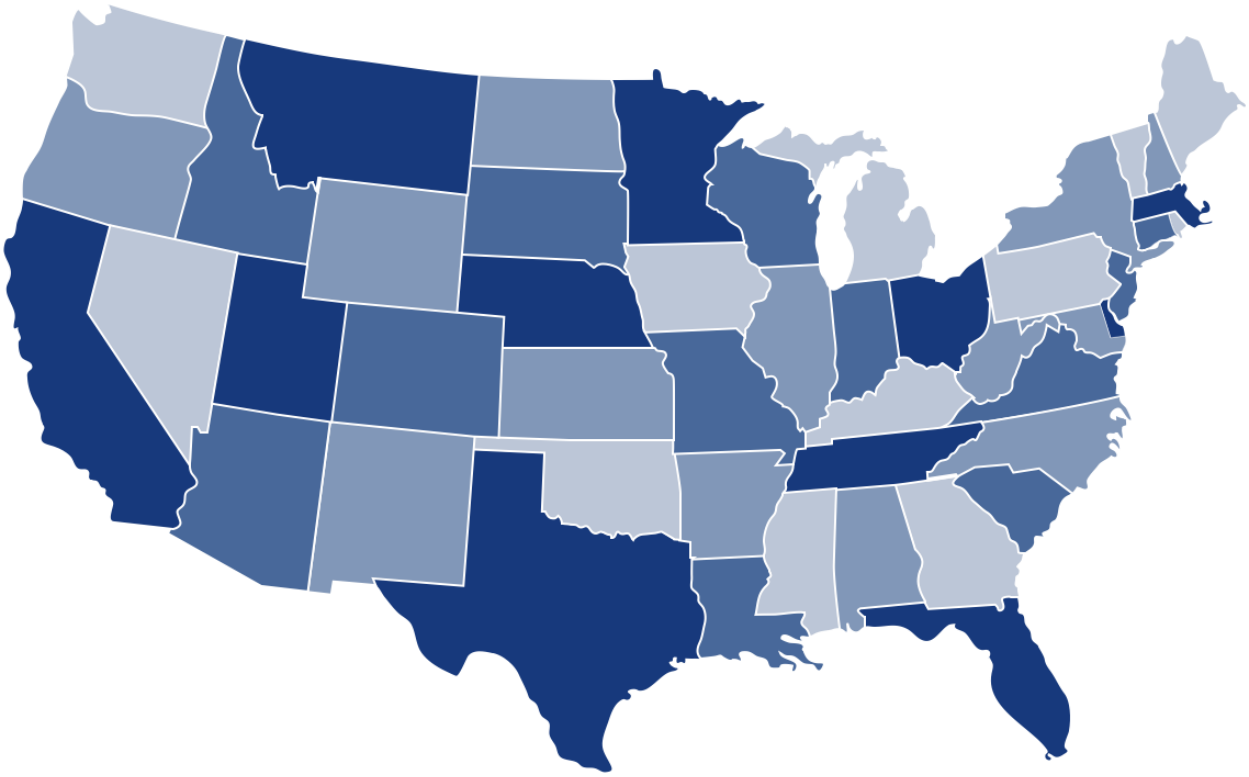


Measuring Broadband America

Technical Appendix to the Tenth MBA Report

FCC's Office of Engineering and Technology

Table of Contents

1 - INTRODUCTION AND SUMMARY 5
2 - PANEL CONSTRUCTION 5
2.1 - USE OF AN ALL VOLUNTEER PANEL 6
2.2 - SAMPLE SIZE AND VOLUNTEER SELECTION 6
2.3 - PANELIST RECRUITMENT PROTOCOL 13
2.4 - VALIDATION OF VOLUNTEERS' SERVICE TIER 15
2.5 - PROTECTION OF VOLUNTEERS' PRIVACY 16
3 - BROADBAND PERFORMANCE TESTING METHODOLOGY 17
3.1 - RATIONAL FOR HARDWARE-BASED MEASUREMENT APPROACH 17
3.2 - DESIGN OBJECTIVES AND TECHNICAL APPROACH 18
3.3 - TESTING ARCHITECTURE 21
Overview of Testing Architecture 21
Approach to Testing and Measurement 22
Home Deployment of the NETGEAR Based Whitebox 23
Home Deployment of the TP-Link Based Whitebox 23
Home Deployment of the SamKnows Whitebox 8.0 23
Internet Activity Detection 23
Test Nodes (Off-Net and On-Net) 24
Test Node Locations 25
Test Node Selection 27
3.4 - TESTS METHODOLOGY 28
3.5 - TEST DESCRIPTIONS 29
Download speed and upload speed 29
Web Browsing 29
UDP Latency and Packet Loss 30

Voice over IP 31
DNS Resolutions and DNS Failures 31
ICMP Latency and Packet Loss 31
Latency Under Load 31
Consumption 36
Cross-Talk Testing and Threshold Manager Service 36
4 - DATA PROCESSING AND ANALYSIS OF TEST RESULTS 37
4.1 - BACKGROUND 37
Time of Day 37
ISP and Service Tier 37
4.2 - DATA COLLECTION AND ANALYSIS METHODOLOGY 40
Data Integrity 40
Legacy Equipment 40
Collation of Results and Outlier Control 42
Peak Hours Adjusted to Local Time 42
Congestion in the Home Not Measured 42
Traffic Shaping Not Studied 42
Analysis of PowerBoost and Other ”Enhancing” Services 43
Consistency of Speed Measurements 43
Latencies Attributable to Propagation Delay 44
Limiting Factors 44
4.3 - DATA PROCESSING OF RAW AND VALIDATED DATA 44
5 - REFERENCE DOCUMENTS 53
5.1 - USER TERMS AND CONDITIONS 53
5.2 - CODE OF CONDUCT 63
5.3 - TEST NODE BRIEFING 65

LIST OF TABLES

Table 1: ISPs, Sample Sizes and Percentages of Total Volunteers 8
Table 2: Distribution of Whiteboxes by State 10
Table 3: Distribution of Whiteboxes by Census Region 12
Table 4: Panelists States Associated with Census Regions 12
Table 5: Design Objectives and Methods 18
Table 6: Overall Number of Testing Servers 24
Table 7: List of tests performed by SamKnows 28
Table 8: Estimated Total Traffic Volume Generated by Test 33
Table 9: Test to Data File Cross-Reference List 46
Table 10: Validated Data Files - Dictionary 46

LIST OF FIGURES

Figure 1: Panelist Recruitment Protocol 14
Figure 2: Testing Architecture 21

1 - INTRODUCTION AND SUMMARY

This Appendix to the Tenth Measuring Broadband America Report,[1] a report on consumer wireline broadband performance in the United States, provides detailed technical background information on the methodology that produced the Report. It covers the process by which the panel of consumer participants was originally recruited and selected for the August 2011 MBA Report, and maintained and evolved over the last ten years. This Appendix also discusses the testing methodology used for the Report and describes how the test data was analyzed.

2 - PANEL CONSTRUCTION

This section describes the background of the study, as well as the methods employed to design the target panel, select volunteers for participation, and manage the panel to maintain the operational goals of the program.

The study aims to measure fixed broadband service performance in the United States as delivered by an Internet Service Provider (ISP) to the consumer's broadband modem. Many factors contribute to end-to-end broadband performance, only some of which are under the control of the consumer's ISP. The methodology outlined here is focused on the measurement of broadband performance within the scope of an ISP's network, and specifically focuses on measuring performance from the consumer Internet access point, or consumer gateway, to a close major Internet gateway point. The actual quality of experience seen by consumers depends on many other factors beyond the consumer's ISP, including the performance of the consumer's in-home network, transit providers, interconnection points, content distribution networks (CDN) and the infrastructure deployed by the providers of content and services. The design of the study methodology allows it to be integrated with other technical measurement approaches that focus on specific aspects of broadband performance (i.e., download speed, upload speed, latency, packet loss), and in the future, could focus on other aspects of broadband performance.

[1] The First Report (2011) was based on measurements taken in March 2011, the Second Report (2012) on measurements taken in April 2012, and the Third (2013) through this, the Tenth (2020) Reports on measurements taken in September of the year prior to the reports' release dates.

2.1 - USE OF AN ALL VOLUNTEER PANEL

During a 2008 residential broadband speed and performance test in the United Kingdom,[2] SamKnows[3] had determined that attrition rates of an all-volunteer panel was lower than a panel maintained with an incentive scheme of monthly payments. Consequently, in designing the methodology for this broadband performance study, the Commission had decided to rely entirely on volunteer consumer broadband subscribers. Volunteers are selected from a large pool of prospective participants according to a plan designed to generate a representative sample of desired consumer demographics, including geographical location, ISP, and speed tier. As an incentive for participation, volunteers are given access to a personal dashboard which allows them to monitor the performance of their broadband service. They are also provided with a measurement device referred to in the study as a "Whitebox," consisting of an off-the-shelf commodity router configured to run custom SamKnows software.[4]

2.2 - SAMPLE SIZE AND VOLUNTEER SELECTION

The Tenth MBA Report relies on data gathered from 2,931 volunteer panelists across the United States. The methodological factors and considerations that influenced the selection of the sample size and makeup include proven practices originating from the first MBA report and test period, and adaptations beyond the first period. Both are described below:

- The panel of U.S. broadband subscribers was initially drawn from a pool of over 175,000 volunteers during a recruitment campaign that ran in May 2010. Since then, to manage attrition and accommodate the evolving range of subscriber demographics (*i.e.*, tiers, technology, population), additional panelists have been recruited through email solicitations by the ISPs as well as through press releases, a web page,[5] social media outreach and blog posts.

---

[2] *See* https://files.samknows.com/~fcc_public/PM_Summer_08.pdf, (last accessed June 21, 2016).

[3] SamKnows is a company that specializes in broadband availability measurement and was retained under contract by the FCC to assist in this study. *See* http://www.samknows.com/.

[4] The Whiteboxes are named after the appearance of the first hardware implementation of the measurement agent. The Whiteboxes remain in consumer homes and continue to run the tests described in this report. Participants may remain in the measurement project as long as it continues and may retain their Whitebox when they end their participation.

[5] https://www.measuringbroadbandamerica.com/

- The volunteer sample was originally organized with a goal of covering major ISPs in the 48 contiguous states across five broadband technologies: DSL, cable, fiber-to-the-home, fixed terrestrial wireless, and satellite.[6]
- Target numbers for volunteers were set across the four Census Regions—Northeast, Midwest, South, and West—to help ensure geographic diversity in the volunteer panel and compensate for differences in networks across the United States.[7]
- A target plan for allocation of Whiteboxes was developed based on the market share of participating ISPs. Initial market share information was based principally on FCC Form 477[8] data filed by participating ISPs for December 2018. This data is further enhanced by the ISPs who brief SamKnows on new products and changes in subscribership numbers which may have occurred after the submission of the 477 data. Speed tiers that comprise the top 80% of a Participating ISP's subscriber base are included. This threshold ensures that we are measuring the ISP's most popular speed tiers and that it is possible to recruit sufficient panelists.
- An initial set of prospective participants was selected from volunteers who had responded directly to SamKnows as a result of media solicitations, as described in detail in Section 2.3. Where gaps existed in the sample plan, SamKnows worked with participating ISPs via email solicitations targeted at underrepresented tiers.
- Since the initial panel was created in 2011, participating ISPs have contacted random subsets of their subscribers by email to replenish cells that were falling short of their desired panel size. Additional recruitment via social media, press releases and blog posts has also taken place.

The sample plan is designed prior to the reporting period and is sent to each ISP by SamKnows. ISPs review this and respond directly to SamKnows with feedback on speed tiers that ought to be included based on the threshold criteria stated above. SamKnows will include all relevant tiers in the final report, assuming a target sample size is available. As this may not be known until after the reporting period is over, a final sample description containing all included tiers is produced and shared with the FCC and ISPs once the reporting period has finished and the data has been processed. Test results from a total of 2,931 panelists were used in the Tenth MBA Report. This

---

[6] At the request of, and with the cooperation of the Department of Commerce and Consumer Affairs, Hawaii, we are also collecting data from the state of Hawaii.

[7] Although the Commission's volunteer recruitment was guided by Census Region to ensure the widest possible distribution of panelists throughout the United States, as discussed below, a sufficient number of testing devices were not deployed to enable, in every case, the evaluation of regional differences in broadband performance. The States associated with each Census Region are described in Table 4.

[8] The FCC Form 477 data collects information about broadband connections to end user locations, wired and wireless local telephone services, and interconnected Voice over Internet Protocol (VoIP) services. *See* https://www.fcc.gov/general/broadband-deployment-data-fcc-form-477 for further information.

figure includes only panelists that are subscribed to the tiers that were tested as part of the sample plan.

The recruitment campaign resulted in the coverage needed to ensure balanced representation of users across the United States. Table 1 shows the number of volunteers with reporting Whiteboxes for the months of September/October 2019 listed by ISP, as well as the percentage of total volunteers subscribed to each ISP. Tables 2 and 3 shows the distributions of the Whiteboxes by State and by Region respectively. This can be compared with the percentage of subscribers per state or region.[9]

Table 1: ISPs, Sample Sizes and Percentages of Total Volunteers

ISP	Sample Size	% of Total Volunteers
CenturyLink	571	19.48%
Charter	250	8.53%
Cincinnati Bell DSL	66	2.25%
Cincinnati Bell Fiber	243	8.29%
Comcast	276	9.42%
Cox	197	6.72%
Frontier DSL	222	7.57%
Frontier Fiber	333	11.36%
Mediacom	188	6.41%
Optimum	162	5.53%
Verizon Fiber	177	6.04%
Windstream	246	8.39%
Total	**2,931**	**100%**

[9] Subscriber data in the Tenth MBA Report is based on the FCC's Internet Access Services Report with data current to December 31, 2017. See *Internet Access Services: Status as of Dec 30, 2017*, Wireline Competition Bureau, Industry Analysis and Technology Division (rel. Nov. 2018), available at https://docs.fcc.gov/public/attachments/DOC-359342A1.pdf.

Table 2: Distribution of Whiteboxes by State

State	Total Boxes	% of Total Boxes	% of Total US Broadband
Alabama	25	0.9%	1.50%
Alaska	0	0.0%	0.23%
Arizona	121	4.1%	1.97%
Arkansas	15	0.5%	0.86%
California	179	6.1%	12.17%
Colorado	86	2.9%	1.71%
Connecticut	56	1.9%	1.13%
Delaware	9	0.3%	0.30%
District of Columbia	4	0.1%	0.27%
Florida	182	6.2%	6.56%
Georgia	85	2.9%	3.18%
Hawaii	17	0.6%	0.47%
Idaho	23	0.8%	0.48%
Illinois	47	1.6%	3.92%
Indiana	43	1.5%	1.92%
Iowa	146	5.0%	0.90%
Kansas	14	0.5%	1.21%
Kentucky	106	3.6%	1.35%
Louisiana	15	0.5%	1.41%
Maine	0	0.0%	0.42%
Maryland	37	1.3%	1.91%
Massachusetts	48	1.6%	2.27%
Michigan	33	1.1%	2.98%
Minnesota	97	3.3%	1.68%
Mississippi	2	0.1%	0.86%
Missouri	67	2.3%	1.78%

Montana	9	0.3%	0.32%
Nebraska	24	0.8%	0.54%
Nevada	24	0.8%	0.91%
New Hampshire	8	0.3%	0.42%
New Jersey	103	3.5%	2.86%
New Mexico	45	1.5%	0.58%
New York	151	5.2%	6.51%
North Carolina	75	2.6%	3.02%
North Dakota	2	0.1%	0.24%
Ohio	316	10.8%	3.55%
Oklahoma	18	0.6%	1.09%
Oregon	78	2.7%	1.26%
Pennsylvania	109	3.7%	3.93%
Rhode Island	7	0.2%	0.31%
South Carolina	15	0.5%	1.45%
South Dakota	3	0.1%	0.26%
Tennessee	13	0.4%	2.01%
Texas	166	5.7%	8.16%
Utah	17	0.6%	0.83%
Vermont	2	0.1%	0.20%
Virginia	87	3.0%	2.44%
Washington	126	4.3%	2.31%
West Virginia	21	0.7%	0.46%
Wisconsin	52	1.8%	1.64%
Wyoming	3	0.1%	0.19%
	2,931		

The distribution of Whiteboxes by Census Region is found in the table on the next page.

Table 3: Distribution of Whiteboxes by Census Region

Census Region	Total Boxes	% Total Boxes	% Total U.S. Broadband Subscribers
Midwest	844	28.8%	21%
Northeast	484	16.5%	18%
South	875	29.9%	37%
West	728	24.8%	24%

The distribution of states associated with the four Census Regions used to define the panel strata are included in the table below.

Table 4: Panelists States Associated with Census Regions

Census Region	States
Northeast	CT MA ME NH NJ NY PA RI VT
Midwest	IA IL IN KS MI MN MO ND NE OH SD WI
South	AL AR DC DE FL GA KY LA MD MS NC OK SC TN TX VA WV
West	AK AZ CA CO HI ID MT NM NV OR UT WA WY

2.3 - PANELIST RECRUITMENT PROTOCOL

Panelists are recruited in the 2011- 2019 panels using the following method:

- Recruitment has evolved since the start of the program. At that time, (2011) several thousand volunteers were initially recruited through an initial public relations and social media campaign led by the FCC. This campaign included discussion on the FCC website and on technology blogs, as well as articles in the press. Currently volunteers are drafted with the help of a recruitment website[10] which keeps them informed about the MBA program and allows them to view MBA data on a dashboard. The composition of the panel is reviewed each year to identify any deficiencies with regard to the sample plan described above. Target demographic goals are set for volunteers based on ISP, speed tier, technology type, and region. Where the pool of volunteers falls short of the desired goal, ISPs send out email messages to their customers asking them to participate in the MBA program. The messages direct interested volunteers to contact SamKnows to request participation in the trial. The ISPs do not know which of the email recipients volunteer. In almost all cases, this ISP outreach allows the program to meet its desired demographic targets.

The mix of panelists recruited using the above methodologies varies by ISP.

A multi-mode strategy was used to qualify volunteers for the 2019 testing period. The key stages of this process were as follows:

1. Volunteers were directed to complete an online form which provided information on the study and required volunteers to submit a small amount of information.
2. Volunteers were selected from respondents to this follow-up email based on the target requirements of the panel. Selected volunteers were then asked to agree to the *User Terms and Conditions* that outlined the permissions to be granted by the volunteer in key areas such as privacy.[11]
3. From among the volunteers who agreed to the User Terms and Conditions, SamKnows selected the panel of participants,[12] each of whom received a Whitebox for self-installation. SamKnows provided full support during the Whitebox installation phase.

The graphic in Figure 1 illustrates the study recruitment methodology.

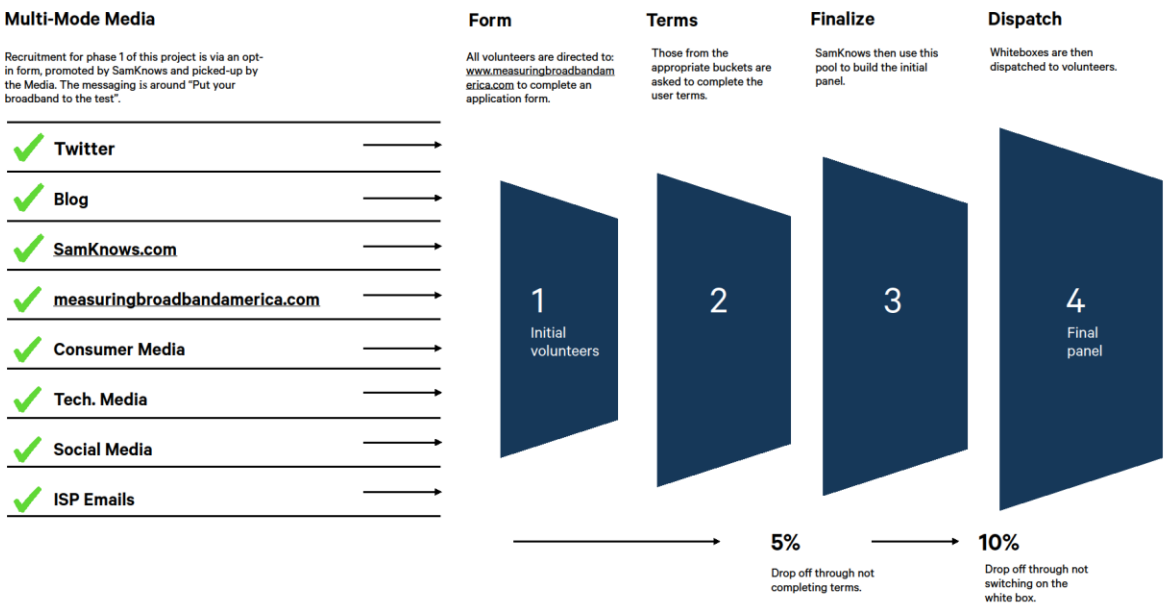
---

[10] The Measuring Broadband America recruitment website is: https://www.measuringbroadbandamerica.com/.

[11] The *User Terms and Conditions* is found in the Reference Documents at the end of this Appendix.

[12] Over 23,000 Whiteboxes have been shipped to targeted volunteers since 2011, of which 6,006 were online and reporting data from the months of September/October 2019.

Figure 1: Panelist Recruitment Protocol



-

2.4 - VALIDATION OF VOLUNTEERS' SERVICE TIER

The methodology employed in this study included verifying each panelist's service tier and ISP against the customer records of participating ISPs.[13] Initial throughput tests were used to confirm reported speeds.

The broadband service tier reported by each panelist was validated as follows:

- When the panelist installed the Whitebox, the device automatically ran an IP address test to check that the ISP identified by the volunteer was correct.
- The Whitebox also ran an initial test which flooded each panelist's connection in order to accurately detect the throughput speed when their deployed Whitebox connected to a test node.
- Each ISP was asked to confirm the broadband service tier reported by each selected panelist.
- SamKnows then took the validated speed tier information that was provided by the ISPs and compared this to both the panelist-provided information, and the actual test results obtained, in order to ensure accurate tier validation.

SamKnows manually completed the following four steps for each panelist:

- Verified that the IP address was in a valid range for those served by the ISP.
- Reviewed data for each panelist and removed data where speed changes such as tier upgrade or downgrade appeared to have occurred, either due to a service change on the part of the consumer or a network change on the part of the ISP.
- Identified panelists whose throughput appeared inconsistent with the provisioned service tier. Such anomalies were re-certified with the consumer's ISP.[14]
- Verified that the resulting downstream-upstream test results corresponded to the ISP-provided speed tiers and updated accordingly if required.

---

[13] Past FCC studies found that a high rate of consumers could not reliably report information about their broadband service, and the validation of subscriber information ensured the accuracy of expected speed and other subscription details against which observed performance was measured. *See* John Horrigan and Ellen Satterwhite, *Americans' Perspectives on Online Connection Speeds for Home and Mobile Devices*, 1 (FCC 2010), available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-298516A1.doc (finding that 80 percent of broadband consumers did not know what speed they had purchased).

[14] For example, when a panelist's upload or download speed was observed to be significantly higher than that of the rest of the tier, it could be inferred that a mischaracterization of the panelist's service tier had occurred. Such anomalies, when not resolved in cooperation with the service provider, were excluded from the Tenth Report, but will be included in the raw bulk data set.

Of the more than 23,000 Whiteboxes that were shipped to panelists since 2011, 6,006[15] units reported sufficient data in September/October 2019, with the participating ISPs validating 4,964 for the reporting period. Of the validated units, 17 percent were reallocated to a different tier following the steps listed above. A total of 2,931 validated units were part of download or upload tiers included in the sample plan and were ultimately included in this report.

A total of 3,075 boxes were excluded for the following reasons:

- 1,763 belonged to users subscribed to plans that were not included in this study
- 263 were excluded due to legacy equipment such as modem that could not fully support the subscribed speeds
- 291 Whiteboxes were legacy models that could not fully support the plan speeds
- 293 belonged to users whose details or subscribed tier could not be successfully validated by the ISP
- 142 Whiteboxes were excluded due to ethernet limitations
- 23 were connected to non-residential plans
- 1 Whitebox was a test unit not to be included in the program
- 7 belonged to employees of ISPs taking part in the MBA program
- And a further 292 were excluded as the test speed profile did not match the product validated by the ISP.

2.5 - PROTECTION OF VOLUNTEERS' PRIVACY

Protecting the panelists' privacy is a major concern for this program. The panel was comprised entirely of volunteers who knowingly and explicitly opted into the testing program. For audit purposes, we retain the correspondence with panelists documenting their opt-in.

All personal data was processed in conformity with relevant U.S. law and in accordance with policies developed to govern the conduct of the parties handling the data. The data were processed solely for the purposes of this study and are presented here and in all online data sets with all personally identifiable information (PII) removed.

A set of materials was created both to inform each panelist regarding the details of the trial, and to gain the explicit consent of each panelist to obtain subscription data from the participating ISPs. These documents were reviewed by the Office of General Counsel of the FCC and the participating ISPs and other stakeholders involved in the study.

[15] This figure represents the total number of boxes reporting during September/October 2019, the month chosen for the Tenth Report. Shipment of boxes continued in succeeding months and these results will be included in the raw bulk data set.

3 - BROADBAND PERFORMANCE TESTING METHODOLOGY

This section describes the system architecture and network programming features of the tests, and other technical aspects of the methods employed to measure broadband performance during this study.

3.1 - RATIONALE FOR HARDWARE-BASED MEASUREMENT APPROACH

Either a hardware or software approach can be used to measure broadband performance. Software approaches are by far the most common and allow for measurements to easily and cost-effectively include a very large sample size. Web-based speed tests fall into this category and typically use Flash applets, Java applets or JavaScript that execute within the user's web browser. These clients download content from remote web servers and measure the throughput. Some web-based performance tests also measure upload speed or round-trip latency.

Other, less common, software-based approaches to performance measurement install applications on the user's computer. These applications run tests periodically while the computer is on.

All software solutions implemented on a consumer's computer, smart phone, or other device connected to the Internet suffer from the following disadvantages:

- The software and computing platform running the software may not be capable of reliably recording the higher speed service tiers currently available.
- The software typically cannot know if other devices on the home network are accessing the Internet when the measurements are being taken. The lack of awareness as to other, non-measurement related network activity can produce inconsistent and misleading measurement data.
- Software measurements may be affected by the performance, quality and configuration of the device.
- Potential bottlenecks, such as Wi-Fi networks and other in-home networks, are generally not accounted for and may result in unreliable data.
- If the device hosting the software uses in-home WIFI access to fixed broadband service, differing locations in the home may impact measurements.
- The tests can only run when the computer is turned on, limiting the ability to provide a 24-hour profile.

- If software tests are performed manually, panelists might only run tests when they experience problems and thus bias the results.

In contrast, the hardware approach used in the MBA program requires the placement of the previously described Whitebox inside the user's home, directly connected to the consumer's service interconnection device (router), via Ethernet cable. The measurement device therefore directly accesses fixed Internet service to the home over this dedicated interface and periodically runs tests to remote targets over the Internet. The use of hardware devices avoids the disadvantages listed earlier with the software approach. However, hardware approaches are much more expensive than the software alternative, are thus more constrained in the achievable panel size, and require correct installation of the device by the consumer or a third party. This is still subject to unintentional errors due to misconfigurations, *i.e.*, connecting the Whitebox incorrectly but these can often be detected in the validation process that follows installation. The FCC chose the hardware approach since its advantages far outweigh these disadvantages.

3.2 - DESIGN OBJECTIVES AND TECHNICAL APPROACH

For this test of broadband performance, as in previous Reports, the FCC used design principles that were previously developed by SamKnows in conjunction with their study of broadband performance in the U.K. The design principles comprise 17 technical objectives:

Table 5: Design Objectives and Methods

#	Technical Objectives	Methodological Accommodations
1	The Whitebox measurement process must not change during the monitoring period.	The Whitebox measurement process is designed to provide automated and consistent monitoring throughout the measurement period.
2	Must be accurate and reliable.	The hardware solution provides a uniform and consistent measurement of data across a broad range of participants.
3	Must not interrupt or unduly degrade the consumer's use of the broadband connection.	The volume of data produced by tests is controlled to avoid interfering with panelists' overall broadband experience, and tests only execute when consumer is not making heavy use of the connection.
4	Must not allow collected data to be distorted by any use of the broadband connection by other applications on the host PC and other devices in the home.	The hardware solution is designed not to interfere with the host PC and is not dependent on that PC.

5	Must not rely on the knowledge, skills and participation of the consumer for its ongoing operation once installed.	The Whitebox is "plug-and-play." Instructions are graphics-based and the installation process has been substantially field tested. Contacts for support are also provided and the outreach once a Whitebox has been dispatched and activated.
6	Must not collect data that might be deemed to be personal to the consumer without consent.	The data collection process is explained in plain language and consumers are asked for their consent regarding the use of their personal data as defined by any relevant data protection legislation.
7	Must be easy for a consumer to completely remove any hardware and/or software components if they do not wish to continue with the MBA program.	Whiteboxes can be disconnected at any time from the home network. As soon as the Whitebox is reconnected the reporting is resumed as before.
8	Must be compatible with a wide range of DSL, cable, satellite and fiber-to-the-home modems.	Whiteboxes can be connected to all modem types commonly used to support broadband services in the U.S., either in a routing or bridging mode, depending on the model.
9	Where applicable, must be compatible with a range of computer operating systems, including, without limitation, Windows XP, Windows Vista, Windows 7, Mac OS and Linux.	Whiteboxes are independent of the PC operating system and therefore able to provide testing with all devices regardless of operating system.
10	Must not expose the volunteer's home network to increased security risk, *i.e.*, it should not be susceptible to viruses, and should not degrade the effectiveness of the user's existing firewalls, antivirus and spyware software.	The custom software in the Whitebox is hardened for security and cannot be accessed without credentials only available to SamKnows. Most user firewalls, antivirus and spyware systems are PC-based. The Whitebox is plugged in to the broadband connection "before" the PC. Its activity is transparent and does not interfere with those protections.
11	Must be upgradeable remotely if it contains any software or firmware components.	The Whitebox can be completely controlled remotely for updates without involvement of the consumer, providing the Whitebox is switched on and connected.
12	Must identify when a user changes broadband provider or package (*e.g.*, by a reverse look up of the consumer's IP address to check provider, and by capturing changes in modem connection speed to identify changes in package).	Ensures regular data pool monitoring for changes in speed, ISP, IP address or performance, and flags when a panelist should notify and confirm any change to their broadband service since the last test execution.

13	Must permit, in the event of a merger between ISPs, separate analysis of the customers of each of the merged ISP's predecessors.	Data are stored based on the ISP of the panelist, and therefore can be analyzed by individual ISP or as an aggregated dataset.
14	Must identify if the consumer's computer is being used on a number of different fixed networks (*e.g.*, if it is a laptop).	The Whiteboxes are broadband dependent, not PC or laptop dependent.
15	Must identify when a specific household stops providing data.	The Whitebox needs to be connected and switched on to push data. If it is switched off or disconnected its absence is detected at the next data push process.
16	Must not require an amount of data to be downloaded which may materially impact any data limits, usage policy, or traffic shaping applicable to the broadband service.	The data volume generated by the information collected does not exceed any policies set by ISPs. Panelists with bandwidth restrictions can have their tests set accordingly.
17	Must limit the possibility for ISPs to identify the broadband connections which form their panel and therefore potentially "game" the data by providing different quality of service to the panel members and to the wider customer base.	ISPs signed a Code of Conduct[16] to protect against gaming test results. While the identity of each panelist was made known to the ISP as part of the speed tier validation process, the actual Unit ID for the associated Whitebox was not released to the ISP so specific test results were not directly assignable against a specific panelist. Moreover, most ISPs had hundreds, and some had more than 1,000, participating subscribers spread throughout their service territory, making it difficult to improve service for participating subscribers without improving service for all subscribers.

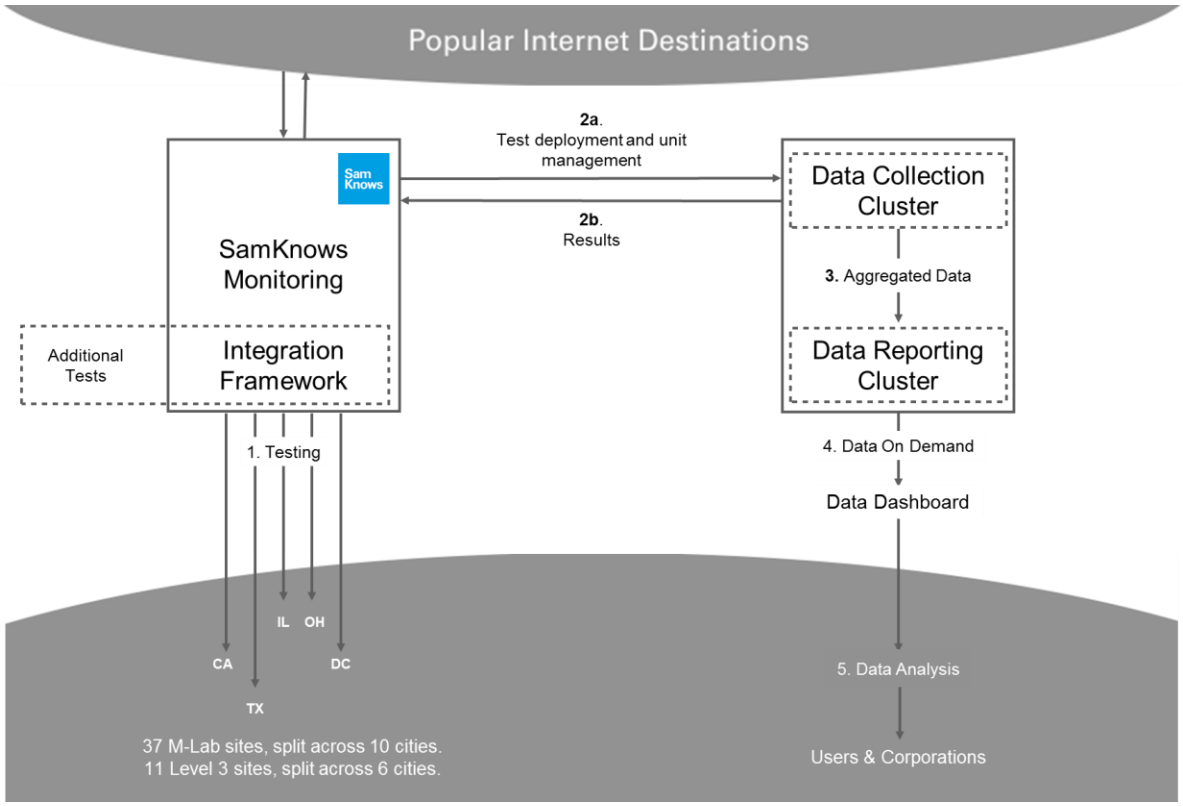
[16] Signatories to the Code of Conduct are: CenturyLink, Charter, Cincinnati Bell, Comcast, Cox, Frontier, Level3, Measurement Lab, Mediacom, NCTA, Optimum, Time Warner Cable, Verizon and Windstream. A copy of the Code of Conduct is included as a Reference Document attached to this Appendix.

3.3 - TESTING ARCHITECTURE

Overview of Testing Architecture

As illustrated in Figure 2, the performance monitoring system comprises a distributed network of Whiteboxes in the homes of members of the volunteer consumer panel. The Whiteboxes are controlled by a cluster of servers, which hosts the test scheduler and the reporting database. The data was collated on the reporting platform and accessed via a reporting interface[17] and secure FTP site. The system also included a series of speed-test servers, which the Whiteboxes called upon according to the test schedule.

Figure 2: Testing Architecture



[17] Each reporting interface included a data dashboard for the consumer volunteers, which provided performance metrics associated with their Whitebox.

Approach to Testing and Measurement

Any network monitoring system needs to be capable of monitoring and executing tests 24 hours a day, seven days a week. Similar to the method used by the television audience measurement industry, each panelist is equipped with a Whitebox, which is self-installed by each panelist and conducts the performance measurements. Since 2011, the project has used three different hardware platforms, described below. The software on each of the Whiteboxes was programmed to execute a series of tests designed to measure key performance indicators (KPIs) of a broadband connection. The tests comprise a suite of applications, written by SamKnows in the programming language C, which were rigorously tested by the ISPs and other stakeholders. The Tenth Report incorporates data from all three types of Whiteboxes and we use the term Whitebox generically. Testing has found that they produce results that are indistinguishable.

During the initial testing period in 2011, the Whitebox provided used hardware manufactured by NETGEAR, Inc. (NETGEAR) and operated as a broadband router. It was intended to replace the panelist's existing router and be directly connected to the cable or DSL modem, ensuring that tests could be run at any time the network was connected and powered, even if all home computers were switched off. Firmware for the Whitebox routers was developed by SamKnows with the cooperation of NETGEAR. In addition to running the latest versions of the SamKnows testing software, the routers retained all of the native functionality of the NETGEAR consumer router.

Following the NETGEAR Whitebox new models were introduced starting with the 2012 testing period. These versions were based upon hardware produced by TP-Link and then later manufactured by SamKnows and operate as a bridge rather than as a router. It connects to the customer's existing router, rather than replacing it, and all hardwired home devices connect to LAN ports on the TP-Link Whitebox. The TP-Link Whitebox / SamKnows Whitebox passively monitors wireless network activity in order to determine when the network is active and defer measurements. It runs a modified version of OpenWrt, an open source router platform based on Linux. All Whiteboxes deployed since 2012 use the TP-Link or SamKnows hardware.

SamKnows Whiteboxes (Whitebox 8.0), introduced in August 2016, have been shown to provide accurate information about broadband connections with throughput rates of up to 1 Gbps.

Home Deployment of the NETGEAR Based Whitebox

This study was initiated by using existing NETGEAR firmware, and all of its features were intended to allow panelists to replace their existing routers with the Whitebox. If the panelist did not have an existing router and used only a modem, they were asked to install the Whitebox according to the usual NETGEAR instructions.

However, this architecture could not easily accommodate scenarios where the panelist had a combined modem/router supplied by their ISP that had specific features that the Whitebox could not provide. For example, some Verizon FiOS gateways connect via a MoCA (Multimedia over Cable) interface and AT&T IPBB gateways provide U-Verse specific features, such as IPTV.

In these cases, the Whitebox was connected to the existing router/gateway and all home devices plugged into the Whitebox. In order to prevent a double-NAT configuration, in which multiple routers on the same network perform network address translation (NAT) and make access to the SamKnows router difficult, the Whitebox was set to dynamically switch to operate as a transparent Ethernet bridge when deployed in these scenarios. All consumer configurations were evaluated and tested by participating ISPs to confirm their suitability.[18]

Home Deployment of the TP-Link Based Whitebox

The TP-Link-based Whitebox, which operates as a bridge, was introduced in response to the increased deployment of integrated modem/gateway devices. To use the TP-Link-based Whitebox, panelists are required to have an existing router. Custom instructions guided these panelists to connect the Whitebox to their existing router and then connect all of their home devices to the Whitebox. This allows the Whitebox to measure traffic volumes from wired devices in the home and defer tests accordingly. As an Ethernet bridge, the Whitebox does not provide services such as network address translation (NAT) or DHCP.

Home Deployment of the SamKnows Whitebox 8.0

The Whitebox 8.0 was manufactured by SamKnows and deployed starting in August 2016. Like the TP-Link device, this Whitebox works as a bridge, rather than a router, and operates in a similar manner. Unlike the NETGEAR and TP-Link hardware, it can handle bandwidths of up to 1 Gbps.

Internet Activity Detection

No tests are performed if the Whiteboxes detect wired or wireless traffic beyond a defined bandwidth threshold. This ensures both that testing does not interfere with consumer use of

[18] The use of legacy equipment has the potential to impede some panelists from receiving the provisioned speed from their ISP, and this impact is captured by the survey.

their Internet service and that any such use does not interfere with testing or invalidate test results.

Panelists were not asked to change their wireless network configurations. Since the TP-Link Whiteboxes and Whitebox 8.0 attach to the panelist's router that may contain a built-in wireless (Wi-Fi) access point, these devices measure the strongest wireless signal. Since they only count packets, they do not need access to the Wi-Fi encryption keys and do not inspect packet content.

Test Nodes (Off-Net and On-Net)

For the tests in this study, SamKnows employed fifty-four core measurement servers as test nodes that were distributed geographically across eleven locations, outside the network boundaries of the participating ISPs. These off-net measurement points were supplemented by additional measurement points located within the networks of some of the ISPs participating in this study, called on-net servers. The core measurement servers were used to measure consumers' broadband performance between the Whitebox and an available reference point that was closest in roundtrip time to the consumer's network address. The distribution of off-net primary reference points operated by M-Lab, Level 3 and Stackpath[19].

On-net secondary reference points operated by broadband providers provided additional validity checks and insight into broadband service performance within an ISP's network. In total, the following 133 measurement servers were deployed for the Tenth Report:

Table 6: Overall Number of Testing Servers

Operated By	Number of Servers
AT&T	9
CenturyLink (inc Qwest)	14
Charter (inc TWC)	18
Comcast	37
Cox	2

[19] Stackpath was added to the list of hosting providers for the MBA project to provide further resilience for the testing platform. Stackpath servers have a minimum 10Gbps – 200Gbps transit / peering links and have been located in the major US cities as per the other hosting providers used for the program.

Frontier	5
Hawaiian Telecom	1
Level 3 (off-net)	13
M-Lab (off-net)	51
Mediacom	1
Optimum	3
Time Warner Cable (now part of Charter)	18
Uhnet (Hawaii)	1
Verizon	2
Windstream	4
Stackpath	10

Test Node Locations

Off-Net Test Nodes

The M-Lab test nodes were located in the following major U.S. Internet peering locations:

- New York City, New York (five locations)
- Chicago, Illinois (five locations)
- Atlanta, Georgia (five locations)
- Miami, Florida (five locations)
- Washington, DC (five locations)
- Mountain View, California (six locations)

- Seattle, Washington (six locations)
- Los Angeles, California (five locations)
- Dallas, Texas (five locations)
- Denver, Colorado (four locations)

The Level 3 nodes were located in the following major U.S. Internet peering locations:

- Chicago, Illinois (two locations)
- Dallas, Texas (two locations)
- New York City, New York (two locations)
- San Jose, California (two locations)
- Washington D.C. (two locations)
- Los Angeles, California (three locations)

The Stackpath nodes were located in the following major U.S. Internet peering locations:

- Ashburn, Virginia (one location)
- Atlanta, Georgia (one location)
- Chicago, Illinois (one location)
- Dallas, Texas (one location)
- Los Angeles, California (one location)
- New York City, New York (one location)
- San Jose, California (one location)
- Seattle, Washington (one location)
- Denver, Colorado (one location)
- Miami, Florida (one location)

On-Net Test Nodes

In addition to off-net nodes, some ISPs deployed their own on-net servers to cross-check the results provided by off-net nodes. Whiteboxes were instructed to test against the off-net M-Lab, Stackpath and Level 3 nodes and the on-net ISP nodes, when available.

The following ISPs provided on-net test nodes:

- CenturyLink[20]
- Charter[21]
- Cincinnati Bell
- Comcast
- Cox
- Frontier
- Mediacom
- Optimum
- Verizon
- Windstream

The same suite of tests was scheduled for these on-net nodes as for the off-net nodes and the same server software developed by SamKnows was used regardless of whether the Whitebox was interacting with on-net or off-net nodes. Off-net test nodes are continually monitored for load and congestion.

While these on-net test nodes were included in the testing, the results from these tests were used as a control set; the results presented in the Report are based only on tests performed using off-net nodes. Results from both on-net and off-net nodes are included in the raw bulk data set that will be released to the public.

Test Node Selection

Each Whitebox fetches a complete list of off-net test nodes and on-net test nodes hosted by the serving ISP from a SamKnows server and measures the round-trip time to each. This list of test servers is loaded at startup and refreshed daily. It then selects the on-net and off-net test nodes with lowest round trip time to test against. The selected nodes may not be the geographically closest node.

Technical details for the minimum requirements for hardware and software, connectivity, and systems and network management are available in the **5.3 - Test Node Briefing** provided in the Reference Document section of this Technical Appendix.

[20] QWest was reported separately from Centurylink in reports prior to 2016. The entities completed merging their test infrastructure in 2016.

[21] Time Warner Cable was reported separately from Charter in reports prior to the Eighth report. The entities completed merging their test infrastructure in early 2018.

3.4 - TESTS METHODOLOGY

Each deployed Whitebox performs the following tests.[22] All tests are conducted with both the on-net and off-net servers except as noted, and are described in more detail in the next section.

Table 7: List of Tests Performed by SamKnows[23]

Metric	Primary Metric(s)
Download speed	Throughput in Megabits per second (Mbps) utilizing three concurrent TCP connections
Upload speed	Throughput in Mbps utilizing three concurrent TCP connections
Web browsing	Total page fetch time and all its embedded resources from a popular website
UDP latency	Average round trip time of a series of randomly transmitted UDP packets distributed over a long timeframe
UDP packet loss	Fraction of UDP packets lost from UDP latency test
Voice over IP	Upstream packet loss, downstream packet loss, upstream jitter, downstream jitter, round trip latency
DNS resolution	Time taken for the ISP's recursive DNS resolver to return an A record[24] for a popular website domain name
DNS failures	Percentage of DNS requests performed in the DNS resolution test that failed
ICMP latency	Round trip time of five evenly spaced ICMP packets
ICMP packet loss	Percentage of packets lost in the ICMP latency test
UDP Latency under load	Average round trip time for a series of evenly spaced UDP packets sent during downstream/upstream sustained tests
Lightweight download speed	Downstream throughput in Megabits per second (Mbps) utilizing a burst of UDP datagrams
Lightweight upload speed	Upstream throughput in Megabits per second (Mbps) utilizing a burst of UDP datagrams

[22] Specific questions on test procedures may be addressed to team@samknows.com.

[23] Other tests may be run on the MBA panel; this list outlines the published tests in the report.

[24] An "A record" is the numeric IP address associated with a domain address such as https://www.fcc.gov.

3.5 - TEST DESCRIPTIONS

The following sub-sections detail the methodology used for the individual tests. As noted earlier, all tests only measure the performance of the part of the network between the Whitebox and the target (*i.e.*, a test node). In particular, the VoIP tests can only approximate the behavior of real applications and do not reflect the impact of specific consumer hardware, software, media codecs, bandwidth adjustment algorithms, Internet backbones and in-home networks.

Download Speed and Upload Speed

These tests measure the download and upload throughput by performing multiple simultaneous HTTP GET and HTTP POST requests to a target test node.

Binary, non-zero content—herein referred to as the payload—is hosted on a web server on the target test node. The test operates for a fixed duration of 10 seconds. It records the average throughput achieved during this 10 second period. The client attempts to download as much of the payload as possible for the duration of the test.

The test uses three concurrent TCP connections (and therefore three concurrent HTTP requests) to ensure that the line is saturated. Each connection used in the test counts the numbers of bytes transferred and is sampled periodically by a controlling thread. The sum of these counters (a value in bytes) divided by the time elapsed (in microseconds) and converted to Mbps is taken as the total throughput of the user's broadband service.

Factors such as TCP slow start and congestion are taken into account by repeatedly transferring small chunks (256 kilobytes, or kB) of the target payload before the real testing begins. This "warm-up" period is completed when three consecutive chunks are transferred at within 10 percent of the speed of one another. All three connections are required to have completed the warm-up period before the timed testing begins. The warm-up period is excluded from the measurement results.

Downloaded content is discarded as soon as it is received, and is not written to the file system. Uploaded content is generated and streamed on the fly from a random source.

The test is performed for both IPv4 and IPv6, where available, but only IPv4 results are reported.

Web Browsing

The test records the averaged time taken to sequentially download the HTML and referenced resources for the home page of each of the target websites, the number of bytes transferred, and the calculated rate per second. The primary measure for this test is the total time taken to download the HTML front page for each web site and all associated images, JavaScript, and stylesheet resources. This test does not measure against the centralized testing nodes; instead

it tests against actual websites, ensuring that the effects of content distribution networks and other performance enhancing factors can be taken into account.

Each Whitebox tests against the following nine websites:[25]

- http://www.edition.cnn.com/
- http://www.bing.com/
- http://www.msn.com/
- http://www.bbc.com/
- http://www.apple.com/
- http://www.ebay.com/
- http://www.m.imdb.com/help/
- http://www.google.com/policies/

The results include the time needed for DNS resolution. The test uses up to eight concurrent TCP connections to fetch resources from targets. The test pools TCP connections and utilizes persistent connections where the remote HTTP server supports them.

The client advertises the user agent as Microsoft Internet Explorer 10. Each website is tested in sequence and the results summed and reported across all sites.

UDP Latency and Packet Loss

These tests measure the round-trip time of small UDP packets between the Whitebox and a target test node.

Each packet consists of an 8-byte sequence number and an 8-byte timestamp. If a response packet is not received within three seconds of sending, it is treated as being lost. The test records the number of packets sent each hour, the average round trip time and the total number of packets lost. The test computes the summarized minimum, maximum, standard deviation and mean from the lowest 99 percent of results, effectively trimming the top (*i.e.*, slowest) 1 percent of outliers.

The test operates continuously in the background. It is configured to randomly distribute the sending of the requests over a fixed interval of one hour (using a Poisson distribution), reporting the summarized results once the interval has elapsed. Approximately two thousand packets are sent within a one-hour period, with fewer packets sent if the line is not idle.

This test is started when the Whitebox boots and runs permanently as a background test. The test is performed for both IPv4 and IPv6, where available, but only IPv4 results are reported.

[25] These websites were chosen based on a list by Alexa, http://www.alexa.com/, of the top twenty websites in October 2010.

Voice over IP

The Voice over IP (VoIP) test operates over UDP and utilizes bidirectional traffic, as is typical for voice calls.

The Whitebox handshakes with the server, and each initiates a UDP stream with the other. The test uses a 64 kbps stream with the same characteristics and properties (*i.e.*, packet sizes, delays, bitrate) as the G.711 codec. 160 byte packets are used. The test measures jitter, delay, and loss.

Jitter is calculated using the Packet Delay Variation (PDV) approach described in section 4.2 of RFC 5481. The 99th percentile is recorded and used in all calculations when deriving the PDV.

DNS Resolutions and DNS Failures

These tests measure the DNS resolution time of an A record query for the domains of the websites used in the web browsing test, and the percentage of DNS requests performed in the DNS resolution test that failed.

The DNS resolution test is targeted directly at the ISP's recursive resolvers. This circumvents any caching introduced by the panelist's home equipment (such as another gateway running in front of the Whitebox) and also accounts for panelists that might have configured the Whitebox (or upstream devices) to use non-ISP provided DNS servers. ISPs provide lists of their recursive DNS servers for the purposes of this study.

ICMP Latency and Packet Loss

These tests measure the round-trip time (RTT) of ICMP echo requests in microseconds from the Whitebox to a target test node. The client sends five ICMP echo requests of 56 bytes to the target test node, waiting up to three seconds for a response to each. Packets that are not received in response are treated as lost. The mean, minimum, maximum, and standard deviation of the successful results are recorded. The number of packets sent and received are recorded too.

Latency Under Load

The latency under load test operates for the duration of the 10-second downstream and upstream speed tests, with results for upstream and downstream recorded separately. While the speed tests are running, the latency under load test sends UDP datagrams to the target server and measures the round-trip time and number of packets lost. Packets are spaced five hundred milliseconds (ms) apart, and a three second timeout is used. The test records the mean, minimum, and maximum round trip times in microseconds. The number of lost UDP packets is also recorded.

This test represents an updated version of the methodology used in the initial August 2011 Report and aligns it with the methodology for the regular latency and packet loss metrics.

Traceroute

A traceroute client is used to send UDP probes to each hop in the path between client and destination. Three probes are sent to each hop. The round-trip times, the standard deviation of the round-trip times of the responses from each hop and the packet loss are recorded. The open source traceroute client "mtr" (https://github.com/traviscross/mtr) is used for carrying out the traceroute measurements.

Lightweight Capacity Test

This test measures the instantaneous capacity of the link using a small number of UDP packets. The test supports both downstream and upstream measurements, conducted independently.

In the downstream mode, the test client handshakes with the test server over TCP, requesting a fixed number of packets to be transmitted back to the client. The client specifies the transmission rate, number of packets and packet size in this handshake. The client records the arrival times of each of the resulting packets returns to it.

In the upstream mode, the client again handshakes with the test server, this time informing it of the characteristics of the stream it is about to transmit. The client then transmits the stream to the server, and the server locally records the arrival times of each packet. At the conclusion of this stream, the client asks the server for its summary of the arrival time of each packet.

With this resulting set of arrival times, the test client calculates the throughput achieved. This throughput may be divided into multiple windows, and an average taken across those, in order to smooth out buffering behavior.

This test uses approximately 99% less data than the TCP speed test and completes in a fraction of the time (100 milliseconds versus 10 seconds). The lightweight capacity test achieves results are within 1% deviation from the existing speed test results on fixed-line connections tested on average.

Table 8: Estimated Total Traffic Volume Generated by Test

The standard test schedule, below, was used across all ISPs, with the exception of Viasat. In 2017, Viasat opted to no longer provide panelists with an increased data allowance to offset the amount of data used by the measurements. This meant that the standard test schedule could no longer be used on Viasat, so a lighter weight test schedule was developed for them.

Standard Test Schedule

Test Name	Test Target(s)	Test Frequency	Test Duration	Est. Daily Volume
Web Browsing	9 popular US websites	Every 2 hours, 24x7	Est. 30 seconds	80 MB
Voice over IP	1 off-net test node	Hourly, 24x7	Fixed 10 seconds at 64k	1.8 MB
	1 on-net test node	Hourly, 24x7	Fixed 10 seconds at 64k	1.8 MB
Download Speed (Capacity – 8x parallel TCP connections)	1 off-net test node	Once 12 am - 6 am Once 6 am - 12 pm Once 12 pm - 6 pm Hourly thereafter	Fixed 10 seconds	107 MB at 10 Mbps
	1 on-net test node	Once 12am-6am, Once 6am-12pm, Once 12pm-6pm, Once 6pm-8pm, Once 8pm-10pm, Once 10pm-12am	Fixed 10 seconds	70 MB at 10 Mbps
Download Speed (Single TCP connection)	1 off-net test node 1 on-net test node	Once in peak hours, once in off-peak hours	Fixed 10 seconds	46 MB at 10 Mbps
Upload Speed (Capacity – 8x parallel TCP connections on terrestrial, 3x on satellite)	1 off-net test node	Once 12am-6am, Once 6am-12pm, Once 12pm-6pm, Hourly thereafter	Fixed 10 seconds	11 MB at 1 Mbps
	1 on-net test node	Once 12am-6am, Once 6am-12pm, Once 12pm-6pm, Once 6pm-8pm, Once 8pm-10pm, Once 10pm-12am	Fixed 10 seconds	7 MB at 1 Mbps

Test Name	Test Target(s)	Test Frequency	Test Duration	Est. Daily Volume
Upload Speed (Single TCP connection)	1 off-net test node 1 on-net test node	Once in peak hours, once in off-peak hours	Fixed 10 seconds	6 MB at 1 Mbps
UDP Latency	2 off-net test nodes (Level3/MLab) 1 on-net test node	Hourly, 24x7 Hourly, 24x7	Permanent Permanent	5.8 MB 2.9 MB
UDP Packet Loss	2 off-net test node 1 on-net test nodes	Hourly, 24x7 Hourly, 24x7	Permanent Permanent	N/A (uses above) N/A (uses above)
Consumption	N/A	24x7	N/A	N/A
DNS Resolution	10 popular US websites	Hourly, 24x7	Est. 3 seconds	0.3 MB
ICMP Latency	1 off-net test node 1 on-net test node	Hourly, 24x7	Est. 5 seconds	0.3 MB
ICMP Packet loss	1 off-net test node 1 on-net test node	Hourly, 24x7	N/A (As IMCP latency)	N/A (uses above)
Traceroute	1 off-net test node 1 on-net test node	Three times a day, 24x7	N/A	N/A
Download Speed IPv6^^	1 off-net test node	Three times a day	Fixed 10 seconds	180 MB at 50 Mbps 72 MB at 20 Mbps 11 MB at 3 Mbps 5.4 MB at 1.5 Mbps
Upload Speed IPv6^^	1 off-net test node	Three times a day	Fixed 10 seconds	172 MB at 2 Mbps 3.6MB at 1 Mbps 1.8MB at 0.5 Mbps
UDP Latency / Loss IPv6^^	2 off-net test nodes (Level3/MLab)	Hourly, 24x7	Permanent	5.8 MB
Lightweight Capacity Test – Download (UDP)	1 off-net test node	Once 12am-6am,	Fixed 1000 packets	9MB

Test Name	Test Target(s)	Test Frequency	Test Duration	Est. Daily Volume
		Once 6am-12pm, Once 12pm-6pm, Hourly thereafter		
Lightweight capacity test – Upload (UDP)	1 off-net test node	Once 12am-6am, Once 6am-12pm, Once 12pm-6pm, Hourly thereafter	Fixed 1000 packets	9MB

Lightweight test schedule (currently Viasat only)

Test Name	Test Target(s)	Test Frequency	Test Duration	Est. Daily Volume
Web Browsing	9 popular US websites	Once 8pm-10-pm	Est. 30 seconds	7MB
Download Speed (Capacity – 8x parallel TCP connections)	1 off-net test node	Once 8pm-10-pm	Fixed 10 seconds	30MB at 10Mbps
Upload Speed (Capacity – 8x parallel TCP connections on terrestrial, 3x on satellite)	1 off-net test node	Once 8pm-10-pm	Fixed 10 seconds	3MB at 1Mbps
UDP Latency	1 off-net test node	Hourly, 24x7	Permanent	1MB
UDP Latency	1 on-net test node	Hourly, 24x7	Permanent	1MB
UDP Packet loss	1 off-net test node	Hourly, 24x7	Permanent	N/A (uses above)
UDP Packet loss	1 on-net test node	Hourly, 24x7	Permanent	N/A (uses above)
Consumption	N/A	24x7	N/A	N/A
DNS Resolution	10 popular US websites	Hourly, 24x7	Est. 3 seconds	0.3MB
ICMP Latency	1 off-net test node 1 on-net test node	Hourly, 24x7	Est. 5 seconds	0.3MB
ICMP Packet Loss	1 off-net test node 1 on-net test node	Hourly, 24x7	N/A (As IMCP latency)	N/A (uses above)
Traceroute	1 off-net test node	Three times a day, 24x7	N/A	N/A

Test Name	Test Target(s)	Test Frequency	Test Duration	Est. Daily Volume
	1 on-net test node			
CDN Performance	Amazon, Apple, Microsoft, Google, Cloudflare, Akamai	Every 2 hours, 24x7	5 seconds	3MB
UDP Latency / Loss **IPv6^**	1 off-net test node	Hourly, 24x7	Permanent	1MB
Lightweight Capacity Test – Download (UDP)	1 off-net test node	Once 12am-6am, Once 6am-12pm, Once 12pm-6pm, Hourly thereafter	Fixed 1000 packets	9MB
Lightweight Capacity Test – Upload (UDP)	1 off-net test node	Once 12am-6am, Once 6am-12pm, Once 12pm-6pm, Hourly thereafter	Fixed 1000 packets	9MB

**Download/upload daily volumes are estimates based upon likely line speeds. All tests will operate at maximum line rate so actual consumption may vary.

^Currently in beta testing.

^^Only carried out on broadband connections that support IPv6.

Tests to the off-net destinations use the nearest (in terms of latency) server from the Level3, M-Lab and StackPath list of test servers. The one exception is the latency and packet loss tests, which operate continuously to Level3, M-Lab and StackPath off-net servers. All tests are also performed to the closest on-net server, where available.

Consumption

This test was replaced by the new data usage test. A technical description for this test is outlined here: https://transition.fcc.gov/oet/mba/Data-Usage-Technical-Methodology_2018-08-24_Final-v1.3.pdf

Cross-Talk Testing and Threshold Manager Service

In addition to the tests described above, for 60 seconds prior to and during testing, a ”threshold manager” service on the Whitebox monitors the inbound and outbound traffic across the WAN interface to calculate if a panelist is actively using the Internet connection. The threshold for

traffic is set to 64 kbps downstream and 32 kbps upstream. Metrics are sampled and computed every 10 seconds. If either of these thresholds is exceeded, the test is delayed for a minute and the process repeated. If the connection is being actively used for an extended period of time, this pause and retry process continues for up to five times before the test is abandoned.

4 - DATA PROCESSING AND ANALYSIS OF TEST RESULTS

This section describes the background for the categorization of data gathered for the Tenth Report, and the methods employed to collect and analyze the test results.

4.1 - BACKGROUND

Time of Day

Most of the metrics reported in the Tenth Report draw on data gathered during the so-called peak usage period of 7:00 p.m. to 11:00 p.m. local time[26]. This time period is generally considered to experience the highest amount of Internet usage under normal circumstances.

ISP and Service Tier

A sufficient sample size is necessary for analysis and the ability to robustly compare the performance of specific ISP speed tiers. In order for a speed tier to be considered for the fixed line MBA Report, it must meet the following criteria:

(a) The speed tier must make up the top 80% of the ISP's subscriber base;
(b) There must be a minimum of 45 panelists that are recruited for that tier who have provided valid data for the tier within the validation period; and
(c) Each panelist must have a minimum of five days of valid data within the validation period.

The study achieved target sample sizes for the following download and upload speeds[27] (listed in alphabetical order by ISP):

Download Speeds:

CenturyLink: 1.5, 3, 7, 10, 12, 20, and 40 Mbps tiers;

[26] This period of time was agreed to by ISP participants in open meetings conducted at the beginning of the program.

[27] Due to the large number of different combinations of upload/download speed tiers supported by ISPs where, for example, a single download speed might be offered paired with multiple upload speeds or vice versa, upload and download test results were analyzed separately.

Charter: 60, 100, and 200 Mbps tiers;
Cincinnati Bell DSL: 5 Mbps tier;
Cincinnati Bell Fiber: 50, 250 and 500 Mbps tier;
Comcast: 60, 150, and 250 Mbps tiers;
Cox: 30, 100, 150, and 300 Mbps tiers;
Frontier DSL: 3, 6, and 12 Mbps tiers;
Frontier Fiber: 50, 75, 100, 150 and 200 Mbps tiers;
Mediacom: 60, 100 and 200 Mbps tiers;
Optimum: 100 and 200 Mbps tiers;
Verizon Fiber: 75, 100 and 1000 Mbps tiers;[28]
Windstream: 3, 10 and 25 Mbps tiers.

Upload Speeds:

CenturyLink: 0.768, 0.896, 2, and 5 Mbps tiers;
Charter: 10, and 20 Mbps tiers;
Cincinnati Bell DSL: 0.768 and 3 Mbps tiers;
Cincinnati Bell Fiber: 10, 100 and 125 Mbps tiers;
Comcast: 5 and 10 Mbps tiers;
Cox: 3, 10, and 30 Mbps tiers;
Frontier DSL: 0.768 Mbps tier;
Frontier Fiber: 50, 75, 100, 150 and 200 Mbps tiers;
Mediacom: 5, 10 and 20 Mbps tiers;
Optimum: 35 Mbps tier;
Verizon Fiber: 75 and 100 Mbps tiers;[29]
Windstream: 1 and 1.5 Mbps tiers.

A file containing averages for each metric from the validated September/October 2019 data can be found on FCC's Measuring Broadband America website.[30] Some charts and tables are divided into speed bands, to group together products with similar levels of advertised performance. The results within these bands are further broken out by ISP and service tier. Where an ISP does not offer a service tier within a specific band or a representative sample could not be formed for tier(s) in that band, the ISP will not appear in that speed band.

---

[28] Verizon's 1 Gbps tier was not included in the final report. 1Gbps tiers may be included in a separate/subsequent report focusing on faster speeds.
[29] Verizon's 1 Gbps tier was not included in the final report. *Id* at n. 28.

[30] See: https://www.fcc.gov/general/measuring-broadband-america-measuring-fixed-broadband.

Results from tests run on speed tiers of 1Gbps were not included in the Tenth Report. This was due to concerns from ISPs that the Whitebox 8.0 could not measure these speeds accurately. An investigation was conducted to establish if this was the case, or if speeds of 1Gbps could be reliably reported on.

Following investigation and testing with one of the ISPs which takes part in the program this conclusion was reached:

The network of the ISP concerned was quite "bursty" in nature, with servers on a 1Gbps network sometimes bursting to 3Gbps. This caused small amounts of packet loss which negatively affected overall speed test results. However once implementing new traffic shaping rules restricting traffic from the server to 1Gbps consistent high speeds were recorded by the Whitebox. The other solution to this specific problem was seen when using a very large number of parallel TCP connections. This investigation established that there is not an issue with the Whitebox 8.0 measuring speeds up to 1Gbps consistently.

4.2 - DATA COLLECTION AND ANALYSIS METHODOLOGY

Data Integrity

To ensure the integrity of the data collected, the following validity checks were developed:

1. *Change of ISP intra-month*: By checking the WHOIS results once a day for the user's IP address, we found units that changed ISP during the month. We only kept data for the ISP where the panelist was active the most.
2. *Change of service tier intra-month*: This validity check found units that changed service tier intra-month by comparing the average sustained throughput observed for the first three days in the reporting period against that for the final three days in the reporting period. If a unit was not online at the start or end of that period, we used the first or final three days when they were actually online. If this difference was over 50 percent, the downstream and upstream charts for this unit were individually reviewed. Where an obvious step change was observed (*e.g.*, from 1 Mbps to 3 Mbps), the data for the shorter period was flagged for removal.
3. *Removal of any failed or irrelevant tests*: This validity check removed any failed or irrelevant tests by removing measurements against any nodes other than the US-based off-net nodes. We also removed measurements using any off-net server that showed a failure rate of 10 percent or greater during a specific one-hour period, to avoid using any out-of-service test nodes.
4. *Removal of any problem Whiteboxes*: We removed measurements for any Whitebox that exhibited greater than or equal to 10 percent failures in a particular one-hour period. This removed periods when the Whitebox was unable to reach the Internet.

Legacy Equipment

In previous reports, we discussed the challenges ISPs face in improving network performance where equipment under the control of the subscriber limits the end-to-end performance achievable by the subscriber.[31] Simply, some consumer-controlled equipment may not be capable of operating fully at new, higher service tiers. Working in open collaboration with all service providers we developed a policy permitting changes in ISP panelists when their installed modems were not capable of meeting the delivered service speed that included several conditions on participating ISPs. First, proposed changes in consumer panelists would only be considered where an ISP was offering free upgrades for modems they owned and leased to the consumer. Second, each ISP needed to disclose its policy regarding the treatment of legacy modems and its efforts to inform consumers regarding the impact such modems may have on

[31] See pgs. 8-9, 2014 Report, pg. 8 of the 2013 Report, as well as endnote 14. http://www.fcc.gov/measuring-broadband-america/2012/july.

their service.

While the issue of DOCSIS 3 modems and network upgrades affect the cable industry today, we may see other cases in the future where customer premises equipment affects the achievable network performance.

In accordance with the above stated policy, 135 Whiteboxes connected to legacy modems were identified and removed from the final data set in order to ensure that the study would only include equipment that would be able to meet its advertised speed. The 95 excluded Whiteboxes were connected to Charter, Comcast, and Cox.

Collation of Results and Outlier Control

All measurement data were collated and stored for analysis purposes as monthly trimmed averages during three time intervals (24 hours, 7:00 p.m. to 11:00 p.m. local time Monday through Friday, 12:00 a.m. to 12:00 a.m. local time Saturday and Sunday). Only participants who provided a minimum of five days of valid measurements and had valid data in each of the three time intervals were included in the September / October 2019 test results. In addition, the top and bottom 1 percent of measurements were trimmed to control for outliers that may have been anomalous or otherwise not representative of actual broadband performance. All results were computed on the trimmed data.[32]

Data was only charted when results from at least 45 separate Whiteboxes was available for individual ISP download speed tiers. Service tiers of 50 or fewer Whiteboxes were noted for possible future panel augmentation.

The resulting final validated sample of data for September/October 2019 included in the MBA Tenth Report was collected from 2,931 participants.

Peak Hours Adjusted to Local Time

Peak hours were defined as weekdays (Mondays through Fridays) between 7:00 p.m. to 11:00 p.m. (inclusive) for the purposes of the study. All times were adjusted to the panelist's local time zone. Since some tests are performed only once every two hours on each Whitebox, the duration of the peak period had to be a multiple of two hours.

Congestion in the Home Not Measured

Download, upload, latency, and packet loss measurements were taken between the panelist's home gateway and the dedicated test nodes provided by M-Lab and Level 3. Web browsing measurements were taken between the panelist's home gateway and nine popular United States-hosted websites. Any congestion within the user's home network is, therefore, not measured by this study. The web browsing measurements are subject to possible congestion at the content provider's side, although the choice of eight popular websites configured to serve high traffic loads reduced that risk.

Traffic Shaping Not Studied

The effect of traffic shaping is not studied in the Tenth Report, although test results were subject to any bandwidth management policies put in place by ISPs. The effects of bandwidth management policies, which may be used by ISPs to maintain consumer traffic rates within

[32] These methods were reviewed with statistical experts by the participating ISPs.

advertised service tiers, may be most readily seen in those charts in the 2016 Report that show performance over 24-hour periods, where tested rates for some ISPs and service tiers flatten for periods at a time.

Analysis of PowerBoost and Other ”Enhancing” Services

The use of transient speed enhancing services marketed under names such as “PowerBoost” on cable connections presented a technical challenge when measuring throughput. These services will deliver a far higher throughput for the earlier portion of a connection, with the duration varying by ISP, service tier, and potentially other factors. For example, a user with a contracted 6 Mbps service tier may receive 18 Mbps for the first 10 MB of a data transfer. Once the “burst window” is exceeded, throughput will return to the contracted rate, with the result that the burst speed will have no effect on very long sustained transfers.

Existing speed tests transfer a quantity of data and divide this quantity by the duration of the transfer to compute the transfer rate, typically expressed in Mbps. Without accounting for burst speed techniques, speed tests employing the mechanism described here will produce highly variable results depending on how much data they transfer or how long they are run. Burst speed techniques will have a dominant effect on short speed tests: a speed test running for two seconds on a connection employing burst speed techniques would likely record the burst speed rate, whereas a speed test running for two hours will reduce the effect of burst speed techniques to a negligible level.

The earlier speed test configuration employed in this study isolated the effects of transient performance enhancing burst speed techniques from the long-term sustained speed by running for a fixed 30 seconds and recording the average throughput at 5 second intervals. The throughput at the 0-5 second interval is referred to as the burst speed and the throughput at the 25-30 second interval is referred to as the actual speed. Testing was conducted prior to the start of trial to estimate the length of time during which the effects of burst speed techniques might be seen. Even though the precise parameters used for burst-speed techniques are not known, their effects were no longer observable in testing after 20 seconds of data transfer.

In the Sixth report we noted that the use of this technology by providers was on the decline. For the Seventh, Eighth, Ninth and Tenth reports, we no longer provide the results of burst-speed since these techniques are now rarely used. The speed test configuration has been altered to shorten the test duration to 10 seconds, as there is no need to run it for 30 seconds any more.

Consistency of Speed Measurements

In addition to reporting on the median speed of panelists, the MBA Report also provides a measure of the consistency of speed that panelists experience in each tier. For purposes of discussion we use the term “80/80 consistent speed” to refer to the minimum speed that was experienced by at least 80% of panelists for at least 80% of the time during the peak periods. The process used in defining this metric for a specific ISP tier is to take each panelist’s set of download or upload speed data during the peak period across all the days of the validated measurement

period and arrange it in increasing order. The speed that corresponds to the 20th percentile represents the minimum speed that the panelist experienced at least 80% of the time. The 20 percentile values of all the panelists on a specific tier are then arranged in an increasing order. The speed that corresponds to the 20th percentile now represents the minimum speed that at least 80% of panelists experienced 80% of the time. This is the value reported as the 80/80 consistent speed for that ISP's tier. We also report on the 70/70 consistent speed for an ISP's tier, which is the minimum speed that at least 70% of the panelists experience at least 70% of the time. We typically report the 70/70 and the 80/80 consistent speeds as a percentage of the advertised speed.

When reporting on these values for an ISP, we weigh the 80/80 or 70/70 consistent speed results (as a percentage of the advertised speed) of each of the ISP's tier based on the number of subscribers to that tier; so as to get a weighted average across all the tiers for that ISP.

Latencies Attributable to Propagation Delay

The speeds at which signals can traverse networks are limited at a fundamental level by the speed of light. While the speed of light is not believed to be a significant limitation in the context of the other technical factors addressed by the testing methodology, a delay of approximately 5ms per 1000 km of distance traveled can be attributed solely to the speed of light (depending on the transmission medium). The geographic distribution and the testing methodology's selection of the nearest test servers are believed to minimize any significant effect. However, propagation delay is not explicitly accounted for in the results.

Limiting Factors

A total of 8,417,695,058 measurements were taken across 144,636,223 unique tests.

All scheduled tests were run, aside from when monitoring units detected concurrent use of bandwidth.

Schedules were adjusted when required for specific tests to avoid triggering data usage limits applied by some ISPs.

4.3 DATA PROCESSING OF RAW AND VALIDATED DATA

The data collected in this program are made available as open data for review and use by the public. Raw and processed data sets, mobile testing software, and the methodologies used to process and analyze data are freely and publicly available. Researchers and developers interested in working with measurement data in raw form will need skills in database management, SQL programming, and statistics, depending on the analysis. A developer FAQ for database configuration and data importing instructions for MySQL and PostgreSQL are available at https://www.fcc.gov/general/database-setup-and-importing-measuring-broadband-america-data-april-2012.

The process flow below describes how the raw collected data was processed for the production of the *Measuring Broadband America Report*. Researchers and developers interested in replicating or extending the results of the Report are encouraged to review the process below and supporting files that provide details.

Raw Data:	Raw data for the chosen period is collected from the measurement database. The ISPs and products that panelists were on are exported to a "unit profile" file, and those that changed during the period are flagged. 2020 Raw Data Links
Validated Data Cleansing:	Data is cleaned. This includes removing measurements when a user changed ISP or tier during the period. Anomalies and significant outliers are also removed at this point. A data cleansing document describes the process in detail. 2020 Data Cleansing Document Link
SQL Processing:	Per-unit results are generated for each metric. Time-of-day averages are computed and a trimmed median is calculated for each metric. The SQL scripts used here are contained in SQL processing scripts available with the release of each report. 2020 SQL Processing Links
Unit Profile:	This document identifies the various details of each test unit, including ISP, technology, service tier, and general location. Each unit represents one volunteer panelists. The unit ID's were randomly generated, which served to protect the anonymity of the volunteer panelists. 2020 Unit Profile link
Excluded Units:	A listing of units excluded from the analysis due to insufficient sample size for that particular ISP's speed tier. 2020 Excluded Units Link
Unit Census Block:	This step identifies the census block (for blocks containing more than 1,000 people) in which each unit running tests is located. Census block is from 2010 census and is in the FIPS code format. We have used block FIPS codes for blocks that contains more than 1,000 people. For blocks with fewer than 1,000 people we have aggregated to the next highest level, *i.e.*, tract, and used the Tract FIPS code, provided there are more than 1,000 people in the tract. In cases where there are less than 1,000 people in a tract we have aggregated to Regional level. 2020 Unit Census Block Link.
Excel Tables & Charts:	Summary data tables and charts in Excel are produced from the averages. These are used directly in the report. 2020 Statistical Averages Links

The raw data collected for each active metric is made available by month in tarred gzipped files. The files in the archive containing active metrics are described in table 9.

Table 9: Test to Data File Cross-Reference List

Test	Validated Data File Name
Download Speed	curr_httpgetmt.csv — IPv4 Tests curr_httpgetmt6.csv — IPv6 Tests
Upload Speed	curr_httppostmt.csv — IPv4 Tests curr_httppostmt6.csv — IPv6 Tests
Web Browsing	curr_webget.csv
UDP Latency	curr_udplatency.csv — IPv4 Tests curr_udplatency6.csv — IPv6 Tests
UDP Packet Loss	curr_udplatency.csv — IPv4 Tests curr_udplatency6.csv — IPv6 Tests
Voice over IP	curr_udpjitter.csv
DNS Resolution	curr_dns.csv
DNS Failures	curr_dns.csv
ICMP Latency	curr_ping.csv
ICMP Packet Loss	curr_ping.csv
Latency under Load	curr_dlping.csv – Downstream latency under load results curr_ulping.csv – Upstream latency under load results
Traceroute	curr_traceroute.csv

Table 10: Validated Data Files - Dictionary

The following Data Dictionary file describes the schema for each active metric test for row level results stored in the files described in table 9.[33] All dtime entries are in the UTC timezone. All durations are in microseconds unless otherwise noted. The location_id field should be ignored.

curr_dlping.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	Target hostname or IP address

[33] This data dictionary is also available on the FCC Measuring Broadband America website, located with the other validated data files available for download.

rtt_avg	Average RTT
rtt_min	Minimum RTT
rtt_max	Maximum RTT
rtt_std	Standard deviation in measured RTT
successes	Number of successes
failiures	Number of failures
location_id	Internal key mapping to unit profile data
curr_dns.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
nameserver	Name server used to handle the DNS request
lookup_host	Hostname to be resolved
response_ip	Field currently unused
rtt	DNS resolution time
successes	Number of successes (always 1 or 0 for this test)
failures	Number of failures (always 1 or 0 for this test)
location_id	Internal key mapping to unit profile data
curr_httpgetmt.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	Target hostname or IP address
address	The IP address of the server (resolved by the client's DNS)
fetch_time	Time the test ran for
bytes_total	Total bytes downloaded across all connections
bytes_sec	Running total of throughput, which is sum of speeds measured for each stream (in bytes/sec), from the start of the test to the current interval
bytes_sec_interval	Throughput at this specific interval (*e.g.*, Throughput between 25-30 seconds)
warmup_time	Time consumed for all the TCP streams to arrive at optimal window size
warmup_bytes	Bytes transferred for all the TCP streams during the warm-up phase

sequence	The interval that this row refers to (*e.g.*, in the US, sequence=0 implies result is for 0-5 seconds of the test)
threads	The number of concurrent TCP connections used in the test
successes	Number of successes (always 1 or 0 for this test)
failures	Number of failures (always 1 or 0 for this test)
location_id	Internal key mapping to unit profile data
curr_httppostmt.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	Target hostname or IP address
address	The IP address of the server (resolved by the client's DNS)
fetch_time	Time the test ran for
bytes_total	Total bytes downloaded across all connections
bytes_sec	Running total of throughput, which is sum of speeds measured for each stream (in bytes/sec), from the start of the test to the current interval
bytes_sec_interval	Throughput at this specific interval (*e.g.*, throughput between 25-30 seconds)
warmup_time	Time consumed for all the TCP streams to arrive at optimal window size
warmup_bytes	Bytes transferred for all the TCP streams during the warm-up phase.
sequence	The interval that this row refers to (*e.g.*, in the US, sequence=0 implies result is for 0-5 seconds of the test)
threads	The number of concurrent TCP connections used in the test
successes	Number of successes (always 1 or 0 for this test)
failures	Number of failures (always 1 or 0 for this test)
location_id	Internal key mapping to unit profile data
curr_ping.csv	ICMP based
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	Target hostname or IP address

rtt_avg	Average RTT
rtt_min	Minimum RTT
rtt_max	Maximum RTT
rtt_std	Standard deviation in measured RTT
successes	Number of successes
failiures	Number of failures
location_id	Internal key mapping to unit profile data
curr_udpjitter.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	Target hostname or IP address
packet_size	Size of each UDP Datagram (bytes)
stream_rate	Rate at which the UDP stream is generated (bits/sec)
duration	Total duration of test
packets_up_sent	Number of packets sent in upstream (measured by client)
packets_down_sent	Number of packets sent in downstream (measured by server)
packets_up_recv	Number of packets received in upstream (measured by server)
packets_down_recv	Number of packets received in downstream (measured by client)
jitter_up	Upstream Jitter measured
jitter_down	Downstream Jitter measured
latency	99th percentile of round trip times for all packets
successes	Number of successes (always 1 or 0 for this test)
failures	Number of failures (always 1 or 0 for this test)
location_id	Internal key mapping to unit profile data
curr_udplatency.csv	UDP based
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	Target hostname or IP address
rtt_avg	Average RTT
rtt_min	Minimum RTT

rtt_max	Maximum RTT
rtt_std	Standard deviation in measured RTT
successes	Number of successes (note: use failures/(successes + failures)) for packet loss)
failiures	Number of failures (packets lost)
location_id	Internal key mapping to unit profile data
curr_ulping.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	Target hostname or IP address
rtt_avg	Average RTT
rtt_min	Minimum RTT
rtt_max	Maximum RTT
rtt_std	Standard deviation in measured RTT
successes	Number of successes
failures	Number of failures
location_id	Internal key mapping to unit profile data
curr_webget.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
target	URL to fetch
address	IP address used to fetch content from initial URL
fetch_time	Sum of time consumed to download HTML content and then concurrently download all resources
bytes_total	Sum of HTML content size and all resources size (bytes)
bytes_sec	Average speed of downloading HTML content and then concurrently downloading all resources (bytes/sec)
objects	Number of resources (images, CSS, …) downloaded
threads	Maximum number of concurrent threads allowed
requests	Total number of HTTP requests made
connections	Total number of TCP connections established
reused_connections	Number of TCP connections re-used

lookups	Number of DNS lookups performed
request_total_time	Total duration of all requests summed together, if made sequentially
request_min_time	Shortest request duration
request_avg_time	Average request duration
request_max_time	Longest request duration
ttfb_total_time	Total duration of the time-to-first-byte summed together, if made sequentially
ttfb_min_time	Shortest time-to-first-byte duration
ttfb_avg_time	Average time-to-first-byte duration
ttfb_max_time	Longest time-to-first-byte duration
lookup_total_time	Total duration of all DNS lookups summed together, if made sequentially
lookup_min_time	Shortest DNS lookup duration
lookup_avg_time	Average DNS lookup duration
lookup_max_time	Longest DNS lookup duration
successes	Number of successes
failures	Number of failures
location_id	Internal key mapping to unit profile data
curr_netusage.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished
wan_rx_bytes	Total bytes received via the WAN interface on the unit (incl. Ethernet and IP headers)
wan_tx_bytes	Total bytes transmitted via the WAN interface on the unit (incl. Ethernet and IP headers)
sk_rx_bytes	Bytes received as a result of active performance measurements
sk_tx_bytes	Bytes transmitted as a result of active performance measurements
location_id	Internal key mapping to unit profile data

curr_lct_dl.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished in UTC

target	Target hostname
address	Target IP address
packets_received	Total number of packets received
packets_sent	Total number of packets sent
packet_size	Packet size
bytes_total	Total number of bytes
duration	Duration of the test in microseconds
bytes_sec	Throughput in bytes/sec
error_code	An internal error code from the test.
successes	Number of successes (always 1 or 0 for this test)
failures	Number of failures (always 1 or 0 for this test)
location_id	Please ignore (this is an internal key mapping to unit profile data)

curr_lct_ul.csv	
unit_id	Unique identifier for an individual unit
dtime	Time test finished in UTC
target	Target hostname
address	Target IP address
packets_received	Total number of packets received
packets_sent	Total number of packets sent
packet_size	Packet size
bytes_total	Total number of bytes
duration	Duration of the test in microseconds
bytes_sec	Throughput in bytes/sec
error_code	An internal error code from the test.
successes	Number of successes (always 1 or 0 for this test)
failures	Number of failures (always 1 or 0 for this test)
location_id	Please ignore (this is an internal key mapping to unit profile data)

5 - Reference Documents

5.1 - USER TERMS AND CONDITIONS

The following document was agreed to by each volunteer panelist who agreed to participate in the broadband measurement study:

End User License Agreement

PLEASE READ THESE TERMS AND CONDITIONS CAREFULLY. BY APPLYING TO BECOME A PARTICIPANT IN THE BROADBAND COMMUNITY PANEL AND/OR INSTALLING THE WHITEBOX, YOU ARE AGREEING TO THESE TERMS AND CONDITIONS.

YOUR ATTENTION IS DRAWN PARTICULARLY TO CONDITIONS 3.5 (PERTAINING TO YOUR CONSENT TO YOUR ISPS PROVIDING CERTAIN INFORMATION AND YOUR WAIVER OF CLAIMS), 6 (LIMITATIONS OF LIABILITY) AND 7 (DATA PROTECTION).

1. Interpretation

1.1. The following definitions and rules of interpretation apply to these terms & conditions.

Connection: the Participant's own broadband internet connection, provided by an Internet Service Provider ("ISP").

Connection Equipment: the Participant's broadband router or cable modem, used to provide the Participant's Connection.

Intellectual Property Rights: all patents, rights to inventions, utility models, copyright and related rights, trademarks, service marks, trade, business and domain names, rights in trade dress or get-up, rights in goodwill or to sue for passing off, unfair competition rights, rights in designs, rights in computer software, database right, moral rights, rights in confidential information (including know-how and trade secrets) and any other intellectual property rights, in each case whether registered or unregistered and including all applications for and renewals or extensions of such rights, and all similar or equivalent rights or forms of protection in any part of the world.

ISP: the company providing broadband internet connection to the Participant during the term of this Program.

Participant/You/Your: the person who volunteers to participate in the Program, under these terms and conditions. The Participant must be the named account holder on the Internet service account with the ISP.

Open Source Software: the software in the Whitebox device that is licensed under an open source license (including the GPL).

Participant's Equipment: any equipment, systems, cabling or facilities provided by the Participant and used directly or indirectly in support of the Services, excluding the Connection Equipment.

Parties: both the Participant and SamKnows.

Party: one of either the Participant or SamKnows.

Requirements: the requirements specified by SamKnows as part of the sign-up process that the Participant must fulfil in order to be selected to receive the Services.

SamKnows/We/Our: the organization providing the Services and conducting the Program, namely:

SamKnows Limited (Co. No. 6510477) of 25 Harley Street, London W1G 9BR

Services / Program: the performance and measurement of certain broadband and Internet services and research program (Broadband Community Panel), as sponsored by the Federal Communications Committee (FCC), in respect of measuring broadband Internet Connections.

Software: the software that has been installed and/or remotely uploaded onto the Whitebox, by SamKnows as updated by SamKnows, from time to time, but not including any Open Source Software.

Test Results: Information concerning the Participant's ISP service results.

Whitebox: the hardware supplied to the Participant by SamKnows with the Software.

1.2. Headings in these terms and conditions shall not affect their interpretation.

1.3. A person includes a natural person, corporate or unincorporated body (whether or not having separate legal personality).

1.4. The schedules form part of these terms and conditions.

1.5. A reference to writing or written includes faxes and e-mails.

1.6.Any obligation in these terms and conditions on a person not to do something includes, without limitation, an obligation not to agree, allow, permit or acquiesce in that thing being done.

2. SamKnows' Commitment to You

2.1 Subject to the Participant complying fully with these terms and conditions, SamKnows shall use reasonable care to:

(a) provide the Participant with the Measurement Services under these terms and conditions;

(b) supply the Participant with the Whitebox and instructions detailing how it should be connected to the Participant's Connection Equipment; and

(c) if requested, SamKnows will provide a pre-paid postage label for the Whitebox to be returned.

(d) comply with all applicable United States, European Union, and United Kingdom privacy laws and directives, and will access, collect, process and distribute the information according to the following principles:

Fairness: We will process data fairly and lawfully;

Specific purpose: We will access, collect, process, store and distribute data for the purposes and reasons specified in this agreement and not in ways incompatible with those purposes;

Restricted: We will restrict our data collection and use practices to those adequate and relevant, and not excessive in relation to the purposes for which we collect the information;

Accurate: We will work to ensure that the data we collect is accurate and up-to-date, working with Participant and his/her ISP;

Destroyed when obsolete: We will not maintain personal data longer than is necessary for the purposes for which we collect and process the information;

Security: We will collect and process the information associated with this trial with adequate security through technical and organizational measures to protect personal data against destruction or loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network.

2.2 In addition, SamKnows shall:

(a) provide Participant with access to a Program-specific customer services email address, which the Participant may use for questions and to give feedback and comments;

(b) provide Participant with a unique login and password in order to access to an online reporting system for access to Participant's broadband performance statistics.

(c) provide Participant with a monthly email with their specific data from the Program or notifying Participant that their individual data is ready for viewing;

(d) provide Participant with support and troubleshooting services in case of problems or issues with their Whitebox;

(e) notify Participant of the end of the FCC-sponsored Program and provide a mechanism for Participant to opt out of any further performance/measuring services and research before collecting any data after termination of the Program;

(f) use only data generated by SamKnows through the Whitebox, and not use any Participant data for measuring performance without Participant's prior written consent; and

(g) not monitor/track Participant's Internet activity without Participant's prior written consent.

2.3 While SamKnows will make all reasonable efforts to ensure that the Services cause no disruption to the performance of the Participant's broadband Connection, including only running tests when there is no concurrent network activity generated by users at the Participant's location. The Participant acknowledges that the Services may occasionally impact the performance of the Connection and agrees to hold SamKnows and their ISP harmless for any impact the Services may have on the performance of their Connection.

3. Participant's Obligations

3.1 The Participant is not required to pay any fee for the provision of the Services by SamKnows or to participate in the Program.

3.2 The Participant agrees to use reasonable endeavors to:

(a) connect the Whitebox to their Connection Equipment within 14 days of receiving it;

(b) not to unplug or disconnect the Whitebox unless (i) they will be absent from the property in which it is connected for more than 3 days and/or (ii) it is reasonably necessary for maintenance of the Participant's Equipment and the Participant agrees that they shall use reasonable endeavors to minimize the length of time the Whitebox is unplugged or disconnected;

(c) in no way reverse engineer, tamper with, dispose of or damage the Whitebox, or attempt to do so;

(d) notify SamKnows within 7 days in the event that they change their ISP or their Connection tier or package (for example, downgrading/upgrading to a different broadband package), to the email address provided by SamKnows;

(e) inform SamKnows of a change of postal or email address by email; within 7 days of the change, to the email address provided by SamKnows;

(f) agrees that the Whitebox may be upgraded to incorporate changes to the Software and/or additional tests at the discretion of SamKnows, whether by remote uploads or otherwise;

(g) on completion or termination of the Services, return the Whitebox to SamKnows by mail, if requested by SamKnows. SamKnows will provide a pre-paid postage label for the Whitebox to be returned;

(h) be an active part of the Program and as such will use all reasonable endeavors to complete the market research surveys received within a reasonable period of time;

(i) not publish data, give press or other interviews regarding the Program without the prior written permission of SamKnows; and

(k) contact SamKnows directly, and not your ISP, in the event of any issues or problems with the Whitebox, by using the email address provided by SamKnows.

3.3 You will not give the Whitebox or the Software to any third party, including (without limitation) to any ISP. You may give the Open Source Software to any person in accordance with the terms of the relevant open source licence.

3.4 The Participant acknowledges that he/she is not an employee or agent of, or relative of, an employee or agent of an ISP or any affiliate of any ISP. In the event that they become one, they will inform SamKnows, who at its complete discretion may ask for the immediate return of the Whitebox.

3.5 THE PARTICIPANT'S ATTENTION IS PARTICULARLY DRAWN TO THIS CONDITION. The Participant expressly consents to having their ISP provide to SamKnows and the Federal Communications (FCC) information about the Participant's broadband service, for example: service address, speed tier, local loop length (for DSL customers), equipment identifiers and other similar information, and hereby waives any claim that its ISPs disclosure of such information to SamKnows or the FCC constitutes a violation of any right or any other right or privilege that the Participant may have under any federal, state or local statute, law, ordinance, court order, administrative rule, order or regulation, or other applicable law, including, without limitation, under 47 U.S.C**.** §§ 222 and 631 (each a "Privacy Law"). If notwithstanding Participant's consent under this Section 3.5, Participant, the FCC or any other party brings any claim or action against any ISP under a Privacy Law, upon the applicable ISPs request SamKnows promptly shall cease collecting data from such Participant and remove from its records all data collected with respect to such Participant prior to the date of such request, and shall not provide such data in any form to the FCC. The Participant further consents to transmission of information from this Program Internationally, including the information provided by the Participant's ISP, specifically the transfer of this information to SamKnows in the United Kingdom, SamKnows' processing of it there and return to the United States.

4. Intellectual Property Rights

4.1 All Intellectual Property Rights relating to the Whitebox are the property of its manufacturer. The Participant shall use the Whitebox only to allow SamKnows to provide the Services.

4.2 As between SamKnows and the Participant, SamKnows owns all Intellectual Property Rights in the Software. The Participant shall not translate, copy, adapt, vary or alter the Software. The Participant shall use the Software only for the purposes of SamKnows providing the Services and shall not disclose or otherwise use the Software.

4.3 Participation in the Broadband Community Panel gives the participant no Intellectual Property Rights in the Test Results. Ownership of all such rights is governed by Federal Acquisition Regulation Section 52.227-17, which has been incorporated by reference in the relevant contract between SamKnows and the FCC. The Participant hereby acknowledges and agrees that SamKnows may make such use of the Test Results as is required for the Program.

4.4 Certain core testing technology and aspects of the architectures, products and services are developed and maintained directly by SamKnows. SamKnows also implements various technical features of the measurement services using particular technical components from a variety of vendor partners including: NetGear, Measurement Lab, TP-Link.

5. SamKnows' Property

The Whitebox and Software will remain the property of SamKnows. SamKnows may at any time ask the Participant to return the Whitebox, which they must do within 28 days of such a request being sent. Once SamKnows has safely received the Whitebox, SamKnows will reimburse the Participant's reasonable postage costs for doing so.

6. Limitations of Liability - THE PARTICIPANT'S ATTENTION IS PARTICULARLY DRAWN TO THIS CONDITION

6.1 This condition 6 sets out the entire financial liability of SamKnows (including any liability for the acts or omissions of its employees, agents, consultants, and subcontractors) to the Participant, including and without limitation, in respect of:

(a) any use made by the Participant of the Services, the Whitebox and the Software or any part of them; and

(b) any representation, statement or tortious act or omission (including negligence) arising under or in connection with these terms and conditions.

6.2 All implied warranties, conditions and other terms implied by statute or other law are, to the fullest extent permitted by law, waived and excluded from these terms and conditions.

6.3 Notwithstanding the foregoing, nothing in these terms and conditions limits or excludes the liability of SamKnows:

(a) for death or personal injury resulting from its negligence or willful misconduct;

(b) for any damage or liability incurred by the Participant as a result of fraud or fraudulent misrepresentation by SamKnows;

(c) for any violations of U.S. consumer protection laws;

(d) in relation to any other liabilities which may not be excluded or limited by applicable law.

6.4 Subject to condition 6.2 and condition 6.3, SamKnows' total liability in contract, tort (including negligence or breach of statutory duty), misrepresentation, restitution or otherwise arising in connection with the performance, or contemplated performance, of these terms and conditions shall be limited to $100.

6.5 In the event of any defect or modification in the Whitebox, the Participant's sole remedy shall be the repair or replacement of the Whitebox at SamKnows' reasonable cost, provided that the defective Whitebox is safely returned to SamKnows, in which case SamKnows shall pay the Participant's reasonable postage costs.

6.6 The Participant acknowledges and agrees that these limitations of liability are reasonable in all the circumstances, particularly given that no fee is being charged by SamKnows for the Services or participation in the Program.

6.7 It is the Participant's responsibility to pay all service and other charges owed to its ISP in a timely manner and to comply with all other ISP applicable terms. The Participant shall ensure that their broadband traffic, including the data pushed by SamKnows during the Program, does not exceed the data allowance included in the Participant's broadband package. If usage allowances are accidentally exceeded and the Participant is billed additional charges from the ISP as a result, SamKnows is not under any obligation to cover these charges although it may choose to do so at its discretion.

7. Data protection - the participation's attention is particularly drawn to this condition.

7.1 The Participant acknowledges and agrees that his/her personal data, such as service tier, address and line performance, will be processed by SamKnows in connection with the program.

7.2 Except as required by law or regulation, SamKnows will not provide the Participant's personal data to any third party without obtaining Participant's prior consent. However, for the avoidance of doubt, the Participant acknowledges and agrees that subject to the privacy polices discussed below, the specific technical characteristics of tests and other technical features associated with the Internet Protocol environment of architecture, including the client's IP address, may be shared with third parties as necessary to conduct the Program and all aggregate statistical data produced as a result of the Services (including the Test Results) may be provided to third parties.

7.3 You acknowledge and agree that SamKnows may share some of Your information with Your ISP, and request information about You from Your ISP so that they may confirm Your service tiers and other information relevant to the Program. Accordingly You hereby expressly waive claim that any disclosure by Your ISP to SamKnows constitutes a violation of any right or privilege that you may have under any law, wherever it might apply.

8. Term and Termination

8.1 This Agreement shall continue until terminated in accordance with this clause.

8.2 Each party may terminate the Services immediately by written notice to the other party at any time. Notice of termination may be given by email. Notices sent by email shall be deemed to be served on the day of transmission if transmitted before 5.00 pm Eastern Time on a working day, but otherwise on the next following working day.

8.3 On termination of the Services for any reason:

(a) SamKnows shall have no further obligation to provide the Services; and

(b) the Participant shall safely return the Whitebox to SamKnows, if requested by SamKnows, in which case SamKnows shall pay the Participant's reasonable postage costs.

8.4 Notwithstanding termination of the Services and/or these terms and conditions, clauses 1, 3.3 and 4 to 14 (inclusive) shall continue to apply.

9. Severance

If any provision of these terms and conditions, or part of any provision, is found by any court or other authority of competent jurisdiction to be invalid, illegal or unenforceable, that provision or part-provision shall, to the extent required, be deemed not to form part of these terms and conditions, and the validity and enforceability of the other provisions these terms and conditions shall not be affected.

10. Entire agreement

10.1 These terms and conditions constitute the whole agreement between the parties and replace and supersede any previous agreements or undertakings between the parties.

10.2 Each party acknowledges that, in entering into these terms and conditions, it has not relied on, and shall have no right or remedy in respect of, any statement, representation, assurance or warranty.

11. Assignment

11.1 The Participant shall not, without the prior written consent of SamKnows, assign, transfer, charge, mortgage, subcontract all or any of its rights or obligations under these terms and conditions.

11.2 Each party that has rights under these terms and conditions acknowledges that they are acting on their own behalf and not for the benefit of another person.

12. No Partnership or Agency

Nothing in these terms and conditions is intended to, or shall be deemed to, constitute a partnership or joint venture of any kind between any of the parties, nor make any party the agent of another party for any purpose. No party shall have authority to act as agent for, or to bind, the other party in any way.

13. Rights of third parties

Except for the rights and protections conferred on ISPs under these Terms and Conditions which they may defend, a person who is not a party to these terms and conditions shall not have any rights under or in connection with these Terms and Conditions.

14. Privacy and Paperwork Reduction Acts

14.1 For the avoidance of doubt, the release of IP protocol addresses of client's Whiteboxes are not PII for the purposes of this program and the client expressly consents to the release of IP address and other technical IP protocol characteristics that may be gathered within the context of the testing architecture. SamKnows, on behalf of the FCC, is collecting and storing broadband performance information, including various personally identifiable information (PII) such as the street addresses, email addresses, sum of data transferred, and broadband performance information, from those individuals who are participating voluntarily in this test. PII not necessary to conduct this study will not be collected. Certain information provided by or collected from you will be confirmed with a third party, including your ISP, to ensure a representative study and otherwise shared with third parties as necessary to conduct the program. SamKnows will not release, disclose to the public, or share any PII with any outside entities, including the FCC, except as is consistent with the SamKnows privacy policy or these Terms and Conditions. See https://www.measuringbroadbandamerica.com/privacy/. The broadband performance

information that is made available to the public and the FCC, will be in an aggregated form and with all PII removed. For more information, see the Privacy Act of 1974, as amended (5 U.S.C. § 552a), and the SamKnows privacy policy.

14.2 The FCC is soliciting and collecting this information authorized by OMB Control No. 3060-1139 in accordance with the requirements and authority of the Paperwork Reduction Act, Pub. L. No. 96-511, 94 Stat. 2812 (Dec. 11, 1980); the Broadband Data Improvement Act of 2008, Pub. L. No. 110-385, Stat 4096 § 103(c)(1); American Reinvestment and Recovery Act of 2009 (ARRA), Pub. L. No. 111-5, 123 Stat 115 (2009); and Section 154(i) of the Communications Act of 1934, as amended.

14.3 *Paperwork Reduction Act of 1995 Notice*. We have estimated that each Participant of this study will assume a one hour time burden over the course of the Program. Our estimate includes the time to sign-up online, connect the Whitebox in the home, and periodic validation of the hardware. If you have any comments on this estimate, or on how we can improve the collection and reduce the burden it causes you, please write the Federal Communications Commission, Office of Managing Director, AMD-PERM, Washington, DC 20554, Paperwork Reduction Act Project (3060-1139). We will also accept your comments via the Internet if you send an e-mail to PRA@fcc.gov. Please DO NOT SEND COMPLETED APPLICATION FORMS TO THIS ADDRESS. You are not required to respond to a collection of information sponsored by the Federal government, and the government may not conduct or sponsor this collection, unless it displays a currently valid OMB control number and provides you with this notice. This collection has been assigned an OMB control number of 3060-1139. THIS NOTICE IS REQUIRED BY THE PAPERWORK REDUCTION ACT OF 1995, PUBLIC LAW 104-13, OCTOBER 1, 1995, 44 U.S.C. SECTION 3507. This notice may also be found at https://www.measuringbroadbandamerica.com/paperwork-reduction-act/.

15. Jurisdiction

These terms and conditions shall be governed by the laws of the state of New York.

SCHEDULE

THE SERVICES

Subject to the Participant complying with its obligations under these terms and conditions, SamKnows shall use reasonable endeavors to test the Connection so that the following information is recorded:

1. Web browsing
2. Video streaming
3. Voice over IP
4. Download speed
5. Upload speed
6. UDP latency
7. UDP packet loss
8. Consumption

9. Availability
10. DNS resolution
11. ICMP latency
12. ICMP packet loss

In performing these tests, the Whitebox will require a variable download capacity and upload capacity per month, which will be available to the Participant in motion 2.3. The Participant acknowledges that this may impact on the performance of the Connection.

1. SamKnows will perform tests on the Participant's Connection by using SamKnows' own data and will not monitor the Participant's content or internet activity. The purpose of this study is to measure the Connection and compare this data with other consumers to create a representative index of US broadband performance.

5.2 - CODE OF CONDUCT

The following Code of Conduct, available at http://data.fcc.gov/download/measuring-broadband-america/2020/Code-of-Conduct-fixed.pdf, was signed by ISPs and other entities participating in the study:

FCC MEASURING BROADBAND AMERICA PROGRAM

FIXED TESTING AND MEASUREMENT

STAKEHOLDERS CODE OF CONDUCT

WHEREAS the Federal Communications Commission of the United States of America (FCC) is conducting a Broadband Testing and Measurement Program, with support from its contractor SamKnows, the purpose of which is to establish a technical platform for the Measuring Broadband America Program Fixed Broadband Testing and Measurement and further to use that platform to collect data;

WHEREAS volunteer panelists have been recruited, and in so doing have agreed to provide broadband performance information measured on their Whiteboxes to support the collection of broadband performance data; and steps have been taken to protect the privacy of panelists to the program's effort to measure broadband performance. WE, THE UNDERSIGNED, as participants and stakeholders in that Fixed Broadband Testing and Measurement, do hereby agree to be bound by and conduct ourselves in accordance with the following principles and shall:

1. At all times act in good faith;
2. Not act, nor fail to act, if the intended consequence of such act or omission is inconsistent with the privacy policies of the program;
3. Not act, nor fail to act, if the intended consequence of such act or omission is to enhance, degrade, or tamper with the results of any test for any individual panelist or broadband provider, except that:

3.1. It shall not be a violation of this principle for broadband providers to:

3.1.1. Operate and manage their business, including modifying or improving services delivered to any class of subscribers that may or may not include panelists among them, provided that such actions are consistent with normal business practices, and

3.1.2. Address service issues for individual panelists at the request of the panelist or based on information not derived from the trial;

3.2. It shall not be a violation of this principle for academic and research purposes to simulate or observe tests and components of the testing architecture, provided that no impact to MBA data or the Internet Service of the subscriber volunteer panelist occurs; and

4. Not publish any data generated by the tests, nor make any public statement based on such data, until such time as the FCC releases data, or except where expressly permitted by the FCC; and

5. Not publish or make use of any test data or testing infrastructure in a manner that would significantly reduce the anonymity of collected data, compromise panelists privacy, or compromise the MBA privacy policy governing collection and analysis of data except that:

5.1. It shall not be a violation of this principle for stakeholder signatories under the direction of the FCC to:

5.1.1. Make use of test data or testing infrastructure to support the writing of FCC fixed Measuring Broadband America Reports;

5.1.2. Make use of test data or testing infrastructure to support various aspects of the testing and architecture for the program including to facilitate data processing or analysis;

5.1.3. Make use of test data or testing infrastructure to support the analysis of collected data or testing infrastructure for privacy risks or concerns, and plan for future measurement efforts;

6. Ensure that their employees, agents, and representatives, as appropriate, act in accordance with this Code of Conduct.

Signatories: _____

Printed: _____

Date: _____

5.3 - TEST NODE BRIEFING

Test Node Briefing

DOCUMENT REFERENCE:

SQ302-002-EN

TEST NODE BRIEFING

Technical information relating to

the SamKnows test nodes

August 2013

Important Notice

Limitation of Liability

The information contained in this document is provided for general information purposes only. While care has been taken in compiling the information herein, SamKnows does not warrant or represent that this information is free from errors or omissions. To the maximum extent permitted by law, SamKnows accepts no responsibility in respect of this document and any loss or damage suffered or incurred by a person for any reason relying on the any of the information provided in this document and for acting, or failing to act, on any information contained on or referred to in this document.

Copyright

The material in this document is protected by Copyright.

1 - SamKnows Test Nodes

In order to gauge an Internet Service Provider's broadband performance at a User's access point, the SamKnows Whiteboxes need to measure the service performance (*e.g.*, upload/download speeds, latency, *etc.*) from the Whitebox to a specific test node. SamKnows supports a number of "test nodes" for this purpose.

The test nodes run special software designed specifically for measuring the network performance when communicating with the Whiteboxes.

It is critical that these test nodes be deployed near to the customer (and their Whitebox). The further the test node is from the customer, the higher the latency and the greater the possibility that third-party networks may need to be traversed, making it difficult to isolate the individual ISP's performance. This is why SamKnows operates so many test nodes all around the world—locality to the customer is critical.

1.1 Test node definition

When referring to "test nodes," we are specifically referring to either the dedicated servers that are under SamKnows' control, or the virtual machines that may be provided to us. In the case of virtual machines provided by Measurement-Lab, Level3, Stackpath and others, the host operating system is under the control of and maintained by these entities and not by SamKnows.

1.2 Test node selection

The SamKnows Whiteboxes select the nearest node by running round-trip latency checks to all test nodes before measurement begins. Note that when we use the term "nearest" we are referring to the test node nearest to the Whitebox from the point of view of network delay, which may not necessarily always be the one nearest geographically.

Alternatively, it is possible to override test node selection based on latency and implement a static configuration so that the Whitebox will only test against the test node chosen by the Administrator. This is so that the Administrator can choose to test any particular test node that is of interest to the specific project and also to maintain configuration consistency. Similarly, test node selection may be done on a scheduled basis, alternating between servers, to collect test data from multiple test nodes for comparison purposes.

1.3 Test node positioning—on-net versus off-net

It is important that measurements collected by the test architecture support the comparison of ISP performance in an unbiased manner. Measurements taken from using the standardized set of "off-net" measurement test nodes (off-net here refers to a test node located outside a specific ISP's network) ensure that the performance of all ISPs can be measured under the same conditions and would avoid artificially biasing results for any one ISP over another. Test nodes located on a particular ISP's network ("on-net" test nodes), might introduce bias with respect to the ISP's own network performance. Thus data to be used to compare ISP performance are collected using "off-net" test nodes, because they reside outside the ISP network.

However, it is also very useful to have test nodes inside the ISP network ("on-net" test nodes). This allows us to:

- Determine what degradation in performance occurs when traffic leaves the ISP network; and
- Check that the off-net test nodes are performing properly (and vice versa).
- By having both on-net and off-net measurement data for each Whitebox, we can have a great deal of confidence in the quality of the data.

2.3 Data that is stored on test nodes

No measurement data collected by SamKnows is stored on test nodes.[34] The test nodes provide a "dumb" endpoint for the Whiteboxes to test against. All measurement performance results are recorded by the Whiteboxes, which are then transmitted from the Whitebox to data collection servers managed by SamKnows..

2 - Test Node Hosting and Locations

SamKnows test nodes reside in major peering locations around the world. Test nodes are carefully sited to ensure optimal connectivity on a market-by-market basis. SamKnows' test

[34] Note that Measurement-Lab runs sidestream measurements for all TCP connections against their test nodes and publishes these data in accordance with their data embargo policy.

infrastructure utilizes nodes made available by Level3, Measurement-Lab, Stackpath and various network operators, as well as under contract with select hosting providers.

2.1 Global Test Nodes

Level3 has provided SamKnows with 11 test nodes to use for the FCC's Measuring Broadband America Program. These test nodes are virtual servers meeting SamKnows specifications. Similarly, Measurement-Lab has also provided SamKnows with test nodes in various cities and countries for use with the Program's fixed measurement efforts. Measurement-Lab provides location hosting for at least three test nodes per site. SamKnows has also contracted with StackPath, a major CDN, to host virtual servers at its 10 US locations. Each location has one node with up to 100Gbps capacity.

Furthermore, SamKnows maintains its own test nodes, which are separate from the test nodes provided by Measurement-Lab and Level3 and Stackpath.

Table 1 below shows the locations of the SamKnows test node architecture supporting the Measuring Broadband America Program.[35] All of these listed test nodes reside outside individual ISP networks and therefore are designated as off-net test nodes. Note, that in many locations there are multiple test nodes installed which may be connected to different providers.

Location	SamKnows	Level3	Measurement-Lab	Stackpath
Atlanta, Georgia			✓	✓
Chicago, Illinois		✓	✓	✓
Dallas, Texas		✓	✓	✓
Los Angeles, California	✓	✓	✓	✓
Miami, Florida			✓	✓
Mountain View, California			✓	

[35] In addition to the test nodes used to support the Measuring Broadband America Program, SamKnows utilizes a diverse fleet of nodes in locations around the globe for other international programs.

New York City, New York	✓	✓	✓	✓
San Jose, California		✓		✓
Seattle, Washington			✓	✓
Washington D.C	✓	✓		
Washington, Virginia			✓	✓
Denver, Colorado			✓	✓

Table 1: Test Node Locations

SamKnows also has access to many test nodes donated by ISPs around the world. These particular test nodes reside within individual ISP networks and are therefore considered on-net test nodes.

ISPs have the advantage of measuring to both on-net and off-net test nodes, which allows them to segment end-to-end network performance and determine the performance of their own network versus third party networks. For example, an ISP can see what impact third party networks have on their end-users Quality of Experience ('QoE') by placing test nodes within their own network and at major National and International peering locations.

Diagram 1 below shows this set-up.

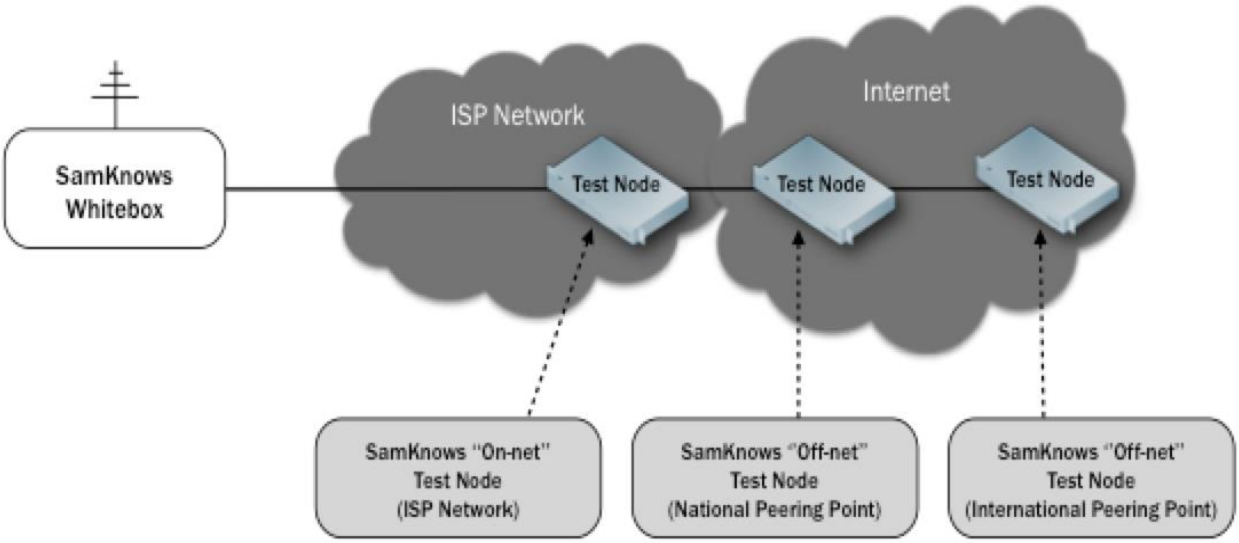


Diagram 1: On-net and Off-net Testing

Both the on-net and off-net test nodes are monitored by SamKnows as part of the global test node fleet. Test node management is explained in more detail within the next section of this document.

3 - Test Node Management

SamKnows test node infrastructure is a critical element of the SamKnows global measurement platform and includes extensive monitoring in place. SamKnows uses a management tool to control and configure the test nodes, while the platform is closely scrutinized using the Nagios monitoring application. System alerts are also in place to ensure the test node infrastructure is always available and operating well within expected threshold bounds.

The SamKnows Operations team continuously checks all test nodes to monitor capacity and overall health. Also included is data analysis to safeguard data accuracy and integrity. This level of oversight not only helps to maintain a healthy, robust platform but also allows us to spot and flag actual network issues and events as they happen. Diagnostic information also supports the Program managers' decision-making process for managing the impact of data accuracy and integrity incidents. This monitoring and administration is fully separate from any monitoring and administration of operating systems and platforms that may be necessary by hosting entities with which SamKnows may be engaged.

3.1 Seamless Test Node Management

SamKnows controls its network of test nodes via a popular open-source management tool called Puppet (**https://puppetlabs.com**). Puppet allows the SamKnows Operations team to easily

manage hundreds of test nodes and ensure that each group of test nodes is configured properly as per each project requirement. Coded in Python, Puppet uses a low-overhead agent installed on each test node that regularly communicates with the controlling SamKnows server to check for updates and ensure the integrity of the configuration.

This method of managing our test nodes allows us to deal with the large number of test nodes without affecting the user's performance in any way. We are also able to quickly and safely make changes to large parts of our test node fleet while ensuring that only the relevant test nodes are updated. This also allows us to keep a record of changes and rapidly troubleshoot any potential problems.

3.2 Proactive Test Node Monitoring

While Puppet handles the configuration and management of the test nodes, Nagios (the most popular online monitoring application) is used by SamKnows to monitor the test nodes. Each test node is configured to send Nagios regular status updates on core metrics such as CPU usage, disk space, free memory, and SamKnows-specific applications. Nagios will also perform active checks of each test nodes where possible, providing us with connectivity information—both via "ping" and connections to any webserver that may be running on the target host.

4 - Test Node Specification and Connectivity

SamKnows maintains a standard specification for all test nodes to ensure consistency and accuracy across the fleet.

4.1 SamKnows test node specifications

All dedicated test nodes must meet the following minimum specifications:

- CPU: Dual core Xeon (2 GHz+)
- RAM: 4 GB
- Disk: 80 GB
- Operating System: CentOS/RHEL 6.x
- Connectivity: Gigabit Ethernet connectivity, with gigabit upstream link.

4.2 Level3 test node specifications

All test nodes provided by level3 meet the following minimum specifications:

- CPU: 2.2 GHz Dual Core
- RAM: 4GB
- Disk: 10 GB

- Operating System: CentOS 6 (64bit)
- Connectivity: 4x1 Gigabit Ethernet (LAG protocol)

4.3 Measurement-Lab Test Node Specifications

All test nodes provided by Measurement-Lab meet the following minimum specifications:

- CPU: 2 GHz 8-core CPU
- RAM: 8 GB
- Disk: 2x100 GB
- OS: CentOS 6.4
- Connectivity: some locations 1 Gbps, some locations 10 Gbps

4.4 Stackpath test node notifications

- CPU Dual Core Xeon (2 GHz+)
- RAM: 8 GB
- Disk: 25 GB root disk
- OS: CentOS 7
- Connectivity: 10 Gbps

4.5 Test Node Connectivity

Measurement test nodes must be connected to a Tier-1 or equivalently neutral peering point. Each test node must be able to sustain 1 Gbps throughput.

At minimum, one publicly routable IPv4 address must be provisioned per-test node. The test node must not be presented with a NAT'd address. It is highly preferable for any new test nodes to also be provisioned with an IPv6 address at installation time.

It is preferred that the test nodes do not sit behind a firewall. If a firewall is used, then care must be taken to ensure that it can sustain the throughput required above.

4.6 Test Node Security

Each of the SamKnows test nodes is firewalled using the IPTables linux firewall. We close any ports that are not required, restrict remote administration to SSH only, and ensure access is only granted from a limited number of specified IP addresses. Only ports that require access from the outside world—for example TCP Port 80 on a webserver—would have that port fully open.

SamKnows regularly checks its rulesets to ensure that there are no outdated rules and that the access restriction is up to date.

SamKnows accounts on each test node are restricted to the systems administration team by default. When required for further work, an authorized SamKnows employee will have an account added.

5 - Test Node Provisioning

SamKnows also has a policy of accepting test nodes provided by network operators providing that

- The test node meets the specifications outlined earlier
- Minimum of 1 Gbps upstream is provided and downstream connectivity to national peering locations

Please note that donated test nodes may also be subject to additional local requirements.

5.1 Installation and Qualification

ISPs are requested to complete an information form for each test node they wish to provision. This will be used by SamKnows to configure the test node on the management system.

SamKnows will then provide an installation script and an associated installation guide. This will require minimal effort from the ISPs involved and will take a very similar form to the package used on existing test nodes.

Once the ISP has completed installation, SamKnows will verify the test node meets performance requirements by running server-to-server tests from known-good servers. These server-to-server measurements will be periodically repeated to verify performance levels.

5.2 Test Node Access and Maintenance

ISPs donating test nodes are free to maintain and monitor the test nodes using their existing toolsets, providing that these do not interfere with the SamKnows measurement applications or system monitoring tools. ISPs must not run resource intensive processes on the test nodes (*e.g.*, packet captures), as this may affect measurements.

ISPs donating test nodes must ensure that these test nodes are only accessed by maintenance staff when absolutely necessary.

SamKnows requests SSH access to the test nodes, with sudo abilities. sudo is a system administration tool that allows elevated privileges in a controlled granular manner. This has greatly helped diagnosis of performance issues with ISP-provided test nodes historically and would enable SamKnows to be far more responsive in investigating issues.

[DOCUMENT ENDS]