ripe.net/