INTERNET-DRAFT
Intended Status: Informational
Expires: April 5, 2014

Marc Linsner
Cisco Systems
Philip Eardley
Trevor Burbridge
BT

October 2, 2013

Large-Scale Broadband Measurement Use Cases
draft-linsner-lmap-use-cases-04

Abstract

Measuring broadband performance on a large scale is important for network diagnostics by providers and users, as well for as public policy. To conduct such measurements, user networks gather data, either on their own initiative or instructed by a measurement controller, and then upload the measurement results to a designated measurement server. Understanding the various scenarios and users of measuring broadband performance is essential to development of the system requirements. The details of the measurement metrics themselves are beyond the scope of this document.

Status of this Memo

Table of Contents

1 Introduction

Large-scale measurement efforts in [LMAP-REQ] describe three use cases to be considered in deriving the requirements to be used in developing the solution. This documents attempts to describe those use cases in further detail and include additional use cases.

1.1 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2 Use Cases

2.1 Internet Service Provider (ISP) Use Case

An ISP, or indeed another network operator, needs to understand the performance of their networks, the performance of the suppliers (downstream and upstream networks), the performance of services, and the impact that such performance has on the experience of their customers. In addition they may also desire visibility of their competitor's networks and services in order to be able to benchmark and improve their own offerings. Largely the processes that ISPs operate (which are based on network measurement) include:

- o Identifying, isolating and fixing problems in the network, services or with CPE and end user equipment. Such problems may be common to a point in the network topology (e.g. a single exchange), common to a vendor or equipment type (e.g. line card or home gateway) or unique to a single user line (e.g. copper access). Part of this process may also be helping users understand whether the problem exists in their home network or with an over-the-top service instead of with their BB product.

- o Design and planning. Through identifying the end user experience the ISP can design and plan their network to ensure specified levels of user experience. Services may be moved closer to end users, services upgraded, the impact of QoS assessed or more capacity deployed at certain locations. SLAs may be defined at network or product boundaries.

- o Benchmarking and competitor insight. The operation of sample panels across competitor products can enable and ISP to assess where they play in the market, identify opportunities where other products operate different technology, and assess the performance

of network suppliers that are common to both operators.

o Understanding the quality experienced by customers. Alongside benchmarking competitors, gaining better insight into the user's service through a sample panel of the operator's own customers. The end-to-end perspective matters, across home /enterprise networks, peering points, CDNs etc.

o Understanding the impact and operation of new devices and technology. As a new product is deployed, or a new technology introduced into the network, it is essential that its operation and impact on other services is measured. This also helps to quantify the advantage that the new technology is bringing and support the business case for larger roll-out.

2.2 Regulators

Regulators in jurisdictions around the world are responding to consumers' adoption of broadband technology solution for traditional telecommunications and media services by reviewing the historical approaches to regulating these industries and services and in some cases modifying existing approaches or developing new solutions.

Some jurisdictions have responded to a perceived need for greater information about broadband performance in the development of regulatory policies and approaches for broadband technologies by developing large-scale measurement programs. Programs such as the U.S. Federal Communications Commission's Measuring Broadband America, U.K. Ofcom's UK Broadband Speeds reports and a growing list of other programs employ a diverse set of operational and technical approaches to gathering data in scientifically and statistical robust ways to perform analysis and reporting on diverse aspects of broadband performance.

While each jurisdiction responds to distinct consumer, industry, and regulatory concerns, much commonality exists in the need to produce datasets that are able to compare multiple broadband providers, diverse technical solutions, geographic and regional distributions, and marketed and provisioned levels and combinations of broadband services.

Regulators role in the development and enforcement of broadband policies also require that the measurement approaches meet a high level of verifiability, accuracy and fairness to support valid and meaningful comparisons of broadband performance

LMAP standards could answer regulators shared needs by providing scalable, cost-effective, scientifically robust solutions to the

measurement and collection of broadband performance information.

2.2.1 Measurement Providers

In some jurisdictions, the role of measuring is provided by a measurement provider. Measurement providers measure a network performance from users to multiple content providers to show a performance of the actual network. Users need to know a performance that are using. In addition, they need to know a performance of other ISP of same location as information for selecting the network. Measurement providers will show the measurement result with measurement methods and measurement parameters.

2.2.2 Benchmarking and competitor insight

An operator may want to check that the results reported by the regulator match its own belief about how its network is performing. There is quite a lot of variation in underlying line performance for customers on (say) a nominal 20Mb/s service, so it is possible for two panels of ~100 probes to produce different results.

An operator may also want more detailed understanding of its competitors, beyond that reported by the regulator - probably by getting a third party to establish a panel of probes in its rival ISPs. Measurements could, for example, help an operator: target its marketing by showing that it's 'best for video streaming' but 'worst for web browsing'; gain detailed insight into the strengths and weaknesses of different access technologies (DSL vs cable vs wireless); understand market segments that it currently doesn't serve; and so on.

The characteristics of large scale measurements that emerge from these examples are very similar to the sub use case above:

1. Averaged data (over say 1 month) is generally ok

2. A panel (subset) of only a few customers is OK

3. Both active and passive measurements are possible, though the former seems easier

4. Regularly scheduled tests are fine (providing active tests back off if the customer is using the line). Scheduling can be done some time ahead ('starting tomorrow, run the following test every day').

5. The performance metrics are whatever the operator wants to benchmark. As well as QoE measures, it may want to measure some

network-specific parameters.

6. As well as the performance of the access link, the performance of different network segments, including end-to-end.

2.3 Fixed and Mobile Service

From a consumer perspective, the differentiation between fixed broadband and mobile (cellular) service is blurring as the applications used are very similar. Hence, similar measurements will take place on both fixed and mobile broadband services.

3 Details of ISP Use Case

3.1 Existing Capabilities and Shortcomings

In order to get reliable benchmarks some ISPs use vendor provided hardware measurement platforms that connect directly to the home gateway. These devices typically perform a continuous test schedule, allowing the operation of the network to be continually assessed throughout the day. Careful design ensures that they do not detrimentally impact the home user experience or corrupt the test results by testing when the user is also using the Broadband line. While the test capabilities of such probes are good, they are simply too expensive to deploy on mass scale to enable detailed understanding of network performance (e.g. to the granularity of a single backhaul or single user line). In addition there is no easy way to operate similar tests on other devices (eg set top box) or to manage application level tests (such as IPTV) using the same control and reporting framework.

ISPs also use speed and other diagnostic tests from user owned devices (such as PCs, tablets or smartphones). These often use browser related technology to conduct tests to servers in the ISP network to confirm the operation of the user BB access line. These tests can be helpful for a user to understand whether their BB line has a problem, and for dialogue with a helpdesk. However they are not able to perform continuous testing and the uncontrolled device and home network means that results are not comparable. Producing statistics across such tests is very dangerous as the population is self-selecting (e.g. those who think they have a problem).

Faced with a gap in current vendor offerings some ISPs have taken the approach of placing proprietary test capabilities on their home gateway and other consumer device offerings (such as Set Top Boxes). This also means that different device platforms may have different

and largely incomparable tests, developed by different company sub-divisions managed by different systems.

3.2 Understanding the quality experienced by customers

Operators want to understand the quality of experience (QoE) of their broadband customers. The understanding can be gained through a "panel", ie a measurement probe is deployed to a few 100 or 1000 of its customers. The panel needs to be a representative sample for each of the operator's technologies (FTTP, FTTC, ADSL...) and broadband options (80Mb/s, 20Mb/s, basic...), ~100 probes for each. The operator would like the end-to-end view of the service, rather than (say) just the access portion. So as well as simple network statistics like speed and loss rates they want to understand what the service feels like to the customer. This involves relating the pure network parameters to something like a 'mean opinion score' which will be service dependent (for instance web browsing QoE is largely determined by latency above a few Mb/s).

An operator will also want compound metrics such as "reliability", which might involve packet loss, DNS failures, re-training of the line, video streaming under-runs etc.

The operator really wants to understand the end-to-end service experience. However, the home network (Ethernet, wifi, powerline) is highly variable and outside its control. To date, operators (and regulators) have instead measured performance from the home gateway. However, mobile operators clearly must include the wireless link in the measurement.

Active measurements are the most obvious approach, ie special measurement traffic is sent by - and to - the probe. In order not to degrade the service of the customer, the measurement data should only be sent when the user is silent, and it shouldn't reduce the customer's data allowance. The other approach is passive measurements on the customer's real traffic; the advantage is that it measures what the customer actually does, but it creates extra variability (different traffic mixes give different results) and especially it raises privacy concerns.

From an operator's viewpoint, understanding customers better enables it to offer better services. Also, simple metrics can be more easily understood by senior managers who make investment decisions and by sales and marketing.

The characteristics of large scale measurements that emerge from these examples:

1. Averaged data (over say 1 month) is generally ok

2. A panel (subset) of only a few customers is OK

3. Both active and passive measurements are possible, though the former seems easier

4. Regularly scheduled tests are fine (providing active tests back off if the customer is using the line). Scheduling can be done some time ahead ('starting tomorrow, run the following test every day').

5. The operator needs to devise metrics and compound measures that represent the QoE

6. End-to-end service matters, and not (just) the access link performance

3.3 Understanding the impact and operation of new devices and technology

Another type of measurement is to test new capabilities and services before they are rolled out. For example, the operator may want to: check whether a customer can be upgraded to a new broadband option; understand the impact of IPv6 before it makes it available to its customers (will v6 packets get through, what will the latency be to major websites, what transition mechanisms will be most is appropriate?); check whether a new capability can be signaled using TCP options (how often it will be blocked by a middlebox? - along the lines of some existing experiments) [Extend TCP]; investigate a quality of service mechanism (eg checking whether Diffserv markings are respected on some path); and so on.

The characteristics of large scale measurements that emerge from these examples are:

1. New tests need to be devised that test a prospective capability.

2. Most of the tests are probably simply: "send one packet and record what happens", so an occasional one-off test is sufficient.

3. A panel (subset) of only a few customers is probably OK, to gain an understanding of the impact of a new technology, but it may be necessary to check an individual line where the roll-out is per customer.

4. An active measurement is needed.

3.4 Design and planning

Operators can use large scale measurements to help with their network planning - proactive activities to improve the network.

For example, by probing from several different vantage points the operator can see that a particular group of customers has performance below that expected during peak hours, which should help capacity planning. Naturally operators already have tools to help this - a network element reports its individual utilisation (and perhaps other parameters). However, making measurements across a path rather than at a point may make it easier to understand the network. There may also be parameters like bufferbloat that aren't currently reported by equipment and/or that are intrinsically path metrics.

With better information, capacity planning and network design can be more effective. Such planning typically uses simulations to emulate the measured performance of the current network and understand the likely impact of new capacity and potential changes to the topology. It may also be possible to run stress tests for risk analysis, for example 'if whizzy new application (or device) becomes popular, which parts of my network would struggle, what would be the impact on other services and how many customers would be affected'. What-if simulations could help quantify the advantage that a new technology brings and support the business case for larger roll-out. This approach should allow good results with measurements from a limited panel of customers.

Another example is that the operator may want to monitor performance where there is a service level agreement. This could be with its own customers, especially enterprises may have an SLA. The operator can proactively spot when the service is degrading near to the SLA limit, and get information that will enable more informed conversations with the customer at contract renewal.

An operator may also want to monitor the performance of its suppliers, to check whether they meet their SLA or to compare two suppliers if it is dual-sourcing. This could include its transit operator, CDNs, peering, video source, local network provider (for a global operator in countries where it doesn't have its own network), even the whole network for a virtual operator.

Through a better understanding of its own network and its suppliers, the operator should be able to focus investment more effectively - in the right place at the right time with the right technology.

The characteristics of large scale measurements emerging from these examples:

1. A key challenge is how to integrate results from measurements into existing network planning and management tools

2. New tests may need to be devised for the what-if and risk analysis scenarios.

3. Capacity constraints first reveal themselves during atypical events (early warning). So averaging of measurements should be over a much shorter time than the sub use case discussed above.

4. A panel (subset) of only a few customers is OK for most of the examples, but it should probably be larger than the QoE use case #1 and the operator may also want to regularly change who is in the subset, in order to sample the revealing outliers.

5. Measurements over a segment of the network ("end-to-middle") are needed, in order to refine understanding, as well as end-to-end measurements.

6. The primary interest is in measuring specific network performance parameters rather than QoE.

7. Regularly scheduled tests are fine

8. Active measurements are needed; passive ones probably aren't

3.5 Identifying, isolating and fixing network problems

Operators can use large scale measurements to help identify a fault more rapidly and decide how to solve it.

Operators already have Test and Diagnostic tools, where a network element reports some problem or failure to a management system. However, many issues are not caused by a point failure but something wider and so will trigger too many alarms, whilst other issues will cause degradation rather than failure and so not trigger any alarm. Large scale measurements can help provide a more nuanced view that helps network management to identify and fix problems more rapidly and accurately. The network management tools may use simulations to emulate the network and so help identify a fault and assess possible solutions.

One example was described in [IETF85-Plenary]. The operator was running a measurement panel for reasons discussed in sub use case #1. It was noticed that the performance of some lines had unexpectedly degraded. This led to a detailed (off-line) investigation which discovered that a particular home gateway upgrade had caused a

(mistaken!) drop in line rate.

Another example is that occasionally some internal network management event (like re-routing) can be customer-affecting (of course this is unusual). This affects a whole group of customers, for instance those on the same DSLAM. Understanding this will help an operator fix the fault more rapidly and/or allow the affected customers to be informed what's happening and/or request them to re-set their home hub (required to cure some conditions). More accurate information enables the operator to reassure customers and take more rapid and effective action to cure the problem.

There may also be problems unique to a single user line (e.g. copper access) that need to be identified.

Often customers experience poor broadband due to problems in the home network - the ISP's network is fine. For example they may have moved too far away from their wireless access point. Perhaps 80% of customer calls about fixed BB problems are due to in-home wireless issues. These issues are expensive and frustrating for an operator, as they are extremely hard to diagnose and solve. The operator would like to narrow down whether the problem is in the home (with the home network or edge device or home gateway), in the operator's network, or with an over-the-top service. The operator would like two capabilities. Firstly, self-help tools that customers use to improve their own service or understand its performance better, for example to re-position their devices for better wifi coverage. Secondly, on-demand tests that can the operator can run instantly - so the call centre person answering the phone (or e-chat) could trigger a test and get the result whilst the customer is still on-line session.

The characteristics of large scale measurements emerging from these examples:

1. A key challenge is how to integrate results from measurements into the operator's existing Test and Diagnostics system.

2. Results from the tests shouldn't be averaged

3. Tests are generally run on an ad hoc basis, ie specific requests for immediate action

4. "End-to-middle" measurements, ie across a specific network segment, are very relevant

5. The primary interest is in measuring specific network performance parameters and not QoE

6. New tests are needed for example to check the home network (ie the connection from the home hub to the set top boxes or to a tablets on wifi)

7. Active measurements are critical. Passive ones may be useful to help understand exactly what the customer is experiencing.

3.6 Comparison with the regulator use case

Today an increasing number of regulators measure the performance of broadband operators. Typically they deploy a few 1000 probes, each of which is connected directly to the broadband customer's home gateway and periodically measures the performance of that line. The regulator ensures they have a set of probes that covers the different ISPs and their different technology types and contract speeds, so that they can publish statistically-reasonable average performances. Publicising the results stimulates competition and so pressurises ISPs to improve broadband service.

The operator use case has similarities but several significant differences from the regulator one:

- Performance metrics: A regulator and operator are generally interested in the same performance metrics. Both would like standardised metrics, though this is more important for regulators.

- Sampling: The regulator wants an average across a representative sample of broadband customers (per operator, per type of BB contract). The operator also wants to measure individual lines with a problem.

- Timeliness: The regulator wants to know the (averaged) performance last quarter (say). For fault identification and fixing, the operator would like to know the performance at this moment and also to instruct a test to be run at this moment (so the requirement is on both the testing and reporting). Also, when testing the impact of new devices and technology, the operator is gaining insight about future performance.

- Scheduling: The regulator wants to run scheduled tests ('measure download rate every hour'). The operator also wants to run one-off tests; perhaps also the result of one test would trigger the operator to run a specific follow-up test.

- Pre-processing: A regulator would like standard ways of processing the collected data, to remove outlier measurements and aggregate results, because this can significantly affect the final

"averaged" result. Pre-processing is not important for an operator.

o Historic data: The regulator wants to track how the (averaged) performance of each operator changes on (say) a quarterly basis. The operator would like detailed, recent historic data (eg a customer with an intermittent fault over the last week).

o Scope: To date, regulators have measured the performance of access lines. An operator also wants to understand the performance of the home (or enterprise) network and of the end-to-end service, ie including backbone, core, peering and transit, CDNs and application /content servers.

o Control of testing and reporting: The operator wants detailed control. The regulator contracts out the measurement caboodle and 'control' will be via negotiation with its contractor.

o Politics: A regulator has to take account of government targets (eg UK government: "Our ambition (by 2015) is to provide superfast broadband (24Mbps) to at least 90 per cent of premises in the UK and to provide universal access to standard broadband with a speed of at least 2Mbps.") This may affect the metrics the regulator wants to measure and certainly affects how they interpret results. The operator is more focused on winning market share.

3.7 Conclusions

There is a clear need from an ISP point of view to deploy a single coherent measurement capability across a wide number of heterogeneous devices both in their own networks and in the home environment. These tests need to be able to operate from a wide number of locations to a set of interoperable test points in their own network as well as spanning supplier and competitor networks.

Regardless of the tests being operated, there needs to be a way to demand or schedule the tests and critically ensure that such tests do not affect each other; are not affected by user traffic (unless desired) and do not affect the user experience. In addition there needs to be a common way to collect and understand the results of such tests across different devices to enable correlation and comparison between any network or service parameters.

Since network and service performance needs to be understood and analysed in the presence of topology, line, product or contract information it is critical that the test points are accurately defined and authenticated.

Finally the test data, along with any associated network, product or contract data is commercial or private information and needs to be protected.

4 Security Considerations

The transport of Controller to MA and MA to Collector traffic must be protected both in-flight and such that each entity is known and trusted to each other.

It is imperative that end user identifying data is protected. Identifying data includes, end user name, time and location of the MA, and any attributes about a service such as service location, including IP address that could be used to re-construct physical location.

5 IANA Considerations

TBD

Appendix A. End User Use Case

End users may want to determine whether their network is performing according to the specifications (e.g., service level agreements) offered by their Internet service provider, or they may want to diagnose whether components of their network path are impaired. End users may perform measurements on their own, using the measurement infrastructure they provide or infrastructure offered by a third party, or they may work directly with their network or application provider to diagnose a specific performance problem. Depending on the circumstances, measurements may occur at specific pre-defined intervals, or may be triggered manually. A system administrator may perform such measurements on behalf of the user. Example use cases of end user initiated performance measurements include:

o An end user may wish to perform diagnostics prior to calling their ISP to report a problem. Hence, the end user could connect a MA to different points of their home network and trigger manual tests. Different attachment points could include their in-home 802.11 network or an Ethernet port on the back of their BB modem.

o An OTT or ISP service provider may deploy a MA within an their service platform to provide the end user a capability to diagnose service issues. For instance a video streaming service may include a manually initiated MA within their platform that has the Controller and Collector predefined. The end user could initiate performance tests manually, with results forwarded to both the

provider and the end user via other means, like UI, email, etc.

Contributors

The information in this document is partially derived from text written by the following contributors:

James Miller jamesmilleresquire@gmail.com

Rachel Huang rachel.huang@huawei.com

Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[LMAP-REQ] Schulzrinne, H., "Large-Scale Measurement of Broadband Performance: Use Cases, Architecture and Protocol Requirements", draft-schulzrinne-lmap-requirements, September, 2012

[IETF85 Plenary] Crawford, S., "Large-Scale Active Measurement of Broadband Networks", http://www.ietf.org/proceedings/85/slides/slides-85-iesg-opsandtech-7.pdf 'example' from slide 18

[Extend TCP] Michio Honda, Yoshifumi Nishida, Costin Raiciu, Adam Greenhalgh, Mark Handley and Hideyuki Tokuda. "Is it Still Possible to Extend TCP?" Proc. ACM Internet Measurement Conference (IMC), November 2011, Berlin, Germany. http://www.ietf.org/proceedings/82/slides/IRTF-1.pdf

Authors' Addresses

Marc Linsner
Marco Island, FL
USA

EMail: mlinsner@cisco.com

Philip Eardley
BT

B54 Room 77, Adastral Park, Martlesham
Ipswich, IP5 3RE
UK

Email: philip.eardley@bt.com

Trevor Burbridge
BT
B54 Room 77, Adastral Park, Martlesham
Ipswich, IP5 3RE
UK

Email: trevor.burbridge@bt.com