

The Need for a New Global Internet Measurement Platform

Abstract

The point of an Internet-wide Quality of Experience (QoE) measurement system is to quantifiably understand the nature, quality, reliability, security, and performance of someone's (or some organization's) Internet service. A QoE measurement system should measure the Internet in a neutral fashion as well, to find and quantify performance issues wherever they may exist and regardless of the party or parties that may be the cause.

In short, such QoE systems help the Internet community at large to determine ways to improve the design, performance, and security of the Internet, and to determine ways to improve end user QoE. This can also support decision-making by a range of actors, such as an end user (i.e. change ISPs, complain to their ISP, troubleshoot their network/software/devices, purchase new equipment/software), an ISP (i.e. add capacity, adjust network parameters, develop/update software or end user devices, develop new services), a regulator (i.e. assess performance across a country, compare the relative development/quality between countries), and application providers (i.e. implement multi-homing, add interconnection capacity, change or add ISPs, reconfigure networks/servers).

Unfortunately, many of today's QoE measurement systems have designs that suffer from designed-in blind spots (limited range of measurement), involve tests that have editorial controls that limit what tests may be run and by whom, are proprietary in nature, lack open access to data, and/or are incapable of measuring emerging end user access capacities (1 Gbps+).

The Internet needs a more open, standards-based system of global QoE measurement. This paper briefly touches on some of the shortcomings of existing systems, and suggests potential goals for a new global Internet QoE measurement platform.

Open and Proprietary Measurement Systems

A range of open and proprietary measurement systems exist today. Open ones tend to be "open" to some degree insofar as how the systems are operated may be publicly documented in whole or in part, and access to the raw data tends also to be openly available. As with open source software, a benefit this openness brings is a wider pool of contributors and extenders of a platform. It also brings the benefit of having more people looking for problems in the design and operation of the platform, which tends to improve quality, security, and other factors over the long term.

In contrast, proprietary measurement platforms do not disclose much about how their systems operate, what data may have been discarded to arrive at a measurement result or report, etc. As a result, in proprietary systems, measurement

data is not openly accessible, nor is the configuration of the tests, and so on. It is difficult to establish trust in proprietary systems as a result of these and related shortcomings. These proprietary systems may also have very small pools of research, development, and operations personnel and so bugs, impairments, and other issues may go unnoticed.

Despite these shortcomings, proprietary measurement systems have received widespread notice and commentary by the press, politicians, policy makers, regulators, and in social media. While this may be due in part to political factors, *proprietary systems fill a vacuum that exists due to a dearth of comparable or superior open systems*. Thus, while some may be quick to criticize proprietary systems, they exist due to a lack of more open alternatives.

Measurement Biases and Blind Spots in Existing Systems

Some of today's measurement systems have a built-in biases and blind spots that may limit both their accuracy and trustworthiness. Some tools focus on consumer-facing ISPs to the exclusion of other network participants (CDNs, transit providers, etc.). Since those participants play key roles in determining customer QoE, ISP-only measurements may present a picture that is at best, incomplete, and at worst, misleading, at least in terms of understanding overall QoE impact. Editorial bias is another issue that negatively impacts some testing methodologies. When tools are offered with the stated goal of proving a preconceived bias about ISPs and the ISP space, they create strong potential for self-selection, skewed results and issues with validity.

For example, some of today's measurement systems have limited themselves to specific and unique situations rather than measuring the full, end-to-end Internet ecosystem. Furthermore, some systems may only measure isolated situations that many times are in control of the measurement system rather than the ISPs they are claiming to measure. True Internet QoE is determined by decisions made by multiple parties, which have a direct and material impact on the QoE measurement.

These biases are counterproductive if the stated goal of a measurement system is truly to measure the Internet in a neutral fashion, to find and quantify performance issues wherever they may exist and regardless of the party or parties that may be the cause, to determine ways to improve the design, performance, and security of the Internet, and to determine ways to improve end user QoE.

While systems to measure ISPs most certainly have value in the proper context, when the focus is on specific parties, or pairs, to the exclusion of all or most others, then a significant measurement blind spot exists in those systems. Thus, the inherent bias in many measurement systems creates measurement blind spots, which might also be considered "limited range of measurement". This is due to the fact that many players in the Internet ecosystem can and do influence end user QoE, but many measurement systems are designed to solely focus on last mile ISPs. Thus,

these systems fail to measure many other parts on the Internet and parties that can and do influence end user QoE.

Administrative Shortcomings of Existing Systems

Some existing systems do not display data visualizations or other statistics on an automated, real-time basis. This creates in some cases a significant time lag between the collection of data and the ability of various stakeholders to take action based upon that data, such as to fix a technical impairment in a router someplace. In other cases systems rely on manual processes to copy data into separate documents, which must be distributed to stakeholders via email or other ad hoc method.

In contrast, some systems will show data visualizations and other statistics on a real-time basis and will automatically create and post/distribute reports (such as monthly/quarterly reports). Some systems even show and track outages, such as BGP-related outages that disrupt network interconnection and performance, and do not rely in any shape or form on manual, administrative action.

Technical Shortcomings of Existing Systems

Most existing systems are incapable of consistently measuring end user connection capacities that exceed 100 Mbps. In some cases this is due to the end point placed in end user homes being unable to measure speeds that exceed 100 Mbps. For a simple example, a system will be unable to measure a connection provisioned at 250 Mbps if the measurement module in the home has a physical 100 Mbps Ethernet interface.

This may also be due to technical shortcomings or constraints in the measurement servers and their associated networks. For example, one part system used for the FCC's Measuring Broadband America platform (outside of the SamKnows system) has for several annual measurement periods been affected by server issues, including software impairments or monitoring gaps, potentially insufficient entrance controls to limit the number of simultaneous tests in order to ensure full and accurate capacity testing, and insufficient interconnection capacities between measurement servers and end users.

In this example, one function that is missing is scheduled tests and active monitoring to determine the health and performance of the measurement system itself, particularly the servers against which end user tests may be targeted. It therefore appears to be uncommon for measurement systems to be calibrated at predetermined intervals to ascertain its maximum capabilities, or that thresholds have been set to ensure measurement load does not exceed those thresholds, though some measurement systems do implement such best practices.

In another, simpler example, the same capacity constraints that affect last mile network links may also affect servers. For example, if a particular measurement server has a single 1 Gbps Ethernet interface but the measurement system is attempting to simultaneously measure multiple homes that have 1 Gbps access

capacities, then the system will measure each home as having less than 1 Gbps, due solely to server-side constraints and a lack of proper test admission control.

Editorial Shortcomings of Existing Systems

Some existing measurement systems have a steering committee or other decision-making group to determine who may create a test, run a test, or access data, among other things. Such a group may for example choose to reject a new measurement test if it is viewed as “competing” with another existing or more politically favored test. Such groups also raise the risk of conflicts of interest or other issues influencing acceptance of new measurement tests, access to raw data, input on analytical reports, and related work.

In contrast, some open measurement systems (e.g. RIPE Atlas) use credit-based systems whereby if a party is providing a certain “value” of system resources to a measurement network, then those credits may be spent in order to run measurements. Beyond that, an open measurement system that did not have a stringent framework for accepting new tests would in essence be free from editorial control (with exceptions to protect the availability and security of the platform, privacy of end users, etc.).

Proposed Design of a New Global Internet QoE Measurement Platform

The following are proposed design goals for a new global measurement system:

1. It should be based in open standards, using RFCs that come out of the Internet Engineering Task Force’s Large-Scale Measurement of Broadband Performance (LMAP) working group and the IP Performance Metrics (IPPM) working group.
2. The design should be openly documented
3. All measurement data should be transparent and openly accessible, with appropriate end user data privacy protections.
4. It should be possible to measure end user connection capacities of at least 1 Gbps at launch, with the ability to measure 10 Gbps, 40 Gbps, 100 Gbps, and 1 Tbps within some set interval of launch.
5. The measurement system should be regularly calibrated to ensure it is not impaired in any way, and actively monitored to ensure that this is the case.
6. The measurement system should be monitored to ensure that measurement tests do not negatively impact network performance and end user QoE due to the tests themselves, such as via careful analysis of aggregate load.
7. Any test that runs over the measurement network must be open source, so that others may understand precisely how it works and may build and improve upon it in the future.
8. There will be no editorial control exerted over the measurements that may run over the network. Thus, there is no steering group or other organization that may decide what person or organization may run measurements, or even what measurements may run (subject to some reasonable criteria such as that the measurements are lawful, non-abusive, protect end user privacy, etc.). The most open and fair system so far in this area is the RIPE Atlas

- system, which is credit-based and has many other attributes that are worth studying.
9. It should be inclusive and able to measure the performance of ALL network interconnections between Autonomous Networks.
 10. It should be able to measure interconnects that handle Internet traffic.
 11. It should be able to measure the performance of Content Delivery Networks, such as path selection, localization, etc.
 12. It should be able to perform application-specific measurements, such as HTTP, HTTPS, DNS, SMTP, POP3, IMAP4 tests, as well as tests of independent application providers (i.e. VoIP, video streaming, gaming).
 13. It should measure any site, application, or system that comprises 5% or more of peak hour Internet usage. This suggests a certain expectation of transparency and openness that is one of the responsibilities that comes with the success of achieving that level of national, regional, or global Internet usage.
 14. It should be able to integrate easily with other open measurement systems that share similar attributes, such as the RIPE Atlas system.
 15. A baseline of measurement tests should be run on a continuous basis in order to establish metrics that are not influenced by end user self-selection bias or other end user behavior.
 16. While the primary goal of the system is measure end user QoE, the system must also measure performance and other network characteristics at points all around the Internet, given that these other aspects of the Internet have a direct or indirect influence on end user QoE.

Submitted by:

Jason Livingood

Comcast - Vice President, Internet Services

email: jason_livingood@cable.comcast.com

phone: +1-215-286-7813

Twitter: <http://www.twitter.com/jlivingood>

LinkedIn: <http://www.linkedin.com/in/jlivingood>